

POSTĘPY FIZYKI

DWUMIESIĘCZNIK
POŚWIĘCONY UPOWSZECHNIANIU
WIEDZY FIZYCZNEJ



TOM XVII

ZESZYT 5

1966

RADA REDAKCYJNA

Przewodniczący — Szczepan Szczęniowski, czł. koresp. PAN (Warszawa), *Członkowie* — Władysław Kapuściński (Warszawa), Karol Majewski (Warszawa), Ludwik Natanson (Warszawa), Henryk Niewodniczański, czł. rzecz. PAN (Kraków), Wojciech Rubinowicz, czł. rzecz. PAN (Warszawa), Leonard Sosnowski, czł. koresp. PAN (Warszawa), Jan Weyssenhoff, czł. rzecz. PAN (Kraków)

KOMITET REDAKCYJNY

Redaktor Naczelny — Jan Weyssenhoff, czł. rzecz. PAN (Kraków)
Członkowie Redakcji — Jan Babecki (Kraków), Antoni Bajorek (Kraków),
Andrzej Kisiel (Kraków)

Adres Redakcji: Kraków, Reymonta 4
tel. 363-80, wewn. 442

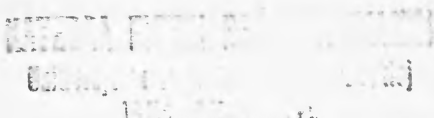
Maszynopisy prac należy nadsyłać w 2 egzemplarzach pod adresem Redakcji. O przyjęciu do druku decyduje Komitet Redakcyjny.

P O L S K I E T O W A R Z Y S T W O F I Z Y C Z N E

POSTĘPY FIZYKI

DWUMIESIĘCZNIK POŚWIĘCONY UPOWSZECHNIANIU
WIEDZY FIZYCZNEJ

TOM XVII, ZESZYT 5


Lublin, ul. Nowotki Nr 8

P A Ń S T W O W E W Y D A W N I C T W O N A U K O W E
1966

PAŃSTWOWE WYDAWNICTWO NAUKOWE
Oddział Kraków, ul. Smoleńsk 14

Nakład 1975 + 145 egz.	Rękopis dostarczono 21. VI. 1966 r.
Ark. wyd. 9,5 ark. druk. 8 $\frac{1}{2}$.	Podpisano do druku 20. X. 1966 r.
Pap. rotogr. 80 g. kl. III. 70×100	Druk ukończono w październiku 1966 r.
Zamówienie 510/66	Cena zł 15.— T-14

DRUKARNIA UNIwersytetu Jagiellońskiego, Kraków, Czapskich 4

H. Alfvén

Królewski Instytut Technologiczny
Sztokholm, Szwecja

Antymateria i ewolucja metagalaktyki *

I. Problem kosmologiczny

1. Wstęp

Okolo czterdzieści lat temu okazało się, że mgławice spiralne są galaktykami, nie różniącymi się w zasadzie ani rozmiarami, ani budową od naszej Galaktyki. Następnie stwierdzono, że linie widmowe odległych galaktyk przesunięte są w kierunku czerwieni. Jeśli zjawisko to interpretować jako efekt Dopplera, wskazuje ono na oddalanie się od nas tych galaktyk z prędkością proporcjonalną do ich odległości (prawo Hubble'a). Wszystkie dotychczas zaobserwowane galaktyki (i całe mnóstwo galaktyk jeszcze nie odkrytych) uważamy za część olbrzymiego układu zwanego układem metagalaktycznym.

W modelowym przedstawieniu ewolucji i budowy układu metagalaktycznego dały się zauważyć w zasadzie dwa odmienne podejścia (patrz np. McVittie (1961)). W teorii „rozszerzającego się wszechświata“ układ metagalaktyczny zostaje utożsamiony z całym wszechświatem. W pewnej chwili początkowej wszechświat miał „powstać“ w postaci bryły gęstej materii, która uległa eksplozji, a galaktyki zostały emitowane we wszystkich kierunkach. W teorii „stacjonarnego wszechświata“, na odwrót, przyjmuje się, że wszechświat winien być jednorodny¹ i znajduje się w stanie stacjonarnym. W teorii tej zakłada się, że materia tworzona jest w sposób ciągły.

W sytuacji, kiedy przeciw obu przedstawionym modelom wysunięto poważne zastrzeżenia, alternatywne podejście zaproponował O. Klein (1953, 1958, 1961, 1962). Jako „stan początkowy“ przyjął on nadzwyczaj rozrzedzoną

* Tłumaczenie artykułu zamieszczonego w czasopiśmie *Reviews of Modern Physics*, 17 652 (1965), dokonane za zgodą Wydawcy.

¹ Teoria ta opiera się na tzw. doskonałej zasadzie kosmologicznej, przyjmującej, że poznanie wszechświata nie zależy ani od miejsca, ani od czasu obserwacji. W każdej chwili stwierdzamy ten sam jednorodny i izotropowy rozkład galaktyk. Patrz też H. Bondi, *Kosmologia*, PWN 1965 (przyp. tłum.).

chmurę złożoną z protonów i elektronów. Część jej w wyniku działania sił grawitacyjnych uległa zagęszczeniu i utworzyła układ metagalaktyczny. Teoria ta, w której nie wprowadzono żadnych nowych praw fizycznych, przyniosła nową interpretację relacji Eddingtona pomiędzy wielkościami atomowymi i kosmologicznymi. Jako niezbędną konsekwencję podstawowych założeń przyjęto później, że owa chmura początkowa stanowiła mieszaninę cząstek i antycząstek w jednakowych ilościach. Argumentów na rzecz takiej symetrii dostarczyło wielu fizyków (wśród nich O. Klein) bezpośrednio po odkryciu pozytonu przed więcej niż trzydziestu laty, podniosła zaś ich wagę fizyka cząstek elementarnych, której bieżące tendencje wskazują na jak najpełniejszą symetrię między cząstkami i antycząstkami. Wydaje się logicznie niezadowalające, by teorie kosmologiczne opierały się jedynie na przyjęciu, że wszechświat zawiera wyłącznie materię, a nie antimaterię.

Podejście Kleina wymaga istnienia skutecznego mechanizmu rozdzielania materii od antimaterii (gdyż fluktuacje statystyczne nie wystarczają w tym celu, jak wskazali Goldhaber (1956) oraz Alpher i Herman (1958)). Wynika z niego poza tym, że na obecnym etapie rozwoju metagalaktyki istnieć winny obszary zawierające materię, jak również inne obszary wypełnione antimaterią, rozdzielone zaś winny być one cienkimi strefami granicznymi.

2. Właściwości antimaterii

Dirakowska teoria elektronu i odkrycie pozytonu legły u podstaw powszechnego przekonania, że do każdej cząstki istnieje odpowiednia antycząstka. Przekonanie to zostało potwierdzone odkryciem antyprotonu. Wydaje się, że również inne cząstki elementarne posiadają odpowiednie antycząstki. Można stąd wywnioskować, że mogą także istnieć „antyatomy“, o tej samej w zasadzie strukturze co zwykle atomy pierwiastków, lecz o jądrach złożonych z antyprotonów i antyneutronów, otoczonych przez pozytony. Właściwości takich antyatomów byłyby takie same jak właściwości atomów „normalnych“. Mogłyby one np. wchodzić w skład związków chemicznych, analogicznych do zwykłych związków chemicznych, jak również wysyłałyby linie widmowe o dokładnie tych samych długościach fal co atomy zwykle. W polach magnetycznych, względnie elektrycznych, powstawałyby zjawiska Zeemana względnie Starka, identyczne z zachodzącymi w normalnej materii przy odwróceniu kierunku pól.

Wynika stąd, że nie jesteśmy w stanie *a priori* wykluczyć możliwości istnienia ciał niebieskich zbudowanych z antimaterii. Badając jedynie światło emitowane przez określone ciało niebieskie nie możemy rozstrzygnąć, czy ciało to składa się z materii czy antimaterii (w przypadku, gdy nie znamy kierunku pola magnetycznego). Podobnie i własności jądrowe materii nie różnią się niczym od analogicznych własności antimaterii; i tak w gwiazdzie zbudowanej z antimaterii energia wydzielana będzie również kosztem reakcji syntezy termonuklearnej.

Jeśli cząstka trafi w antycząstkę, wtedy może nastąpić anihilacja. Elektron i pozyton anihilują się wzajemnie, czemu towarzyszy emisja dwóch (zazwyczaj) fotonów o energii 0,5 MeV każdy. Proton trafiając w antyproton wytwarza pewną liczbę mezonów, które gwałtownie rozpadają się na elektrony, pozytony, fotony i neutrina, tak że w efekcie po kilku mikrosekundach otrzymujemy kilka pozytonów i elektronów o energiach rzędu 10^8 eV, oraz kilka fotonów i neutrin o energiach tego samego rzędu. Jeśli elektrony zbliżą się do pozytonów, może nastąpić dalsza ich anihilacja.

Obok zachowania się w polu elektrycznym lub magnetycznym o *znanym* kierunku, proces anihilacji stanowi jedyny sposób rozróżnienia pomiędzy materią a antymaterią. Jeśli więc chcemy się upewnić, czy pewne ciało niebieskie składa się z materii czy też z antymaterii, winniśmy albo zbadać np. efekt Zeemana dla linii widmowej, gdy wysyłający ją atom znajduje się w polu magnetycznym o znanym kierunku, bądź też stwierdzić, co się stanie, gdy w ciało to trafi bryła materii (względnie antymaterii).

Zatem w wyniku obserwacji *bezpośrednich* nie możemy na obecnym etapie rozstrzygnąć, czy oddalone ciało niebieskie składa się z materii, czy też z antymaterii i nie da się wykluczyć możliwości, że np. połowa ciał niebieskich we wszechświecie składa się z antymaterii. (Argumenty pośrednie za i przeciw istnieniu antymaterii są rozpatrzone w rozdziale II.6 i IV.)

3. Fizyka i kosmologia

Normalna fizyka opiera się w zasadzie na obserwacjach laboratoryjnych. W wielu przypadkach prawa wydedukowane teoretycznie z doświadczeń laboratoryjnych zostały następnie potwierdzone w obserwacjach astronomicznych. Nie ma takiej obserwacji astronomicznej, która by wyraźnie wymagała wprowadzenia nowych praw fizycznych. Nie ma ani jednego zjawiska astronomicznego, które by wyraźnie nie zgadzało się z teorią kwantową czy też z teorią względności. Nie oznacza to oczywiście, że możemy być pewni, iż zwykłe prawa przyrody (które oczywiście obejmują i ogólną teorię względności) zachowują swoją słusność dla tak dużych rozmiarów liniowych i przedziałów czasu, z jakimi mamy do czynienia w teoriach kosmologicznych. Gdy dochodzimy do takich zagadnień, jak możliwe ograniczenia w przestrzeni i czasie naszego wszechświata, nie można się dziwić ewentualnej konieczności wprowadzenia nowych praw fizycznych. Raczej niechętnie przyjmujemy jednak nowe prawa fizyczne, zanim nie upewnimy się, że zaobserwowanego zjawiska nie da się objaśnić przy użyciu praw, które już w wystarczający sposób sprawdzono na innych drogach.

Oznacza to, że kosmologię, która nie wprowadza żadnych nowych praw przyrody, winniśmy stawiać wyżej od takiej, która opiera się na pewnych założeniach *ad hoc*. Jako ogólną zasadę przyjąć należy, że założenia *ad hoc* należy unikać, dopóki nie okaże się, że wszystkie inne sposoby tłumaczenia są wykluczone. Z tego punktu widzenia nie widać żadnego racjonalnego powodu dla przyjęcia hipotezy "ciągłego powstawania materii".

Istnienia antymaterii we wszechświecie nie można uznać za założenie *ad hoc*. Doskonała symetria między cząstkami i antycząstkami jest ważnym wynikiem fizyki cząstek elementarnych, a jeśli ma istnieć logiczne powiązanie pomiędzy kosmologią a podstawami fizyki, niepodobna nie przyjąć choćby w skali kosmicznej symetrii pomiędzy materią a antymaterią. To właśnie przyjmowane zazwyczaj założenie, że wszechświat zbudowany jest z materii i nie zawiera antymaterii, zaklasyfikować należy do założeń *ad hoc*. Głównym celem niniejszego artykułu jest wykazanie, że brak jakiegokolwiek rozstrzygającego poparcia na rzecz tej rozpowszechnionej tezy.

W tym miejscu należy dodać, że nawet jeśli nie wprowadza się nowych praw fizycznych, kosmologia nie może uniknąć pewnych założeń o „początkowym stanie” wszechświata. Korzystnie jest oczywiście, jeśli założenia o stanie początkowym są możliwie proste.

4. Model metagalaktyki Kleina

W modelu Kleina przyjmuje się za „stan początkowy” bardzo rzadką plazmę, stanowiącą jednorodną mieszaninę materii i antymaterii. Gęstość jest nadzwyczaj mała, o wiele mniejsza od obecnej średniej gęstości w układzie metagalaktycznym. Stan początkowy jest więc bardzo prosty, a model wykazuje idealną symetrię względem materii i antymaterii. Ze względu na małą gęstość szybkość anihilacji jest nader powolna. Plazma zajmuje objętość znacznie większą od obecnych rozmiarów układu metagalaktycznego. Pod wpływem niestabilności grawitacyjnej pewna część plazmy zaczyna się zagęszczać. Wydzieleniu może ulec przy tym duża ilość początkowej energii grawitacyjnej. Dla prostoty można przyjąć, że zagęszczająca się plazma ma kształt kulisty. W dalszym ciągu bada się ewolucję takiej kuli pod działaniem sił grawitacyjnych, przy czym zakłada się słuszność jedynie zwykłych praw fizyki (z włączeniem, oczywiście, teorii względności). Okazuje się, że ewolucja zagęszczającej się kuli prowadzi do stanu zbliżonego do obecnego stanu układu metagalaktycznego.

W zasadzie można sobie wyobrazić większą liczbę zagęszczeń w plazmie początkowej, co prowadzi do wytworzenia innych układów metagalaktycznych we wszechświecie. W rozdziale V omówimy części wszechświata znajdujące się na zewnątrz naszego układu metagalaktycznego.

II. Właściwości ambiplazmy

1. Wstęp

Kleinowski model metagalaktyki zwrócił uwagę na znaczenie badania właściwości plazmy stanowiącej mieszaninę materii i antymaterii. Plazmę taką nazwiemy *ambiplazmą* (od słowa ambi = oba). Pewne jej właściwości omówimy w następujących punktach tego rozdziału.

Ambiplazma może w ogólności stanowić mieszaninę wszystkich pierwiastków i antypierwiastków. Dla uproszczenia obrazu ograniczymy się do przypadku, gdy istnieją tylko cztery składowe n_i^\pm o stężeniach liczonych na jednostkę objętości: n_p^+ protonów, n_p^- antyprotonów, n_e^+ pozytonów i n_e^- elektronów. Podobnie jak wszelkie rodzaje plazmy kosmicznej plazma ta poddana jest działaniu pola magnetycznego (pole = B), oraz grawitacyjnego g . Warunek quasi-neutralności brzmi:

$$n_p^+ + n_e^+ = n_p^- + n_e^- . \quad (\text{II.1})$$

Trzema ważnymi, specjalnymi przypadkami są:

Materia:

$$n_p^- = 0, \quad n_e^+ = 0, \quad n_p^+ = n_e^- ; \quad (\text{II.2})$$

Antymateria:

$$n_p^+ = 0, \quad n_e^- = 0, \quad n_p^- = n_e^+ ; \quad (\text{II.3})$$

„Symetryczna ambiplazma“:

$$n_p^+ = n_p^-, \quad n_e^+ = n_e^- . \quad (\text{II.4})$$

W omawianym modelu metagalaktyka składa się początkowo z symetrycznej ambiplazmy. Jednym z głównych zadań jest zbadanie, jak można ją rozdzielić na materię i antymaterię. Ważne jest również zbadanie, jakie rodzaje promieniowania może ona wysyłać.

2. Procesy anihilacyjne

Proton i antyproton mogą anihilować się wzajemnie. Produktami anihilacji będą mezony, neutrino i wysokoenergetyczne kwanty gamma. Po rozpadach mezonów π i μ powstaną średnio po dwa elektrony i dwa pozytony w jednym procesie anihilacyjnym, a energia ich będzie około 100 MeV.

Wprowadźmy „klasyczny promień elektronu“:

$$d = e^2(mc^2)^{-1} = 2,8 \cdot 10^{-13} \text{ cm}$$

i

$$\sigma_0 = \pi d^2 = 2,5 \cdot 10^{-25} \text{ cm}^2. \quad (\text{II.5})$$

Średnie czasy życia T_p^+ i T_p^- dla protonu i antyprotonu można teraz zapisać:

$$T_p^+ = (n_p^- c \sigma_0)^{-1} k_p \quad (\text{II.6})$$

$$T_p^- = (n_p^+ c \sigma_0)^{-1} k_p . \quad (\text{II.7})$$

Podobnie jest dla elektronu i pozytonu:

$$T_e^- = (n_e^+ c \sigma_0)^{-1} k_e , \quad (\text{II.8})$$

$$T_e^+ = (n_e^- c \sigma_0)^{-1} k_e . \quad (\text{II.9})$$

Mamy

$$c \sigma_0 = 0,75 \cdot 10^{-14} \text{ cm}^3 \text{ sek}^{-1}. \quad (\text{II.10})$$

W powyższych wzorach k_p jest to funkcja, która jest rzędu jedności aż do energii relatywistycznych. Wielkość k_e osiąga jedność przy energiach nie re-

latywistycznych, a wzrasta logarytmicznie dla energii relatywistycznych. Wartości

$$T_0 = (nc\sigma_0)^{-1} \quad (\text{II.11})$$

podajemy w tabelicy I. Z uwagi na to, że k_p i k_e są rzędu jedności, wartość T_0 określa rząd wielkości czasu życia. Dane w tabelicy I wskazują na to, że jeśli mamy jednorodną mieszaninę materii i antymaterii przy dzisiejszej gęstości średniej dla metagalaktyki ($\sim 10^{-6} \text{ cm}^{-3}$), to anihilacja zachodzi bardzo wolno ($T_0 = 4,3 \cdot 10^{12}$ lat). Jeśli natomiast gęstość wzrośnie do 10^{-3} cm^{-3} , wtedy czas życia zmaleje do wielkości rzędu miliardów lat, co jest porównywalne ze zwykłą skalą czasu w teoriach kosmologicznych. Dla ośrodka międzygwiazdowego (wewnątrz galaktyki) przyjmujemy gęstość 1 cząstki/cm³, co oznacza, że antycząstka ulegnie tam rozpadowi po około 5 milionach lat.

Tabela I

Czasy życia ze względu na anihilację

n	10^{-6}	10^{-4}	10^{-2}	1 cząstka w cm^3
T_0	$4,3 \cdot 10^{12}$	$4,3 \cdot 10^{10}$	$4,3 \cdot 10^8$	$4,3 \cdot 10^6$ lat

3. Promieniowanie wysyłane przez ambiplazmę

Elektrony i pozytony, powstałe w wyniku anihilacji proton-antyproton, posiadają energie rzędu 10^8 eV. W znanej nam plazmie kosmicznej występuje pole magnetyczne i nie ma żadnego powodu, by sądzić, że w tym przypadku będzie inaczej. W polu magnetycznym B elektrony i pozytony poruszają się będą po torach spiralnych, wysyłając promieniowanie synchrotronowe. Czas zaniku wynosi:

$$T_s = \frac{5 \cdot 10^8}{B^2} \frac{1}{1 + (W/W_0)} \text{ sek} \quad (\text{II.12})$$

(patrz np. Alfvén i Fälthammar (1963), str. 69), gdzie W oznacza energię kinetyczną, a $W_0 = m_e c^2$. Jeśli $W = 100 \text{ MeV} = 200 W_0$, to dostajemy

$$T_s = 2,5 \cdot 10^6 B^{-2} \text{ sek} = 0,08 B^{-2} \text{ lat}. \quad (\text{II.13})$$

O ile $T_s \ll T_0$, to elektrony i pozytony zdążą wypromieniować w postaci promieniowania synchrotronowego większą część swojej energii przed anihilacją (patrz tabela II). Jeśli, z drugiej strony, $T_s \gg T_0$, wtedy emisja promieniowania synchrotronowego jest niewielka. Jeśli np. gęstość materii elektronowo-pozytonowej wynosi 10^{-2} cm^{-3} , wtedy warunek na to, by pokaźna część energii została wysłana w postaci promieniowania synchrotronowego, jest $B > 10^{-5} \text{ Gs}$.

latywistycznych, a wzrasta logarytmicznie dla energii relatywistycznych. Wartości

$$T_0 = (nc\sigma_0)^{-1} \quad (\text{II.11})$$

podajemy w tabelicy I. Z uwagi na to, że k_p i k_e są rzędu jedności, wartość T_0 określa rząd wielkości czasu życia. Dane w tabelicy I wskazują na to, że jeśli mamy jednorodną mieszaninę materii i antymaterii przy dzisiejszej gęstości średniej dla metagalaktyki ($\sim 10^{-6} \text{ cm}^{-3}$), to anihilacja zachodzi bardzo wolno ($T_0 = 4,3 \cdot 10^{12}$ lat). Jeśli natomiast gęstość wzrośnie do 10^{-3} cm^{-3} , wtedy czas życia zmaleje do wielkości rzędu miliardów lat, co jest porównywalne ze zwykłą skalą czasu w teoriach kosmologicznych. Dla ośrodka międzygwiazdowego (wewnątrz galaktyki) przyjmujemy gęstość 1 cząstki/cm³, co oznacza, że antycząstka ulegnie tam rozpadowi po około 5 milionach lat.

Tabela I

Czasy życia ze względu na anihilację

n	10^{-6}	10^{-4}	10^{-2}	1 cząstka w cm^3
T_0	$4,3 \cdot 10^{12}$	$4,3 \cdot 10^{10}$	$4,3 \cdot 10^8$	$4,3 \cdot 10^6$ lat

3. Promieniowanie wysyłane przez ambiplazmę

Elektrony i pozytony, powstałe w wyniku anihilacji proton-antyproton, posiadają energie rzędu 10^8 eV. W znanej nam plazmie kosmicznej występuje pole magnetyczne i nie ma żadnego powodu, by sądzić, że w tym przypadku będzie inaczej. W polu magnetycznym B elektrony i pozytony poruszają się będą po torach spiralnych, wysyłając promieniowanie synchrotronowe. Czas zaniku wynosi:

$$T_s = \frac{5 \cdot 10^8}{B^2} \frac{1}{1 + (W/W_0)} \text{ sek} \quad (\text{II.12})$$

(patrz np. Alfvén i Fälthammar (1963), str. 69), gdzie W oznacza energię kinetyczną, a $W_0 = m_e c^2$. Jeśli $W = 100 \text{ MeV} = 200 W_0$, to dostajemy

$$T_s = 2,5 \cdot 10^6 B^{-2} \text{ sek} = 0,08 B^{-2} \text{ lat.} \quad (\text{II.13})$$

O ile $T_s \ll T_0$, to elektrony i pozytony zdążą wypromieniować w postaci promieniowania synchrotronowego większą część swojej energii przed anihilacją (patrz tablica II). Jeśli, z drugiej strony, $T_s \gg T_0$, wtedy emisja promieniowania synchrotronowego jest niewielka. Jeśli np. gęstość materii elektronowo-pozytonowej wynosi 10^{-2} cm^{-3} , wtedy warunek na to, by pokaźna część energii została wysłana w postaci promieniowania synchrotronowego, jest $B > 10^{-5} \text{ Gs}$.

Tablica II

Częstość przy maksimum energetycznym widma i czas zaniku dla promieniowania synchrotronowego

B	10^{-6}	10^{-5}	10^{-4} Gs
T_s	$8 \cdot 10^{10}$	$8 \cdot 10^8$	$8 \cdot 10^6$ lat
ν_{\max}	10^5	10^6	10^7 Hz

Maksimum w widmie promieniowania synchrotronowego osiągnięte zostaje przy częstości:

$$\nu_{\max} = (eB/2\pi mc)(W/W_0)^2 = 3 \cdot 10^6 (W/W_0)^2 B. \quad (\text{II.14})$$

Przyjmując $W/W_0 = 200$ mamy:

$$\nu_{\max} = 10^{11} B \text{ sek}^{-1}. \quad (\text{II.15})$$

Wypada zauważyć, że jeśli cząstki spędzają większą część swego czasu w polu B_0 , a zaledwie 1% czasu przebywają w polu $B_1 = 10B_0$, wtedy połowa energii wypromieniowanej wysłana zostaje w tym silniejszym polu. Zarówno więc w obliczeniach czasu życia, jak i maksimum widma, posługiwać się należy raczej polem maksymalnym niż średnim.

W porównaniu z promieniowaniem synchrotronowym promieniowanie hamowania można zazwyczaj zaniedbać.

Jeśli zatem przyjmiemy

$$T_s \ll T_0 k_e, \quad (\text{II.16})$$

wtedy anihilacja w końcowym efekcie prowadzi do emisji czterech kwantów gamma o energii ~ 200 MeV, oraz do emisji około 300 MeV w postaci fal radiowych. Ponadto zachodzi emisja neutrin.

Na każdy kwant promieniowania gamma przypada energia promieniowania radiowego około 75 MeV, co odpowiada $1,2 \cdot 10^{-11}$ J.

Przy obecnej technice pomiarowej (patrz Kraushaar i współpr. (1965)) można wykryć strumień fotonów o energii 200 MeV, posiadający natężenie około 1 foton $\text{m}^{-2}\text{sek}^{-1} = (10^{-4}\text{cm}^{-2}\text{sek}^{-1}) = 3,2 \cdot 10^{-11} \text{ W m}^{-2}$. Z drugiej strony można wykryć strumień energii w postaci promieniowania częstości radiowych o natężeniu $10^{-18} \text{ W m}^{-2}$ (szerokość pasma 10^8 , czułość $10^{-26} \text{ W m}^{-2}\text{Hz}^{-1}$). (Patrz np. Ryle (1963)). Tak więc czułość przyrządów do wykrywania fal radiowych z ambiplazmy przewyższa więcej niż 10^7 razy czułość detektorów γ .

Kosmiczna ambiplazma może się zatem ujawnić jako źródło fal radiowych. Mierzalny strumień kwantów γ pojawiłby się dopiero wtedy, gdyby strumień energii w zakresie radiowym przekraczał $8 \cdot 10^{-11} \text{ W m}^{-2}$, albo przy szerokości pasma 10^8 Hz — $8 \cdot 10^{-19} \text{ W m}^{-2}\text{Hz}^{-1}$.

Podajmy dla porównania, że najsilniejsze gwiazdy radiowe wysyłają $3 \cdot 10^{-22} \text{ W m}^{-2}\text{Hz}^{-1}$. Gdyby gwiazda radiowa składała się z ambiplazmy, i gdyby zarówno kwanty γ , jak i fale radiowe nie były na swej drodze ku nam pochła-

niane, wtedy jeszcze natężenie promieniowania γ byłoby 10^3 razy za słabe, by można było je wykryć za pomocą stosowanej obecnie techniki pomiarowej.

Pozytony i elektrony również anihilują, emitując fotony o energii przewyższającej 0,5 MeV. Tych ostatnich jest mniej więcej tyle, co elektronów i pozytonów. Ponadto ambiplazma jest źródłem neutrin, których jednak nie da się wykryć². Ambiplazma może być wreszcie źródłem światła i wszelkich innych rodzajów promieniowania, charakterystycznych dla plazmy kosmicznej.

W nieskończonej warstwie płaskiej o grubości Δ energia całkowita przypadająca na 1 cm² wynosi:

$$W_M = 2m_p c^2 n_p \Delta. \quad (\text{II.17})$$

Energia powyższa zostaje wypromieniowana w ciągu czasu $T_p = (n_p c \sigma_0)^{-1} k_p$. Strumień energii Φ na jednostkę czasu wynosi:

$$\Phi = W_M / T_p = 2m_p c^3 \sigma_0 n_p^2 \Delta \cdot k_p^{-1} \quad (\text{II.18})$$

albo

$$\Phi = 2,2 \cdot 10^{-17} n_p^2 \Delta k_p^{-1} \text{erg cm}^{-2} \text{sek}^{-1}. \quad (\text{II.19})$$

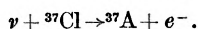
Pamiętajmy, że wszystko to, co wyprowadziliśmy, pozostaje słuszne jedynie wtedy, gdy $T_s \ll T_0 k_e$. Jeśli natomiast mamy $T_s \ll T_0 k_e$, to więcej energii wysyłane jest w postaci kwantów γ niż fal radiowych, ale emisję γ wcale nie jest łatwiej wykryć, dopóki nie będzie spełniony warunek $T_s > 10^7 T_0 k_e$.

4. Ambiplazma w polu grawitacyjnym

W rozdziale tym abstrahujemy od anihilacji i promieniowania, dążąc jedynie do zbadania plazmowych własności mieszaniny materii i antymaterii. Załóżmy, że w objętości jednostkowej mamy n_p^+ protonów, n_p^- antyprotonów, n_e^+ pozytonów i n_e^- elektronów. Gęstość liczbowa jest tak znikoma, że procesu anihilacji można nie brać pod uwagę. Zaniedbujemy również pole magnetyczne. Pole grawitacyjne g działa na ambiplazmę w kierunku $-z$. W ogólnym przypadku spowoduje ono spolaryzowanie ambiplazmy, tak że powstanie pole elektryczne E . Przyjmujemy, że wszystkie zmienne zależą wyłącznie od współrzędnej z . Dla każdej składowej plazmy otrzymujemy równanie:

$$\frac{\partial(nkT)}{\partial z} = (-g \cdot m \pm eE)n, \quad (\text{II.20})$$

² Zadanie to nie jest niewykonalne i w zasadzie można wykryć dość silny strumień neutrin kosmicznych posługując się olbrzymimi detektorami, zawierającymi CCl_4 lub C_2Cl_4 , w których pod wpływem neutrin kosmicznych z chloru powstaje promieniotwórczy argon:



W kwestii tej odsyłam zainteresowanych do artykułów Zacepina i Kuzmina (1964), Rudermana (1965) i Schwartz'a (1965), a w języku polskim — Jateczaka (1965) i Kuchowicza (1964, 1965). Bibliografia uzupełniająca znajduje się na końcu artykułu (przyp. tłum.).

gdzie k oznacza stałą Boltzmann, T — temperaturę, m — masę, e — ładunek cząstki. Ograniczmy się do przypadku plazmy izotermicznej, $T = \text{const}$. Kładąc $gm_p/kT = \mu_p$, $gm_e/kT = \mu_e$ i $eE/kT = F$ otrzymujemy:

$$(n_p^-)^{-1}(\partial n_p^-/\partial z) = -\mu_p - F, \quad (\text{II.21})$$

$$(n_p^+)^{-1}(\partial n_p^+/\partial z) = -\mu_p + F, \quad (\text{II.22})$$

$$(n_e^-)^{-1}(\partial n_e^-/\partial z) = -\mu_e - F, \quad (\text{II.23})$$

$$(n_e^+)^{-1}(\partial n_e^+/\partial z) = -\mu_e + F. \quad (\text{II.24})$$

Do tych czterech równań dochodzi jeszcze warunek quasi-neutralności:

$$n_p^+ + n_e^+ = n_p^- + n_e^-. \quad (\text{II.25})$$

Łatwo można otrzymać rozwiązanie powyższego układu równań:

$$n_p^- = \tau_p^{-1} \{v_p^- v_p^+ [1 + (v_e^+/v_p^+) \tau_p]\}^{1/2} [1 + (v_e^-/v_p^-) \tau_p]^{-1/2}, \quad (\text{II.26})$$

$$n_p^+ = \tau_p^{-1} \{v_p^+ v_p^- [1 + (v_e^-/v_p^-) \tau_p]\}^{1/2} [1 + (v_e^+/v_p^+) \tau_p]^{-1/2}, \quad (\text{II.27})$$

$$n_e^- = \tau_e^{-1} \{v_e^- v_e^+ [1 + (v_p^+/v_e^+) \tau_p^{-1}]\}^{1/2} [1 + (v_p^-/v_e^-) \tau_p^{-1}]^{-1/2}, \quad (\text{II.28})$$

$$n_e^+ = \tau_e^{-1} \{v_e^+ v_e^- [1 + (v_p^-/v_e^-) \tau_p^{-1}]\}^{1/2} [1 + (v_p^+/v_e^+) \tau_p^{-1}]^{-1/2}, \quad (\text{II.29})$$

gdzie $\tau_p = \exp(\mu_p z)$ i $\tau_e = \exp(\mu_e z)$, a v oznacza wartość n dla $z = 0$. Zanedbaliśmy tutaj μ_e w porównaniu z μ_p .

Jeśli nie ma wcale antymaterii ($v_p^- = v_e^+ = 0$), wtedy

$$n_p^+ = n_e^- = n \quad (\text{II.30})$$

1

$$n = v/\tau_p^{1/2} = v \exp(-\frac{1}{2}\mu_p z).$$

Wytworzone zostaje stałe pole elektryczne:

$$E = \frac{1}{2}(m_p g/e). \quad (\text{II.31})$$

Grawitacja nie prowadzi wcale do rozdzielenia składowej ciężkiej i lekkiej, wytwarza jedynie pole elektryczne (pole Rosselanda).

Podobnie jest w przypadku czystej antymaterii, kiedy to kierunek pola elektrycznego zostaje odwrócony.

Jeśli plazma zawiera jednakowe ilości materii i antymaterii („symetryczna ambiplazma“), wtedy mamy:

$$n_p^- = n_p^+ = v_p/\tau_p = v_p \exp(-\mu_p z), \quad (\text{II.32})$$

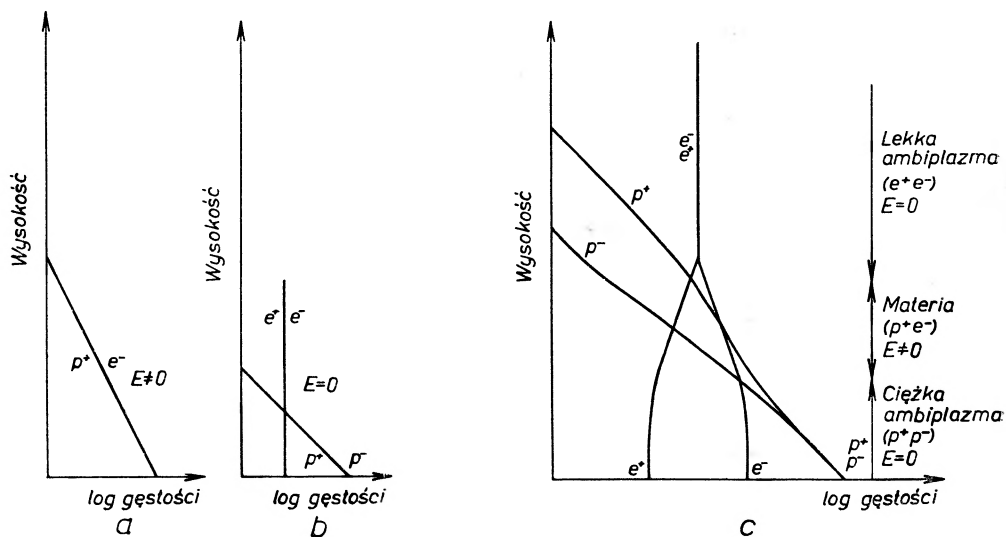
$$n_e^- = n_e^+ = v_e/\tau_e = v_e \exp(-\mu_e z). \quad (\text{II.33})$$

Pole elektryczne wynosi zero i nie ma żadnego związku pomiędzy składową ciężką i lekką. Wysokość skali dla gazu elektronowo-pozytonowego jest 1840 razy większa niż dla gazu protonowo-antyprotonowego. Całkowita liczba protonów-antyprotonów nie zależy od całkowitej liczby elektronów-pozytonów.

W asymetrycznej ambiplazmie, która np. zawiera więcej materii niż antymaterii, istnieją trzy typowe obszary. W obszarze najniższym i najwyższym

znajduje się symetryczna ambiplazma i brak pola elektrycznego, natomiast w obszarze pośrednim występuje pole Rosselanda i prawie nie ma antymaterii.

Prawie wszystkie antyprotony zebrane są w obszarze najniższym wraz z równą liczbą protonów. Przyjeliśmy na wstępie, że w całej objętości jest więcej materii niż antymaterii; nadmiar protonów zajmuje obszar pośredni wraz z elektronami. W górnym obszarze zgromadzone są praktycznie wszystkie pozytony wraz z równą liczbą elektronów.



Rys. 1. Ambiplazma w polu grawitacyjnym. (a) Wylącznie materia (albo też wyłącznie antymateria). (b) Symetryczna ambiplazma. (c) Asymetryczna ambiplazma

Jeśli zatem asymetryczną ambiplazmę (zawierającą, na przykład, więcej materii niż antymaterii) umieścimy w polu grawitacyjnym, ujawni się dążenie do rozszczepienia na trzy oddzielne części (rys. 1). W obszarze najniższym zgromadzi się symetryczna ambiplazma zawierająca cząstki ciężkie, w obszarze najwyższym — symetryczna ambiplazma złożona z cząstek lekkich. Oddzielone są one obszarem pośrednim, w którym skoncentrowany jest prawie cały nadmiar jednej składowej (w naszym przypadku — materii), (patrz tablica III).

Tablica III

Parametr μ wzrostu z wysokością z , określonego zależnością $n_0 \exp(\mu z)$. () oznacza nieduży ułamek procentowy; (()) oznacza bardzo mały ułamek procentowy

Obszar	Antyprotony	Protony	Elektrony	Pozytony	Pole elektryczne
Najniższy	$-\mu_p$	$-\mu_p$	$(-\mu_e)$	$((-\mu_e))$	0
Pośredni	$(-\frac{3}{2}\mu_p)$	$-\frac{1}{2}\mu_p$	$-\frac{1}{2}\mu_p$	$(+\frac{1}{2}\mu_p)$	$\frac{1}{2}m_p g/e$
Najwyższy	$((-\mu_p))$	$(-\mu_p)$	$-\mu_e$	$-\mu_e$	0

W zwykłej plazmie niełatwo jest rozdzielić jony dodatnie i elektrony. Przedstawiona analiza wykazała, że ambiplazma zachowuje się pod tym względem odmiennie. Pod działaniem sił grawitacyjnych składowa lekka wykazuje tendencję do oddzielenia od składowej ciężkiej. Ponadto, jeśli w ambiplazmie znajduje się nadmiar bądź materii, bądź też antymaterii, wykazuje on tendencję do oddzielenia się od reszty ambiplazmy. W rozważaniach powyższych nie braliśmy pod uwagę pól magnetycznych. Wiadomo, że pole magnetyczne nie zmienia stanu równowagi atmosfery izotermicznej.

5. Rozdzielenie materii i antymaterii

Sama grawitacja nie jest w stanie doprowadzić do rozdzielenia początkowo symetrycznej ambiplazmy (zawierającej jednakowe ilości materii i antymaterii) na materię i antymaterię. Jest to jednak w zasadzie możliwe przy włączeniu efektów elektromagnetycznych (Alfvén i Klein (1962)).

Aby wykazać to, przyjmijmy, że w rozważanej w poprzednim rozdziale symetrycznej ambiplazmie płynie prąd elektryczny w kierunku przeciwnym do kierunku g . Będzie on usuwał ładunki dodatnie z górnych warstw atmosfery, ponieważ zaś w najwyższym obszarze znajdują się głównie pozytony i elektrony, usuwane więc będą stamtąd przede wszystkim pozytony. W podobny sposób prąd ten będzie usuwał ładunki ujemne z obszaru najniższego, ponieważ zaś w obszarze tym znajdują się przede wszystkim cząstki ciężkie, efekt spowodzi się do usuwania antyprotonów. Działanie prądu sprowadza się zatem do usuwania antymaterii, pozostawiając asymetryczną ambiplazmę z nadmiarem materii, ta zaś będzie przemieszczać się w kierunku obszaru pośredniego.

W podobny sposób prąd płynący w kierunku zgodnym z kierunkiem g spowoduje powstanie nadmiaru antymaterii. Prąd płynący w kierunku pionowym może być wynikiem efektów hydromagnetycznych, spowodowanych ruchem namagnesowanej ambiplazmy podczas ewolucji metagalaktyki, może też wiązać się z powstawaniem galaktyk.

Można by wysunąć zastrzeżenia, że proces tego rodzaju sprzeczny jest z ogólnymi zasadami mechaniki statystycznej, gdyż prowadzi od stanu nie uporządkowanego (jednorodna mieszanina materii i antymaterii) do stanu o wyższym uporządkowaniu. Zastrzeżenie to nie jest słuszne, gdyż ostatecznym źródłem energii jest grawitacja całego układu. Weźmy dla porównania wodę spływającą z góry. Przy użyciu elektrowni wodnej można przekształcić jej energię grawitacyjną w elektryczną, zaś energię dielektryczną można zastosować do elektrolizy wody i rozłożenia jej na wodór i tlen. Proces rozważany przez nas jest podobny. Wydzielenie materii czy też antymaterii z symetrycznej ambiplazmy można uważać za wynik elektrolizy. Ilość M materii (bądź też antymaterii) wydzielonej przez prąd o natężeniu I płynący w ciągu czasu T dana jest wyrażeniem:

$$M = (M_H/e)IT \quad (\text{II.34})$$

w zgodności ze zwykłymi prawami elektrolizy.

Podobny proces opisany został we wcześniejszej pracy (Alfvén i Klein (1962)). Także w tym przypadku prąd płynie prostopadle do powierzchni rozdziału pomiędzy ciężką i lekką ambiplazmą. Na powierzchni rozdziału gromadzi się materia bądź też antymateria, zależnie od kierunku prądu.

Wiadomo, że w poruszającej się, namagnesowanej plazmie powstają prądy, procesy tego rodzaju winny więc być zjawiskiem normalnym. Jednakże całkowita ilość materii (czy też antymaterii), którą można w ten sposób rozdzielić, jest dla obwodu zamkniętego ograniczona. Weźmy prąd o natężeniu $I = \pi r^2 i$ płynący wzdłuż osi walca kołowego o promieniu r . Pole magnetyczne na powierzchni tego walca wynosi:

$$B = 2I/cr. \quad (\text{II.35})$$

Podstawiając to wyrażenie do wzoru (II.34) otrzymujemy:

$$M = m_H TcrB/2e = 0,5 \cdot 10^{-4} TrB. \quad (\text{II.36})$$

Przyjmuje się zazwyczaj, że pole magnetyczne w przestrzeni kosmicznej nie przekracza 10^{-5} czy też 10^{-4} Gs. Skoro omawiane przez nas zjawiska odbywają się w obszarach, w których gęstość energii jest pokaźna, pola magnetyczne mogą być także silne. Wartość T nie powinna przekraczać 10^{17} sek ($= 3 \cdot 10^9$ lat). Przyjmując $B < 10^{-4}$ Gs i $T < 10^{17}$ sek, otrzymujemy:

$$M < 0,5 \cdot 10^9 r \text{ g}. \quad (\text{II.37})$$

Rozdzielanie w skali galaktyk ($r = 10^{23}$ cm) nie wystarcza więc do wydzielenia nawet jednej masy słonecznej.

Proces ten jest o wiele bardziej wydajny w małej skali. Jeśli gęstość ośrodka wynosi nm_p , wtedy jego masa jest $nm_p r^3$. Do całkowitego rozdzielenia materii od antymaterii dojdzie zatem, jeśli

$$nm_p r^3 = 0,5 \cdot 10^{-4} TrB, \quad (\text{II.38})$$

względnie

$$r^2 = [(0,5 \cdot 10^{-4} TB)/m_p n] = 3 \cdot 10^{19} (BT/n). \quad (\text{II.39})$$

Czas T ograniczony jest anihilacją. Wprowadzając wielkość T_0 ze wzoru (II.11) otrzymujemy:

$$r = 0,63 \cdot 10^{17} B^{1/2} n^{-1}. \quad (\text{II.40})$$

Przyjmując $B < 10^{-4}$ i $n = 10^{-2}$ dostajemy: $r < 10^{17}$ cm i $M < m_p n (10^{17})^3 < 10^{25}$ g.

Widać stąd, że turbulencja „w małej skali“, w obszarach wielkości jednej dziesiątej roku świetlnego czy nawet mniejszych, doprowadzić może do wysokiego stopnia rozdzielenia materii od antymaterii. W każdym takim elementarnym obszarze masa wydzielona nie musi być wielka, jednak całkowita jej ilość może stanowić znaczny ułamek całej masy ambiplazmy. Otwartym problemem pozostaje jeszcze możliwość rozdzielenia na skalę galaktyczną.

6. O współistnieniu materii i antymaterii

Średnią gęstość w metagalaktyce szacuje się na 10^{-6} cząstek w cm^3 . Zgodnie z danymi w tabeli I jednorodna ambiplazma o takiej gęstości posiada bardzo duży czas życia. Gęstość materii w galaktyce jest rzędu 1 cząstki w cm^3 . Oznacza to, że gdyby nasza galaktyka zbudowana była z jednorodnej ambiplazmy, dawno już uległaby anihilacji. Antymateria w galaktyce istnieć może tylko wtedy, gdy jest oddzielona od materii. Wypada zatem omówić procesy na powierzchni rozdziału pomiędzy materią i antymaterią.

Jeśli wylewamy wodę na gorącą blachę kuchenną o temperaturze nieco powyżej 100°C , ulega ona szybkiemu wyparowaniu. Jeśli jednak płyta kuchenna jest silnie rozżarzona, wtedy woda może pozostać na niej w ciągu pięciu do dziesięciu minut (zwłaszcza, jeśli płyta jest lekko wklęsła). W wyniku parowania wody powstaje cienka warstwa pary, która izoluje kroplę od gorącego podłoża. W podręcznikach fizyki zjawisko to nosi nazwę „zjawiska Leidenfrosta”. Odgrywa ono rolę w technice kotłów wysokotemperaturowych.

Jeśli materia zetknie się z antymaterią, możliwe jest podobne zjawisko. Procesy anihilacyjne na powierzchni rozdziału pomiędzy materią i antymaterią wytwarzają bardzo gorącą warstwę, która oddziela materię od antymaterii w ten sam sposób, w jaki para wodna w warstwie Leidenfrosta oddziela wodę od rozżarzonej płyty.

Rozważmy prosty przypadek, gdy w jednorodnym polu magnetycznym B skierowanym wzdłuż osi z znajduje się materia na lewo od płaszczyzny $x = -\Delta/2$, a antymateria na prawo od płaszczyzny $x = +\Delta/2$. Warstwa pomiędzy $-\Delta/2$ i $\Delta/2$ zawiera ambiplazmę. Wszystkie trzy środowiska znajdują się w stanie równowagi hydrostatycznej, tak więc ciśnienia gazu w nich są jednakowe. Z uwagi na to, że ciśnienie może być przenoszone pomiędzy ośrodkami za pośrednictwem sił elektromagnetycznych, zderzenia pomiędzy cząstkami nie odgrywają istotnej roli.

W ambiplazmie, w której zachodzi anihilacja, o ciśnieniu gazu decydują w głównej mierze elektrony i pozytony o energiach rzędu 10^8 eV, powstające w anihilacji p_+p_- . Jeśli gęstość liczbowa tych cząstek o energii 10^8 eV oznaczymy literą n , a ciśnienie gazu w obszarach zawierających materię względnie antymaterię będzie $p = n_0 eV_0$, wtedy:

$$n_e = n_0 V_0 / V, \quad (\text{II.41})$$

gdzie $eV = 10^8$ eV = $1,6 \cdot 10^{-7}$ erg. Jeżeli przyjmiemy wartości typowe dla przestrzeni międzygwiazdnej: $n_0 = 1$ cząstka w cm^3 i $eV_0 = 1$ eV, dostaniemy $n_e = 10^{-8}$ cm^{-3} . Grubość warstwy Leidenfrosta musi przynajmniej kilka, powiedzmy ν razy, przekraczać promień krzywizny ρ toru elektronu o energii 10^8 eV w polu magnetycznym B . Stąd otrzymujemy $\Delta \geq \nu \rho = \nu V / B$. Całkowita liczba elektronów na jednostkę powierzchni wynosi $N = \Delta \cdot n_e = 10^{-8} \nu V / B$. Biorąc na przykład $\nu = 10$ i $V = 10^8$ eV = $0,33 \cdot 10^6$ j.e.s. oraz $B = 10^{-5}$ Gs otrzymujemy $N = 3000$ cząstek na cm^2 . Ponieważ zgodnie z tabelą II cząstki

ulegają odnowieniu raz na $8 \cdot 10^8$ lat w wyniku nowej anihilacji, wysłany musi być jeden kwant γ na $8 \cdot 10^8 / 3000 \approx 0,3 \cdot 10^6$ lat. Promieniowanie gamma z warstwy Leidenfrosta o przyjętych wyżej własnościach jest zatem do zaniedbania. Natężenie emisji radiowej wynosi natomiast $3 \cdot 10^3 \cdot 10^8 \cdot 1,6 \cdot 10^{-12} \cdot (2,5 \cdot 10^{18})^{-1} = 2 \cdot 10^{-17}$ erg cm^{-2} sek^{-1} . Dla porównania podajmy, że strumień fal radiowych da się wykryć, jeśli przekracza on wartość 10^{-18} W $m^{-2} = 10^{-15}$ erg sek^{-1} cm^{-2} (patrz rozdz. II.3). Tak więc emisja radiowa warstwy pośredniej jest też do zaniedbania.

Nasz prosty wywód wykazał możliwość występowania obszarów, w których oddzielnie występowałyby materia i antymateria. Byłyby one rozdzielone cienkimi warstwami równoległymi do kierunku pola magnetycznego. W warstwach tych występowałyby bardzo gorąca ambiplazma, której natężenie promieniowania byłoby jednak zbyt słabe, by pozwolić na jej wykrycie. Kwestią otwartą pozostaje jeszcze stabilność takiej konfiguracji. Ponadto w kierunku równoległym do pola magnetycznego warstwa dzieląca nie może zachowywać tej samej struktury. Geometria obszaru dzielącego materię od antymaterii musi prawdopodobnie być ściśle powiązana z kształtem pola magnetycznego.

Anihilacja w warstwie Leidenfrosta prowadzi do pozornego odpychania pomiędzy materią i antymaterią. Przyjmijmy, że procesy rozdzielania na małą skalę (takie, jakie omówiliśmy w rozdz. II.5) doprowadziły do wytworzenia szeregu małych obszarów zajętych przez materię i przemieszanych z podobnymi obszarami, w których znajduje się antymateria. Wtedy dochodzi do odpychania pomiędzy obszarami materii i antymaterii, odpychanie to nie występuje jednak pomiędzy obszarami, w których znajduje się materia tego samego rodzaju. Możliwe, że w ten sposób nastąpi łączenie niewielkich obszarów materii i antymaterii w coraz większe. W taki sposób rozdzielanie w małej skali prowadzić może w końcowym efekcie do oddzielenia dużych ilości czystej materii i antymaterii? Dokładniejsze zbadanie tego typu procesów posiada podstawowe znaczenie dla zrozumienia rozdzielania materii od antymaterii. Należyta analiza teoretyczna tego procesu wydaje się jednak trudna, podjęcie zaś obecnie badań doświadczalnych wydaje się niemożliwe.

7. Optyczne metody wykrycia ambiplazmy (Uzupełnienie tłumacza)

Wydaje się, że w tym miejscu warto szerzej omówić ciekawy pomysł Własowa (1965), który zwrócił uwagę na możliwość zastosowania zwykłej spektroskopii optycznej do wykrycia obszarów, zajętych przez ambiplazmę. Alfvén wspomina o tej pracy w przypisku dodanym w korekcie, my natomiast możemy w tym miejscu przytoczyć szerzej pewne sugestie Własowa.

Od szeregu lat znany jest fizykom układ zwany pozytonium, złożony z elektronu i pozytonu, powiązanych siłami podobnie jak elektron z protonem. Na wzór pozytonium można sobie wyobrazić układ protonium, złożony z protonu i antyprotonu. Układy tego rodzaju mogą powstawać w obszarach, w których występuje ambiplazma, przy czym prawdopodobieństwo tworzenia się

ich będzie nawet dość znaczne. Z uwagi na duże masy protonu i antyprotonu tworzenie protonium będzie najbardziej prawdopodobne w stanach o dużej wartości orbitalnej i głównej liczby kwantowej. W stanach tych prawdopodobieństwo anihilacji protonium jest znacznie mniejsze od prawdopodobieństwa przejść radiacyjnych. Po utworzeniu protonium w stanie silnie wzbudzonym, wysłana zostaje cała kaskada fotonów, po czym następuje anihilacja w stanie o niskich liczbach kwantowych. Promieniowanie, wysłane w przejściu pomiędzy sąsiednimi stanami o głównej liczbie kwantowej rzędu 30, leży w obszarze podczerwieni. Podobnie dzieje się i z pozytonium, które przed anihilacją też zdąży wypromieniować szereg fotonów. Seria Lymana leży w ultrafiolecie, seria Balmera — w bliskiej podczerwieni.

Dokładne wartości długości fal, emitowanych przy wzmiankowanych przejściach, podaje Własow, nam jednak wystarcza to, że leżą one w ultrafiolecie, obszarze światła widzialnego i podczerwieni. Dla promieniowania takiego atmosfera ziemska jest w wystarczającym stopniu przezroczysta, można więc szukać wszystkich wymienionych linii za pomocą zwykłych teleskopów. Natężenie strumienia kwantów optycznych, wysyłanych przez pozytonium i protonium, ocenić można w oparciu o dane astronomii promieni gamma — górna granica tego strumienia będzie rzędu 0,0054 kwantów na cm^2 na sekundę.

Zdaniem Własowa miękkie promieniowanie rentgenowskie, pochodzące z centralnych obszarów Galaktyki (o natężeniu ok. 5 kwantów na cm^2 na sekundę na powierzchni Ziemi) pochodzi z przejścia radiacyjnego protonium ze stanu o głównej liczbie kwantowej $n = 3$ do stanu o $n = 2$.

III. Ewolucja układu metagalaktycznego

1. Proces kurczenia się kuli wypełnionej ambiplazmą

Chcąc bardziej szczegółowo skonkretyzować „warunki początkowe“, rozważmy kulę, zawierającą jednorodną mieszaninę protonów i antyprotonów (Alfvén i Klein (1962)). W rozważaniach tych nie obchodzi nas temperatura. Nie może ona być równa zero, gdyż rozważać będziemy zderzenia pomiędzy protonami i antyprotonami. Ze względu na znaczenie pola magnetycznego, jego wartość początkowa też będzie różna od zera. Istotne jest też, by już w chwili początkowej występował pewien stopień turbulencji.

W wyniku zderzeń pomiędzy protonami i antyprotonami i anihilacji wytworzone zostają elektrony, pozytony i fotony. Można przyjąć, że już w stanie początkowym istnieją elektrony, pozytony i fotony, nie odgrywa to jednak zasadniczej roli, gdyż cząstki te w każdym razie powstają w rezultacie anihilacji.

Wychodząc ze stanu o nadzwyczaj niskiej gęstości kula kurczy się jednostajnie pod wpływem własnej siły ciężenia, proporcjonalnej do odległości r od środka. Na początku kurczenia się efekty relatywistyczne można zaniedbać

i prędkość cząstki w odległości r od środka jest prędkością spadku swobodnego:

$$v = \eta r, \quad (\text{III.1})$$

gdzie

$$\eta = \frac{2}{3}(t_0 - t)^{-1}. \quad (\text{III.2})$$

W wyniku spadku swobodnego gęstość ρ zmienia się w sposób następujący:

$$\rho = (6\pi\gamma)^{-1}(t_0 - t)^{-2}, \quad (\text{III.3})$$

gdzie γ jest stałą grawitacyjną. Stała t_0 jest to czas, w którym wszystkie cząstki osiągnęłyby środek, gdyby kureczenie się trwało bez zmian do tego czasu.

W fazie kureczenia się wytwarzane zostaje promieniowanie elektromagnetyczne w wyniku zderzeń naładowanych elektrycznie cząstek oraz wskutek przyspieszania elektronów w polu magnetycznym. W dalszym ciągu ilość promieniowania będzie wzrastać na skutek rozpraszania na poruszających się elektronach. Te źródła promieniowania nie odgrywają jednak większej roli w porównaniu z procesem anihilacji i dlatego pominiemy je w naszych wstępnych rozważaniach. Natomiast anihilacja stanowi źródło promieniowania, którego gęstość wewnątrz kuli pozostaje stała, przynajmniej dotąd, dopóki można pominąć spadek masy spowodowany anihilacją. Stałe źródło promieniowania wytwarza strumień promieniowania skierowany na zewnątrz od środka układu. (Zaniedbujemy tutaj czas potrzebny do ustalenia tego stanu). Pod ciśnieniem promieniowania ruch do środka zostaje zatrzymany i kierunek jego zmienia się na przeciwny, co według Hubble'a charakteryzuje obecny etap rozwoju metagalaktyki.

Energia wytwarzana we wnętrzu kuli o promieniu r wskutek anihilacji proton-antypoton wynosi:

$$P = (4\pi/3)r^2 n_p 2m_p (c^2/T_p) = (8\pi/3k_p) m_p c^2 n_p^2 \sigma_0 r^2, \quad (\text{III.4})$$

(patrz też rozdz. II, równ. (II.6)), a wypadkowy stacjonarny wypływ pędu na zewnątrz, liczony na cm^2 (z pominięciem opóźnienia czasowego przy jego ustaleniu) dany jest przez:

$$F = P/4\pi r^2 c = (1/6k_p)(c^2 \rho^2 \sigma_0 / m_p) r. \quad (\text{III.5})$$

W wyniku anihilacji powstają mezony, neutrina i wysokoenergetyczne kwanty gamma (z rozpadu π^0). Po rozpadach mezonów π i μ przypadają po cztery elektrony i pozytony na jeden proces anihilacyjny; każdy z nich unosi około 100 MeV. Będą się one poruszać po linii spiralnej w polu magnetycznym i wysyłać promieniowanie synchrotronowe aż do chwili anihilacji połączonej z emisją kwantów γ . Gdyby cały strumień energii wychodzącej F stanowiło promieniowanie elektromagnetyczne o niskiej energii, wtedy efekt ciśnienia jego na gaz dany byłby przez thomsonowski przekrój czynny $\frac{2}{3}\sigma_0$ dla elektronów. Jednakże dla kwantów γ o energii osiągającej 1 MeV przekrój czynny małe jest mniej więcej o rząd wielkości, zanim w wyniku produkcji par nie wzrosnie

przy energiach powyżej 10 MeV. Dalej, znaczna część energii wysyłana jest w postaci neutrin. Jeśli zatem siłę działającą na elektron przyjmiemy równą $f_e = \epsilon \frac{2}{3} \sigma_0 F$, uwzględniając powyższe okoliczności, to winniśmy spodziewać się wartości ϵ w przedziale pomiędzy 0,1 a 1. Dla F określonego wzorem (III.5) daje to:

$$f_e = (4\epsilon/9k_p)(\rho^2/m_p)\sigma_0^2 c^2 r. \quad (\text{III.6})$$

Ze względu na istnienie pola magnetycznego, siła ta działa na cały gaz. Pod wpływem ciśnienia promieniowania element objętości dV w odległości r od środka poddany zostaje działającej w kierunku na zewnątrz sile $f_r dV$, gdzie

$$f_r = (8\epsilon \rho^2 n_e / 9k_p m_p) \sigma_0 c^2 r. \quad (\text{III.7})$$

W przeciwnym kierunku działa siła grawitacyjna, która liczona na jednostkę objętości wynosi:

$$f_g = -\gamma(4\pi/3)\rho^2 r. \quad (\text{III.8})$$

Tak więc otrzymujemy następujące równanie ruchu:

$$\rho \left[\frac{\delta v}{\delta t} + \frac{\delta v}{\delta r} \right] = f_r + f_g, \quad (\text{III.9})$$

gdzie v oznacza prędkość radialną, którą uważamy za wielkość dodatnią, gdy skierowana jest na zewnątrz, a za ujemną, gdy skierowana jest do środka. Kładziemy znowu $v = \eta r$, przy czym η nie zależy od r . Wielkość η dana jest wyrażeniem (III.2) dla wczesnych okresów czasu i winna przejść w parametr Hubble'a dla $t =$ chwila obecna. Korzystając z (III.7) i (III.8) otrzymujemy:

$$\dot{\eta} + \eta^2 = \frac{4}{3} \pi \gamma \rho \left[\frac{2}{3} \pi (\epsilon/k_p) (c^2 d^4 / \gamma m_p) n_e - 1 \right]. \quad (\text{III.10})$$

Do równania tego dochodzi jeszcze równanie określające zmianę ρ spowodowaną dodatkowymi efektami kinematycznymi oraz anihilacją, jak również odpowiednie równanie dla zmiany n_e w czasie. Zanim równania te wypiszemy, wprowadzimy następujące dogodne jednostki (jednostki te wprowadził Klein w związku ze swoją interpretacją zależności Eddingtona, później zaś okazało się, że nadają się one również do opisu procesów anihilacji):

$$\text{jednostka długości:} \quad c^2 d^2 / \gamma m_p = 6,3 \cdot 10^{26} \text{ cm}, \quad (\text{III.11})$$

$$\text{jednostka czasu:} \quad c^2 d^2 / \gamma m_p = 6,7 \cdot 10^8 \text{ lat}, \quad (\text{III.12})$$

$$\text{jednostka masy:} \quad c^4 d^2 / \gamma^2 m_p = 8 \cdot 10^{54} \text{ g} = 4 \cdot 10^{21} M_{\text{st\u00f3n}}, \quad (\text{III.13})$$

$$\text{jednostka g\u00e9sto\u015bci masy:} \quad \gamma m_p^2 / c^2 d^4 = 3,2 \cdot 10^{-26} \text{ g/cm}^3. \quad (\text{III.14})$$

W dalszym ci\u0105gu jako jednostk\u0119 g\u00e9sto\u015bci liczby cz\u0105stek u\u017cyjemy jednostk\u0119 g\u00e9sto\u015bci masy podzielon\u0105 przez m_p , co daje jednostk\u0119 g\u00e9sto\u015bci liczbowej cz\u0105stek:

$$\gamma m_p / c^2 d^4 = 1,9 \cdot 10^{-2} \text{ cm}^{-3}. \quad (\text{III.15})$$

Stosując te same co poprzednio symbole, pamiętając tylko, że odnoszą się one już do wielkości bezwymiarowych, wyrażonych w odpowiednich, wprowadzonych wyżej jednostkach, zapisujemy równanie (III.10) w postaci:

$$\dot{\eta} + \eta^2 = \frac{4}{3} \pi \left(\frac{2}{3} \pi \frac{\epsilon}{k_p} \eta_e - 1 \right). \quad (\text{III.16})$$

Równanie opisujące zmianę ϱ w czasie ma postać:

$$\dot{\varrho} = -3\eta\varrho - (\pi/2k_p)\varrho^2, \quad (\text{III.17})$$

podczas gdy analogiczne równanie dla zmiany n_e w czasie ma postać:

$$\dot{n}_e = -3\eta n_e - (\pi/k_e)n_e^2 + (\pi/2k_p)\varrho^2. \quad (\text{III.18})$$

Ostatni człon reprezentuje tutaj elektrony powstające w wyniku anihilacji protonów. Gdy ruch skierowany do wnętrza ulegnie zahamowaniu, energia kinetyczna przekształca się w energię promienistą. Oznacza to, że f_r winno wzrastać.

2. Omówienie wyprowadzonych równań

Podczas wyprowadzania równań (III.16), (III.17) i (III.18) posłużyliśmy się pewnymi upraszczającymi założeniami, które mogły przyczynić się do wprowadzenia błędów. I tak zaniedbaliśmy rozproszenie promieniowania. Gdy skierowany na zewnątrz strumień energii ulega rozproszeniu wstecznemu przez cząstki poruszające się w kierunku do środka, zachodzi przesunięcie dopplerowskie i część energii kinetycznej cząstek zostaje zamieniona na energię promieniowania. Efekt ten da się w pewien sposób uwzględnić poprzez podwyższenie wartości ϵ (wydajności ciśnienia promieniowania).

Poważne trudności wnosi pole magnetyczne B . Od wartości tej zależy promieniowanie synchrotronowe wysokoenergetycznych elektronów i pozytonów, a co za tym idzie, i wartość k_e . Ponadto występuje pewne opóźnienie w czasie pomiędzy anihilacją a pojawieniem się ciśnienia promieniowania, i opóźnienie to zależy od B . Do tej wartości opóźnienia dodać należy jeszcze i drugą, uwarunkowaną skończoną wartością prędkości światła.

Jeśli wreszcie pragniemy porównać przedstawiony tu model z aktualnym stanem metagalaktyki, musimy uwzględnić to, że metagalaktyka w rzeczywistości daleka jest od jednorodności. Już w dość wczesnym stadium rozwoju procesy kondensacji prowadzą do wytworzenia się galaktyk i gwiazd: wiąże się to z rozdzielaniem materii od antimaterii. Wpływa to na szybkość anihilacji, a więc i na skuteczność ciśnienia promieniowania.

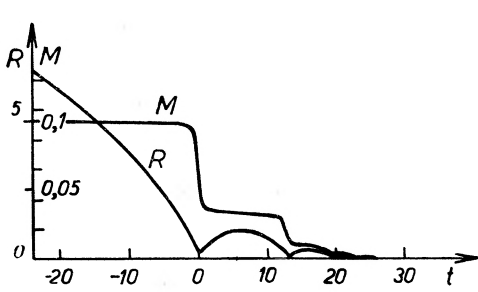
Powyższa analiza wskazuje na to, że gdybyśmy byli nawet w stanie udoskonalic w wysokim stopniu nasze obliczenia, pozostanie i tak pokaźna liczba niewystarczająco znanych parametrów. Jeśli chcemy więc scałkować równania dla porównania teorii z doświadczeniem, najbardziej odpowiednie wydaje się wykonanie całkowań dla szeregu wartości parametrów.

3. Scalkowanie równań

Bonnevier (1964) przecałkował wyprowadzone równania numerycznie, dobierając szereg wartości trzech parametrów: K , k_p i τ . Pierwszy z tych parametrów:

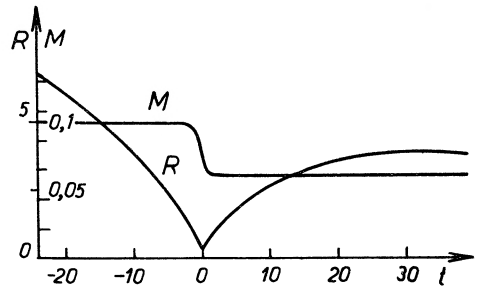
$$K = \frac{2\pi\epsilon}{3k_p} \left\{ \frac{k_e}{4k_p} \left[1 + \left(1 + \frac{8k_p}{k_e} \right)^{1/2} \right] \right\} \quad (\text{III.19})$$

stanowi miarę wartości ϵ . Drugi parametr został już zdefiniowany (patrz rozdz. II). τ oznacza opóźnienie czasowe pomiędzy anihilacją, a chwilą, od której począwszy wytworzone promieniowanie wywiera ciśnienie na elektrony.



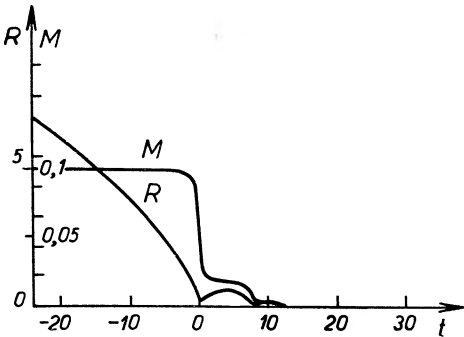
$$K=2, k_p=\pi/2, \tau=0$$

a



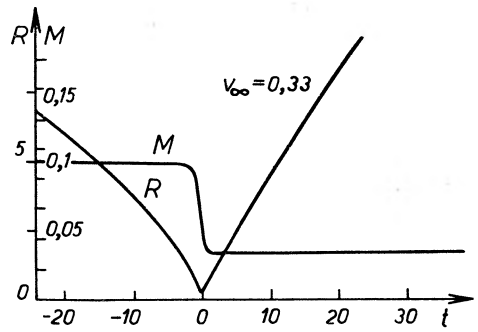
$$K=8, k_p=\pi/2, \tau=0$$

b



$$K=2, k_p=1, \tau=0$$

c



$$K=2, k_p=1, \tau=0,4$$

d

Rys. 2. Zmiana w czasie promienia R i masy M metagalaktyki przy różnych wartościach stałych K , k_p i τ . K jest proporcjonalne do stosunku siły pochodzącej od ciśnienia promieniowania do siły grawitacyjnej; k_p jest proporcjonalne do przekroju czynnego na anihilację protonów. Stała τ jest to opóźnienie czasowe. Jednostki: dla czasu — $6,7 \cdot 10^8$ lat; dla długości — 210 Mpc; dla masy — $4 \cdot 10^{21} M_{\text{Słońc}}$. Z rysunków widać, w jaki sposób początkowe kurczenie się przejść może w rozszerzanie. Wartości parametrów: (a) $K=2$, $k_p=\pi/2$, $\tau=0$; (b) $K=8$, $k_p=\pi/2$, $\tau=0$; (c) $K=2$, $k_p=1$, $\tau=0$; (d) $K=2$, $k_p=1$, $\tau=0,4$

Część wyników Bonneviaera przedstawiona jest na rys. 2. Widoczne są dwa rodzaje rozwiązań. Na rys. 2a i 2c metagalaktyka kurczy się w sposób nieograniczony, przy czym na proces kurczenia nakładają się oscylacje. Prawie cała ambiplazma ulega ostatecznej anihilacji. Rozwiązaniu przedstawionemu na rys. 2d odpowiada przeskok od kurczenia do rozszerzania się, przy czym nawet w czasie nieskończenie odległym ($t \rightarrow \infty$) zachowana zostanie skończona prędkość i skończona masa. Istnieje również przypadek pośredni, przedstawiony na rys. 2b, kiedy to przy bardzo dużych wartościach t prędkość rozszerzania się spada do zera, a wartość masy dąży do wielkości skończonej.

Ze względu na to, że w dobie obecnej metagalaktyka znajduje się w fazie rozszerzania, musimy znajdować się w pewnym miejscu na prawo od minimalnego promienia R metagalaktyki. Ponieważ przypadek 2d najlepiej zdaje się odpowiadać warunkom rzeczywistym, wypada przyjąć zespół parametrów, dający to właśnie rozwiązanie.

4. Porównanie z teorią rozszerzającego się wszechświata

Podobnie jak w teorii rozszerzającego się wszechświata, w teorii Kleina przyjmuje się, że poczerwienienie światła z odległych galaktyk jest efektem dopplerowskim. Ekstrapolacja tego wyniku w kierunku przyszłości jest zbliżona w obu tych teoriach. Ekstrapolacja wstecz jest również podobna dla pewnego przedziału czasowego. Podczas gdy teoria rozszerzającego się wszechświata przyjmuje ekstrapolację wstecz aż do znikomej wartości promienia (niemal równej zeru w naszej skali), w teorii Kleina zakłada się istnienie pewnego punktu zwrotnego, w którym włączenie się ciśnienia promieniowania spowodowało przejście od kurczenia się do rozszerzania. W punkcie tym gęstość materii była większa od gęstości obecnej, ale tylko o kilka rzędów wielkości.

5. Obliczenie średniej gęstości w metagalaktyce

Kurcząca się kula osiągnie maksymalną gęstość nieco przed osiągnięciem minimalnych rozmiarów. Maksymalna gęstość liczbowa cząstek wynosi $n_{\max} = \rho_{\max}/m_p$. Z równań naszych znajdujemy

$$n_{\max} = A(\gamma m_p/c^2 d^4) = A(\gamma m_p m_e^4 c^6/e^8) = 1,9 \cdot 10^{-2} A \text{ cm}^{-3}, \quad (\text{III.20})$$

gdzie A oznacza pewną stałą, która dla krzywych Bonneviaera na rys. 2 waha się pomiędzy 0,5 a 1,6.

Aktualną gęstość średnią w metagalaktyce szacuje się na $(0,3-3) \cdot 10^{-30} \text{ g cm}^{-3}$ (patrz McVittie (1961)). Oznacza to $n = (0,2-2) \cdot 10^{-6} \text{ cm}^{-3}$, co jest o cztery do pięciu rzędów wielkości niższe od teoretycznego maksimum. Ponieważ zgodnie z naszkicowanym tu obrazem przekroczyliśmy już punkt zwrotny i trwa rozszerzanie, winniśmy spodziewać się aktualnej gęstości średniej znacznie niższej od gęstości maksymalnej.

6. Obliczenie stałej Hubble'a

W podobny sposób możemy obliczyć parametr Hubble'a η_H . Wielkość η ulega zmianom w trakcie ewolucji układu. W punkcie zwrotnym równa się ona zeru, wkrótce jednak potem osiąga maksimum, a następnie opada asymptotycznie do zera. Oto wartość maksymalna:

$$\eta_{\max} = B(\gamma m_p / c \bar{d}^2) = B(\gamma m_p m_e^2 c^3 / e^4) = 1,7 \cdot 10^{-9} B \text{ lat}^{-1}, \quad (\text{III.21})$$

gdzie B jest stałą, której wartość dla krzywych Bonneviera waha się od 0,9 do 1,9. Po przejściu maksimum wielkość η gwałtownie maleje. Ponieważ $\eta = v/R$, spadek ten spowodowany jest częściowo zmniejszeniem szybkości rozszerzania pod wpływem sił grawitacyjnych, w głównej jednak mierze spowodowany jest on wzrostem R .

Wartość stałej Hubble'a z obserwacji jest rzędu

$$\eta_H = 10^{-10} \text{ lat}^{-1}. \quad (\text{III.22})$$

Wartość ta jest mniej więcej o rząd wielkości mniejsza od teoretycznego maksimum, odpowiada więc wielkości osiągniętej w pewnej chwili po przejściu przez maksimum.

7. Zależność pomiędzy średnią gęstością a stałą Hubble'a

Jeśli zaniedbać zmianę prędkości rozszerzania się, jak również odstęp czasu pomiędzy punktem zwrotnym a punktem, w którym η osiąga maksimum, wtedy otrzymujemy następującą zależność pomiędzy dziś obserwowaną wartością η_H , a wartością maksymalną η_{\max} :

$$\eta_H = \eta_{\max} \cdot (R_0/R), \quad (\text{III.23})$$

gdzie R/R_0 oznacza stosunek pomiędzy liniowymi rozmiarami metagalaktyki R w dobie dzisiejszej, a jej rozmiarami R_0 w punkcie zwrotnym. Jeśli przez θ_{\max} oznaczymy gęstość maksymalną (w punkcie zwrotnym), a przez θ — gęstość obecną, to otrzymujemy następujący związek:

$$\theta = \theta_{\max} (R_0/R)^3. \quad (\text{III.24})$$

Z obu powyższych zależności otrzymujemy następujący przybliżony związek pomiędzy θ i η_H :

$$\theta / \eta_H^3 = \theta_{\max} / \eta_{\max}^3. \quad (\text{III.25})$$

W myśl teorii przedstawionej w rozdz. III.5 oraz III.6 maksymalna wartość η_H powinna być rzędu odwrotności jednostki czasu określonej związkami (III.12), a θ_{\max} winna być wielkością rzędu jednostki gęstości masy, zdefiniowanej równaniem (III.14). Z równań (III.12) i (III.14) otrzymujemy więc:

$$\theta / \eta_H^3 = (\gamma m_p^2 / e^2 \bar{d}^4) (c \bar{d}^2 / \gamma m_p)^3 \quad (\text{III.26})$$

względnie

$$\theta / \eta_H^3 = c \bar{d}^2 / \gamma^2 m_p = 0,32 \cdot 10^{24} \text{ gsek}^3 \text{ cm}^{-3}.$$

Wartości obserwacyjne wynoszą: $\theta = (3-30) \cdot 10^{-31} \text{ g cm}^{-3}$,
 $\eta_H = (100-150) \text{ km/M parsek} = (3,2-4,8) \cdot 10^{-18} \text{ sek}^{-1}$, co daje:

$$10^{22} < \frac{\theta}{\eta_H^3} < 10^{23}.$$

Wartość ta niewiele odbiega od wartości teoretycznej.

IV. O istnieniu antymaterii

1. Gwiazdy i galaktyki zbudowane z antymaterii

Jak już stwierdziliśmy w rozdziale III, teoria nasza prowadzi do takich wartości średniej gęstości i stałej Hubble'a, które dość dobrze zgadzają się z wynikami obserwacji. W rozdziale II wskazaliśmy na to, że obserwacje nie dają wystarczających podstaw po temu, by wykluczyć możliwość istnienia antymaterii. W tym rozdziale omówimy dokładniej problem istnienia antymaterii.

Na wstępie warto zauważyć, że gdyby nawet zostało wykazane, że w metagalaktyce nie ma wcale antymaterii, przedstawiony tu model w zasadniczych swych zarysach pozostałby słuszny. Jeśli przyjmiemy, że początkowo plazma zawiera np. 60% materii i 40% antymaterii, wtedy ewolucja metagalaktyki przebiegać będzie w przybliżeniu tak, jak to przedstawiliśmy. Przy maksymalnej gęstości w pobliżu punktu zwrotnego większa część antymaterii ulegnie anihilacji wraz z jednakową ilością materii, pozostała zaś część zniknie w podobny sposób podczas kolejnych kondensacji galaktyk. W rezultacie pozostanie metagalaktyka, zawierająca tylko materię, w ilości równej $60 - 40 = 20\%$ jej masy początkowej. Tego typu teoria mogłaby z pomyślnym wynikiem współzawodniczyć z innymi teoriami kosmologicznymi, nie spełnia ona jednak naszego żądania pełnej symetrii między materią i antymaterią. Z tego też powodu nie wrócimy już do niej.

Wiadomo, że na Ziemi nie ma antymaterii. Wiemy i o tym, że Księżyc nie jest zbudowany z antymaterii, gdyż podczas lądowania pojazdów kosmicznych na jego powierzchni nie zaobserwowano żadnych rzucających się w oczy zjawisk. Słońce wyrzuca plazmę, która dociera do Ziemi podczas burz magnetycznych i zorzy polarnej. Nie zauważono podczas tego żadnych wyraźnych zjawisk anihilacyjnych, możemy być więc pewni, że i Słońce składa się ze zwykłej materii. To samo dotyczy z dużym prawdopodobieństwem wszystkich ciał w układzie słonecznym.

Jeśli zatem układ słoneczny powstał w trakcie tej ewolucji, o której przed chwilą mówiliśmy, materia i antymateria musiały już wtedy być rozdzielone, tak że należałoby się spodziewać występowania antymaterii w innych miejscach w metagalaktyce.

Nie da się w pewny sposób rozstrzygnąć, czy obserwowane przez nas gwiazdy składają się z materii czy też z antymaterii. Rozpowszechnione jest przekonanie,

że zbudowane są one z materii, brak jednak na to dowodu. Jak już zauważyliśmy w rozdziale I.2, w wyniku analizy światła wysyłanego przez gwiazdę nie możemy stwierdzić czy składa się ona z materii, czy też z antymaterii. Komety względnie meteory dochodzące do nas z dużych odległości poruszają się po orbitach eliptycznych albo parabolicznych; ani jednego przypadku orbity hiperbolicznej nie znamy z wystarczającą pewnością.

Z analizy promieniowania kosmicznego, jak to wykazemy w rozdziale IV.3, też nie można wyciągnąć wyraźnych wniosków.

Jeślibyśmy zatem twierdzili, że co druga gwiazda obserwowana przez nas na niebie składa się z antymaterii, nie da się tego obecnie udowodnić, ale równie niemożliwe jest też obalenie tego twierdzenia. Twierdzenie takie byłoby w zgodzie z przyjętą przez nas symetrią pomiędzy materią i antymaterią, nie stanowi ono jednak koniecznej konsekwencji tego założenia, możliwe jest bowiem, że każda oddzielna galaktyka składa się z materii bądź z antymaterii. W tym przypadku nasza galaktyka może składać się wyłącznie z materii, nie da się jednak uniknąć wniosku, że, być może, co druga z obserwowanych przez nas galaktyk składa się z antymaterii.

2. O możliwości wykrycia antymaterii

Jak już wspominaliśmy w poprzednim rozdziale, za pomocą analizy widmowej nie da się rozstrzygnąć, czy jakieś ciało niebieskie zbudowane jest z materii czy też z antymaterii. Ambioplazma przejawia swą obecność poprzez emisję fal radiowych, nie musi ona jednak koniecznie wysyłać dającego się wykryć promieniowania γ . Wszystkie gwiazdy radiowe mogą zatem w zasadzie czerpać swoją energię z procesu anihilacji. Jak widać jednak z tablicy II, emisja winna zachodzić raczej przy niskich częstościach, chyba że pola magnetyczne będą nadzwyczaj silne. Ponieważ elektrony wysokoenergetyczne mogą powstawać również na innych drogach, gwiazdy radiowe nie stanowią wystarczającej oznaki istnienia antymaterii.

Pewne ciała niebieskie, np. kwazary, wysyłają, jak się zdaje, takie olbrzymie ilości energii, że jest nieprawdopodobne, by dostarczały jej reakcje jądrowe. W niektórych przypadkach jedynym możliwym źródłem energii wydaje się kompletna anihilacja materii i antymaterii.

Nagłego wyzwolenia ogromnej ilości energii podczas wybuchu gwiazdy supernowej nie udało się również dotychczas w zadowalający sposób wytłumaczyć. Gdyby istniały gwiazdy zbudowane z antymaterii, wtedy zderzenie pomiędzy gwiazdą zwykłą i „antygwiazdą“ mogłoby stanowić przyczynę wybuchu supernowej.

3. Promieniowanie kosmiczne a antymateria

Zazwyczaj przyjmuje się, że promieniowanie kosmiczne ulega przyspieszeniu w wyniku procesów elektromagnetycznych w przestrzeni kosmicznej. Jeśli część docierającego do nas promieniowania powstaje w obszarach wypełnionych antymaterią, wtedy składać się ona powinna z antyprotonów.

W części niskoenergetycznej promieniowanie kosmiczne nie zawiera na pewno antyprotonów. Zdaniem Petersa (wiadomość prywatna) słuszne jest to do 10^{12} eV. Na razie nie da się rozstrzygnąć, czy pierwotne promieniowanie kosmiczne o energii powyżej 10^{12} eV zawiera antyprotony, czy też nie.

Zwykle przyjmuje się, że międzygwiazdne pole magnetyczne jest rzędu (10^{-6} — 10^{-5}) Gs. Weźmy pole $3 \cdot 10^{-6}$ Gs, wtedy proton o energii 10^{12} eV posiada promień larmorowski rzędu 10^{15} cm = 10^{-3} lat świetlnych. Oznaczmy przez L odległość od nas do najbliższego obszaru zajętego przez antymaterię, w którym powstają promienie kosmiczne. Możemy być pewni, że spełniona jest nierówność $L > 10^{15}$ cm, ale to niewiele nam mówi, gdyż odległość ta jest rzędu rozmiarów układu planetarnego, brak zaś powodów, aby sądzić, że zawiera on antymaterię. Jeśli uwzględnimy rozproszenie promieniowania kosmicznego w przestrzeni międzygwiazdnej, możemy od razu podwyższyć wartość L o kilka rzędów wielkości, sporne jest jednak, o ile. Niewykluczone, że biorąc długość rozpraszania jako L ($\approx 4 \cdot 10^{18}$ cm), dojdziemy aż do najbliższych gwiazd, stwierdzając, że nie składają się one z antymaterii — ale to wszystko jest niepewne.

W każdym razie brak podstaw do przyjęcia, że niskoenergetyczne promienie kosmiczne ($< 10^{12}$ eV) wytwarzane są poza naszą galaktyką, tak więc dane o promieniowaniu kosmicznym nie wykluczają tego, by inne galaktyki zawierały antymaterię, lub składały się z niej.

4. Rozdzielanie materii od antymaterii i tworzenie galaktyk

W trakcie ewolucji metagalaktyki powinny powstawać galaktyki. Proces ten może stanowić wynik grawitacyjnej niestabilności w ośrodku metagalaktycznym. Ponieważ niestabilność ośrodka wzrasta ze wzrostem gęstości, powstawanie galaktyk zachodzić winno zwłaszcza wtedy, gdy ośrodek odznacza się maksymalną gęstością, co ma miejsce niedługo przed osiągnięciem punktu zwrotnego.

Niewykluczone jest, że galaktyki powstały już po rozdzieleniu materii od antymaterii. W tym przypadku każda z nich zawierałaby albo materię, albo antymaterię. Jednakże, jak pokazano w rozdziale II.5, napotykamy na pewne trudności w znalezieniu odpowiedniego procesu na dużą skalę, podczas gdy łatwiej można sobie wyobrazić procesy rozdzielania w małej skali. Należałoby zatem oczekiwać, że warunki dla rozdzielania będą bardziej korzystne w zagęszczeniach w rodzaju galaktyk, niż w układzie bardziej jednorodnym, który istniał przed tworzeniem galaktyk. Tak więc, jeśli przedstawiony w rozdziale II.5 proces odpowiada w głównej mierze za rozdzielanie, to należałoby się raczej spodziewać, że każda galaktyka zawiera i materię i antymaterię. Nie jest to sprzeczne z wynikami obserwacji.

Jeśli materia i antymateria współistnieją w pewnej galaktyce, rozkład ich może przybierać najrozmaitsze formy. I tak można przyjąć, że rozkład ich jest zupełnie przypadkowy, a zatem w dowolnej części galaktyki co druga gwiazda składa się z antymaterii. Jeśli tak jest naprawdę, to proces rozdzielania musi zachodzić stosunkowo późno, kiedy ogólna struktura galaktyki już jest ustalona.

Inna możliwość polega na tym, że rozdzielanie materii od antymaterii wiąże się z ogólną strukturą galaktyk. Przyjmijmy, że ambiplazma, z której powstaje galaktyka, zawiera nadmiar materii. W myśl tego, co powiedzieliśmy w rozdziale II.4 (patrz rys. 2) należy oczekiwać, że pod działaniem siły ciężkości ambiplazma rozdzieli się na trzy obszary: obszar środkowy, zawierający ciężką ambiplazmę, który możemy utożsamić z jądrem galaktyki; obszar pośredni, zawierający tylko materię, który stanowi być może tę część galaktyki, w której żyjemy; i wreszcie obszar zewnętrzny, zawierający lekką ambiplazmę (elektrony i pozytony), który możemy utożsamić z obszarem galaktycznego halo, skąd wysyłane są fale radiowe. W jądrze galaktyki następuje w dalszym ciągu rozdzielanie materii od antymaterii, tak że nie wszystko ulegnie anihilacji.

A oto alternatywny obraz dla przypadku, gdy plazma tworząca początkowo galaktykę zawiera jednakowe ilości materii i antymaterii, lecz jedna część galaktyki jest wzbogacona o materię, druga zaś jej połowa — o antymaterię. Znowu jądro składa się z ciężkiej ambiplazmy. Spośród dwóch ramion spiralnych galaktyki jedno winno składać się z materii, drugie — z antymaterii. Galaktykę winny otaczać dwie chmury lekkiej ambiplazmy. Niekiedy mogą się one ujawniać jako dwie radiogwiazdy; konfiguracja taka występuje dość często.

Należy wreszcie pamiętać, że punkt zwrotny, w którym następuje przejście od kurczenia do rozszerzania się metagalaktyki, pojawia się jako rezultat osiągnięcia odpowiednio dużej wartości przez ciśnienie promieniowania anihilacyjnego. W punkcie zwrotnym metagalaktyka zasadniczo była nieprzezroczysta dla promieniowania. Dziś już tak nie jest, gdyż gęstość zmalała o około cztery rzędy wielkości. Większa część promieniowania zdażyła już uciec z metagalaktyki. Szybkość anihilacji w metagalaktyce spadła również gwałtownie z uwagi na rozszerzenie się, oraz z uwagi na rozdzielanie materii od antymaterii. Możemy spodziewać się promieniowania synchrotronowego z odległych części metagalaktyki, obserwujemy bowiem to, co się tam wcześniej działo.

V. Czy istnieją inne metagalaktyki?

W świetle teorii Kleina możliwe jest istnienie innych metagalaktyk we wszechświecie. O naszej metagalaktyce sądzi się, że stanowi ona wynik grawitacyjnej niestabilności w jednorodnej ambiplazmie, zajmującej objętość o wiele rzędów wielkości większą niż objętość metagalaktyki. W ośrodku tym mogły istnieć inne obszary, w których tworzyły się metagalaktyki. Geometria może być w dużej skali euklidesowa, a wszechświat — nieskończony i w przestrzeni i w czasie. Jednakże podczas ewolucji metagalaktyki potencjał grawitacyjny mógł osiągnąć tak duże wartości, że geometria przestrzeni w silnym stopniu odchyliła się lokalnie od euklidesowości. Zachodzi to zwłaszcza w punkcie zwrotnym, kiedy to spowodowane przez grawitację przesunięcie prążków widmowych ku czerwieni staje się ogromne, co oznacza, że osiągamy graniczne rozmiary rzędu promienia Schwarzschilda.

Tak więc pod pewnymi względami jesteśmy świadkami nawrotu do idei rozwinętych przez Charliera (1908 i 1921). Według niego wszechświat może być nieskończony i w przestrzeni i w czasie, a jego całkowita masa może też być nieskończona, mimo że średnia gęstość będzie zerem. Możliwe jest to, jak wykazał Charlier, jeśli przyjąć nieskończony szereg układów o wzrastającej wielkości.

Warto zastanowić się nad tym, czy można zaobserwować inne metagalaktyki, jeśli zaś byłoby to możliwe, to w jakiej postaci jawiłyby się one naszym oczom.

Gdy metagalaktyka znajduje się we wczesnej fazie kurczenia się, zawiera olbrzymie ilości ambiplazmy. W trakcie anihilacji wysyłane jest promieniowanie γ i synchrotronowe. O ile oba rodzaje promieniowania dochodzą do nas bez pochłaniania, to o wiele łatwiej jest wykryć drugie z nich. Częstość wyznaczona jest przez pole magnetyczne, o którym bardzo mało wiemy. Prawdopodobne jest, że maksimum energii wysyłane jest przy raczej niskich częstościach (np. znacznie niższych od jednego megaherca).

Gdy metagalaktyka osiąga stan o maksymalnej gęstości, to wzrasta ilość emitowanej przez nią energii, jeśli zaś pole magnetyczne staje się silniejsze — co wydaje się prawdopodobne — to częstość też rośnie. Jednocześnie należy oczekiwać, że rozpocznie się powstawanie galaktyk, a także powstawanie gwiazd w galaktykach. Proces ten prowadzi do emisji przy częstościach optycznych, coraz to silniejszej, dopóki metagalaktyka nie osiągnie takiego stanu, który w przybliżeniu odpowiada obecnemu stanowi naszej metagalaktyki. Ostatnia faza obejmuje powszechne rozbieganie się galaktyk podczas rozszerzania się całego układu metagalaktycznego.

W pobliżu punktu zwrotnego ilość emitowanej energii P , wyzwalanej w wyniku anihilacji, jest rzędu

$$P = Mc^2/\tau = 8 \cdot 10^{54} (3 \cdot 10^{10})^2 / 2 \cdot 10^{16} = 4 \cdot 10^{59} \text{ erg/sek} = 10^{26} P_{\odot},$$

gdzie dla masy metagalaktyki i jednostki czasu τ przyjęliśmy wprowadzone w rozdz. III.1 jednostki. P_{\odot} oznacza ilość energii (w ergach na sek) wypromieniowywaną w jednostce czasu przez nasze Słońce.

Dla porównania: z silnego źródła promieniowania radiowego odbieramy strumień energii $10^{-11} \text{ erg cm}^{-2} \text{ sek}^{-1}$, co odpowiada mocy 10^{59} erg/sek dla źródła znajdującego się w odległości $10^{85} \text{ cm} = 10^{17} \text{ lat}$ świetlnych.

Jednakże duży ułamek wydzielonej energii ulegnie autoabsorpcji w samej metagalaktyce. Istotnie, właśnie ciśnienie promieniowania doprowadza do powstania punktu zwrotnego. Ponadto, w pobliżu punktu zwrotnego w wyniku efektów relatywistycznych metagalaktyka mniej lub bardziej odgranicza się od świata zewnętrznego, tak że wypływa z niej tylko nieznaczną część wyzwolonej energii.

Ponadto, zwłaszcza we wczesnej fazie ewolucji, podstawowa część promieniowania synchrotronowego może być wysyłana przy tak niskich częstościach, że nie przedostanie się ona przez jonosferę. Można więc obserwować ją tylko poza atmosferą ziemską, z pojazdów kosmicznych. Tymczasem nawet ośrodek

międzygwiazdowy może pochłaniać, czy nawet obciąć w pełni część widma. I tak gęstość materii międzygwiazdowej rzędu jednego elektronu w cm^3 odpowiada częstości plazmy około 10^4 sek^{-1} ; promieniowanie o częstości niższej nie przejdzie.

Jeśli obserwujemy inną metagalaktykę, znajdującą się w tej samej fazie ewolucyjnej co nasza, wtedy jej jasność powierzchniowa powinna być zbliżona do średniej jasności nocnego nieba λ_0 , uwarunkowanej świeceniem wszystkich obiektów z naszej metagalaktyki, ale spoza naszej własnej galaktyki (byłaby to więc jasność powierzchniowa nieba, powstała w wyniku jakiegoś rozmazania obiektów pozagalaktycznych). Trudno oczywiście zauważyć taki obiekt. W punkcie zwrotnym gęstość masy w metagalaktyce mniej więcej 10^3 do 10^4 razy przewyższałaby obecną gęstość w naszej metagalaktyce. Jeśli przyjąć, że całkowita ilość światła wysyłanego przez gwiazdy pozostaje niezmienna, będziemy mieli jasność powierzchniową około $10^2 \lambda_0$ do $10^3 \lambda_0$. Przyjmując, że metagalaktyka zawiera 10^{10} galaktyk, widzimy, że będzie ona wyglądała jak galaktyka, jeśli odległość do niej jest 10^5 razy większa niż odległość do galaktyki. Oznaczałoby to, że można optycznie zaobserwować metagalaktykę, jeśli znajduje się ona w odległości nie większej niż 10^{14} lat świetlnych. Jednakże w pobliżu punktu zwrotnego nawet fale o częstościach optycznych nie będą w stanie wydostać się z metagalaktyki.

Jak widać z tego co powiedziano, istnieją pewne trudności przy obserwowaniu innej metagalaktyki. Na wczesnym etapie rozwoju może ona wysyłać przede wszystkim fale radiowe niskiej częstości, które mogą być obcinane przez plazmę w ośrodku międzygwiazdowym w pobliżu nas. W późnej fazie ewolucyjnej jasność powierzchniowa takiej metagalaktyki nie będzie przewyższać jasności powierzchniowej naszej własnej metagalaktyki, trudno więc ją będzie zauważyć. Ogromna ilość energii anihilacyjnej wydzielonej w pobliżu punktu zwrotnego dałaby się wykryć, gdyby metagalaktyka nie była zbyt daleko od nas, ale efekty relatywistyczne odcinają metagalaktykę od zewnętrznych obserwatorów na pewną część tego okresu.

Przesunięcie dopplerowskie Δ prążka widmowego, wysłanego przez gwiazdę znajdującą się w obcej metagalaktyce składa się z czterech części

$$\Delta = \Delta_c + \Delta_g + \Delta_r - \Delta_g^0.$$

Tutaj przesunięcie Δ_c pochodzi od kurczenia czy też rozszerzania się metagalaktyki, Δ_g jest to przesunięcie grawitacyjne ku czerwieni, spowodowane przez siły grawitacyjne w obcej metagalaktyce, Δ_r uwarunkowane jest względnym ruchem środka ciężkości obcej metagalaktyki w stosunku do naszego Słońca, wreszcie Δ_g^0 pochodzi od potencjału grawitacyjnego Słońca związanego z jego położeniem w naszej metagalaktyce.

Ze względu na to, że inną metagalaktykę można zaobserwować tylko bardzo blisko jej punktu zwrotnego, jest bardzo prawdopodobne, że Δ_c jest raczej wielkością małą. Dalej, Δ_g musi być znacznie większe od Δ_g^0 , gdyż nasza meta-

galaktyka znajduje się już raczej dość daleko od swego punktu zwrotnego. Przesunięcie dopplerowskie redukuje się więc do

$$\Delta = \Delta_g + \Delta_r.$$

Nic nie wiemy o Δ_r . Wartość Δ_g zależy od odstępów czasu, który minął od przejścia przez punkt zwrotny. W pobliżu punktu zwrotnego oczekiwać należy bardzo silnego przesunięcia grawitacyjnego prążków ku czerwieni.

Podsumowanie

Pełna symetria pomiędzy cząstkami elementarnymi i ich antycząstkami pozwala przypuszczać, że wszechświat zawiera jednakowe ilości materii i antymaterii.

Brak danych obserwacyjnych wskazujących nieodwołalnie na to, że wszystkie ciała niebieskie składają się z materii; równie dobrze co drugie ciało niebieskie może być zbudowane z antymaterii. Obszary, w których występuje materia, mogą być oddzielone cienkimi warstwami, w których zachodzi mieszanie materii i antymaterii, od obszarów wypełnionych antymaterią.

Zbadano własności ambiplazmy (składającej się ze zjonizowanej materii i antymaterii). Okazuje się, że ambiplazma w polu magnetycznym powinna być źródłem promieniowania synchrotronowego, nie da jednak promieniowania γ , które można by wykryć. Możliwe, że pewne — a może nawet wszystkie — gwiazdy radiowe zawierają ambiplazmę, a ich źródłem energii jest proces anihilacji.

Pod łącznym działaniem sił grawitacyjnych i elektromagnetycznych ambiplazma może ulec rozdzieleniu na materię i antymaterię.

Uwaga autora dodana w korekcie. Interesującą metodę wykrycia ambiplazmy zaproponował ostatnio N. A. Własow (1965).

Widmo promieniowania radiowego, pochodzące z ambiplazmy, ulegającej anihilacji w polu magnetycznym, obliczyli ostatnio Bonnevier, Ekspong i Yamdagni. Wydaje się, że charakter tego widma dość dobrze zgadza się z ogólnym kształtem widma gwiazd radiowych (wiadomość prywatna).

Podziękowania

Pragnę wyrazić wdzięczność profesorowi Kleinowi za wprowadzenie mnie do jego teorii kosmologicznych oraz za zwrócenie mojej uwagi na znaczenie badań nad właściwościami ambiplazmy.

Współpracowałem z dr B. Bonnevierem, który łaskawie pomagał mi przy wielu obliczeniach. Pani Gun Giuli pomagała w przygotowaniu rękopisu.

Tłumaczyła Janina Pochmurska

Antimatter and the Development of the Metagalaxy

Translated from *Reviews of Modern Physics*, 17, 652 (1965).

Literatura

- H. Alfvén, C. G. Fälthammar, *Cosmical Electrodynamics* (Oxford University Press, Oxford), 2 wyd., 1963.
- H. Alfvén, O. Klein, *Matter-Antimatter Annihilation and Cosmology*, Ark. Fys., **23**, 187—194 (1962).
- R. A. Alpher, R. Herman, *On Nucleon-Antinucleon Symmetry in Cosmology*, Science **128**, 904 (1958).
- B. Bonnevier, *On the Early Development of the Metagalactic System*, Ark. Fys. **27**, 310 (1964).
- C. V. L. Charlier, *Wie eine unendliche Welt aufgebaut sein kann*, Ark. Mat. Astron. Fys., **4**, 24 (1908).
- C. V. L. Charlier, *How an Infinite World May Be Built Up*, Ark. Mat. Astron. Fys., **16**, 22 (1921).
- M. Goldhaber, *Speculations on Cosmogony*, Science **124**, 218 (1956).
- O. Klein, *Some Cosmological Considerations in Connection with the Problem of the Origin of the Elements*, Mém. Soc. Roy. Sci. Liège **14**, 42—51 (1953), (specjalny tom pt. *Les Processus Nucléaires dans les Astres*, także tłum. ros. *Jadziernyje Processy w Zwiezdach*, 1957).
- O. Klein, *La structure et l'évolution de l'univers* (Institut International de Physique, Solvay, Bruxelles), 33—51 (1958).
- O. Klein, *Einige Probleme der allgemeinen Relativitätstheorie*, str. 58—72 w zbiorze *Werner Heisenberg und die Physik unserer Zeit*, Braunschweig 1961.
- O. Klein, *Mach's Principle and Cosmology in Their Relation to General Relativity*, s. 293—302 w zbiorze *Recent Developments in General Relativity*, Pergamon Press, Oxford i PWN, Warszawa 1962.
- W. Kraushaar, G. W. Clark, G. Garmire, H. Helmken, P. Higbie, M. Agogino, *Astrophys. J.* **141**, 845 (1965).
- G. C. McVittie, *Fact and Theory in Cosmology*, Londyn 1961.
- G. C. McVittie, *Problems of Extragalactic Research*, 15 Sympozjum IAU (Międz. Unii Astron.), s. 441 (1961).
- M. Ryle, *Developments in Radio Source Work Since 1960*. Intern. lecture commission V, URSI, Tokio 1963.
- N. A. Własow, *Opticzeskij metod razwiedki antiwieszczestwa wo Wsielennoj*, *Astron. Żurn.* **41**, 893 (1964); także raport nr 622 Inst. im. Kurczatowa.

Bibliografia uzupełniająca (zwłaszcza w jęz. polskim)

I. Pozycje ogólne

- H. Bondi, *Wszechświat nieznanym*, WP Warszawa 1964.
- H. Bondi, *Kosmologia*, PWN, Warszawa 1965.
- G. Gamow, *Materia, Ziemia i Niebo*, PWN, Warszawa 1963.
- W. Ł. Ginzburg, *Promienie kosmiczne w pobliżu Ziemi i we Wszechświecie*, *Postępy Fizyki* **13**, 507—48 (1962).
- W. Ł. Ginzburg, S. I. Syrowatski, *Proischożdenie kosmiczeskich łuczey*, Izd. AN SSSR, Moskwa 1963.
- Mezony, gravitacja, antymateria*, PWN, Warszawa 1962.

II. Pozycje dotyczące problemu detekcji neutrin kosmicznych (do rozdz. II.3)

- J. N. Bahcall, *Observational Neutrino Astronomy*, Science **147**, 115—20 (1965); także tłum. ros. *Usp. Fiz. Nauk* **87**, 143—54.
- J. Jateczak, *Astrofizyka neutrinowa*, *Postępy Astronomii* **13**, 63—84 (1965).

- J. Jatzak, *Neutrina kosmiczne*, Postępy Fizyki **16**, 279—93 (1965).
- B. Kuchowicz, *Neutrina*, Post. Techn. Jąd., **8**, 730—40 (1964).
- B. Kuchowicz, *Nowe propozycje doświadczeń neutrinowych*, Post. Techn. Jąd., **9**, 1036—50 (1965).
- M. A. Ruderman, *Astrophysical Neutrinos*, Rep. Progress in Physics **28**, 411—62 (1965).
- M. Schwartz, *Neutrino Physics*, Rep. Progr. Phys., **28**, 61—75 (1965).
- G. T. Zacepin, W. A. Kuzmin, *Niekotoryje problemy nejtrinnoj fiziki*, Wiestnik Akad. Nauk SSSR, **11**, 50—55 (1964).

Laurie M. Brown *

Poprzez kwarki do symetrii cząstek **

Za każdym razem, kiedy fizycy myślą, że znaleźli podstawowe cegiełki materii, odkrywają, że cegiełki te same mają strukturę. Obecnie wydaje się, że tymi cegiełkami mogą być kwarki¹ — z trzech różnych rodzajów kwarków można zbudować wszystkie znane mezony oraz bariony i uporządkować je w symetryczne zestawy złożone z ośmiu lub dziesięciu cząstek.

Kiedy redaktorzy „Physics Today“ poprosili mnie, lub raczej wyzwali mnie, abym napisał artykuł wyjaśniający unitarne symetrie cząstek fizykom, którzy nie są specjalistami w dziedzinie cząstek elementarnych, zawahałem się. Ale wzięwszy pod uwagę fizyka, którego dzieci mogą zapytać: „Tatusiu, z czego są zrobione cząstki elementarne?“ zdecydowałem ostatecznie, że napiszę ten artykuł. Pytanie to — tak niewinne, tak jasne, tak pełne zdrowej ciekawości — zasługuje na odpowiedź i to, o ile tylko można, na tym samym poziomie.

Spróbowałem więc dążyć do tego celu po najbardziej bezpośredniej drodze, która wiedzie poprzez kwarki. Termin „kwark“, wprowadzony przez Murraya Gell-Manna, jest użyty tutaj na oznaczenie dowolnego modelu zawierającego jeden lub więcej unitarnych tripletów. W takim modelu symetrie i inne własności cząstek wynikają z symetrii i innych własności kwarków, które są ich składnikami. Jedną wielką przeszkodę stanowi to, że nikt do tej pory nie zaobserwował kwarku. Może nie szukano ich wystarczająco dokładnie. Może nie istnieją. Niemniej jednak kwarki stanowią najprostszy, konkretny sposób zrozumienia cząstek elementarnych.

Stwierdzenie, że jądro atomowe składa się z nukleonów: protonów i neutronów *jest* bardzo sensowne. Najlżejszym jądrem *jest* proton, masy jąder są z grubsza całkowitymi wielokrotnościami masy nukleonu, ładunki jąder są całkowite, a powłoki jądrowe są zapełnione w taki sposób, jakiego można się spodziewać, gdy w grę wchodzi dwa rodzaje cząstek o połówkowym spinie.

* Jest to trzeci artykuł Laurie Browna o unitarnych symetriach w „Physics Today“ (patrz „Musings in the Musey Room“ w kwietniu 1965 i „Tyger Hunting in the Everglades“ w kwietniu 1964). Laurie Brown jest profesorem fizyki na uniwersytecie Northwestern, dokąd przybył w 1950 roku po uzyskaniu stopnia doktora u Richarda Feynmana w Cornell. Brown interesuje się cząstkami elementarnymi i elektrodynamiką kwantową.

** Tłumaczenie artykułu zamieszczonego w czasopiśmie *Physics Today*, 19, No 2, 44 (1966), dokonane za zgodą Wydawcy.

¹ W terminologii angielskiej: *quarks* (przyp. tłum.).

Być może mniej ważne jest to, że można wybijać nukleony z jąder; w elektromagnetycznym procesie tworzenia par elektron i pozyton są „wybijane“ z fotonu wysokiej energii, choć foton nie „składa się“ z pary elektron-pozyton.

Tak się składa, że neutron i proton mają prawie tę samą masę, a siły działające w jądrach są prawie niezależne od ładunku. Wynika stąd, że stany jądrowe są dobrze opisywane za pomocą liczb kwantowych, które można powiązać (uwzględniając zakaz Pauliego) z liczbami kwantowymi nukleonów. Postępujemy przy tym tak, że zaczynamy od zupełnej symetrii zakładając równe masy protonu i neutronu i zaniehbując oddziaływanie elektromagnetyczne, a potem uwzględniamy poprawki na te, łamiące symetrię, efekty.

Tak więc, przypisując dubletowi nukleonowemu jednostkową liczbę nukleonową N i połówkowy izospin I (przy czym proton ma $I_3 = 1/2$, a neutron $I_3 = -1/2$) tak, że ładunek jest $Q = I_3 + N/2$, mamy następujące addytywne wewnętrzne liczby kwantowe stanów jądrowych:

$$A = \sum_i N_i,$$

$$Z = \sum_i Q_i,$$

$$I = \sum_i I_i,$$

gdzie \sum_i obejmuje wszystkie nukleony, a dozwolone stany I , razem z dynamicznymi liczbami kwantowymi momentu pędu i energii, są ograniczone zakazem Pauliego.

To wszystko jest bardzo dobrze znane i chyba nie wymaga przypomnienia. Ale przypuśćmy, że nigdy nie obserwowalibyśmy nukleonów, a tylko jądra o liczbie masowej dwa lub więcej. Czy wynaleźlibyśmy wtedy nukleony? Oczywiście, że tak! I przy pomocy tych quasi-nukleonów bądź spurionów, czy też elementów symetrii, czy jakkolwiek nazwę moglibyśmy dla nich wymyślić, rozumielibyśmy fizykę jądrową tak samo dobrze, albo tak samo źle, jak teraz z neutronami i protonami. W naszym realnym świecie, ponieważ siły jądrowe są stosunkowo słabe, byłoby oczywiście trudno zmusić jądra do oddziaływania tak, aby nigdy nie wytwarzały rzeczywistych protonów i neutronów. (Przy wyższych energiach wytwarzamy również mezony, dziwne baryony i antybaryony, ale stwierdzenie, że jądra składają się z nich, jest mniej sensowne.)

Podsumujmy powody, dla których powszechnie uważa się, że jądra są zbudowane z nukleonów:

1. Addytywne wewnętrzne liczby kwantowe jąder (a również fakt, że spin jądrowy jest całkowity lub połówkowy) mogą być wytłumaczone przy pomocy modelu nukleonowego. To znaczy, że nukleony tworzą *najmniejszy zespół elementów symetrii*, z kombinacji których mogą wynikać wewnętrzne symetrie stanów jądrowych.

2. Energie wiązania jąder są wystarczająco niskie w stosunku do odstepu energetycznego między nukleonem, a jego pierwszym stanem wzbudzonym (nukleon plus pion), aby było możliwe wzbudzenie lub rozbitcie jąder bez wzbudzenia nukleonów.

Jasne jest, że według tych samych kryteriów można również powiedzieć, że atomy są zbudowane z jąder i elektronów, drobiny składają się z atomów itd.

Rzeczywiste istnienie atomów, jąder i drobin jest na pewno jednym z najważniejszych pytań zadawanych w fizyce i odpowiedziano na nie pozytywnie. Ale objaśnianie zespołów utworzonych z tych obiektów przy pomocy ich własności jest, do pewnego stopnia, oddzielnym problemem. Takie objaśnienie wymaga tylko, aby miały one sens jako quasi-objekty.

Model, ale jakiego rodzaju?

W poszukiwaniu elementarnych cegiełek, z których świat jest zbudowany, fizycy stopniowo dzielili materię na coraz mniejsze części — i odkryli, że na każdym etapie podziału pozostaje przynajmniej jedna część posiadająca godną uwagi strukturę. Leptony chwilowo wydają się być proste (choć niezupełnie zrozumiałe), ale nukleony, liczne inne bariony i mezony² — czyli te cząstki, które silnie oddziałują i są zbiorowo nazywane hadronami³ — posiadają złożoną strukturę. Naturalnie, próbujemy powiązać ze sobą te złożone struktury poprzez model, w którym występują bardziej elementarne struktury, chociaż nie mamy żadnej pewności, że te bardziej elementarne struktury istnieją niezależnie od modelu. To jest, w samej rzeczy, konserwatywne i tradycyjne podejście do zagadnienia, które dawało rezultaty wiele razy w historii fizyki. Ciekawe będzie przekonać się, czy poskutkuje jeszcze raz.

W ciągu dwu ostatnich dziesięcioleci naliczyliśmy około sto czy ileś tam hadronów innych niż jądra atomowe. Każdemu z nich można przypisać liczbę barionową B , która jest uogólnieniem liczby nukleonowej N — i okazuje się, że B jest ściśle zachowywane. Wydaje się ona różnić od również ściśle zachowywanego ładunku Q tym, że nie jest źródłem żadnego znanego pola. Hadrony występują w multipletach ładunkowych bądź izospinowych, które są zdegenerowane pod względem masy, jeśli rozważamy tylko silne oddziaływania. Hadrony posiadają i inną addytywną liczbę kwantową, która jest zachowywana, kiedy wchodzi w grę tylko silne i elektromagnetyczne oddziaływania; tą liczbą kwantową jest albo dziwność S , albo hiperładunek $Y = B + S$. Cząstki i antycząstki zawsze mają przeciwne co do znaku liczby Q , B , S , Y .

Jeśli hadrony są zbudowane z pewnego minimalnego zespołu bardziej elementarnych struktur, to struktury te muszą mieć przynajmniej następujące własności:

(1) co najmniej jeden połówkowy multiplet *spinowy*; najmniejszym możliwym jest dublet o spinie $1/2$,

² Chodzi tu o mezony π i K (przyp. tłum.).

³ Od greckiego słowa $ΑΙΡΟΣ$ — moczny (przyp. tłum.).

(2) co najmniej jeden połówkowy multiplet *izospinowy*, przy czym dublet o izospinie $1/2$ jest najmniejszym z możliwych,

(3) co najmniej jeden element o liczbie barionowej różnej od zera,

(4) co najmniej jeden element o dziwności różnej od zera.

Pierwsze trzy wymagania są spełnione przez same nukleony, które mają spin $1/2$, izospin $1/2$ i liczbę barionową 1. W konsekwencji pierwszym złożonym modelem jaki zaproponowano dla hadronów była fizyka jądrowa!

Model Fermiego-Yanga-Sakaty

Po odkryciu w 1947 roku pierwszego hadronu o liczbie barionowej równej zero czyli pionu, naturalne (jak również wspaniałe) było spostrzeżenie w roku 1949 przez Enrico Fermiego i C. N. Yanga, że pion ma liczby kwantowe odpowiadające bardzo silnie związanemu układowi nukleon-antynukleon w stanie o zerowym spinie, zerowym kręcie orbitalnym i jednostkowym izospinie i że ten model tłumaczyłby również ujemną wewnętrzną parzystość pionu, która jest raczej zagadkowa.

I znowu, po odkryciu dziwnych hadronów, naturalne było, że Skoichi Sakata uznając ważność warunku (4) wprowadził trzeci element symetrii mający połówkowy spin, zerowy izospin, jednostkową liczbę barionową i oprócz tego dziwność $S = -1$ (tak, aby hiperładunek Y był równy zeru). Utożsamił on ten element z barionem Λ tworzącym triplet Sakaty (P , N , Λ) wraz z neutronem i protonem. Ten triplet elementów posiada liczby kwantowe właściwe dla zbudowania wszystkich hadronów. Na przykład pseudoskalarny dublet mezonowy K (K^- , K^0) buduje się silnie wiążąc antycząstkę Λ z dubletem nukleonowym (proton i neutron) w stanie o zerowym spinie i kręcie orbitalnym.

W granicy symetrii, gdzie zaniedbuje się różnice mas wewnątrz tripletu Sakaty, uogólnieniem niezależności ładunkowej sił jądrowych może być wymaganie symetrii oddziaływania pomiędzy trzema elementami tripletu Sakaty.

Mówiąc bardziej precyzyjnie, sformułowanie ładunkowej niezależności sił jądrowych przy pomocy niezmienniczości izospinowej jest równoważne wprowadzeniu dwuwymiarowej przestrzeni zespolonej, w której są reprezentowane wektory stanów izospinowych. Niezależność ładunkowa oznacza, że oddziaływanie jest niezmiennicze względem dowolnego sztywnego obrotu osi w tej przestrzeni. Niezależność ładunkową i hiperładunkową wyraża się przy pomocy trójwymiarowej zespolonej przestrzeni ładunkowo-hyperładunkowej, a to znaczy, że oddziaływanie pomiędzy P , N , Λ jest niezmiennicze względem dowolnego sztywnego obrotu osi w tej trójwymiarowej przestrzeni. Jeśliby tak było i P , N oraz Λ byłyby zdegenerowane pod względem masy, wtedy na przykład mezony π i K byłyby również zdegenerowane pod względem masy.

Grupę symetrii niezmienniczości ładunkowej tworzą wszystkie unitarne macierze 2×2 o wyznaczniku równym jedności. Nosi ona nazwę specjalnej grupy unitarnej w dwóch wymiarach lub $SU(2)$ (jest to również grupa symetrii momentu pędu). Grupą symetrii niezmienniczości ładunkowo-hyperładunkowej

jest specjalna grupa unitarna w trzech wymiarach czyli $SU(3)$. Kiedykolwiek występują wszystkie trzy elementy tripletu Sakaty, jak na przykład w hiperjadrach (które są jądrami zawierającymi hiperony Λ związane z nukleonami), $SU(3)$ jest tą grupą symetrii, której używa się dla klasyfikowania stanów. Kiedy zaś wchodzi w grę tylko dwa z tych trzech elementów, wtedy symetria jest typu $SU(2)$, która jest podgrupą $SU(3)$. Oczywiście są trzy możliwe podgrupy $SU(2)$ wyższej symetrii $SU(3)$. Nazwano je podgrupami spinu U , spinu V i spinu I (izospinu).

Kłopot z modelem Sakaty

Symetria leżąca u podstaw modelu Sakaty jest typu $SU(3)$ i model ten całkiem dobrze przewiduje jakościowe cechy nukleonów, jąder i hiperfragmentów, ale daje on zniekształcony obraz barionów — ważnych stanów o jednostkowej liczbie barionowej. Posuńmy model Sakaty o parę stopni dalej i zobaczymy dlaczego tak jest.

Oznaczmy dowolny element modelu Sakaty przez S (sakaton). Przypuśćmy, że zgodnie z Sakatą utożsamiamy ten triplet z fizycznymi cząstkami: protonem, neutronem i hiperonem Λ . Jak dotąd jest to wysoce zadowolające ponieważ są one najlżejszymi barionami, jakie znamy. Zauważmy teraz, że mocno związany układ złożony z cząstki S i jej antycząstki \bar{S} przy zerowym spinie i zerowym kręcie orbitalnym daje dziewięć stanów o zerowej liczbie barionowej i takimże całkowitym kręcie orbitalnym oraz ujemnej parzystości (ponieważ S i \bar{S} mają przeciwne parzystości wewnętrzne). W granicy symetrii tych dziewięć stanów jest zdegenerowanych i jeśli zaniedbamy wszystkie efekty łamiące symetrię oprócz tego, wyjątkowo silnego, który istnieje w S , a mianowicie dużej różnicy mas pomiędzy singletem izospinowym Λ a dubletem izospinowym (P, N) — to stany ($\bar{S}S$) lub ich liniowe kombinacje zgrupują się gdzie trzeba, w izospinowe multiplety odpowiadające pseudoskalarnym mezonom π, K, η i X^0 . To znaczy, odpowiadają one znanym mezonom pseudoskalarnym, z których wszystkie mają masy mniejsze niż 1 BeV. (W języku specjalistów stany ($\bar{S}S$) tworzą dziewięciowymiarową reprezentację grupy symetrii $SU(3) \times SU(3)$. Z punktu widzenia $SU(3)$, z drugiej strony, tworzymy iloczyn reprezentacji $\bar{3}$ przez reprezentację 3 otrzymując w ten sposób sprzężone ze sobą reprezentacje $SU(3)$ o wymiarach 1 i 8. Procedura mnożenia jest analogiczna do kombinowania reprezentacji krętu przy użyciu „modelu wektorowego“.)

Izospinowy singlet X^0 (rezonans $\eta\pi\pi$ przy 960 MeV) jest singletową reprezentacją $SU(3)$, a inne mezony tworzą dobrze znany oktet pseudoskalarny. Najbardziej oczywisty sposób łamania symetrii, ten, którego już używaliśmy, jest już wystarczający dla otrzymania wzoru Gell-Manna — Okubo dla oktetu. Wzór ten (jeśli π, K i η oznaczają teraz masy) jest następujący:

$$4K = 3\eta + \pi.$$

Ten związek stwierdzono doświadczalnie dla *kwadratów* mas, ale to jest problem, którego nie będziemy tu dyskutować. Należy jednak wskazać na fakt, że ten najprostszy model łamania symetrii (ten „ludowy model“ jak go nazywano) nie daje niczego takiego jak poprawna masa mezonu X^0 . Fakt ten sugeruje, że w modelu Sakaty (a także w opisanym poniżej modelu kwarków, który jest prawie identyczny z nim dla mezonów) singlet i oktet muszą mieć różne funkcje falowe; stąd z kolei wynika, że tą wewnętrzną symetrią, o którą nam chodzi jest raczej $SU(3)$ niż $SU(3) \times SU(3)$.

Przechodząc teraz do barionów można zaproponować ich tworzenie ze stanów ($\bar{S}SS$). Prostą możliwością stanowi kombinacja pseudoskalarne mezonu ($\bar{S}S$) z innym S . Na przykład barion Σ z $T = 1$, $Y = 0$ może być pomyślany jako pion i hiperon Λ związane w stanie $P_{1/2}$ (aby mieć dodatnią parzystość). W rzeczywistości π , K i η związane w stanach $P_{1/2}$ z Λ tworzą cały oktet barionowy razem z *oryginalnymi* P , N i Λ . To jest zagadkową i niepokojącą cechą modelu Sakaty. Gdyby tylko to było zagadkowe, mogłoby się to okazać zagadką z rodzaju tych, które wskazują drogę do nowych zasad fizycznych. Jednak, ponieważ nasze dotychczasowe sukcesy zależały od przyjęcia niezmienniczości $SU(3)$, musimy pójść dalej i zażądać, aby również nukleony wiązały się w stanie $P_{1/2}$ z elementami pseudoskalarne oktetu mezonowego — a to prowadzi do wielu stanów, których po prostu nie obserwuje się. Dobrymi przykładami tej wady modelu są stany barionowe o dodatniej dziwności, utworzone z mezonów K (nie \bar{K}) i nukleonów; stany te nie istnieją.

Ośmiokrotna droga

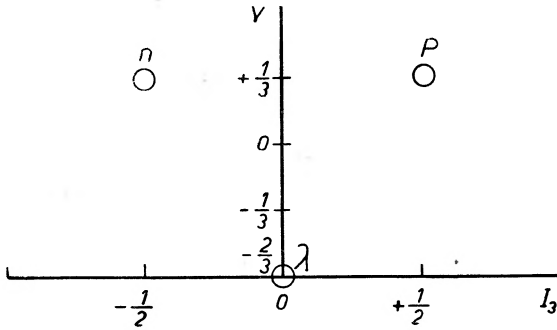
Wobec tego Gell-Mann i Yuval Ne'eman zaproponowali niezależnie alternatywny schemat symetrii $SU(3)$ znany jako „droga ośmiokrotna“. Zawiera on te cechy modelu Sakaty, które odniosły sukces, a równocześnie unika jego trudności. Według ośmiokrotnej drogi cząstki elementarne tworzą, w granicy zupełnej symetrii $SU(3)$, reprezentacje grupy symetrii $SU(3)$, a pozostają możliwe do rozróżnienia, kiedy oddziaływania łamiące symetrię są „włączone“. *Brakuje* jednak najniższej nietrywialnej reprezentacji o wymiarze 3 (na przykład sakatonu), a *fundamentalną* reprezentację 3 zastępuje się *regularną* reprezentacją 8. (*Regularną* lub *dołączoną* reprezentację grupy Lie, taką jak $SU(n)$, utożsamiamy z tą reprezentacją, która ma taki sam wymiar i własności transformacyjne co generatory grupy. Jeśli rozważamy grupę $SU(2)$ krętu, to fundamentalna reprezentacja ma spin 1/2 i wymiar 2, a regularna — spin 1 i wymiar 3. Krętową analogię drogi ośmiokrotnej stanowiłaby „droga trzykrotna“, którą można zrealizować jako grupę rzeczywistych obrotów w trójwymiarowej przestrzeni. Tak byłoby, gdybyśmy wyeliminowali fundamentalną (spinorową) reprezentację grupy $Su(2)$ i inne połówkowe kręty.)

Przy pomocy drogi ośmiokrotnej sklasyfikowano z powodzeniem dobrze ustalone niższe stany mezonów i barionów, a to dało szeroki zakres wglądów i częściowych sukcesów w zrozumieniu dynamiki tych silnych, elektromagne-

tycznych i słabych procesów, w których uczestniczą hadrony. W języku $SU(3)$ można zwięźle wytłumaczyć dlaczego model Sakaty nie potrafi dać oktetu barionowego; zauważmy, że szereg możliwych reprezentacji otrzymanych przez składanie \bar{S} , S i S nie zawiera reprezentacji 8.

Model kwarków

Postępowanie polegające na rezygnacji z reprezentacji fundamentalnej na rzecz regularnej albo, inaczej, na odrzuceniu reprezentacji spinorowych, a zatrzymaniu tylko tensorowych nosi czasem nazwę „dzielenia grupy przez jej środek“. Czyniąc tak w pewnym sensie pożeramy rękę, która nas karmi. Wyrzucanie fundamentalnego tripletu użytego dla przyrządzenia potrawy porównywano również z pewnymi procedurami stosowanymi we francuskiej kuchni.



Rys. 1. Diagram $SU(3)$ dla kwarków. W modelu kwarków z trzech ich rodzajów (p , n i λ) o ładunkach $1/3$ lub $2/3$ można zbudować wszystkie znane mezony i bariony

Najprostszą bazę modelu tripletowego stanowią *kwarki* (nazwane tak przez Gell-Manna), których liczby kwantowe są podane w tabelicy I (aby otrzymać liczby kwantowe antykwarków należy odwrócić wszystkie znaki w tabelicy). Jeden sposób reprezentowania kwarków pokazano na rys. 1.

Tablica I
Liczby kwantowe kwarków

	I	I_3	S	B	Y	Q
p	1/2	1/2	0	1/3	1/3	2/3
n	1/2	-1/2	0	1/3	1/3	-1/3
λ	0	0	-1	1/3	-2/3	-1/3

Kwarkowemu dubletowi izospinowemu (p , n) i singletowi λ przypisano nazwy, które przypominają nam triplet Sakaty. Zauważmy jednak, że liczba barionowa B jest wielokrotnością jednej trzeciej, co nadaje hiperładunkowi Y i ładunkowi elektrycznemu Q również charakter wielokrotności jednej trzeciej.

Kiedy nie ma słabych oddziaływań tak, że Y jest zachowane i w granicy symetrii, gdzie p , n i λ są zdegenerowane pod względem masy, λ jest „dziwne“ z punktu widzenia Y , ale zauważmy, że p jest równie „dziwne“ z punktu widzenia Q . Prawdę mówiąc, trzeci zachowany ładunek Q' może być zdefiniowany jako $Q' = U_3 + Y/2$, gdzie p jest singletem U -spinowym, a (n, λ) tworzy U -spinowy dublet, którego elementy mają odpowiednio $U_3 = 1/2$ i $-1/2$. Zatem dla n, λ, p otrzymujemy $Q' = 2/3, -1/3, -1/3$ tak, że z punktu widzenia Q' to właśnie n jest „dziwne“.

Kwarkowe recepty

Przez kombinację kwarków z antykwarkami w spinowo singletowym stanie S otrzymuje się oktet i singlet pseudoskalaranych mezonów dokładnie tak jak w przypadku Sakaty (patrz tablica II). Ładunki wychodzą prawidłowo, ponieważ Q i Y kwarków są po prostu przesunięte o jedną trzecią względem Q i Y sakatonu, a ponieważ Q i Y dla antykwarków są przesunięte w drugą stronę, więc całkowite przesunięcie znika.

Tablica II

Mezony utworzone z kwarków ($B = 0$)

Y	I	Funkcja falowa kwarków	Mezony $J^P = 0^-$	Mezony $J^P = 1^-$
0	1	$\bar{p}n, (\bar{p}p - \bar{n}n)/\sqrt{2}, \bar{n}p$	π^-, π^0, π^+	ρ^-, ρ^0, ρ^+
1	1/2	$\bar{\lambda}n, \bar{\lambda}p$	K^0, K^+	K^{*0}, K^{*+}
-1	1/2	$\bar{p}\lambda, \bar{n}\lambda$	K^-, K^0	K^{*-}, \bar{K}^{*0}
0	0	$(\bar{p}p + \bar{n}n - 2\bar{\lambda}\lambda)/\sqrt{6}$	η^0	—
0	0	$(\bar{p}p + \bar{n}n + \bar{\lambda}\lambda)/\sqrt{3}$	X^0	—
0	0	$(\bar{p}p + \bar{n}n)/\sqrt{2}$	—	ω^0
0	0	$\bar{\lambda}\lambda$	—	Φ^0

Mezony w modelu kwarków są stanami związanymi kwarków q i antykwarków \bar{q} . Oznaczamy te stany przez $(\bar{q}q)$, gdzie $q = (p, n, \lambda)$, a $\bar{q} = (\bar{p}, \bar{n}, \bar{\lambda})$. W tablicy oznaczenie $\bar{p}p$ wskazuje na iloczyn funkcji falowych \bar{p} i p . Najniższe stany mają zerowy kręt, a ponieważ q i \bar{q} mają przeciwne parzystości wewnętrzne, więc są to stany z $J^P = 0^-$ dla spinu kwarków $s = 0$ (mezony pseudoskalarne), a $J^P = 1^-$ dla spinu kwarków $s = 1$ (mezony wektorowe). Aktualne masy wzrastają od góry do dołu tablicy.

Obecny przepis na tworzenie barionów jest inny niż w przypadku Sakaty, ponieważ, aby otrzymać jednostkową liczbę barionową, musimy użyć kombinacji trzech kwarków, a nie jak w przypadku Sakaty, dwu sakatonów i antysakatonu (patrz tab. III). (Teoretycy grupowi z przyjemnością odnotowują, że $3 \times 3 \times 3$ zawiera 8; w samej rzeczy całkowity rozkład jest $1 + 8 + 8 + 10$.)

Tablica III

Bariony utworzone z kwarków ($B = 1$)

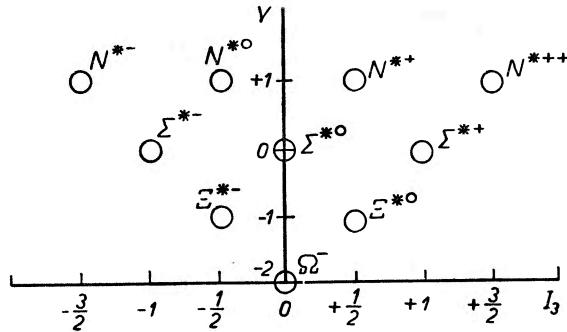
Y	I	Funkcja falowa kwarków $\begin{array}{ c c } \hline \square & \square \\ \hline \end{array}$	Bariony $J^P = (\frac{1}{2})^+$
1	1/2	nnp, ppn	N, P
0	0	$\lambda(np - pn)/\sqrt{2}$	Λ
0	1	$\lambda nn, \lambda(np + pn)/\sqrt{2}, \lambda pp$	$\Sigma^-, \Sigma^0, \Sigma^+$
-1	1/2	$\lambda \lambda n, \lambda \lambda p$	Ξ^-, Ξ^0
Y	I	Funkcja falowa kwarków $\begin{array}{ c c c c } \hline \square & \square & \square & \square \\ \hline \end{array}$	Bariony $J^P = (\frac{3}{2})^+$
1	3/2	nnn, nnp, npp, ppp	$\Delta^{*-}, \Delta^{*0}, \Delta^{*+}, \Delta^{*++}$
0	1	$\lambda nn, \lambda np, \lambda pp$	$\Sigma^{*-}, \Sigma^{*0}, \Sigma^{*+}$
-1	1/2	$\lambda \lambda n, \lambda \lambda p$	Ξ^{*-}, Ξ^{*0}
-2	0	$\lambda \lambda \lambda$	Ω^-

Bariony w modelu kwarków są stanami związanymi trzech kwarków (qqq), a antybariony (nie pokazane) stanami związanymi trzech antykwarków ($\bar{q}\bar{q}\bar{q}$). W tej tablicy zapis nnp na przykład oznacza iloczyn dwóch funkcji falowych n i jednej funkcji falowej p . Najniższe stany barionowe mają zerowy kręt orbitalny i wobec tego dodatnią parzystość. Dziesięć stanów o spinie kwarku $s = 3/2$ ma symetryczną strukturę spinową i kwarkową (zaznaczoną trzema kwadratami na jednej linii). Osiem stanów o spinie kwarku $s = 1/2$ ma mieszaną strukturę spinową i kwarkową (zaznaczoną trzema kwadratami nie leżącymi na jednej linii). W odroźnieniu od podanych przykładów bariony $(\frac{1}{2})^+$ nie mają, ogólnie biorąc, funkcji falowych, które można by napisać w postaci iloczynu funkcji falowej kwarku i funkcji falowej spinu. Aktualne masy wzrastają od góry do dołu tablicy.

W granicy symetrii $SU(3)$ oczekiwalibyśmy degeneracji stanów o danej przestrzeni i strukturze spinowej. Wybierając jako przykład najbardziej symetryczny przypadek, a mianowicie trzy λ w stanach S względem siebie i przy równoległych wszystkich spinach, otrzymujemy liczby kwantowe $I = 0, S = -3$ i $Q = -1$ oraz spin $3/2$. To jest słynna Ω^- ! Ω^- , ze swoją całkowitą symetrią spinową i izospinową jest reprezentantem klasy dziesięciu elementów: deкуплетu (bądź decymetu) $SU(3)$ lub reprezentacji 10, która jest przedstawiona na rys. 2.

Ponieważ deкуплет barionowy składa się z dziesięciu cząstek o spinie i parzystości $(\frac{3}{2})^+$, których istnienie jest dobrze ustalone, możemy to odnotować jako jeszcze jeden triumf ośmiokrotnej drogi. Ale co będzie z oktetem barionowym, który okazał się taki szkodliwy dla modelu Sakaty? Aby pomóc w zrozumieniu jak ten oktet tutaj pasuje, zbadajmy izospinową strukturę deкуплетu. Ω^- , mając dziwność $S = -3$, składa się z trzech kwarków λ i wobec tego ma $I = 0$. Ξ^* o dziwności $S = -2$ ma dwa kwarki λ i jeden lub drugi element dubletu (p, n), a wobec tego ma $I = 1/2$. Σ^* ma dwa elementy dubletu

i wobec tego może mieć $I = 0$ lub $I = 1$, ale bez wahania wybieramy $I = 1$ z następującego powodu: tworzymy nieredukowalną reprezentację $SU(3)$, która, jak wiemy, musi mieć określoną strukturę względem wymiany kwarków. Ponieważ wszystkie stany orbitalne i spinowe są identyczne dla dekapletu,



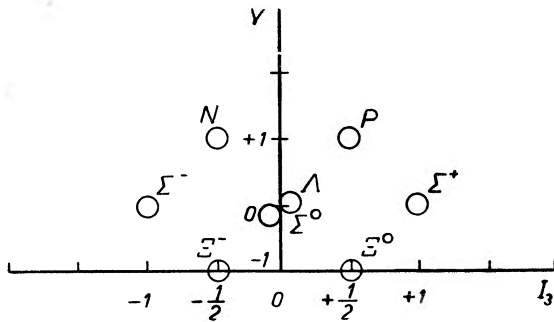
Rys. 2. Barionowy dekaplet $SU(3)$ dla barionów o spinie i parzystości $3/2^+$. Dziesiątym barionem, u dołu trójkąta, jest Ω^- , przewidziany przez teorię $SU(3)$

więc wszystkie stany izospinowe muszą posiadać tę samą symetrię; a ponieważ Ω^- jest w sposób oczywisty symetryczna izospinowo, musimy więc wybrać stan izospinowo symetryczny dla Σ^* , a tym stanem jest stan z $I = 1$. Podobnie N^* , który zawiera trzy elementy dubletu, musi być w stanie izospinowo symetrycznym, to jest w stanie z $I = 3/2$.

Czytelnik mógł zauważyć, że nie wypowiedzieliśmy się wyraźnie na temat całkowitej symetrii stanu trójkwarkowego. Jeśli kwarki są zwykłymi fermionami (a nie na przykład parafermionami), to całkowita symetria musi być oczywiście antysymetryczna. Ponieważ skonstruowaliśmy symetryczną funkcję falową spinowo-izospinową, więc ciężar odpowiedzialności za antysymetrię spoczywa na przestrzennej funkcji falowej, którą zbudowaliśmy tak, aby zawierała tylko fale S . To jest nieładne, ale nie niemożliwe; rozważmy na przykład trzy ortogonalne jednocząstkowe funkcje falowe S i utwórzmy ich antysymetryzowany iloczyn. Nie jest to niemożliwe, ale trudno zrozumieć, dlaczego to miałyby być stan leżący szczególnie nisko. Jeszcze trudniej jest przyjąć, że ta sama szczególna przestrzenna funkcja falowa (jakakolwiek jest) łączy się z inną reprezentacją $SU(3)$. Przyjąwszy jednak, że tak jest, jakie inne można utworzyć spinowo-izospinowe symetryczne funkcje falowe, aby je mnożyć przez tę przestrzenną funkcję falową?

Chociaż stan całkowicie antysymetryczny izospinowo jest skądinąd możliwy, to musiałby on towarzyszyć całkowicie antysymetrycznemu stanowi spinowemu, a ten jest niemożliwy, ponieważ są tylko dwie niezależne funkcje falowe o spinie $1/2$. Istnieją natomiast stany o mieszanej symetrii spinowej (antysymetryczne tylko dla jednej pary) i stany o mieszanej symetrii izospinowej takie, że można utworzyć stan wypadkowy o całkowitej symetrii. Oczywiście nie ma stanu o trzech λ , który miałby taki charakter; stan o dwóch λ ma $I = 1/2$,

stany z jednym λ mają $I = 0$ i $I = 1$, a stan bez λ ma $I = 1/2$, ponieważ ma on tę samą strukturę, co stan spinowy o mieszanej symetrii (który jest również utworzony z trzech połówkowych spinów), a ten ma spin $1/2$. Reprezentacją, którą utworzyliśmy, jest oktet barionowy $J^P = 1/2^+$ przedstawiony na rys. 3.



Rys. 3. Barionowy oktet $SU(3)$. Diagram ten przedstawia bariony o spinie i parzystości $1/2^+$

Ponieważ do tej pory ograniczyliśmy się do stanów S unikając w ten sposób problemu oddziaływania spin—orbita, mogliśmy efektywnie traktować spin jako wewnętrzną grupę symetrii $SU(2)$ i skonstruować pewne reprezentacje grupy $SU(2) \times SU(3)$. W rzeczywistości wyczerpaliśmy te możliwe trójkwarkowe reprezentacje tej grupy, które mogą być stowarzyszone z daną antysymetryczną przestrzenną funkcją falową zawierającą tylko fale S .

Grupy o wyższej symetrii

Przeszło rok temu Feza Gürsey i Luigi Radicati oraz Bunji Sakita zaproponowali znacznie dalej sięgające uogólnienie. Rozważali oni dozwolone stany hadronowe jako reprezentacje (w granicy symetrii) grupy $SU(6)$. To uogólnienie ma kwarkową interpretację, w której jest sześć równoważnych kwarków, przy czym dwie możliwości ustawienia spinu p , n i λ (do góry i na dół) są traktowane osobno. Ta interpretacja sprowadza się do przyjęcia spinowej i izospinowej niezależności sił jądrowych. Cztery krętowe podstany barionu z $J = 3/2$, takiego jak na przykład Ω^- , są uważane za cztery elementy większego multipletu $SU(6)$. Cały dekuplet barionowy, składający się z 40 stanów łączy się następnie z oktetem barionowym składającym się z 16 stanów, tworząc jedną reprezentację $SU(6)$, a mianowicie reprezentację 56-stanową.

Atrakcyjną cechą tej propozycji jest to, że daje ona sposób na powiązanie między sobą statycznych własności oktetu i dekupletu barionowego razem z ich momentami magnetycznymi i przesunięciami masowymi przy łamaniu symetrii. Zakładając zaś oddziaływania niezmiennicze względem $SU(6)$ pomiędzy różnymi elementami $SU(6)$, może ona dać kompletną dynamiczną podstawę dla silnych oddziaływań. Co najważniejsze, wszystko to odbywa się

bez wprowadzania explicite kwarków, które, pomimo swych pedagogicznych zalet, mogą nie istnieć.

Relatywistyczne wersje tej teorii, które można, znowu pedagogicznie, rozpatrywać jako oparte raczej na kwarkach opisywanych przez czteroskładnikowe, dirakowskie funkcje falowe niż na dwuskładnikowych, nierelatywistycznych kwarkach typu Pauliego, zostały zaproponowane przez szereg autorów po początkowych sukcesach $SU(6)$. Wersje te ukazują się pod nazwami $\tilde{U}(12)$, $M(12)$ i $SL(6, c)$. Szczegóły tych teorii, które są jeszcze nieco niejasne, nie potrzebują na razie obchodzić czytelnika nie-specjalisty.

Powracając teraz do mezonów, przypomnijmy, że oktet pseudoskalarny i singlet X^0 były rozpatrywane w modelu kwarków $SU(3)$ jako dziewięć stanów utworzonych przez złożenie kwarku i antykwarku w spinowo-singletowych stanach S . Postępując tak jak dla barionów, tworzymy również dziewięć spinowo-tripletowych stanów S (co daje 27 stanów według klasyfikacji $SU(6)$). Te dziewięć mezonów o jednostkowym spinie tworzy wektorowe, mezonowe reprezentacje $SU(3)$ o wymiarach 1 i 8, a równocześnie tworzy razem z ośmioma stanami pseudoskalarnego oktetu mezonowego 35-wymiarową, regularną reprezentację $SU(6)$. (Regularna reprezentacja $SU(n)$ ma zawsze wymiar n^2-1 . Zauważono, że jeśli mezony tworzą regularną reprezentację danej grupy, to mamy gwarancję, że oddziaływania Yukawy z barionami mogą być tak sformułowane, aby były niezmiennicze względem tej grupy, ponieważ iloczyn danej reprezentacji przez reprezentację sprzężoną do niej zawiera zawsze reprezentację regularną.) Wliczając $X^0(960)$, które pozostaje singletem $SU(6)$, mamy 36 stanów mezonowych, które można uważać za stowarzyszone (w granicy symetrii) z najniższymi stanami S układu kwark—antykwark w modelu kwarków, jeśli rozszczepienie mas jest związane z efektami łamiącymi symetrię. Z drugiej strony X^0 może mieć inną funkcję falową S .

Końcowe uwagi o symetriach

W tym bardzo krótkim przeglądzie wskazano, można mieć nadzieję, pewne punkty, które są istotne dla zrozumienia unitarnych symetrii cząstek elementarnych, ale braki jego będą zupełnie jasne dla każdego, kto się zapoznał z tą dziedziną. Po pierwsze większość interesującej fizyki pozostała poza tym artykułem. Lista tego, co w nim pominięto, byłaby dłuższa od samego artykułu; ważniejsze z pominiętych problemów są następujące:

1. Wielu wierzy, że najpiękniejszą cechą unitarno-symetrycznych teorii nie są przewidywania wymiarów reprezentacji, ale elegancki sposób, w jaki teorie te pozwalają na łamanie symetrii przez oddziaływania średnio silne, elektromagnetyczne i słabe, dając widma masowe i reguły wyboru dla przejść. To jest oczywiście sedno fizyki.

2. Pytania o ewentualne rzeczywiste istnienie kwarków lub innych tripletów, o to, jakie inne własności mogą one posiadać i jak można ich szukać, nie były dyskutowane.

3. Nie poruszono sprawy przypisania wyższych reprezentacji barionowych i mezonowych do reprezentacji $SU(6)$ innych niż 1, 35 i 56, po części dlatego, że sytuacja doświadczalna nie jest jeszcze jasna.

Nic z powyższego nie pochodzi od autora, oprócz ewentualnych błędów, które mogły wśliznąć się do artykułu. Miejmy nadzieję, że da on nie-specjalistom pewną podstawę dla oceny nowych odkryć, które mogą być dokonane. Jako przykład tego, czego ewentualnie możemy się spodziewać, niech posłuży następujące rozumowanie będące trywialnym przedłużeniem powyższych idei:

Składając kwark i antykwark w stanie P przewiduje się oktety i singlety $SU(3)$ po jednym dla $J^P = 0^+, 1^+ \text{ i } 2^+$, otrzymane przez dodanie jednostkowego spinu kwarku do jednostkowego krętu orbitalnego. Stany tych multipletów z $I = 0$ (przez analogię z odpowiednimi stanami pozytronium) powinny być parzyste względem sprzężenia ładunkowego i rozpadać się przede wszystkim na $\pi^+ + \pi^-$. W dodatku powinien istnieć oktet (a może i singlet) o zerowym spinie kwarku, a więc z $J^P = 1^+$, który (przez tę samą analogię) powinien mieć stan z $I = 0$, nieparzysty względem sprzężenia ładunkowego czyli z zabronionym rozpadem na dwa piony. Rozwijający się schemat wyższych rezonansów mezonowych wydaje się, według autora, dawać dużą szansę na spełnienie tych przewidywań.

Definicje

1. Liczba barionowa B i dziwność S nie mogą być zdefiniowane inaczej, jak tylko jako liczby, które mogą być stowarzyszone z cząstkami tak, aby obowiązywała zasada ich zachowania. Liczba barionowa jest zawsze zachowywana addytywnie, a dziwność, która może nie być zachowana w słabych oddziaływaniach, jest addytywnie zachowywana w oddziaływaniach silnych i elektromagnetycznych. Mezony są hadronami o liczbie barionowej 0; innym hadronom przypisujemy liczby barionowe, które są całkowite i dodatnie (dla cząstek) lub całkowite i ujemne (dla antycząstek). Dziwność została wprowadzona w 1953 roku przez Gell-Manna i Nishijimę dla opisanego stowarzyszonej produkcji pewnych hadronów („dziwnych“ cząstek) w silnym oddziaływaniu, po którym następuje ich rozpad poprzez słabe oddziaływanie. Wzór Gell-Manna—Nishijimy daje $Q = I_3 + 1/2(B + S)$. Trzy składowe izospinu i hiperładunek $Y = B + S$ mogą być utożsamiane z czterema spośród ośmiu generatorów grupy $SU(3)$; cztery inne odnoszą się do zmiany dziwności. Kwarki mogą mieć ułamkową liczbę barionową i ułamkowy ładunek elektryczny.

2. *Efektami łamiącymi symetrię* nazywamy te, które w kontekście danej symetrii naruszają tę symetrię. Na przykład niezależność ładunkowa sił jądrowych jest łamana przez oddziaływanie elektromagnetyczne, przez różnicę mas między neutronem a protonem (która sama może być efektem elektromagnetycznym) i przez słabe oddziaływanie.

3. Mezony o spinie 0 i 1 są często nazywane podług własności ich pól względem przestrzennego obrotu i odbicia w układzie spoczynkowym. Mezony o spinie — parzystości $J^P = 0^+, 0^-, 1^+ \text{ i } 1^-$ zwiemy odpowiednio mezonami *skalarnymi*, *pseudoskalarnymi*, *pseudowektorowymi* i *wektorowymi*.

Tłumaczył Krzysztof Rybicki

Andrzej Kotański

Instytut Fizyki Uniwersytetu Jagiellońskiego

Podstawowe pojęcia z teorii grup stosowane w fizyce cząstek elementarnych

1. Wstęp

W ostatnich kilku latach w związku z trudnościami zbudowania dynamicznej teorii oddziaływań cząstek elementarnych zwrócono uwagę na symetrie tych oddziaływań. Okazało się, że mimo braku teorii (niektórzy fizycy sądzą nawet, że jej skonstruowanie jest w ogóle niemożliwe) można wyjaśnić wiele podstawowych własności cząstek przy pomocy rozważań grupowo-teoretycznych. Dla przykładu, wyprowadzono liczne związki między masami cząstek, ich momentami magnetycznymi (a więc np. z grupy SU_6 stosunek momentów magnetycznych neutronu i protonu wypadł równy $-\frac{2}{3}$, co zgadza się z doświadczeniem), a także przekrojami czynnymi na oddziaływania. Spowodowało to wzrost zainteresowania teorią grup wśród fizyków.

W niniejszym artykule omówione są w sposób poglądowy podstawowe pojęcia z teorii grup, które są ostatnio często używane w fizyce cząstek. I tak na początku podajemy odpowiedniości między takimi pojęciami, jak unitarność i nieredukowalność reprezentacji czy też iloczyn prosty, a znanymi pojęciami z mechaniki kwantowej. W dalszym ciągu opisujemy, w jaki sposób wyciąga się wnioski o oddziaływaniach z założenia pewnej symetrii. Ostatnia część zawiera wreszcie krótkie omówienie i fizyczną interpretację następujących grup: translacji wzdłuż prostej, obrotów w zwykłej przestrzeni, SU_2 oraz grupy Poincarégo.

Czytelnik zechce wybaczyć, jeśli popularyzacja tych zagadnień została nieraz dokonana kosztem braku ścisłości matematycznej.

2. Opis cząstek swobodnych

Stany opisywane przez mechanikę kwantową podlegają przekształceniom różnego rodzaju. W wyniku ich działania otrzymujemy nowe stany

$$|a'\rangle = P|a\rangle. \quad (2.1)$$

Jeśli stany $|\alpha\rangle$ tworzą układ zupełny (w jakiejś podprzestrzeni), to nowe stany $|\alpha'\rangle$ dadzą się wyrazić jako kombinacje liniowe stanów $|\alpha\rangle$

$$|\alpha'\rangle = \sum_{\alpha} D(P)_{\alpha'\alpha} |\alpha\rangle. \quad (2.2)$$

Często zdarza się, że układ ortonormalny jest przekształcony przez transformację P w układ, który znów jest ortonormalny. Warunek ten można zapisać przy pomocy macierzy D określonej w równaniu (2.2) — macierz D musi być wtedy unitarna.

Jeśli przekształcenia P tworzą grupę, tzn:

- (a) $P_1 P_2 = P_3$ (złożenie dwu przekształceń jest znów przekształceniem tego samego typu),
- (b) istnieje P^{-1} (element odwrotny do każdego przekształcenia),
- (c) istnieje P_0 (przekształcenie jednostkowe, tj. nie zmieniające żadnego stanu),

to macierze $D(P)$ stanowią *unitarną reprezentację liniową* tej grupy.

Możemy więc następująco zestawić odpowiadające sobie własności reprezentacji grup i własności stanów z mechaniki kwantowej:

zasada superpozycji — liniowość reprezentacji
zachowywanie normy — unitarność reprezentacji.

Jednym słowem liniowość i unitarność reprezentacji są ściśle związane z podstawowymi założeniami mechaniki kwantowej i dlatego w dalszym ciągu będziemy się interesować tylko reprezentacjami liniowymi, unitarnymi.

Cząstki opisujemy przy pomocy stanów, w których mogą się one znajdować. Dlatego z punktu widzenia danej grupy przekształceń z cząstką będzie związana pewna reprezentacja unitarna.

Zastanówmy się teraz, w jaki sposób można opisać układ dwu cząstek nie oddziałujących ze sobą. W mechanice kwantowej używamy do tego iloczynu dwu funkcji falowych:

$$\psi_1 \psi_2 \quad \text{albo} \quad |\alpha_1\rangle |\alpha_2\rangle. \quad (2.3)$$

Wtedy na przykład prawdopodobieństwo znalezienia tych cząstek w pewnym miejscu jest iloczynem prawdopodobieństw znalezienia tam cząstki pierwszej i drugiej. Taki stan dwu cząstek będzie się zachowywać przy przekształceniach w sposób następujący

$$|\alpha'_1\rangle |\alpha'_2\rangle = \sum_{\alpha_1 \alpha_2} D^{(1)}(P)_{\alpha'_1 \alpha_1} D^{(2)}(P)_{\alpha'_2 \alpha_2} |\alpha_1\rangle |\alpha_2\rangle. \quad (2.4)$$

Jednym słowem przekształceniu odpowiada teraz pewna macierz D , która jest *iloczynem prostym* macierzy związanych z pojedynczymi cząstkami, co oznaczamy:

$$D = D^{(1)} \times D^{(2)}. \quad (2.5)$$

Macierz, która jest iloczynem prostym jakichś macierzy, daje się na ogół przedstawić w postaci sumy prostej innych macierzy. Związana z nią repre-

zentacja nazywa się wówczas *redukowalną*. Jeśli więc stan złożony z dwu cząstek jest opisywany przez reprezentację redukowalną, to przez analogię można uznać, że stan opisywany przez reprezentację *nieredurowalną* odpowiada układowi *elementarnemu* (tzn. nie złożonemu z części składowych).

Mamy w ten sposób dalsze odpowiedniości:

cząstki nieoddziałujące — iloczyn prosty reprezentacji
układ elementarny — reprezentacja nieredurowalna.

Zwróćmy uwagę, że termin „elementarny“ jest użyty tutaj w sensie względnym; układ, który jest elementarny z punktu widzenia pewnej grupy może nie być elementarnym z punktu widzenia innej.

Ponieważ układy złożone z dwu lub więcej cząstek opisywane są przez iloczyny proste reprezentacji, zaś w teorii grup wygodniej jest operować reprezentacjami nieredurowalnymi, ważną rolę gra tu pojęcie redukcji iloczynu prostego na sumę prostą reprezentacji nieredurowalnych. Otóż jeśli mamy macierz

$$D'_{\alpha_1\alpha_2\alpha_1\alpha_2} = D'_{\alpha_1\alpha_1} D'_{\alpha_2\alpha_2} \quad \text{czyli} \quad D = D^{(1)} \otimes D^{(2)}, \quad (2.6)$$

to przez transformację unitarną (staramy się zwykle, by była ona ortogonalna) można ją doprowadzić do postaci blokowej:

$$\begin{pmatrix} D^{(3)} & 0 & \dots \\ 0 & D^{(4)} & \\ \vdots & & \ddots \end{pmatrix} \quad (2.7)$$

i przy tym wymiary występujących tu macierzy spełniają związek:

$$\dim D^{(1)} \cdot \dim D^{(2)} = \dim D^{(3)} + \dim D^{(4)} + \dots \quad (2.8)$$

Oznaczmy macierz przekształcającą iloczyn prosty (2.6) na sumę prostą (2.7) przez C . Wówczas macierz:

$$\sum_{\alpha_1\alpha_2} \sum_{\alpha_1\alpha_2} C_{\alpha',\alpha_1\alpha_2} D'_{\alpha_1\alpha_1} D'_{\alpha_2\alpha_2} C_{\alpha_1\alpha_2,\alpha}^{-1} \quad (2.9)$$

jest już w postaci quasi-diagonalnej (2.7). Analogicznie trzeba też przekształcić stany

$$|a\rangle = \sum_{\alpha_1\alpha_2} C_{a,\alpha_1\alpha_2} |\alpha_1\rangle |\alpha_2\rangle \quad (2.10)$$

tak, by nadal zachodził związek (2.2). Jeśli wszystkie reprezentacje $D^{(i)}$ występujące w (2.7) są nieredurowalne, to proces redukcji uważamy za zakończony, jeśli zaś nadto pierwotne reprezentacje $D^{(1)}$ i $D^{(2)}$ są również nieredurowalne, to elementy macierzy C nazywamy *współczynnikami Clebscha-Gordana* lub *współczynnikami Wignera*.

Redukcję iloczynu prostego zapisujemy symbolicznie

$$D^c \otimes D^{c'} = \bigoplus_i D^{ci}, \quad (2.11)$$

wymieniając po prawej stronie reprezentacje występujące w sumie prostej. Rozkład (2.11) nazywamy *szeregiem Clebscha-Gordana*. Jeśli po prawej stronie wzoru (2.11) żadna reprezentacja nie powtarza się i własność tę ma iloczyn prosty każdych dwu reprezentacji nieredukowalnych danej grupy, to grupę tę nazywamy *prosto redukowalną*. Jeśli zaś tak nie jest, trzeba w jakiś sposób rozróżniać identyczne reprezentacje występujące w szeregu Clebscha-Gordana. Używamy do tego tzw. *parametrów zwyrodnienia*.

W grupach używanych w fizyce każda reprezentacja unitarna daje się przedstawić w postaci sumy prostej pewnych reprezentacji nieredukowalnych. Twierdzenie to jest prawdziwe dla grup zwartych, a nawet dla niektórych grup niezwartych, na przykład dla grupy translacji czy grupy Poincarégo (tj. grupy obrotów i translacji w czasoprzestrzeni). Wskutek tego reprezentacje nieredukowalne stanowią bazę wszystkich reprezentacji, zaś operowanie nimi jest prostsze. Dlatego w zastosowaniach fizycznych umiejętność redukowania iloczynów prostych jest bardzo ważna.

Fizyka stosującego pewną grupę do opisu cząstek interesują przede wszystkim następujące informacje:

1. sens fizyczny przekształceń grupy,
2. jakie nieredukowalne reprezentacje unitarne posiada dana grupa,
3. jakie reprezentacje występują w iloczynie prostym (szereg Clebscha-Gordana),

4. wartości liczbowe współczynników Clebscha-Gordana.

Omówimy teraz krótko te zagadnienia.

W przypadku przekształceń zależnych od pewnych parametrów ciągłych wystarczy zbadać sens fizyczny *generatorów*, tzn. operatorów związanych z niezależnymi przekształceniami odpowiadającymi nieskończonej małym wartościom parametrów. Na przykład dla grupy obrotów w zwykłej przestrzeni generatory odpowiadają obrotom wokół trzech osi układu współrzędnych.

Na czym polega znajomość reprezentacji? Trzeba wiedzieć, jakie są reprezentacje (a więc czym są numerowane) i jakie stany należą do tych reprezentacji. Operatory, których wartości własne numerują reprezentacje (tzw. *operatory Casimira*), stanowią układ niezmienników, które dadzą się zbudować z generatorów. Można to uzasadnić w następujący sposób. Niech pewien stan transformuje się według reprezentacji D^x . Po pewnym przekształceniu P stan ten znów musi należeć do reprezentacji D^x (wynika to z faktu, że iloczynowi przekształceń odpowiada iloczyn reprezentacji). Znaczy to, że jeśli przed przekształceniem stan był stanem własnym operatorów Casimira do jakichś wartości własnych, to po tym przekształceniu musi on znów być ich stanem własnym do tych samych wartości własnych. Stąd wniosek, że operatory Casimira muszą komutować z przekształceniami P , a więc muszą być niezmiennikami. Na przykład dla grupy obrotów mamy tylko jeden operator Casimira: $J^2 = J_x^2 + J_y^2 + J_z^2$, który komutuje z generatorami J_x, J_y, J_z .

Natomiast operatory, których wartości własne numerują stany w reprezentacji, stanowią układ komutujących ze sobą generatorów. Gdyby bowiem

ze sobą nie komutowały, to stan nie mógłby być stanem własnym jednocześnie ich wszystkich.

Wreszcie dwie ostatnie informacje (o postaci szeregu C-G i o współczynnikach C-G) są istotne, gdy interesujemy się stanami dwu- i wielocząstkowymi.

3. Uwzględnienie oddziaływania

Bezpośrednio dostępne obserwacji są jedynie stany asymptotyczne, tzn. stany cząstek swobodnych znajdujących się daleko od obszaru oddziaływania. Stany, które pojawiają się po reakcji wiążemy ze stanami początkowymi przy pomocy macierzy S :

$$|\text{końc.}\rangle = S|\text{pocz.}\rangle, \quad (3.1)$$

w której tkwi cała informacja o oddziaływaniu. Dlatego interesują nas elementy macierzy S :

$$\langle \alpha_1 | S | \alpha_2 \rangle, \quad (3.2)$$

które nazywamy amplitudami rozpraszania.

Zobaczmy teraz, co się dzieje z amplitudą rozpraszania, gdy oddziaływanie jest niezmiennicze względem pewnej grupy przekształceń (a więc, gdy macierz S komutuje z generatorami tej grupy). Mamy wtedy:

$$PSP^{-1} = S. \quad (3.3)$$

Zatem

$$\begin{aligned} \langle \alpha_1 | S | \alpha_2 \rangle &= \langle \alpha_1 | P^{-1} P S P^{-1} P | \alpha_2 \rangle = \langle \alpha_1 | D^{(1)}(P)^+ S D^{(2)}(P) | \alpha_2 \rangle = \\ &= \sum_{\alpha'_1 \alpha'_2} D^{(1)}(P)_{\alpha_1 \alpha'_1}^+ D^{(2)}(P)_{\alpha_2 \alpha'_2} \langle \alpha'_1 | S | \alpha'_2 \rangle = P \langle \alpha_1 | S | \alpha_2 \rangle. \end{aligned} \quad (3.4)$$

Obiekt należący do pewnej reprezentacji zachowuje się przy transformacji P w następujący sposób:

$$P\Phi(a) = \Sigma D_{aa'} \Phi(a'), \quad (3.5)$$

zaś w naszym przypadku (3.4) mamy zachowanie się typu:

$$P\Phi(a) = \Phi(a), \quad (3.6)$$

czyli element macierzy S transformuje się według reprezentacji jednostkowej. Innymi słowy w iloczynie $D^{(1)+} \otimes D^{(2)}$ musi się znajdować reprezentacja jednostkowa. W teorii grup dowodzi się, że o ile $D^{(1)}$ i $D^{(2)}$ są nieredukowalne, to reprezentacja jednostkowa pojawi się w iloczynie $D^{(1)+} \otimes D^{(2)}$ tylko wtedy, gdy reprezentacje te są równoważne.

Jeśli teraz oznaczymy nasz stan

$$|\alpha_i\rangle = |C_i n_i\rangle, \quad (3.7)$$

gdzie C_i są wartościami własnymi operatorów Casimira (a więc podają, do jakiej reprezentacji należy ten stan), zaś n_i — wartościami własnymi opera-

torów numerujących stany w reprezentacji, to wzór (3.4) będziemy mogli przepisać w postaci:

$$\langle C_1 n_1 | S | C_2 n_2 \rangle = \sum_{n_1' n_2'} D^{(1)}(P)_{n_1 n_1'}^+ D^{(2)}(P)_{n_2 n_2'} \langle C_1 n_1' | S | C_2 n_2' \rangle. \quad (3.8)$$

Ponieważ zaś reprezentacje $D^{(1)}$ i $D^{(2)}$ mają być równoważne, otrzymujemy: $C_1 = C_2$. Dla grupy obrotów będzie to oznaczać zachowanie całkowitego krętu, gdyż taki właśnie sens ma tam operator Casimira.

Można, co więcej, pokazać, że wartości własne operatorów numerujących stany muszą być równe w stanach początkowym i końcowym:

$$\langle C_1 n_1 | S | C_2 n_2 \rangle = \delta_{c_1 c_2} \delta_{n_1 n_2} \langle C_1 || S || C_2 \rangle, \quad (3.9)$$

a więc że $n_1 = n_2$ (dla grupy obrotów jest to zachowanie rzutu krętu na oś z) i że ostatni czynnik po prawej stronie (zwany amplitudą niezmienniczą) we wzorze (3.9) nie zależy od n_i . Wzór (3.9) jest szczególnym przypadkiem bardzo ważnego *twierdzenia Eckarta-Wignera*, stosowanego często przy badaniu amplitud rozpraszania i rozpadów.

Zobaczmy jeszcze, co się dzieje, jeśli macierz S nie jest niezmiennicza, lub też jest niezmiennicza tylko w przybliżeniu (na przykład obecnie sądzi się powszechnie, że macierz S jest w przybliżeniu niezmiennicza względem grupy SU_3).

Otóż wnioski z niezmienniczości wyciąga się przede wszystkim korzystając z:

- a) szeregu Clebscha-Gordana,
- b) twierdzenia Eckarta-Wignera.

W procesach spotykanych w fizyce mamy zwykle więcej niż jedną cząstkę na wejściu lub na wyjściu, a więc stany początkowy i końcowy transformują się według reprezentacji redukowalnych. Trzeba je więc zredukować (przy użyciu szeregu C-G i współczynników C-G), a potem dopiero można stosować twierdzenie Eckarta-Wignera.

Jeśli teraz macierz S nie komutuje z generatorami grupy, nie zmienia to naszego postępowania w pierwszym etapie: i tak musimy dokonać redukcji stanu początkowego i końcowego. Natomiast przy stosowaniu twierdzenia Eckarta-Wignera musimy wprowadzić pewne modyfikacje.

Każdą funkcję (czy też operator) można rozłożyć na składniki transformujące się tak jak funkcje należące do reprezentacji nieredukowalnych. W szczególności

$$S = \sum_C \sum_{nC} S_n^C. \quad (3.10)$$

Wystarczy więc rozważyć element macierzowy

$$\langle C_1 n_1 | S_n^C | C_2 n_2 \rangle. \quad (3.11)$$

W nim wyrażenie $S_n^C | C_2 n_2 \rangle$ transformuje się według iloczynu $D^C \otimes D^{C_2}$. Przejście opisane wyrażeniem (3.11) jest możliwe tylko wtedy, gdy $D^{C_1} \in D^C \otimes D^{C_2}$.

Dla grupy obrotów proces taki można opisać następująco: oddziaływanie S zmienia kręt C_2 o C , w wyniku czego otrzymujemy kręt C_1 .

Dla elementu macierzewego (3.11) twierdzenie Eckarta-Wignera przybiera postać:

$$\langle C_1 n_1 | S_n^C | C_2 n_2 \rangle = \langle C_1 n_1 | C n_1 C_2 n_2 \rangle \langle C_1 | S^C | C_2 \rangle, \quad (3.12)$$

przy czym pierwszy czynnik po prawej stronie jest współczynnikiem C-G. Jeśli, w szczególności, C oznacza reprezentację jednostkową, to z równania (3.12) otrzymamy poprzednio podaną postać twierdzenia E-W (3.9).

4. Przykłady grup stosowanych w fizyce

a) Grupa translacji na prostej

Jest to jedna z najprostszych grup. Mamy tu tylko jeden generator transformacji infinitezymalnych P . Ogólnej transformacji (przesunięciu o odcinek a) odpowiada operator unitarny:

$$e^{iaP}, \quad (4.1)$$

gdzie a jest pewną liczbą rzeczywistą, zaś P jest operatorem hermitowskim.

Reprezentacje numerowane są (rzeczywistymi) wartościami własnymi p operatora P . Wszystkie reprezentacje są jednowymiarowe, a więc rozróżnianie stanów należących do jakiejś reprezentacji nie jest konieczne. Explicite,

$$D^p(a) = e^{ipa}. \quad (4.2)$$

Iloczyn prosty dwu reprezentacji dany jest wzorem

$$D^{p_1} \otimes D^{p_2} = D^p, \quad (4.3)$$

gdzie $p = p_1 + p_2$. Szereg C-G składa się więc tylko z jednego wyrazu i grupa jest, oczywiście, prosto redukowalna.

Współczynniki C-G są wszystkie równe jedności.

Fizycznie niezmienniczość względem translacji daje prawo zachowania pędu, zaś wartości własne p mają sens składowej pędu. Ma się rozumieć, jeśli chcemy otrzymać prawo zachowania fizycznego pędu, musimy rozważyć grupę translacji w przestrzeni trój- lub czterowymiarowej, której struktura nie różni się jednak od struktury omawianej tu grupy jednowymiarowej.

Twierdzenie E-W zastosowane do procesu z dwoma cząstkami na wejściu i dwoma na wyjściu

$$\langle p_1 p_2 | S | p_3 p_4 \rangle \quad (4.4)$$

daje (po redukcji iloczynów reprezentacji):

$$p_1 + p_2 = p_3 + p_4 \quad (4.5)$$

i podobnie dla innych procesów.

b) Grupy obrotów trójwymiarowych O_3

Mamy w tym przypadku trzy generatory: J_x, J_y, J_z . Reprezentacje unitarne numerowane są przez wartości własne operatora Casimira $J^2 = J_x^2 + J_y^2 + J_z^2$, które oznaczamy przez $j(j+1)$. Liczba j może przebiegać liczby całkowite nieujemne, mamy więc ciąg reprezentacji nieredukowalnych. Reprezentacja odpowiadająca danemu j ma wymiar $2j+1$.

Stany w reprezentacji są numerowane kompletem komutujących ze sobą generatorów. W naszym przypadku zawiera on tylko jeden generator (na przykład J_z), którego wartości własne oznaczamy literą m . Dla danego j liczba m przebiega wartości od $-j$ do j . Macierz reprezentacji oznaczamy symbolem:

$$D^j(\alpha, \beta, \gamma)_{mm'}, \quad (4.6)$$

gdzie α, β, γ są parametrami charakteryzującymi ogólny obrót (na przykład kątami Eulera). Fizycznie stany należące do reprezentacji j opisują różne orientacje spinu cząstki o spinie j .

Szereg $C-G$ można zapisać w postaci:

$$D^{j_1} \otimes D^{j_2} = \bigoplus_{j=|j_1-j_2|}^{j_1+j_2} D^j, \quad (4.7)$$

przy czym po prawej stronie każde D^j występuje w sumie prostej dokładnie raz (grupa jest prosto redukowalna).

Współczynniki $C-G$ są stabilizowane, oznaczamy je na przykład symbolem:

$$\langle j_1 m_1, j_2 m_2 | j m \rangle. \quad (4.8)$$

Twierdzenie E-W

$$\langle j m | S | j' m' \rangle = \delta_{jj'} \delta_{mm'} S^j \quad (4.9)$$

mówi w tym przypadku, że w oddziaływaniach niezmienniczych względem obrotów zachowywany jest całkowity kręt i jego rzut na pewien ustalony kierunek.

Grupa obrotów jest też izomorficzna z grupą izospinu. Mamy przy tym następujące odpowiedniości: całkowity kręt — całkowity izospin, rzut krętu — rzut izospinu. W tej interpretacji stany należące do pewnej reprezentacji opisują stany cząstki z różnymi rzutami izospinu, na przykład proton i neutron są dwoma stanami ładunkowymi (izospinowymi) nukleonu, a mezony π^+, π^0, π^- są różnymi stanami ładunkowymi mezonu π . Oczywiście stany ładunkowe pewnej cząstki muszą mieć identyczne wszystkie liczby kwantowe oprócz rzutu izospinu (i oczywiście ładunku).

c) Grupa SU_3

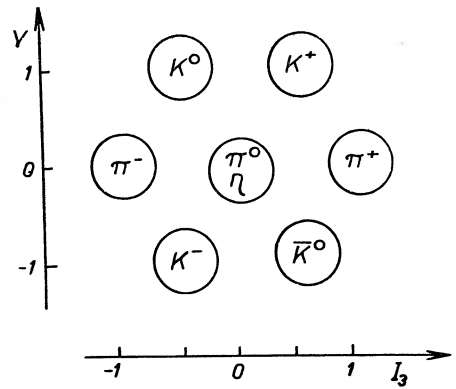
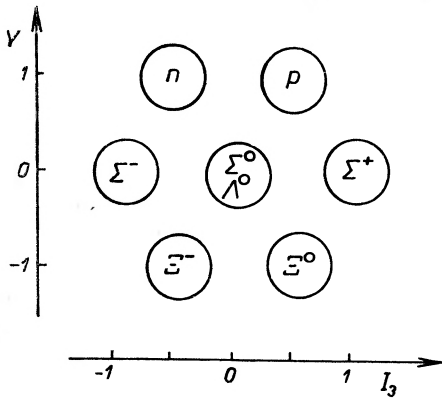
W tym przypadku mamy 8 generatorów, z których można skonstruować dwa operatory Casimira. Ich wartości własne oznaczamy λ, μ . Nie mają one prostego sensu fizycznego, przeto zwykle reprezentacje grupy SU_3 oznaczają się podając ich wymiary (w przypadku, gdy więcej niż jedna reprezentacja ma ten

sam wymiar, rozróżniamy je przy pomocy kresek, gwiazdek, itp.). Zresztą spośród ciągu reprezentacji tej grupy

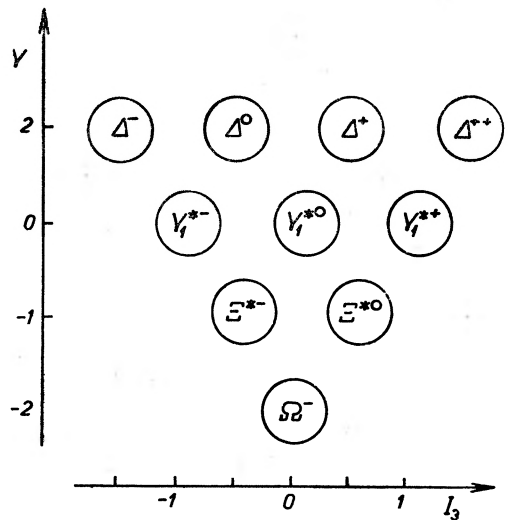
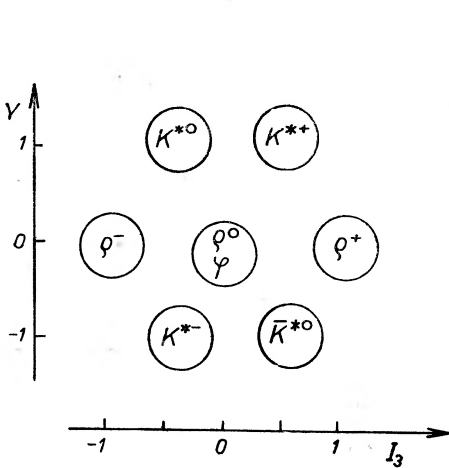
1, 3, 3*, 6, 6*, 8, 10, 10*, 15, 15*, 15', 15'*, 21, 21*, 24, 24*, 27, ...

tylko niewielka liczba jest realizowana w rzeczywistości, a to 1, 8, 10, 10* i być może 27. Niektórzy fizycy sądzą, że mogą również istnieć cząstki należące do reprezentacji 3 i 3* (tzw. kwarki), których jednak dotychczas nie wykryto.

Dwa generatory komutujące ze sobą, których wartości własne numerują stany w reprezentacji, odpowiadają hiperładunkowi Y i rzutowi izospinu I_3 . Okazuje się jednak, że w tym wypadku nie wystarczają one do numeracji stanów w niektórych reprezentacjach (co w pewien sposób wiąże się z faktem, że grupa SU_3 nie jest prosto redukowalna) i dlatego trzeba używać do tego celu dodatkowo całkowitego izospinu I .



Rys. 1 a, b



Rys. 2 a, b

Szereg C-G jest tutaj dość skomplikowany, nie ma prostego wzoru ogólnego (choćby stosunkowo prosto można wypisać go w każdym szczególnym przypadku). Na przykład często występująca redukcja ma postać:

$$8 \otimes 8 = 1 \oplus 8 \oplus 8 \oplus 10 \oplus 10^* \oplus 27.$$

Widać z tego przykładu, że grupa ta nie jest prosto redukowalna.

W reprezentacjach grupy obrotów znajdujemy różne stany spinowe cząstek (na przykład dwa stany spinowe protonu). W reprezentacjach grupy izospinu z kolei mamy różne stany izospinowe (ładunkowe) cząstek (na przykład dwa stany ładunkowe nukleonu: proton i neutron). Grupa SU_3 zajmuje jeszcze wyższe stanowisko w tej hierarchii: na przykład do jednej reprezentacji o wymiarze 8 zaliczamy oprócz nukleonu również hiperony Λ , Σ i Ξ . Bardziej znane reprezentacje grupy SU_3 są przedstawione na rysunkach.

d) Grupa Poincarégo

Zawiera ona grupę Lorentza (obrotów w czasoprzestrzeni) oraz grupę translacji czterowymiarowych.

Reprezentacje unitarne numerowane są przez kwadrat czteropędu m^2 oraz przez całkowity spin s . Stan w danej reprezentacji ustalony jest przez podanie czteropędu p_μ (z warunkiem $p_\mu p^\mu = m^2$) oraz rzutu spinu s_3 na wybraną oś.

Szereg C-G jest tu dość skomplikowany i grupa Poincarégo nie jest prosto redukowalna — trzeba wprowadzać parametry zwyrodnienia. Na przykład dla układu dwu cząstek o niezerowych masach m_1 i m_2 mamy:

$$D^{(m_1 s_1)} \otimes D^{(m_2 s_2)} = \bigoplus_{m_1+m_2}^{\infty} dM \bigoplus_{l=0, s=|s_1-s_2|}^{\infty} \bigoplus_{s_1+s_2} \bigoplus_{l+s} D^{(M, J)}_{(ls)}. \quad (4.10)$$

W podanym powyżej wzorze parametry zwyrodnienia mają sens fizyczny całkowitego spinu s układu dwu cząstek oraz ich względnego krętu orbitalnego l . Natomiast M^2 jest kwadratem sumy ich czteropędów

$$M^2 = (p_\mu^1 + p_\mu^2)^2, \quad (4.11)$$

co w układzie środka masy jest zarazem kwadratem ich całkowitej energii.

Często korzysta się też z innego wyboru parametrów zwyrodnienia. Mogą być nimi na przykład rzuty spinów indywidualnych cząstek na ich kierunki trójpędów (tzw. *helicities*). Szereg C-G nazywany jest tu nieraz rozwinięciem na fale cząstkowe.

Współczynniki C-G podamy przykładowo w dwu wersjach, dla uproszczenia w układzie środka masy. Najpierw w bazie $l-s$:

$$\begin{aligned} \left(\begin{array}{cc} m_1 s_1 & m_2 s_2 \\ p_1 s_{1z} & p_2 s_{2z} \end{array} \middle| \begin{array}{c} MJ \\ pJ_z \end{array}, ls \right) &= N \delta(\vec{p}_1 + \vec{p}_2 - \vec{p}) \delta[(p_1 + p_2)^2 - M^2] \\ &\sum_{s_{1z}} \sum_{s_{2z}} \langle s_1 s_{1z}, s_2 s_{2z} | s s_z \rangle \langle l s_z, s s_z | J J_z \rangle Y_{ls}^1(\vec{e}), \end{aligned} \quad (4.12)$$

gdzie e jest kierunkiem pędu względnego cząstek. Analogiczny współczynnik w bazie *helicity* ma postać:

$$\left(\begin{matrix} m_1 s_1 & m_2 s_2 \\ p_1 \lambda_1 & p_2 \lambda_2 \end{matrix} \middle| \begin{matrix} MJ \\ pJ_z \end{matrix}, \lambda_1 \lambda_2 \right) = N' \delta(\vec{p}_1 + \vec{p}_2 - \vec{p}) \delta(p_1 + p_2)^2 - M^2) D^J(\vec{e})_{J_z, \lambda_1 - \lambda_2}. \quad (4.13)$$

Y jest tu funkcją kulistą, zaś D^J — macierzą reprezentacji grupy obrotów. *Twierdzenie Eckarta-Wignera* oznacza w przypadku grupy Poincarégo, że:

- całkowita energia w układzie środka masy (M) jest zachowywana,
- całkowity kręt J jest zachowywany,
- suma wektorowa pędów jest zachowywana,
- rzut całkowitego krętu na wybraną oś nie zmienia się przy oddziaływaniu,
- amplitudy niezmiennicze (dające się wyrazić przez przesunięcia fazowe) zależą tylko od całkowitej energii M , całkowitego krętu J oraz parametrów zwyrodnienia (l i s).

5. Zakończenie

Prześledziliśmy na przykładach grup translacji, obrotów, SU_3 i Poincarégo znaczenie takich podstawowych pojęć jak reprezentacja, szereg Clebscha-Gordana, twierdzenie Eckarta-Wignera. Z konieczności rozważania te były dość ogólnikowe. Bardziej szczegółowe opracowania znaleźć można w podanej poniżej literaturze, która obejmuje prace znacznie bardziej zaawansowane.

Fundamental Concepts of the Group Theory Applied in the Elementary Particle Physics

Abstract: Fundamental concepts of the group theory are shortly discussed and compared with the corresponding concepts of the quantum mechanics (e.g. irreducibility with elementarity). Consequences of a symmetry for free and interacting particles are described. Finally, the notions such as the Casimir operators, the Clebsch-Gordan coefficients, and the Clebsch-Gordan series are discussed for the following groups: translations on a straight line, three-dimensional rotations, SU_3 , and the Poincaré group.

Literatura

Ogólne:

- E. P. Wigner, *Group Theory*, N. York 1959.
 R. E. Marshak, E. C. G. Sudarshan, *Introduction to Elementary Particle Physics*, N. York 1961 (tłumaczenie rosyjskie, Moskwa 1962).
 P. T. Matthews, *The Relativistic Quantum Theory of Elementary Particle Interactions*, Rochester 1957 (tłumaczenie rosyjskie, Moskwa 1959).

Grupa obrotów:

- A. R. Edmonds, *Angular Momentum in Quantum Mechanics*, Princeton 1957.
M. E. Rose, *Elementary Theory of Angular Momentum*, N. York 1957.
E. P. Wigner, *Group Theory*, N. York 1959.

Grupa SU_3 :

- H. J. Lipkin, *Unitary Symmetry for Pedestrians*, preprint Argonne Nation. Lab. 1963.
J. J. de Swart, *Rev. Mod. Phys.*, **35**, 916 (1963).
G. Racah, *Group Theory and Spectroscopy*, Institute for Advanced Studies, 1951.

Grupa Poincarégo:

- E. P. Wigner, *Ann. Math.*, **40**, 149 (1939).
H. Joos, *Fortschr. Phys.*, **10**, 65 (1962).

Janusz Kryłow

Instytut Fizyki PAN
Warszawa

Wpływ dyslokacji na własności elektryczne germanu II

1. Dane doświadczalne

W roku 1954 G. L. Pearson, W. T. Read, F. J. Morin [1] wyginali kryształy germanu typu n o oporności $15 \Omega\text{cm}$ z domieszką As w temperaturze 650°C do promienia krzywizny 5 cm , dookoła osi $[110]$. Otrzymali oni poślizg w płaszczyznach $(\bar{1}11$ i $1\bar{1}1)$. Gęstość dyslokacji określona metodą figur trawienia na płaszczyźnie (111) wynosiła $3 \cdot 10^6 \text{ cm}^{-2}$. Na próbkach tak przygotowanych mierzono przewodnictwo i stałą Halla, porównując te wielkości oraz wyliczoną ruchliwość hallowską z wielkościami dla próbki grzanej w warunkach deformacji, lecz niedeformowanej. Wyniki pomiarów przedstawia rys. 1 i 2.

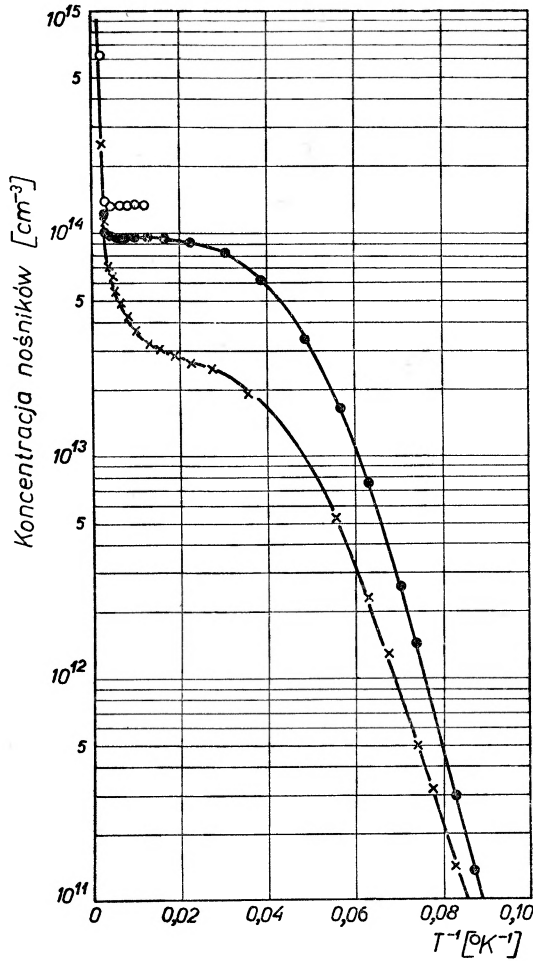
W roku 1959 R. L. Logan, G. L. Pearson, D. A. Kleinman [2] przeprowadzili ponownie pomiary na kryształach germanu typu n o oporności $10 \Omega\text{cm}$ z domieszką Sb. Deformację przeprowadzono w ten sposób, że dzięki odpowiedniej orientacji kryształu uzyskano jeden układ poślizgu i jeden kierunek linii dyslokacji. Pozwoliło to na pomiary σ , R i μ_H przy przepływie prądu równoległym i prostopadłym do linii dyslokacji. Wyniki pomiarów przedstawia rys. 3, 4 i 5.

W roku 1963 R. M. Broudy [3] przeprowadził analogiczne pomiary oraz dodatkowe pomiary magnetooporu dla próbek o maksymalnej jednorodności rozkładu dyslokacji. Stwierdził on istnienie różnic ilościowych w porównaniu z pracami poprzedników uwarunkowanych niejednorodnością ich próbek. Ogólne tendencje są te same:

a) zmniejszenie koncentracji w próbkach deformowanych rosnące z obniżeniem temperatury,

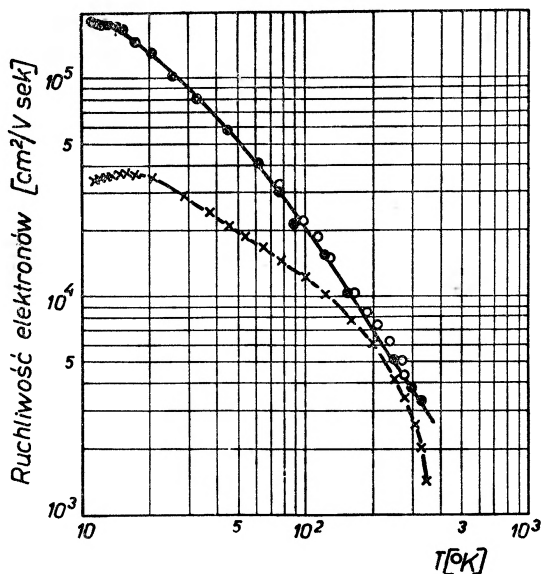
b) zmniejszenie ruchliwości nie tylko dla przepływu prądu $\vec{J} \perp \vec{D}$ (\vec{D} kierunek linii dyslokacji), lecz również dla $\vec{J} \parallel \vec{D}$ (rys. 6), co nie zostało wykryte w pracy [2] ze względu na niejednorodność próbek,

c) rys. 7 przedstawia magnetoopór ρ_H/ρ_0 , dla próbki G-13 nie deformowanej, dla tej samej próbki deformowanej z $\vec{J} \perp \vec{D}$ i $\vec{J} \parallel \vec{D}$ w temperaturze 78°K w funkcji kąta Φ , utworzonego przez kierunek pola magnetycznego \vec{H} i kierunek normalnej do neutralnej płaszczyzny wygięcia. Wektor \vec{H} zawsze leży w płaszczyźnie $\perp \vec{J}$

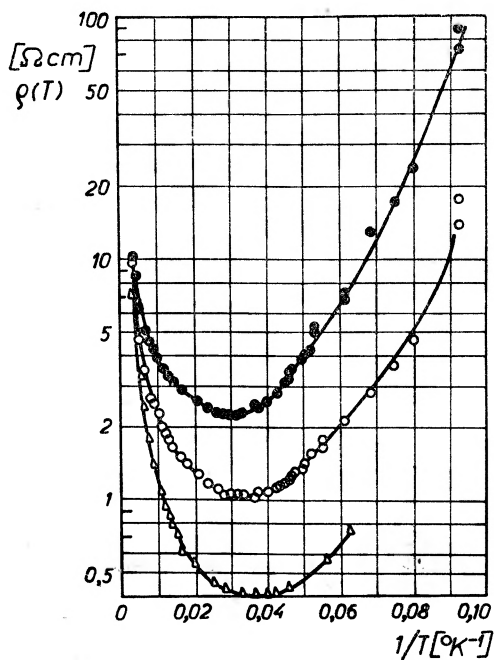


Rys. 1. Gęstość elektronów $1/T$, ○ — próbka kontrolna; ● — próbka grzana; × — próbka deformowana [1]

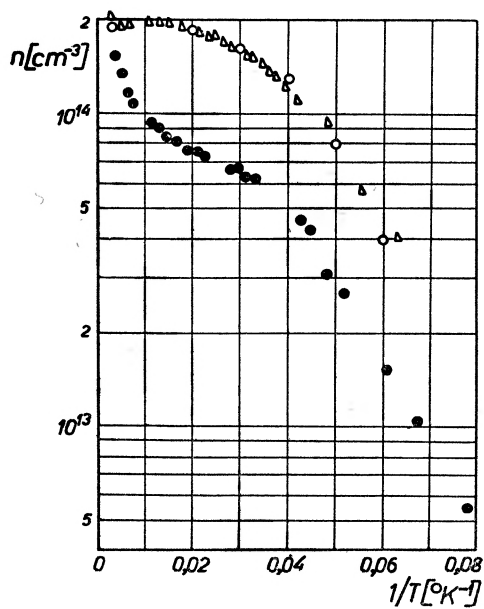
(rys. 10). Rys. 8, to samo dla próbki wyciętej z dolnej części A próbki rys. 7 i rys. 9 dla próbki wyciętej ze środkowej części C próbki rys. 7 — rys. 10. Rysunki te ilustrują wpływ niejednorodności próbki na wielkość magnetooporu. Część A posiada rozkład dyslokacji bardziej jednorodny i lepiej uporządkowany niż C i niż cała próbka.



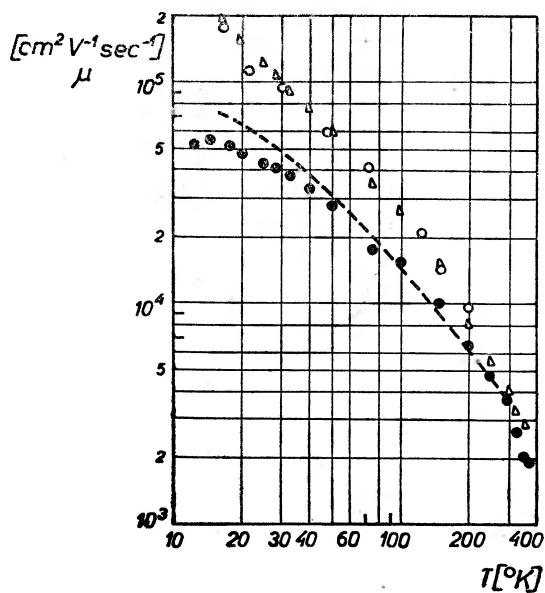
Rys. 2. Ruchliwość elektronów w funkcji temperatury dla próbek — niedeformowanej, nie-deformowanej grzanej w warunkach deformacji i deformowanej monokryształu germanu, \circ — próbka kontrolna; \bullet — próbka grzana; \times — próbka deformowana [1]



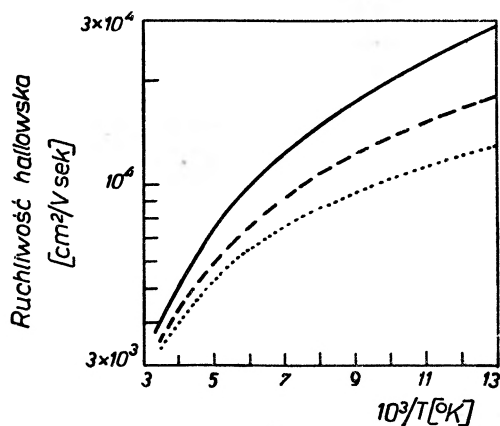
Rys. 3. Oporność monokryształów germanu w funkcji $1/T$ mierzona dla kryształów nie-deformowanego oraz deformowanego przy przepływie prądu równoległym i prostopadłym do linii dyslokacji \bullet — $\vec{I} \perp \vec{D}$; \circ — $\vec{I} \parallel \vec{D}$; \triangle — próbka kontrolna [2]



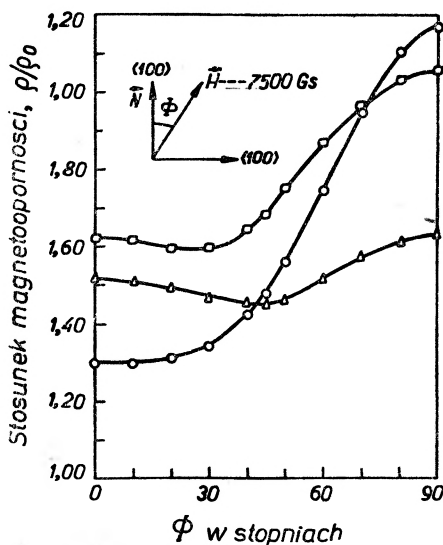
Rys. 4. Temperaturowa zależność koncentracji nośników dla kryształów rys. 3, Δ — próbka kontrolna; \bullet — próbka wygięta; \circ — punkty wyliczone [2]



Rys. 5. Temperaturowa zależność ruchliwości dla kryształów rys. 3, Δ — próbka kontrolna; \circ — $\vec{I} \parallel \vec{D}$; \bullet — $\vec{I} \perp \vec{D}$; — — — krzywa teoretyczna [2]



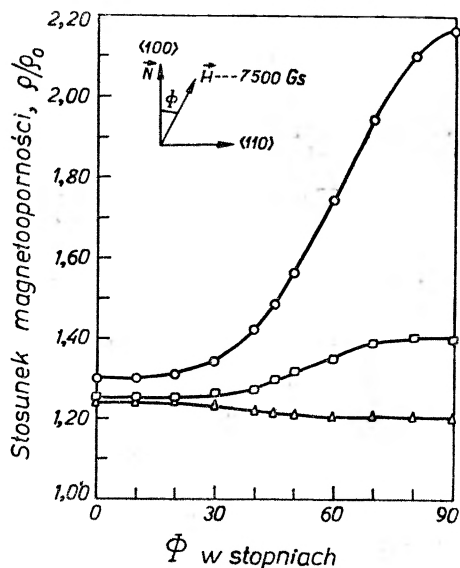
Rys. 6. Ruchliwość hallowska dla $\vec{I} \parallel \vec{D}$ i $\vec{I} \perp \vec{D}$ — próbka niewyginana; — — — $\vec{I} \parallel \vec{D}$; ... $\vec{I} \perp \vec{D}$ [3]



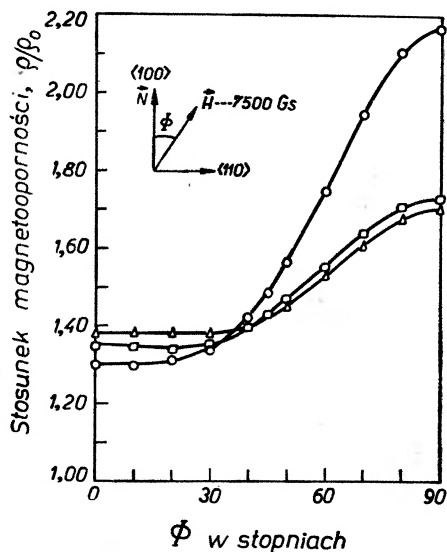
Rys. 7. Magnetoopór ρ_H/ρ_0 dla $G = 13$, \circ — próbka niewyginana; \square — $\vec{I} \parallel \vec{D}$; \triangle — $\vec{I} \perp \vec{D}$

Wzrost gęstości, jednorodności rozkładu i uporządkowania dyslokacji w części A daje silne zmniejszenie i bardziej wyraźną zależność temperaturową przewodnictwa, czemu towarzyszy silny wzrost i bardziej wyraźna zależność temperaturowa stałej Halla w porównaniu z wynikiem dla całej próbki.

Tablica I podaje warunki deformacji do promienia wygięcia = 10 cm, odpuszczania oraz gęstości dyslokacji różnych części szeregu próbek. Maksymalną jednorodność otrzymujemy dla próbki 3. Pozostałe próbki wykazują

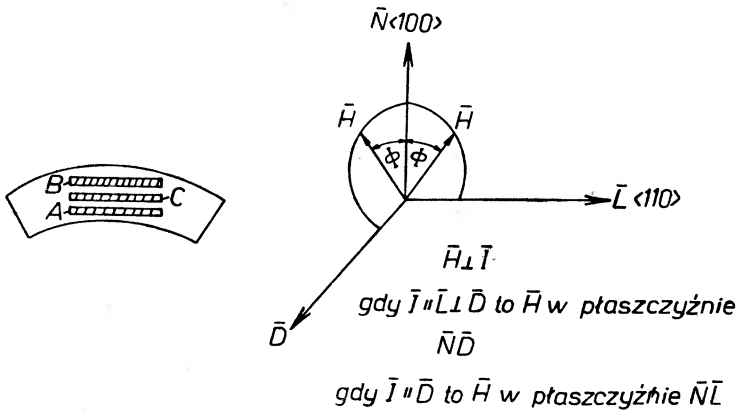


Rys. 8. Magnetoopór ρ_H/ρ_0 dla sekcji A próbki G—13, ○ — próbka niewyginana; □ — $\vec{I} \parallel \vec{D}$; △ — $\vec{I} \perp \vec{D}$ [3]



Rys. 9. Magnetoopór ρ_H/ρ_0 dla sekcji C próbki G—13, ○ — próbka niewyginana; □ — $\vec{I} \parallel \vec{D}$; △ — $\vec{I} \perp \vec{D}$ [3]

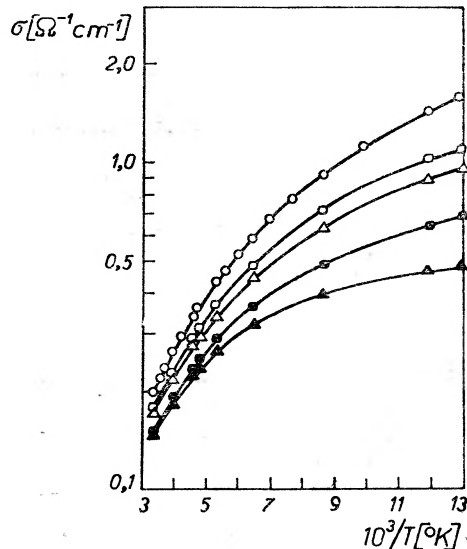
duże niejednorodności. Rys. 11 i 12 pokazują temperaturową zależność przewodnictwa różnych części niejednorodnie deformowanych próbek, a rys. 13 i 14 temperaturową zależność stałej Halla tych próbek.



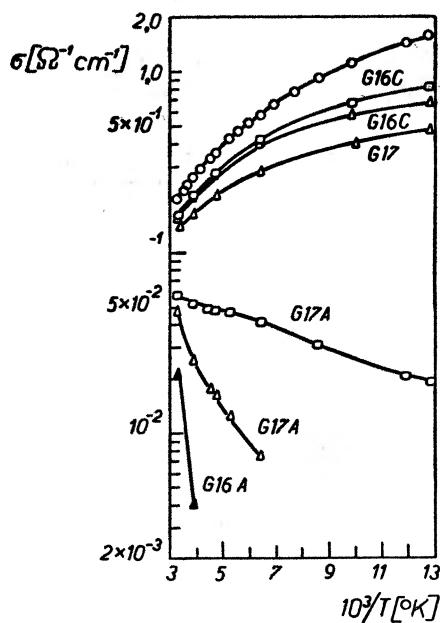
Rys. 10. Sekcja A, B, C wyginanych próbek oraz orientacja wektora \vec{H} dla pomiarów magnetooporu [3]

Tablica I

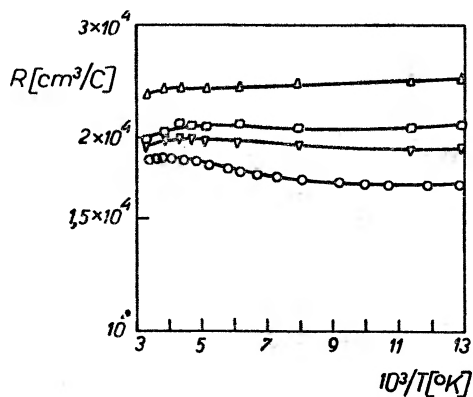
	Deformacja		Odpuszczanie		N — gęst. dysl. — 10^7 cm^{-2}		Uwagi
	Temp. °C	Czas min	Temp. °C	Czas min	Zewn. A i B	Wewn. C	
1.	800	3	—	—	2,4	~ 0,9	G-15 (rys. 11, 13)
2.	550	100	—	—	3,8	~ 1,4	
3.	550	100	800	60	< 1	< 1	G-16 (rys. 12, 14)
4.	550	5	—	—	~ 6,7	~ 1,5	
5.	550	5	550	95	5,6	1,8	G-17 (rys. 12, 14)
6.	550	5	800	60	~ < 1 do ~ 3	~ < 1 do ~ 2	G-18



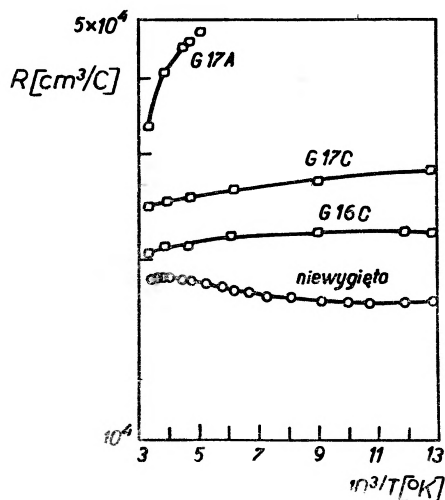
Rys. 11. Przewodnictwo próbki niezdeformowanej oraz przewodnictwa sekcji A, B, C próbki G — 15, ○ — próbka niewygięta; △ — sekcja A — $\vec{I} \perp \vec{D}$; □ — sekcja A — $\vec{I} \parallel \vec{D}$; □ — sekcja C — $\vec{I} \parallel \vec{D}$; □ — sekcja C — $\vec{I} \perp \vec{D}$; ■ — sekcja B — $\vec{I} \parallel \vec{D}$; ▲ — sekcja B — $\vec{I} \perp \vec{D}$ [3]



Rys. 12. Przewodnictwo niezdeformowanej próbki oraz przewodnictwo sekcji A, C próbek G16 i G17, \circ — próbka niewygięta; \square — $I \parallel \vec{D}$; \triangle — $I \perp \vec{D}$ [3]



Rys. 13. Stała Halla próbki niezdeformowanej oraz sekcji A, B i C próbki G—15, \circ — próbka niewygięta; \square — sekcja A; \triangle — sekcja B; ∇ — sekcja C [3]



Rys. 14. Stała Halla próbki niezdeformowanej oraz sekcji A, C próbek G — 16 i G — 17 [3]

2. Teoria Reada [4]-[6]

2.1. Koncentracja

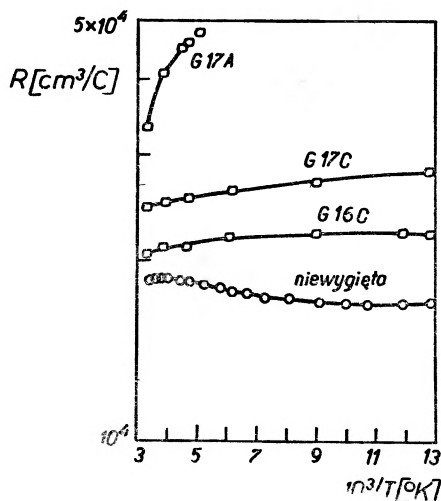
Dyslokację w sieci diamentu przedstawimy jako liniowy szereg atomów pozbawionych jednego elektronu. Poziom energetyczny pozostającego elektronu leży prawdopodobnie w pasmie walencyjnym i nie wykryto donorowego działania dyslokacji.

W typie *n* germanu obserwujemy po wprowadzeniu dyslokacji malenie koncentracji rosnące z obniżeniem temperatury (rys. 1 i 4).

Zakładamy zatem, że poziom energetyczny elektronu, który dyslokacja wychwytuje uzupełniając brakujący elektron, leży w pasmie przewodnictwa poniżej poziomów domieszek donorowych i poziomu Fermiego. Oznaczmy go przez E_2 . Przez c oznaczmy odległość pomiędzy zerwanymi wiązaniami, a przez a — odległość pomiędzy elektronami wychwytanymi przez dyslokację. Stopień obsadzenia dyslokacji przez elektrony

$$f = \frac{c}{a}$$

Ze względu na $E_2 < E_F$ obsadzanie dyslokacji przez elektrony daje zmniejszenie energii swobodnej. W miarę oksadzania wzrost f wywołuje wzrost E_s — energii elektrostatycznego oddziaływania elektronów wychwytanymi przez dyslokację. Wzrost ten kompensuje malenie energii swobodnej wywołane wychwytem elektronów na poziom $E_2 < E_F$ i w końcu uniemożliwia dalsze zapewnianie elektronami wolnych miejsc na dyslokacji. Stopień obsadzenia $f(E_2, T, c)$ jest funkcją temperatury, położenia poziomu E_2 oraz c , i w tempe-



Rys. 14. Stała Halla próbki niezdeformowanej oraz sekcji A, C próbek G — 16 i G — 17 [3]

2. Teoria Reada [4]-[6]

2.1. Koncentracja

Dyslokację w sieci diamentu przedstawimy jako liniowy szereg atomów pozbawionych jednego elektronu. Poziom energetyczny pozostającego elektronu leży prawdopodobnie w pasmie walencyjnym i nie wykryto donorowego działania dyslokacji.

W typie n germanu obserwujemy po wprowadzeniu dyslokacji malenie koncentracji rosnące z obniżeniem temperatury (rys. 1 i 4).

Zakładamy zatem, że poziomy energetyczny elektronu, który dyslokacja wychwytuje uzupełniając brakujący elektron, leży w pasmie przewodnictwa poniżej poziomów domieszek donorowych i poziomu Fermiego. Oznaczmy go przez E_2 . Przez c oznaczmy odległość pomiędzy zerwanymi wiązaniami, a przez a — odległość pomiędzy elektronami wychwytanymi przez dyslokację. Stopień obsadzenia dyslokacji przez elektrony

$$f = \frac{c}{a}$$

Ze względu na $E_2 < E_F$ obsadzanie dyslokacji przez elektrony daje zmniejszenie energii swobodnej. W miarę ośadzania wzrost f wywołuje wzrost E_s — energii elektrostatycznego oddziaływania elektronów wychwytaných przez dyslokację. Wzrost ten kompensuje malenie energii swobodnej wywołane wychwytem elektronów na poziomie $E_2 < E_F$ i w końcu uniemożliwia dalsze zapełnianie elektronami wolnych miejsc na dyslokacji. Stopień obsadzenia $f(E_2, T, c)$ jest funkcją temperatury, położenia poziomu E_2 oraz c , i w tempe-

raturze pokojowej $f = 0,03$, w temperaturze ciekłego azotu $f = 0,11$. W zakresie temperatur przewodnictwa domieszkowego elektrony wychwytywane przez dyslokację pochodzą od otaczających ją donorów.

Ujemny ładunek linii dyslokacji $= q/a = qf/c$ skompensowany jest przestrzennym ładunkiem dodatnim $\pi R^2(N_D - N_A)q$ — donorów zjonizowanych w cylindrze o promieniu R otaczającym dyslokację

$$\pi R^2(N_D - N_A) = \frac{f(E_2, T, c)}{c} \quad (1)$$

$$R^2 = \frac{f(E_2, T, c)}{\pi c(N_D - N_A)} \quad (2)$$

$$\langle n \rangle = n - \pi R^2 N n \quad (3)$$

$\langle n \rangle$ — koncentracja elektronów w zdeformowanym kryształ, N — gęstość dyslokacji.

$$\frac{\langle n \rangle - n}{n} = \varepsilon = \frac{f(E_2, T, c) N}{c(N_D - N_A)} \quad (4)$$

W równaniu (4) ε jest wielkością mierzoną doświadczalnie poprzez pomiar $\langle n \rangle$ i n — koncentracji elektronów w zdeformowanym i nie zdeformowanym kryształ (rys. 1 i 4). $N_D - N_A$ w temperaturowym zakresie jonizacji domieszek przyjmujemy równe n . Temperaturowa zależność ε dana równaniem (4) oraz teoretyczna zależność $f(E_2, T, c)$ pozwalają nam znaleźć E_2 , c i N oraz potwierdzić słuszność teoretycznych wyliczeń $f(E_2, T, c)$ [2].

Naturalnie pożądane byłoby bezpośrednio określenie parametrów N i c pozwalające na dokładniejsze określenie $f(E_2, T, c)$ i porównanie z teorią. c — wpływa na wyniki w nieznacznym stopniu. Dla gęstości dyslokacji N określonej metodami trawienia niemożliwe jest otrzymanie wyników zgodnych z teorią i należałoby zastosować jakąś inną metodę określania N .

Dla zilustrowania

$$c = 3,5 \text{ \AA} \quad \text{daje} \quad E_2 = 0,179 \text{ eV} \quad \text{i} \quad N = 5,4 \times 10^7$$

$$c = 6,9 \text{ \AA} \quad \text{daje} \quad E_2 = 0,2 \text{ eV} \quad \text{i} \quad N = 5,1 \times 10^7$$

Dla

$$f = 0,1: \quad N_D - N_A = 10^{15} \text{ cm}^{-3}; \quad R = 2830 \text{ \AA} \quad \text{i} \quad a = 40 \text{ \AA}$$

2.2. Stopień obsadzenia

$$\Delta F = E_2 + E_s(f) + \delta E_s - E_F - TS \quad (5)$$

$\Delta F(f)$ — całkowity wzrost energii swobodnej układu/elektron, E_2 — energia poziomu dyslokacyjnego mierzona od wierzchołka pasma walencyjnego, $E_s(f)$ — elektrostatyczna energia związana z pracą utworzenia cylindra ładunku przestrzennego z linią równomiernie rozmieszczonych elektronów w środku, $\delta E_s(f)$ — energia niejednorodności rozkładu elektronów, E_F — energia poziomu Fermiego w niezdeformowanym kryształ, mierzona od wierzchołka pasma walencyjnego.

2.2.1. Przybliżenie minimum energii (niskie temperatury)

Zakładamy $fE_0 = f \frac{q^2}{\kappa a}$ = energii elektrostatycznego oddziaływania elektronów $\gg kT$. W tych warunkach dla danego f mamy jedno możliwe rozmieszczenie równoodległych elektronów, dające minimum energii, δE_s energia związana z nierównomiernym rozkładem = 0, znika również entropia. Przybliżenie jest słuszne w temperaturze zera bezwzględnego.

Zależność (5) przybiera postać

$$F\Delta F(f) = E_2 + E_s(f) - E_F \quad (6)$$

$$E_s(f) = E_e + E_c + E_{ec} + E_{ce} = \frac{q^2}{\kappa a} \left\{ \ln \frac{R}{a} - 0,866 \right\} = fE_0 \frac{3}{2} \left[\ln \frac{f}{f_c} - 0,866 \right] \quad (7)$$

$$E_0 = \frac{q^2}{\kappa a} \quad f_c = c\pi^{1/3} [N_D - N_A]^{1/3} \quad \text{to jest} \quad f = f_c \quad \text{przy} \quad R = a$$

E_e jest energią oddziaływania elektronów wychwypanych przez dyslokację = $= \frac{1}{2} \sum \psi_i q_i$, gdzie $\psi_i = \sum_j \frac{q_j}{r_{ij}}$.

Przy założeniu

1. liniowego rozmieszczenia ładunku elektronów na dyslokacji o gęstości q/a , co słuszne jest dla

$$a \ll [N_D - N_A]^{-1/3} \quad \text{i} \quad R \gg a \equiv f \gg f_c \quad (7a)$$

2. równomiernego rozłożenia ładunku przestrzennego, to jest

$$R \gg [N_D - N_A]^{-1/3} \quad (7b)$$

E_c — energię ładunku dodatniego wyliczamy dla cylindra naładowanego dodatnio z ujemnie naładowaną nicią w środku, rozwiązując równanie Poissona'a.

E_{ec} — energia oddziaływania elektronów z polem ładunku dodatniego

$$E_{ec} = \frac{1}{2} q\psi_c(0)$$

$\psi_c(0)$ — potencjał ładunku dodatniego na linii dyslokacji.

E_{ce} — energia oddziaływania ładunku dodatniego z polem ujemnym naładowanej linii dyslokacji

$$E_{ce} = E_{ec}.$$

Żądanie minimum energii daje

$$E_F - E_2 = E^*(f) = \frac{d}{df} fE_s(f), \quad (8)$$

skąd znajdujemy $f(T)$ uwzględniając temperaturową zależność $E_F(T)$. W wysokich temperaturach przybliżenie to nie jest słuszne, ponieważ: a) mamy odchylenie od równomiernego rozkładu i pojawia się człon δE_s , b) $f < f_c$, c) koncentracja dziur staje się duża i dają one przyczynek do ekranującego ładunku dodatniego. Zauważmy, że założenia nie są spełnione również dla dużych $[N_D - N_A]^{1/3}$.

2.2.2. Przybliżenie Fermiego (wysokie temperatury)

Zakładamy $fE_0 \ll kT$. Człon δE_s związany z przemieszczeniem elektronów od położenia jednorodnego rozkładu zaczyna odgrywać rolę. Znajdujemy minimalną wartość energii swobodnej zaniedbując $\delta E_s \ll TS$ i przyjmując maksymalną wartość entropii. Zależność (5) przybiera postać

$$\Delta F = E_2 + E_s - E_F - TS. \quad (9)$$

Maksymalna entropia odpowiada nieskoordynowanym zajęciom miejsc, to znaczy prawdopodobieństwo znalezienia elektronu w danym miejscu nie zależy od zajęcia sąsiednich miejsc

$$fS(f) = -k/M \ln W = -k[f \ln f + (1-f) \ln(1-f)] \quad (10)$$

$$a \quad W = \frac{M!}{N!(M-N)!}$$

N jest liczbą elektronów zajmujących N różnych miejsc spośród M miejsc możliwych. Dla $fE_0 \ll kT$ rozkład odbiega mało od jednorodnego i większość z W dróg wypełnienia N elektronami M miejsc ma tę samą energię. Dodając $-fTS(f)$ do $[E_2 + E_s - E_F]f$ i minimizując mamy

$$E^*(f) = E_F(T) - E_2 + kT \ln \left\{ \frac{1-f}{f} \right\} \quad (11)$$

$$f = \frac{1}{1 + \exp\{(E - E_F)/kT\}}$$

$E = E_2 + E^*(f)$ = energii elektronu dodanego do dyslokacji w punkcie o potencjale ψ , E_2 jest energią elektronu przy $f \rightarrow 0$, E^* — energią elektrostatycznego oddziaływania elektronu z pozostałymi elektronami na dyslokacji.

2.2.3. Przybliżenie Boltzmana (średnie temperatury)

Dla przypadku $fE_0 = kT$ uwzględniamy energię δE_s związaną z przemieszczeniami elektronów z położenia równomiernego rozkładu. Zakładamy przybliżenie Einsteina, to znaczy elektron porusza się w polu innych elektronów i energia jego jest tylko funkcją przemieszczenia. $x_i = nc < a$ jest przemieszczeniem elektronu z położenia równomiernego rozkładu, δE_n — energią związaną z przemieszczeniem x_i ,

$$\delta E_n = \frac{1}{2} K x_i^2 = (2,40 E_0 f^3) n^2 \quad (12)$$

$$\delta E_s = \overline{\delta E_n} = (2,40 E_0 f^3) \overline{n^2}.$$

W celu wyliczenia $\overline{n^2}$ stosujemy statystykę Boltzmana. Po minimizacji energii swobodnej dla znalezienia $f(T)$ otrzymujemy następujące równanie

$$E^* = E_F(T) - E_2 - \delta F^* \quad (13)$$

$$\delta F^* = \delta F + 3\delta E_s$$

$$\delta F = -kT \ln \sum_{-\infty}^{+\infty} \exp(-zn^2)$$

$$\delta E_s = -kT \frac{d \ln \sum_{-\infty}^{+\infty} \exp(-zn^2)}{d \ln z}$$

$$z = \frac{2,40 E_0 f^3}{kT}$$

Ograniczenia:

1. W niskich temperaturach mamy korelację przemieszczeń. Uwzględnienie oddziaływania najbliższych sąsiadów daje

$$\delta E_s = E_0 f^3 \bar{n}^2.$$

2. W wysokich temperaturach przy niskich f , x_i jest bardzo duże i pewne rozkłady liczone jako równe są powtarzalne, przybliżenie traci sens.

2.2.4. Przybliżenie oscylatora harmonicznego

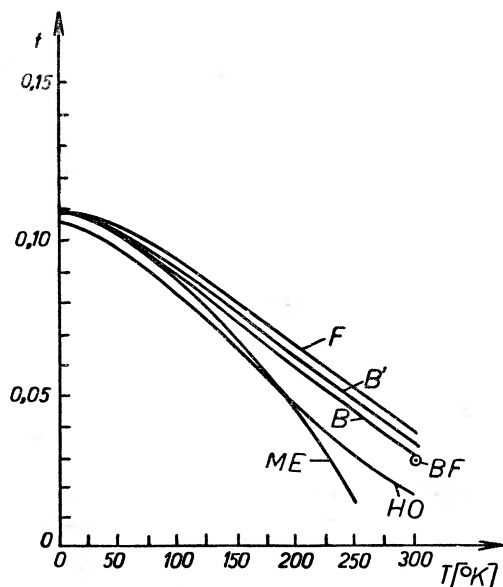
Gdy c jest małe, funkcje falowe poziomów dyslokacyjnych przykrywają się tworząc pasmo. Elektron wychwytyany przez dyslokację traktujemy jako oscylator pozostający w polu innych elektronów dyslokacji.

$$\delta F^* = \frac{3h\nu}{2} \left(\bar{n} + \frac{5}{6} \right) + kT \ln(1 + \bar{n})$$

$$h\nu = f^{3/2} \sqrt{4,80 E_2 E_b}$$

$$\bar{n} = \frac{1}{\exp(h\nu/kT) - 1}$$
(14)

E_2 — minimalna energia elektronu w pasmie, E_b — szerokość pasma.



Rys. 15. Wykres $f(T)$ dla różnych przybliżeń (5): F — przybliżenie Fermiego, B — przybliżenie Boltzmana, B' — przybliżenie Boltzmana z uwzględnieniem korelacji, ME — przybliżenie minimum energii, HO — przybliżenie oscylatora harmonicznego

Wstawiając δF^* do równania (13) mamy równanie określające $f(T)$. Rys. 15 podaje zależność $f(T)$ dla omawianych modeli. Dokładność danych doświadczalnych jest zbyt mała dla ich rozróżnienia i na ogół przy wyliczaniu ε w (4) przyjmujemy przybliżenie Fermiego jako najprostsze. Szersze omówienie problemu daje praca [8].

2.3. Ruchliwość

Zmiany ruchliwości w kryształach z dyslokacjami wywołane są przez dwa czynniki:

1. Gdy zmiana potencjału odpychającego dyslokacji ψ o wielkość kT/q zachodzi na odległości małej, w porównaniu z promieniem cylindra ładunku przestrzennego, elektrony nie mogą przenikać do cylindrów dyslokacyjnych. Nie mogą one dryfować w kierunku pola prostopadłego do linii dyslokacji, drogi ich ulegają zakrzywieniu. Dystorsja taka będąca wynikiem zwiększania pola w obszarach ładunku przestrzennego otaczającego dyslokację zmniejsza ruchliwość. Problem ten można traktować jak zagadnienie przewodnictwa w kryształach zawierającym nieprzewodzące cylindry o promieniu R , prostopadłe do kierunku przepływu prądu [7].

2. Elektrony ulegają rozpraszaniu na cylindrach ładunku przestrzennego. Zakładamy, iż rozpraszanie to jest elastyczne, przy czym składowa pędu równoległa do kierunku linii dyslokacji nie ulega zmianie. Efekt ten występuje, gdy odległość pomiędzy dyslokacjami jest mała w porównaniu ze średnią drogą swobodną. W temperaturze pokojowej w *Ge* jest on do zaniechania.

2.3.1. Przypadek dystorsji

Rozważmy przypadek dystorsji:

A. Pole elektryczne jest przyłożone w kierunku x .

Linie dyslokacji mają kierunek z .

\vec{E} jest lokalnym polem, a J lokalnym prądem

$$J = q\mu n \underline{E},$$

gdzie μ jest ruchliwością w niezdeformowanym kryształach. Wewnątrz cylindrów ładunku przestrzennego dyslokacji $J = 0$, gdyż $n = 0$.

Oznaczmy przez $\langle J_x \rangle$ — średni prąd płynący przez próbkę, a przez $\langle E_x \rangle$ — przyłożone pole.

$$\langle J_x \rangle = q\mu \langle n E_x \rangle = q\mu \langle n \rangle \langle E_x \rangle_n = q\mu n (1 - \varepsilon) \langle E_x \rangle_n,$$

gdzie $\langle n \rangle = n(1 - \varepsilon)$ — koncentracja elektronów w zdeformowanym kryształach. $\varepsilon = \pi R^2 N$ — objętość kryształu zajęta przez cylindry dyslokacyjne. Symbol

$\langle \rangle_n$ oznacza uśrednianie po obszarze poza cylindrami, $\sigma^* = \frac{\langle J_x \rangle}{\langle E_x \rangle}$ jest przewodnictwem mierzonym w zdeformowanym kryształach.

$$\sigma^* = \sigma \frac{\langle n \rangle \langle E_x \rangle_n}{\langle E_x \rangle} = \sigma \frac{\langle n \rangle}{n} g(\varepsilon) = q\mu^* \langle n \rangle \quad (14)$$

$\mu^* = \mu g(\varepsilon)$ jest mierzona ruchliwością, $g(\varepsilon)$ określono z badań przewodnictwa modelu (papier oporowy zawierający dziury, odpowiadające nieprzewodzącym cylindrom dyslokacyjnym).

Dla małych ε : $g(\varepsilon) = 1 - \varepsilon$.

Podobne rozważania dla przypadku:

A' linii dyslokacji w kierunku $z \parallel D_z$, dla pola magnetycznego w kierunku $z - H_z \parallel D_z$ i pola elektrycznego w kierunku $x - E_x \perp H_z \parallel D_z$, dają nam brak zmian ruchliwości w kryształach deformowanym i nie deformowanym.

Dla przypadku:

C' linii dyslokacji w kierunku $y \parallel D_y$, dla pola magnetycznego w kierunku $z - H_z \perp D_y \perp E_x$ i pola elektrycznego w kierunku $x - E_x \perp D_y \perp H_z$ mamy

$$\mu_H^* = \mu_H g(\varepsilon),$$

gdzie μ_H^* jest ruchliwością hallowską w zdeformowanym kryształach. Stała Halla dla obu tych przypadków jest identyczna i wynosi

$$R^* = \frac{\mu_H}{\mu} \frac{1}{ec \langle n \rangle}.$$

2.3.2. Równanie transportu dla kryształu z dyslokacjami

Przejdźmy do rozpraszania elektronów przez dyslokacje.

Przy założeniach:

1. sferycznych powierzchni izoenergetycznych, 2. braku degeneracji i 3. elastyczności rozpraszania z zachowaniem składowej pędu równoległej do linii dyslokacji i małej gęstości dyslokacji, równanie transportu

$$q \left(\vec{E} + \frac{1}{c} \vec{V} \times \vec{H} \right) \vec{V}_p f' + \int_s (f'(\vec{P}') - f'(\vec{P})) W_t ds' + \\ + \int_C [f'(\vec{P}') - f'(P)] W_1(\theta) dC' = 0 \quad (15)$$

daje się rozwiązać w przybliżeniu czasów relaksacji.

$f'(\vec{P}) = f(P) + f_1(\vec{P}, \vec{E}) + f_2(\vec{P}, \vec{E}, \vec{H})$ jest stopniem wypełnienia stanów energetycznych w punkcie P przestrzeni pędów w warunkach przepływu prądu. $f(P)$ oznacza analogiczną wielkość w warunkach równowagi, S jest powierzchnią stałej energii w przestrzeni pędów, C — kołem na powierzchni S , stałej składowej pędu wzdłuż linii dyslokacji, W_t — prawdopodobieństwem przejścia elektronu ze stanu \vec{P}_i do stanu \vec{P}_j dla rozpraszania na drganiach sieci, $W_1(\theta)$ — analogiczną wielkością dla rozpraszania na dyslokacjach, θ jest to kąt pomiędzy rzutami \vec{P}_i i P_j , na powierzchni ax , dla linii dyslokacji leżącej wzdłuż osi y . Dla przypadku A'

$$E_x \perp D_z \parallel H_z \\ f_1 = \frac{q E_x \tau f}{k T} V_x$$

Przy czym

$$\frac{1}{\tau} = \frac{1}{\tau_1} + \frac{1}{\tau_t},$$

gdzie τ_t jest średnim czasem swobodnym dla rozpraszania na drganiach sieci, a τ_1 jest średnim czasem swobodnym dla rozpraszania na dyslokacjach, stałym na kole C .

$$f_2 = \frac{\omega q}{kT} E_x f \tau^2 V_y \quad \omega = qH/mc \quad (16)$$

$$\mu = \frac{q}{kT} \langle v \rangle l_\mu F'_1 \left(\frac{l_1}{l_t} \right)$$

$$\langle v \rangle \text{ — średnia prędkość termiczna} = \sqrt{\frac{8kT}{\pi m}}$$

$$\frac{1}{l_\mu} = \frac{3}{l_t} + \frac{8}{\pi l_1}$$

$$F'_1 \left(\frac{l_1}{l_t} \right) = \frac{1}{2} \left(3 \frac{l_1}{l_t} + \frac{8}{\pi} \right) \int_0^{\pi/2} \frac{\sin^3 \theta}{l_1/l_t + \sin \theta} d\theta$$

$$\mu_H = \frac{q}{m} \frac{l_H}{\langle v \rangle} F'_2 \quad (17)$$

$$\frac{1}{l_H} = \frac{1}{l_t} + \frac{\pi}{4l_1}$$

$$F'_2 = \frac{1}{l_H} \int_0^{\pi/2} \frac{\sin^3 \theta d\theta}{\left(\frac{1}{l_t} + \frac{\sin \theta}{l_1} \right)^2} \int_0^{\pi/2} \frac{\sin^3 \theta d\theta}{\frac{1}{l_t} + \frac{\sin \theta}{l_1}}.$$

Read [6] stabelaryzował funkcje F_1 i F_2 w zależności od l_1/l_t . Dla przypadku:

$$C': E_x \perp D_y \perp H_z$$

czas relaksacji nie istnieje. Rozwiązanie otrzymuje się w postaci

$$f_2(P) = -\frac{\omega q}{kT} E_x f \left(1 - \frac{\tau}{2\tau_1} \right) \tau_t \tau v_y$$

$$\mu_H = \frac{3q l_2}{m \langle v \rangle} \quad (18)$$

$$l_2 = \int_0^{\pi/2} \frac{\cos^2 \theta \sin \theta d\theta}{\frac{1}{l_t} + \frac{\sin \theta}{l_1}} - \frac{1}{2l_1} \int_0^{\pi/2} \frac{\sin^2 \theta d\theta}{\left(\frac{1}{l_t} + \frac{\sin \theta}{l_1} \right)^2}.$$

Dla przypadku:

B' : $E_x \parallel D_x \perp H_z$ przy założeniach zachowania składowej pędu wzdłuż dyslokacji, ruchliwość w zdeformowanych kryształach nie ulega zmianie [2].

Dla wszystkich przypadków stała Halla dana jest przez:

$$R = \frac{\mu_H}{\mu} \frac{1}{ec \langle n \rangle}.$$

Wyniki dla poszczególnych przypadków są zebrane w tablicy II.

Tablica II

	σ	σ_{\perp}/σ_0	$\sigma_{\perp}/\sigma_{\parallel}$	R/R_0	μ/μ_0
A $D_z; H_z = 0$ I_x $\perp D_z$	$e\mu \langle n \rangle g(\varepsilon) F_1 \left(\frac{l_1}{l_t} \right)$	$(1 - \varepsilon) g(\varepsilon) F_1 \left(\frac{l_1}{l_t} \right)$	$g(\varepsilon) F_1 \left(\frac{l_1}{l_t} \right)$	—	$g(\varepsilon) F_1 \left(\frac{l_1}{l_t} \right)$
A' $D_z; H_z \parallel D_z; I_x$ $\perp D_z \parallel H_z$	$e\mu \langle n \rangle F_2 \left(\frac{l_1}{l_t} \right)$	$(1 - \varepsilon) F_2 \left(\frac{l_1}{l_t} \right)$	$F_2 \left(\frac{l_1}{l_t} \right)$	$\frac{1}{1 - \varepsilon}$	$F_2 \left(\frac{l_1}{l_t} \right)$
B $D_x; H_z = 0$ $I_x \parallel D$	$e\mu \langle n \rangle$	—	$\frac{\sigma_{\parallel}}{\sigma_0}$ $1 - \varepsilon$	—	1
B' $D_x; H_z \perp D_x$ $I_x \parallel D_x \perp H_z$	$e\mu \langle n \rangle$	—	$1 - \varepsilon$	$\frac{1}{1 - \varepsilon}$	1
C $D_y; H_z = 0$ I_x $\perp D_y$	$e\mu \langle n \rangle g(\varepsilon) F_1 \left(\frac{l_1}{l_t} \right)$	$(1 - \varepsilon) g(\varepsilon) F_1 \left(\frac{l_1}{l_t} \right)$	$g(\varepsilon) F_1 \left(\frac{l_1}{l_t} \right)$	—	$g(\varepsilon) F_1 \left(\frac{l_1}{l_t} \right)$
C' $D_y; H_z \perp D_y$ $I_x \perp D_y \perp H_z$	$e\mu \langle n \rangle g(\varepsilon) F_3 \left(\frac{l_1}{l_t} \right)$	$(1 - \varepsilon) g(\varepsilon) F_3 \left(\frac{l_1}{l_t} \right)$	$g(\varepsilon) F_3 \left(\frac{l_1}{l_t} \right)$	$\frac{1}{1 - \varepsilon}$	$g(\varepsilon) F_3 \left(\frac{l_1}{l_t} \right)$

1. Tablica opiera się na pracach Reada, przyjmuje niezmienną ruchliwość dla przepływu prądu równoległego do dyslokacji,

2. Dla $H \neq 0$ μ/μ_0 — dotyczy ruchliwości hallowskich,

$$3. F_1 \left(\frac{l_1}{l_t} \right) = \frac{3}{l_t} l_{\mu} F_1' \left(\frac{l_1}{l_t} \right) \quad \text{wzór (16),}$$

$$F_2 \left(\frac{l_1}{l_t} \right) = \frac{1}{l_t} l_H F_2' \left(\frac{l_1}{l_t} \right) \quad \text{wzór (17),}$$

$$F_3 \left(\frac{l_1}{l_t} \right) = \frac{3}{l_t} l_3 \quad \text{wzór (18),}$$

σ_0, R_0, μ_0 — przewodnictwo, stała Halla i ruchliwość niezdeformowanego kryształu,

σ_{\perp} — przewodnictwo dla przepływu prądu prostopadłego do dyslokacji,

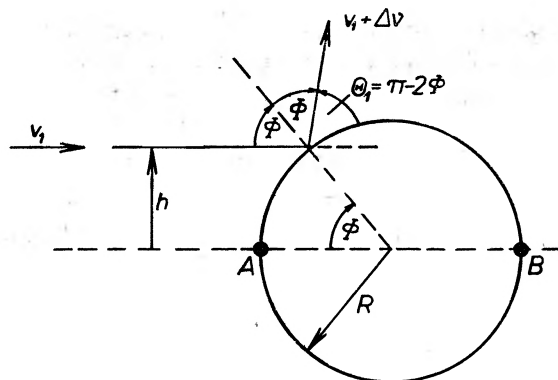
σ_{\parallel} — przewodnictwo dla przepływu prądu równoległego do dyslokacji.

2.3.3. Średni czas swobodny dla rozpraszania na dyslokacjach

Zakładamy, że elektron porusza się w płaszczyźnie xz prostopadłej do dyslokacji z prędkością \vec{V}_1 .

Ze względu na stałość V_y przy rozpraszaniu, zmiana prędkości: $\Delta V = \Delta V_1$.

Prawdopodobieństwo zderzenia elektronu z dyslokacją wynosi $2RN V_1 = \frac{1}{\tau_c}$



Rys. 16. Odbicie elektronu od nieprzenikliwego cylindra ładunku przestrzennego otaczającego dyslokację [6]

τ_c jest czasem pomiędzy zderzeniami. Rys. 16 ilustruje odbicie elektronu padającego na dyslokację z prędkością \vec{V}_1 . Zakładamy, że wszystkie h są jednakowo prawdopodobne. Prawdopodobieństwo odchylenia elektronu o kąt θ_1 zależy od h i wynosi

$$\frac{d h}{d \theta} \frac{1}{R} = \frac{1}{2} \cos \varphi = \frac{1}{2} \sin \frac{1}{2} \theta_1.$$

Zmiana prędkości dla kąta θ_1 wynosi:

$$\Delta \bar{V}_{\theta_1} = -\bar{V}_1 (1 - \cos \theta_1).$$

Średnia zmiana prędkości

$$\langle \Delta \vec{V} \rangle = \frac{1}{R} \int_0^R \Delta V(h) dh = -\bar{V}_1 \int_0^{\pi/2} (1 - \cos \theta_1) \frac{1}{2} \sin \frac{1}{2} \theta_1 d\theta_1 = -\frac{4}{3} \bar{V}_1$$

$$\frac{d \bar{V}}{dt} = \frac{\langle \Delta \bar{V} \rangle}{\tau_c}. \quad (19)$$

Ponieważ składowa y prędkości nie ulega zmianie

$$\frac{d \bar{V}_1}{dt} = -\frac{\bar{V}_1}{\tau_1}. \quad (20)$$

Z równania (19) i (20)

$$\frac{1}{\tau_1} = -\frac{4}{3\tau_c} = \frac{V_1}{l_1} = \frac{8}{3} R N V_1.$$

Znajdźmy $W_1(\theta)$ występujące w (15).

Z definicji

$$\frac{1}{\tau_c} = \int_c^{\infty} W_1(\theta) dC \quad (21)$$

$W_1(\theta_1)$ — prawdopodobieństwo odchylenia elektronu o kąt θ_1 proporcjonalne do $\frac{1}{2} \sin \frac{1}{2} \theta_1$.

(21) określa stałą proporcjonalności i $W_1 = \frac{1}{4} R\tau_c \sin \frac{1}{2} \theta_1$.

3. Sprawdzenie teorii Reada

W pracy (2), w oparciu o powyższą teorię, porównano krzywą doświadczalną z teoretyczną (rys. 5). Niezgodność przypisano przybliżonej ocenie funkcji $g(\varepsilon)$ uwzględniającej efekt dystorsji. $\varepsilon(T)$ i gęstość dyslokacji przyjęto z pomiarów temperaturowej zależności koncentracji, co jest poważnym brakiem.

Dokładne porównanie z doświadczeniem wymagałoby przynajmniej sprawdzenia wszystkich powyżej przedstawionych przypadków. Ze względu na wymaganą dużą dokładność pomiarów nie jest to rzeczą prostą, trudne jest również otrzymanie zdeformowanych kryształów z dobrze zdefiniowanym jednorodnym rozkładem dyslokacji. Prace należy prowadzić na bardzo czystych kryształach, aby móc zaniedbać rozpraszanie na domieszkach, którego to przypadku teoria Reada nie uwzględnia. W świetle pracy [3] założenie elastyczności rozpraszania jest niesłuszne. W wyniku prowadzonych doświadczeń na jednorodnych, wolnych od domieszek kryształach autor zaobserwował następujące niezgodności z teorią Reada:

1. Stosunek ruchliwości $\mu_{||}/\mu_0 = R_{||}\sigma_{||}/R_0\sigma_0 < 1$,
2. $\sigma_{||}/\sigma_{\perp}$; $R(T)$; $\sigma_{||}$ i σ_{\perp} nie spełniają teorii Reada,
3. Magnetoopor deformowanych próbek MR dla $\bar{H}||\langle 110 \rangle$ jest mniejszy od magnetooporu MR_0 dla nie deformowanych próbek, oraz MR deformowanych próbek dla $\bar{H}||\langle 100 \rangle > MR_0$.

Zjawisko to ilustrują rys. 7—9.

$\mu_{||}$, $\sigma_{||}$ — ruchliwość i przewodnictwo w deformowanych kryształach dla przepływu prądu równoległego do linii dyslokacji,

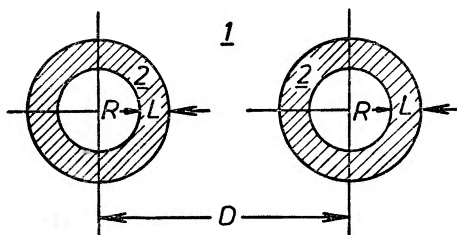
μ_{\perp} , σ_{\perp} — analogicznie dla przepływu prądu prostopadłego do linii dyslokacji,

μ_0 , σ_0 , R_0 — ruchliwość, przewodnictwo i stała Halla niedeformowanych kryształów.

Bardzo ważny jest następujący fakt, że gdy $\mu_{||} < \mu_0$ to $MR_{\langle 100 \rangle} > MR_0$.

Wzrost magnetooporu eliminuje wytłumaczenie malenia ruchliwości, polegające na przyjęciu, oprócz rozpraszania dyslokacyjnego, rozpraszania na jakichś defektach powstających w wyniku deformacji. Zmalenie ruchliwości zgodnie z zależnością $MR = 0,38(\mu_H)^2$ winno wywołać silne zmalenie MR . Fakt wzrostu MR przy maleniu μ można wyjaśnić zakładając w kryształach istnienie obszarów o normalnej ruchliwości materiału niezdeformowanego obok obszarów o zmniejszonej ruchliwości. W celu fenomenologicznego opisanie zjawiska autor przyjmuje teorię dwupasmową Wilsona zakładającą ruch nośników o różnych ruchli-

wościach w dwóch obszarach kryształu. Jeden z obszarów to normalny kryształ, a drugi — obszar o zmniejszonej ruchliwości rozpatruje w przybliżeniu jako ruch nośników w dwóch pasmach. Sprezycujemy bliżej pojęcie dwóch obszarów. Obszar ładunku przestrzennego otaczający dyslokację, pozbawiony nośników podobnie jak w teorii Reada, jest otoczony obszarem 2 tzw. „ Φ ” o zmniejszonej ruchliwości w stosunku do ruchliwości w obszarze 1 niezdeformowanego kryształu. Grubość obszaru Φ przyjmujemy równą długości drogi swobodnej w niezdeformowanym kryształe (rys. 17).



Rys. 17. Model dwóch obszarów dla rozpraszania elektronów na dyslokacjach. R — jest promieniem ładunku przestrzennego otaczającego dyslokację, L — grubością obszaru Φ równą średniej drodze swobodnej, D — odległość dyslokacji [3]

Odpowiednie zależności dwupasmowej teorii Wilsona przybierają postać:

$$\sigma_{\parallel}/\sigma_0 = [1 - \varepsilon - \varphi(1 - \theta)] \quad (22)$$

$$R_{\parallel}R_0 = \frac{1 - \varepsilon - (1 - \theta)^2 \varphi}{[1 - \varepsilon - (1 - \theta)\varphi]^2} \quad (23)$$

$$\frac{\varrho_H}{\varrho_s} - 1 = (\mu_1 H)^2 \frac{\varphi_H \theta_H (1 - \theta_H)^2 (1 - \varepsilon - \varphi_H)}{[1 - \varepsilon - \varphi_H (1 - \theta_H)]^2} \quad (24)$$

$$\sigma_{\perp}/\sigma_{\parallel} = [1 + \varepsilon + \varphi(1 - \theta)]^{-1}. \quad (25)$$

Dla

$$n_1 = n_0(1 - \varepsilon - \varphi)$$

$$n_2 = n_0 \varphi$$

ε — objętość obszaru ładunku przestrzennego, φ — objętość obszaru Φ , n_0 — koncentracja nośników w niezdeformowanym materiale, n_1 — koncentracja nośników w obszarze 1 zdeformowanego kryształu, w porównaniu z n_0 zmniejszona o koncentrację nośników wychwypanych przez dyslokację, analogicznie jak w teorii Reada i koncentrację nośników w obszarze Φ , n_2 — koncentracja nośników w obszarze 2 tzw. Φ zdeformowanego kryształu, $\theta = \frac{\mu_2}{\mu_1}$; μ_1 — ruchliwość w obszarze 1 równa ruchliwości μ_0 niezdeformowanego kryształu, μ_2 — ruchliwość w obszarze 2, ϱ_s — oporność powstająca w wyniku złożenia oporności dwóch obszarów.

Dla $H \neq 0$

$$\varrho_s = \frac{1}{\sigma_{\parallel}(H)}; \quad \sigma_{\parallel}(H) = \sigma_1(H) + \sigma_2(H)$$

ρ_H/ρ_s — jest opornością dodatkową w wyniku mieszania prądów dwóch obszarów o różnych ruchliwościach dla $H \neq 0$. W celu bezpośredniego sprawdzenia z doświadczeniem korzystając z (24) otrzymujemy:

$$M_{\parallel} = MR + 1 = \frac{\rho_H}{\rho_{\parallel}(v)} = \frac{\rho_{\parallel}(H)}{\rho_{\parallel}(0)} (1 + A), \quad (26)$$

ρ_H — oporność zdeformowanego kryształu przy $H \neq 0$, $\rho_{\parallel}(0)$ — oporność dla $H = 0$ zdeformowanego kryształu,

$$\rho_{\parallel}(0) = \frac{1}{\sigma_{\parallel}(0)}; \quad \sigma_{\parallel}(0) = \sigma_{1\parallel}(0) + \sigma_{2\parallel}(0),$$

$\rho_{\parallel}(H) = \rho_s$ — oporność dla $H \neq 0$ zdeformowanego kryształu, dla przepływu prądu równoległego do linii dyslokacji, uwzględniająca zmiany oporu po przyłożeniu pola, właściwe dla obszaru 1 i 2, $A = \frac{\rho_H}{\rho_s} - 1$ czynnik uwzględniający mieszanie prądów obszaru 1 i 2.

(26) daje

$$P = \frac{\rho_H/\rho_{\parallel}(0)}{\rho_0(H)/\rho_0(v)} = \frac{M_{\parallel}}{M_0} = \frac{1 - \varepsilon + \varphi(1 - \theta)}{1 - \varepsilon + \varphi_H(1 - \theta_H)} [1 + A(H)]. \quad (27)$$

Przy przyłożonym polu φ i θ są funkcjami H , ponieważ funkcjami H są ruchliwości μ_1 i μ_2 . Zauważmy, że (24) równa się zeru oraz $M_{\parallel} = M_0$, gdy $\varphi = 0$ lub $\theta = 1$ lub $\theta = 0$, to znaczy że cały efekt wywołany jest istnieniem dwóch obszarów.

Wzory (22)—(27) są bezpośrednio porównywalne z doświadczeniem.

Po przyjęciu:

1. dla funkcji f statystyki Fermiego teorii Reada,
2. dla odległości zerwanych wiązań $c = 4 \times 10^{-8}$ cm,
3. dla θ wartości 0,3 najbardziej zgodnej z doświadczeniem.

Z pomiarów $R_{\parallel}/R_0(T)$ określamy ε , φ i gęstość dyslokacji N oraz sprawdzamy temperaturowe zależności określone przez pozostałe równania. Należy dodać, że położenie poziomu dyslokacyjnego E_2 jest również parametrem, który dobieramy tak, aby osiągnąć maksymalną zgodność z doświadczeniem. Wartość jego wynosi dla Ge 0,52 eV poniżej pasma przewodnictwa, co różni się znacznie od wartości określonej w pracy [2].

Zależność (26) przy przyjęciu doświadczalnej krzywej $M_0(H, \varphi)$ — (φ — kąt pomiędzy kierunkiem pola magnetycznego i normalną do płaszczyzny neutralnej zdeformowanego kryształu, rys. 10) tłumaczy fakt:

$$M_{\langle 110 \rangle} < M_{0\langle 110 \rangle}$$

$$M_{\langle 100 \rangle} > M_{0\langle 100 \rangle}$$

oraz z zależności $M_0(H, \varphi)$ daje zależność $M(H, \varphi)$ zgodną z doświadczeniem.

Zastosowanie teorii Wilsona do kryształu zdeformowanego jest dość dużym przybliżeniem. Praca Broudego pokazuje konieczność modyfikacji teorii Reada przez uwzględnienie nieelastyczności rozpraszania, co autor uwzględnia feno-

menologicznie wprowadzając obszar zmniejszonej ruchliwości otaczający dyslokację. Praca jest bardzo cenna, ponieważ pokazała, że efekty niejednorodności deformacji w znacznym stopniu maskują efekty dyslokacji oraz podała kryteria i warunki jednorodnej deformacji. Różnice wyników tej pracy i prac [1] i [2] nie są wywołane tylko niejednorodnością deformacji, w znacznym stopniu są one wywołane dużą zawartością domieszek w próbkach prac [1] i [2]. Z tych względów w chwili obecnej wyników prac [1] i [2] nie można uważać za dokładne.

4. Pasma dyslokacyjne

4.1. Dyslokacje krawędziowe

W. L. Boncz-Brujewicz i W. B. Głasko [9, 10] rozpatrzyli w sposób kwantowomechaniczny poziomy dziurowe dyslokacji, która wychwytała pewną ilość elektronów. Rozwiązali oni efektywne równanie falowe dla tego przypadku, przy założeniach:

1. Energia potencjalna $V(r, \varphi)$ uwarunkowana jest dwoma czynnikami:
 - a) ekranowanym polem elektrostatycznym naładowanej dyslokacji V_c ,
 - b) potencjałem związanym z deformacją sieci V_d .

Przy wyliczeniu V_c rozpatrywano dyslokację jako równomiernie naładowaną linię nieskończonej długości, co jest niesłuszne dla zbyt małych i zbyt dużych odległości od dyslokacji. Potencjał dyslokacji przyjęto jako potencjał tej linii o wielkości ładunku na jednostkę długości zależnej od temperatury, z uwzględnieniem ekranowania przez swobodne nośniki. W przypadku braku zwyrodnienia i aproksymacji Debeya

$$V_c(r) = \frac{2e}{2\pi\kappa} K_0\left(\frac{r}{L}\right); \quad L = \sqrt{\frac{\kappa kT}{4\pi n e^2}}$$

κ — stała dielektryczna, n — koncentracja nośników w normalnym pasmie. Potencjał V_d dla dyslokacji krawędziowych zaniedbano jako dający mały przyczynek w stosunku do V_c .

2. Układ nie jest lokalnie neutralny, pojawia się człon $V_r > 0$ uwzględniający odpychanie elektronów wychwypanych przez dyslokację.

3. Pole sił jest periodyczne wzdłuż dyslokacji, co jest spełnione przy założeniach dużej długości linii dyslokacji w stosunku do średniej termicznej długości fali elektronu i równomiernego rozmieszczenia domieszek wzdłuż dyslokacji.

4. Mamy do czynienia z prostą strukturą pasmową.

Przy tych założeniach rozwiązując równanie falowe określono w zależności od parametrów:

$$E_0 = \frac{\hbar^2}{2mL^2} \quad \text{i} \quad g^2 = \frac{n_d e^2}{2\pi\kappa E_0}$$

($n_d = \frac{q}{e} = \frac{f}{a}$ = ilość elektronów na jednostkę długości dyslokacji) położenie

krawędzi i gęstości jednowymiarowych pasm dyslokacyjnych. Równanie falowe posiada widmo dyskretne dla wszystkich sensownych wartości g^2 . Dla $g^2 = 1$ ilość pasm jest nieduża, dla $g^2 \gg 1$ mamy całą serię pasm dyslokacyjnych. Ilość ich jest $\simeq \frac{g^2}{2}$. Zauważmy, że poprzez zależność g^2 od temperatury i koncentracji nośników zarówno położenie krawędzi, jak i gęstość pasm dyslokacji zależą od temperatury i koncentracji nośników.

Gęstość stanów w pasmach dyslokacyjnych jest około 10^{-6} razy mniejsza od gęstości stanów w normalnych pasmach. Szerokość pasm jest duża i mogą się one przykrywać z pasmami normalnymi.

4.2. Dyslokacje śrubowe

Dla dyslokacji śrubowych potencjał $V_c = 0$, ponieważ dyslokacje te nie posiadają zerwanych wiązań. Energia oddziaływania nośników z dyslokacją uwarunkowana jest deformacją sieci około dyslokacji i potencjałem Vd . Potencjał ten warunkuje powstanie jednowymiarowych pasm dyslokacyjnych stanów związanych z dyslokacją i rozpraszanie nośników poruszających się w normalnych pasmach. W pracy [10] podano wartości własne i położenia krawędzi pasm dyslokacyjnych dla tego przypadku.

W celu uwzględnienia rozpraszania szukamy rozwiązania równania falowego w postaci:

$$X \sim e^{ikz} + L(\varphi) \frac{e^{ikr}}{\sqrt{r}}$$

$(L(\varphi))^2$ jest efektywną długością rozpraszania.

Przy założeniu niezależności rozpraszania na poszczególnych dyslokacjach o koncentracji N , na średnią ilość cząstek o prędkości $\frac{\hbar k}{m}$ rozproszonych na jednostkę czasu do elementu kąta $d\varphi$ otrzymujemy

$$N|L(\varphi)|^2 \frac{\hbar k}{m} d\varphi.$$

Znajomość $L(\varphi)$ pozwala nam wyliczyć czas relaksacji i całkę zderzeń równania transportu.

Przy założeniu zachowania składowej kwasi-pędu równoległej do dyslokacji i przyjęciu przy uśrednianiu τ statystyki Boltzmana otrzymujemy

$$\mu = \frac{e\tau_0}{m} \sqrt{\frac{kT}{W_0}} \int_0^\infty x^4 \operatorname{cth} \left(\frac{\pi}{2x} \sqrt{\frac{W_0}{kT}} \right) e^{-x^2} dx \quad (28)$$

dla $kT \ll W_0$

$$\mu \simeq \frac{3\sqrt{\pi}}{8} \frac{e\tau_0}{m} \sqrt{\frac{kT}{W_0}}$$

dla $kT \gg W_0$

$$\mu \simeq \frac{2e\tau_0 kT}{\pi M W_0}$$

$$\tau_0 = \frac{2m}{\pi \hbar N}; \quad W_0 = \frac{2m\alpha}{\hbar^2}$$

α — stała występująca w wyrażeniu na Vd — parametr teorii.

Statystyka zapelnienia pasm dyslokacyjnych

Na podstawie teorii W. L. Boncz-Brujewicza, J. W. Gulajew [11] określili stopień zapelnienia pasm dyslokacyjnych w funkcji temperatury i położenia poziomu Fermiego dla silnie domieszkowanych próbek typu n w obszarze niskich temperatur:

$$n_d(T) = n_d(0) - \sum_{\nu} N_z e^{-\frac{\mu - E_{\nu}(n_d(T))}{kT}} \quad (29)$$

$n_d(T)$ jest ilością elektronów wychwytyanych na jednostkę długości dyslokacji w temperaturze T , $N_z = 2 \left(\frac{2\pi M_z kT}{\hbar^2} \right)^{1/2}$ — efektywną ilość stanów w jednowymiarowym pasmie dyslokacyjnym, M_z — efektywną masę dziury w kierunku linii dyslokacji, ν oznacza sumowanie po ilości pasm dyslokacyjnych, μ — położenie poziomu Fermiego, $E_{\nu}(n_d(T))$ — położenie dna ν -tego pasma dyslokacyjnego. $E_{\nu}(n_d(T))$ jest powiązane z $\lambda_{\nu}(n_d(T))$ wartościami własnymi efektywnego równania falowego dla dziur wychwytyanych przez dyslokacje i z V_r — zmianą kulombowskiej energii odpychania przy wychwycie jednej dziury, następującą zależnością

$$E_{\nu}(n_d(T)) = \Delta E + \lambda_{\nu}(n_d(T)) - V_r(n_d(T))$$

ΔE — szerokość przerwy energetycznej.

Zależność (29) otrzymano przy zaniechaniu oddziaływania wychwytyanych dziur, do których zastosowano statystykę idealnego, niezwyrodniałego gazu kwasi-cząstek w równowadze. $f_{\nu k}$ — stopień zapelniania stanu k , ν -tego pasma dyslokacyjnego przyjęto równym $\frac{1}{1 + \exp \left\{ \frac{W_{\nu k} - \mu}{kT} \right\}}$; $W_{\nu k}$ — przedstawia zmianę

energii jednostki długości dyslokacji przy wychwycie jednej dziury do pasma ν na poziom k . Dla średnich temperatur, zgodnie z wynikami pracy W. L. Boncz-Brujewicza, mamy jedno pasmo dyslokacyjne. Wychodząc z doświadczalnie otrzymanej temperaturowej zależności $n_d(T)$ i zależności (29) określono $E_1(T) = 8,10^{-4} T$ eV, co daje $E_1(50^{\circ}\text{K}) = 0,04$ eV i $E_1(200^{\circ}\text{K}) = 0,161$ eV, od dna pasma przewodnictwa dla monokryształu Ge z $N_D = 2,4 \times 10^{14}$ cm $^{-3}$, $N_A = 5 \cdot 10^{18}$ cm $^{-3}$, $E_c - E_D = 0,0097$ eV, $N = 5 \times 10^7$ cm $^{-3}$, $m_e = m_n = 0,25 m_e$ i $n_d(0) = 2,4 \cdot 10^6$ cm $^{-1}$.

Temperaturowa zależność położenia dna pasma dyslokacyjnego w tym zakresie temperatur tłumaczy się temperaturową zależnością V_r . W niższych temperaturach w wyrażeniu na E_p , zaczyna odgrywać rolę czynnik $\lambda(T)$; również rośnie ilość pasm dyslokacyjnych.

Zakończenie

Jednym z ciekawych problemów związanych z przewodnictwem w kryształach z dyslokacjami byłoby doświadczalne stwierdzenie przewodnictwa w pasmach dyslokacyjnych. W niskich temperaturach, gdy przewodnictwo w normalnych pasmach jest bardzo małe, efekt ten prawdopodobnie może być uchwycony (koncentracja elektronów w pasmach dyslokacyjnych wynosi $\sim 10^{13} \text{ cm}^{-3}$), o ile ruchliwość w pasmach dyslokacyjnych nie jest zbyt mała. Dla dyslokacji równoległych do przyłożonego pola elektrycznego efekt Halla powinien zniknąć, gdyż ruch elektronów jest jednowymiarowy, wzdłuż linii dyslokacji. Dla dyslokacji prostopadłych do pola elektrycznego przewodnictwo znika z tych samych względów. Ciekawe jest rozbudowanie teorii podanej przez Gulajewa, dotyczącej zmian koncentracji wywołanych przez dyslokacje z uwzględnieniem wzajemnego oddziaływania nośników wychwytyanych przez dyslokację oraz określenie położenia pasm dyslokacyjnych z uwzględnieniem tego oddziaływania, wpływu domieszek otaczających dyslokację i wpływu deformacji sieci związanej z dyslokacjami.

Teoria zmian ruchliwości w kryształach z dyslokacjami oparta o założenie elastycznego rozpraszania nośników, zbudowana przez Reada, w świetle prac Broude'go nie opisuje zjawiska rozpraszania na dyslokacjach. Należałoby, opierając się o prace Boncz-Brujewicza, zbudować teorię nieelastycznego rozpraszania z uwzględnieniem zmiany rozkładu fononów w otoczeniu dyslokacji, nierówności cylindrów dyslokacyjnych, rozkładu domieszek w otoczeniu dyslokacji, deformacji sieci w otoczeniu dyslokacji. Próba takiego ujęcia jest teoria Boncz-Brujewicza rozpraszania na dyslokacjach śrubowych.

W kryształach z dyslokacjami obserwowano w zależności $R(H)$, $iMR(H)$ dla typu n -Ge maksimum, co może być wyjaśnione podobnie jak i w typie p -Ge istnieniem dwóch rodzajów nośników w pasmach dyslokacyjnych. Wymaga to jednak obecnie dalszych badań doświadczalnych potwierdzających obserwowany efekt. Teoria magnetooporu dla kryształu z dyslokacjami nie istnieje.

Zaznaczyć należy, iż klasyczne ujęcie Reada dotyczące stopnia obsadzenia dyslokacji wymaga uzupełnienia na przypadek małej koncentracji domieszek, w przypadku tym stopień obsadzenia dyslokacji określa położenie poziomu Fermiego i E_p jest funkcją f , również wyliczenie E_g podane przez Reada jest niesłuszne. W przypadku dostatecznej koncentracji domieszek dla spełnienia warunków teorii Reada w wysokich temperaturach należy wnieść poprawkę na ekranowanie dyslokacji przez swobodne dziury, to samo dotyczy przypadku, gdy $E_2 < E_g/2$. W przypadku przybliżenia rozpraszania elastycznego

w niskich temperaturach decydującą rolę odgrywa czynnik $g(\varepsilon)$, czynnik ten jest określony dla małych ε , w przypadku niskich temperatur ε jest duże i należałoby znaleźć $g(\varepsilon)$ dla dużych ε .

W zakończeniu należy wspomnieć, iż istnieje szereg prac dotyczących rekombinacji na dyslokacjach [12].

Przeglądu prac o wpływie dyslokacji na własności elektryczne dokonał w roku 1960 W. Bardsley (*Progress in Semiconductors* **4**, 155 (1960)) i w roku 1963 F. L. Vogel (*Progress in Semiconductors* **6**, 3 (1963)).

E. M. Kuzniecowa (*Fiz. Tw. Tiała* **3**, 1986 (1961)) rozszerzyła prace Reada i Boncz-Brujewicza rozwiązując równanie falowe dla potencjału elektrostatycznego dyslokacji z uwzględnieniem ekranowania zarówno przez domieszki, jak i nośniki swobodne. W tym przybliżeniu w temperaturze 100°K rozpraszanie na dyslokacjach o gęstości $N_D = 10^7 \text{ cm}^{-2}$ jest takie jak na domieszkach o koncentracji $N_D = 10^{14} \text{ cm}^{-3}$, w temperaturze 150°K ruchliwość uwarunkowana przez dyslokacje jest o rząd większa od ruchliwości określonej przez

drżania cieplne sieci. W temperaturze pokojowej $\frac{\mu_{\text{sieci}}}{\mu_{\text{dysl.}}} = 10^{-3}$ dla rozpraszania na potencjale elektrostatycznym dyslokacji i $\frac{\mu_{\text{sieci}}}{\mu_{\text{dysl.}}} = 10^{-4}$ dla rozpraszania na

potencjale deformacji dyslokacji (W. Dexter i F. Seitz *Phys. Rev.* **86**, 770, 964 (1952); W. Shockley, I. Bardeen *Phys. Rev.* **80**, 72 (1950); J. Koechler *Phys. Rev.* **75**, 1 (1949)).

Problem składania rozpraszania na dyslokacjach w przybliżeniu potencjału deformacji i rozpraszania na domieszkach i drżaniach sieci rozpatrywał T. Fukuroi (*Sci. Rep. Res. Insts. Tōhoku Univ. (Japan)*, **9**, 267 (1957)). M. A. Ruwinskij (*Ukrayin. fiz. Zh.*, **9**, 1015 (1964)), uwzględniając potencjał deformacji sieci, potencjał oddziaływania ekscytonu z jądrem dyslokacji, z elektronami zlokalizowanymi na dyslokacji, polaryzację sieci oraz możliwość zlokalizowania kilku elektronów na różnych pasmach dyslokacyjnych rozwiązał równania falowe dla dyslokacji śrubowych. Metodą analogiczną do użytej w pracach Boncz-Brujewicza i Kuźniecowej otrzymał odpowiednie funkcje falowe oraz wartości własne elektronów dyslokacji oddziaływających z ekscytonem. W tym ujęciu rozpraszanie elastyczne zmienia pęd ekscytonu nie zmieniając stanu elektronów na dyslokacjach. W pracy podano średnią drogę swobodną, czas relaksacji i efektywny przekrój takiego rozpraszania ekscytonu. Jest on większy niż przekroje rozpraszania nieelastycznego. Rozpraszanie nieelastyczne odpowiada przegrupowaniu energetycznym elektronów dyslokacji w wyniku oddziaływania z ekscytonem. Największy przekrój czynny posiada reakcja wzbudzenia widma ciągłego translacyjnego ruchu elektronów w pasmie dyslokacyjnym. Procesy związane z dyskretnymi przejściami elektronów dyslokacji w pasmach dyslokacyjnych mają mniejsze przekroje czynne. Należy zauważyć, że w pracach tego typu obok potencjału naładowanej linii, potencjału deformacji i potencjału atmosfery Cottrella należałoby uwzględnić potencjał linii dyslokacji zmieniony przez domieszki związane z zerwanymi wiązaniami

atomów dyslokacji. Istnieją prace zajmujące się badaniami osadzania na dyslokacjach atomów domieszek czy defektów generowanych termicznie, w których stwierdza się zmiany własności elektrycznych dyslokacji oddziaływającej z domieszkami, np. L. S. Milewskij (*Fiz. tverdogo Tela*, **4**, 825 (1962)); W. J. Allen (*J. of Electronics*, **1**, 580 (1956)).

Ciekawe byłyby tu pomiary zmian ruchliwości w pasmie dyslokacyjnym w zależności od stopnia obsadzania dyslokacji przez domieszki. Naturalnie, w tym przypadku, zmienia się też położenie pasm dyslokacyjnych, ich stopień obsadzenia i oddziaływanie nośników wychwytywanych przez dyslokację.

Warto wspomnieć, że możliwość istnienia pasm dyslokacyjnych przewidywał w roku 1953 W. Shockley (*Phys. Rev.*, **91**, 228 (1953)).

Klasyczna teoria Reada znalazła rozwinięcie w pracach K. Kamady (*J. Phys. Soc. Japan*, **15**, 998 (1960)) i J. J. Dugi, R. K. Willardsona i A. C. Beera (*J. Appl. Phys.*, **30**, 1798 (1959); **33**, 169 (1962)). K. Kamada mierzył współczynniki magnetooporu w równaniu F. Seitz'a. Zmiany tych współczynników dla kryształu zdeformowanego w stosunku do niezdeformowanego wyjaśnił składaniem rozpraszania na dyslokacjach według Reada i rozpraszania na drganiach sieci, co prowadzi do zmiany symetrii tensora czasów relaksacji i zmiany współczynników. Ponadto dla $I \perp D || H$ rozwiązując równanie transportu w przybliżeniu Reada znajduje $f_3 \propto EH^2$ i magnetoopór zdeformowanego kryształu.

J. J. Duga i współautorzy badając zdeformowane kryształy InSb interpretowali wyniki przy pomocy teorii Reada z uwzględnieniem, obok rozpraszania na dyslokacjach i drganiach sieci, rozpraszania na domieszkach. Dla badań wpływu dyslokacji na własności galwanomagnetyczne kryształów duże znaczenie mają teorie wpływu niejednorodności w kryształach na zjawiska galwanomagnetyczne (C. Herring, *J. Appl. Phys.*, **31**, 1939 (1960); H. L. Juretschke, R. Landauer, J. A. Swanson, *J. Appl. Phys.*, **27**, 838 (1956); A. C. Beer, *Galvanomagnetic effects in semiconductors* — Suppl. 4 Solid State Phys. (1963)).

Uwaga I

Stopień obsadzenia dyslokacji przez elektrony $f(E_2, T, C)$ dla różnych materiałów zależy od położenia poziomu Fermiego E_F (np. równanie (8)). Wartości $f = 0,11$ dla temperatury azotu i $f = 0,03$ dla temperatury pokojowej wyliczono dla $N_D - N_A = N_A = 10^{15} \text{ cm}^{-3}$ ($N_D - N_A$)^{-1/3} = $L = 1000 \text{ \AA}$ $E_c - E_2 = 0,225 \text{ eV}$ $E_D = 0,01 \text{ eV}$ $c = 4 \text{ \AA}$. Dla tychże wartości wyliczono krzywe rys. 15 [4,5]. W przypadku oscylatora harmonicznego przyjęto $E_b = 0,1 \text{ eV}$. E_2 — energią poziomu dyslokacyjnego mierzymy od wierzchołka pasma walencyjnego na dyslokacji. E_v — wierzchołek pasma walencyjnego na dyslokacji ulega przesunięciu w zależności od potencjału elektrostatycznego w obszarze ładunku przestrzennego. Przesunięcie E_v określone jest przez $E^*(f)$. $E^*(f)$ określone na str. 508 jako energia elektrostatycznego oddziaływania elektronu z pozostałymi elektronami dyslokacji jest ściślej precyzując energią wzajemnego oddziaływania elektronu z pozostałymi elektronami dyslokacji plus energia oddziaływania elektronu z obszarem ładunku dodatniego plus energia oddziaływania ładunku obszaru dodatniego z elektronem dyslokacji plus energia oddziaływania elektronu z najbliższym dodanym elektronem.

Wylczenie E^* opiera się na założeniach strony 507 i w wysokich temperaturach nie jest słuszne. Do ograniczeń wymienionych na str. 507 należy dodać wzrost a z temperaturą i niespełnienie warunku $a \ll (N_D - N_A)^{-1/3} = L$ odległości niezjonizowanych donorów i $a \ll R$, malenie R i niespełnienie warunku $R \gg L$. Dla dużych koncentracji niezjonizowanych donorów $N_D - N_A$, oprócz niespełnienia warunku $a \ll L$, promień R [równanie (2)] staje się mały i założenie cylindra ładunku dodatniego traci sens.

The Effect of Dislocation on the Electric Properties of Germanium II

Abstract: The paper reviews measurements of the resistance, magnetoresistance, Hall constant and mobility for germanium single crystals plastically deformed. Measurements had been made at various mutual orientations of the dislocation axis, current and magnetic field. It is quite difficult to obtain uniformly deformed single crystals of germanium and to eliminate the influence of impurities.

In the light of recent work it appears that the Read classical theory of elastic scattering of carriers by dislocations is incorrect. The quantum-mechanical theory of Bonch-Bruевич forseees the existence of one dimensional dislocation bands in the forbidden gap.

Literatura

- [1] G. L. Pearson, W. T. Read, F. J. Morin, Phys. Rev., **93**, 666 (1954).
- [2] R. L. Logan, G. L. Pearson, D. A. Kleinman, J. appl. Phys. (USA), **30**, 885 (1959).
- [3] R. M. Broudy, Advances in Phys. (GB), **12**, 135 (1963).
- [4] W. T. Read, Phil. Mag., **45**, 775 (1954).
- [5] W. T. Read, Phil. Mag., **45**, 1119 (1954).
- [6] W. T. Read, Phil. Mag., **46**, 111 (1955).
- [7] H. J. Juretschke, R. Landauer, J. A. Swanson, J. appl. Phys. (USA), **27**, 838 (1956).
- [8] R. M. Broudy, J. W. McClure, J. appl. Phys. (USA), **31**, 1511 (1960).
- [9] W. L. Boncz-Brujewicz, W. B. Głasko, Fiz. Tverdogo Tela., **3**, 36 (1961).
- [10] W. L. Boncz-Brujewicz, Fiz. Tverdogo Tela., **3**, 47 (1961).
- [11] J. W. Gulajew, Fiz. Tverdogo Tela., **3**, 1094 (1961).
- [12] T. Figielski, Phys. Status Solidi, **9**, 555 (1965).

Henryk Rzewuski

Instytut Badań Jądrowych
Świerk

Defekty metatrwałe w germanie

1. Wstęp

Wpływ promieniowania jądrowego na ciało stałe jest na ogół procesem złożonym i stanowi od wielu lat przedmiot obszernych badań. Największe zainteresowanie budzą trwale zmiany wywołane bombardowaniem, polegające na zaburzeniu struktury krystalicznej bombardowanych ciał.

W przypadku ciężkich cząstek jądrowych, takich jak n , p , d i α , wywoływane zaburzenia sieci, zwane ogólnie defektami, mają zazwyczaj złożoną „budowę“ i dotychczasowa ich analiza, nawet dla kryształów monoatomowych, jest bardzo trudna.

Dlatego też badania ostatnich lat w tej dziedzinie koncentrują się raczej na badaniu defektów o najprostszej możliwej strukturze, w szczególności defektów punktowych lub tzw. defektów Frenkla. Przez defekt taki rozumie się zwykle parę luka-atom międzywęzłowy znajdujące się blisko siebie.

Procesy radiacyjne zależą od wielu czynników, do najważniejszych jednak pośród nich zaliczyć należy dwa podstawowe parametry, a mianowicie *energie progową* E_{L0} , zwaną również energią Wiegnera, oraz *przekrój czynny na wytwarzanie defektów* $\sigma(E_B)$, będący funkcją energii cząstek bombardujących.

Dokładne określenie wartości E_{L0} związane jest z dużymi trudnościami zarówno teoretycznymi, jak i doświadczalnymi. Oszacowania teoretyczne, zresztą nieliczne, podają bardzo szerokie przedziały wartości E_{L0} , które dla wartości krańcowych różnią się między sobą o czynnik „kilka razy“. Rozrzut wartości doświadczalnych dla przypadku germanu jest nie większy niż o czynnik dwa.

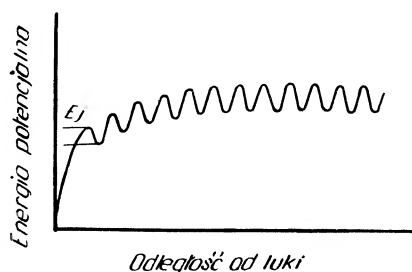
2. Energia wiązania

Pierwsze oszacowanie teoretyczne wartości E_{L0} podał Seitz [1], przyjmując $E_{L0} \simeq 4E_B$, gdzie E_B jest energią sublimacji kryształu, otrzymując na $E_{L0} \simeq 25$ eV.

Dokładniejsze rachunki wykonane zostały w kilka lat później przez Huntingtona [2] dla miedzi oraz przez Kohna [3] dla germanu. O stopniu trudności może świadczyć zakres wartości dla E_{L0} podany w dosyć szerokich granicach. Dla Cu Huntington podaje $18,5 \text{ eV} < E_{L0} < 43 \text{ eV}$ dla kierunku (111), zaś Kohn dla Ge $7 \text{ eV} < E_{L0} < 15 \text{ eV}$ również dla kierunku (111).

Warto może zatrzymać się przez chwilę przy założeniach, jakie przyjął Kohn w swoich obliczeniach energii progowej E_{L0} dla Ge.

Przyjmijmy bardzo uproszczony model przedstawiający zależność energii potencjalnej atomu „frenklowskiego“ (stanowiącego defekt punktowy) w sieci krystalicznej od jego położenia względem luki (rys. 1).



Rys. 1. Zależność energii potencjalnej defektu punktowego w sieci krystalicznej od położenia atomu wybitego względem luki

Według założeń Kohna na wartość E_{L0} składają się głównie dwie wielkości,

1. energia konieczna do zerwania wiązań z czterema najbliższymi sąsiadami;
2. przyrost energii potencjalnej sieci wynikający z utworzenia defektu punktowego i zaburzenia sieci w jego otoczeniu.

Zaniedbuje on natomiast wartość E_j , energię konieczną do pokonania przez wybity atom pierwszej bariery potencjału. Przyjmując na energię jednego wiązania $E_w \simeq 1,85 \text{ eV}$ (z energii spójności), otrzymuje się na dolną granicę $E_{L0} = 4E_w = 7,5 \text{ eV}$.

Jedno z nowszych oszacowań energii wiązania w kryształach o strukturze typu diamentu i blendy cynkowej podał Manca [4].

Energia pojedynczego wiązania dla półprzewodników dwuskładnikowych (A, B) liczona w oparciu o zasadę Paulinga [5] równa jest sumie oddziaływań: niepolarnego $D_{np}(A, B)$ i polarnego $D_p(A, B)$. Pierwszy składnik jest średnią geometryczną energii wiązań niepolarnych obu atomów, a mianowicie

$$D_{np}(A, B) = [D(A, A) \cdot D(B, B)]^{1/2}. \quad (1)$$

Drugi składnik jest proporcjonalny do kwadratu różnicy elektroujemności X_A i X_B obu atomów, a mianowicie

$$D_p(A, B) = (X_A - X_B)^2. \quad (2)$$

Wychodząc z doświadczalnych wartości energii wiązania związków (A, B) wyznaczono energie pojedynczego wiązania dla pierwiastków grupy II, III i IV.

Dla Ge otrzymano na energię pojedynczego wiązania wartość

$$E_s = 37,5 \text{ (Kcal/mol)} = 1,63 \text{ eV/cząsteczkę.}$$

Przyjmując powyższą wartość E_s możemy następnie otrzymać odpowiednio: na energię spójności $E_b = 2E_s = 3,26 \text{ eV}$, na energię wiązania atomu w kryształach $E_w = 2E_b = 4E_s = 6,52 \text{ eV}$ (dla procesów adiabatycznych).

W przypadku procesu nieodwracalnego, jak np. dla zderzeń jądrowych, energia usunięcia atomu z sieci wyniesie $E'_w = 4E_b$, a więc w naszym przypadku $E'_w \simeq 13 \text{ eV}$.

Pierwsze pomiary energii progowej E_{L0} dokonane zostały dla germanu przez Klontza [6]. Próbkę germanu typu *n* bombardowano elektronami z akceleratora Van de Graaffa mierząc zmiany przewodnictwa elektrycznego w funkcji energii padających elektronów. Ekstrapolowana na krzywej doświadczalnej wartość energii progowej elektronów, przy której tworzą się jeszcze defekty sieci, wynosi $E_B \simeq 0,63 \text{ MeV}$. Odpowiada to progowej energii wybicia dla germanu $E_{L0} \simeq 30 \text{ eV}$.

Loferski i Rappaport [7] zastosowali o wiele czulszą metodę pomiaru opartą na zmianach czasu życia nośników mniejszościowych w zależności od liczby wprowadzanych defektów. Pomiary dla germanu i krzemu dały wartość $E_{L0} \simeq 13 \text{ eV}$, a więc znacznie bliższą oszacowań teoretycznych Kohna.

Dla innych materiałów podajemy niżej za Dinesem [8] zestawienie wyników doświadczalnych i teoretycznych dotyczących wartości energii progowej wybicia atomu z węzła sieci krystalicznej.

Tablica 1

Wartość energii progowej E_{L0}

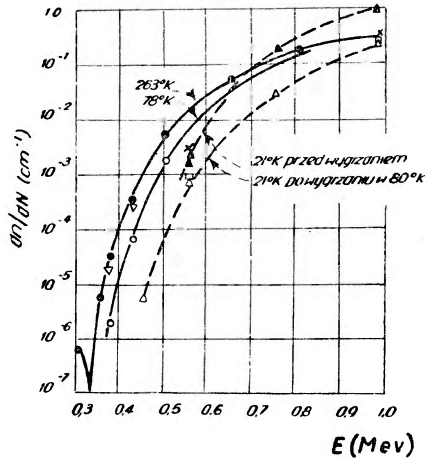
Materiał bombardowany	Mierzona wielkość	Energia progowa (eV)		Odsyłacze
		dośw.	teoret.	
German typu <i>n</i> $T = -196^\circ\text{C}$	oporność	30		Klontz [6]
German typu <i>n</i> $T = T_{\text{pok.}}$	czas życia	13		Loferski i Rappaport [7]
Krzem typu <i>p</i> $T = T_{\text{pok.}}$	czas życia	13		Loferski i Rappaport [7]
German			~ 10	Kohn [3]
Grafit	oporność	25		Eggen [9]
Miedź	oporność	25		Eggen i Laubenstein [10]
Miedź			17—34	Huntington [2]
Stop żelazo—miedź	nasyc. magnetyzacji	27		Denney [11]
Cu ₃ Au	uporządkowanie	~ 10		Dugdale [12]

3. Wydajność wytwarzania defektów $\sigma(E_B)$

Wydajność wytwarzania defektów $\sigma(E_B)$ (lub przekrój czynny na ich wytwarzanie) jest funkcją trzech parametrów: a) energii cząstki bombardującej — E_B , b) temperatury próbki — T , c) natężenia wiązki cząstek bombardujących — Φ .

Zajmiemy się obecnie dwoma pierwszymi parametrami E_B i T . Na rys. 2 podana jest zależność $\sigma(E_B)$ dla kilku temperatur T , wzięta z pracy Browna i Augustyniaka [13]. Krzywe przedstawiają przebieg wydajności wytwarzania defektów $\sigma(E_B)$, a ściślej mówiąc zmianę koncentracji nośników większościowych — $\frac{dn}{dN}$, do której wydajność $\sigma(E_B)$ jest proporcjonalna, dla germanu bombardowanego elektronami monoenergetycznymi z akceleratora Van de Graaffa, w zależności od energii E_B tych elektronów.

W doświadczeniu mierzono przewodnictwo elektryczne σ oraz stałą Halla R_H .



Rys. 2. Temperaturowa zależność wydajności wytwarzania defektów (stanów akceptorowych) w germanie w funkcji energii bombardujących elektronów

Faktem, na który należy zwrócić uwagę, jest wpływ temperatury T na położenie krzywych $\sigma(E_B)$. Jak widać, wyższym temperaturom odpowiadają wyższe wartości wydajności $\sigma(E_B)$. Podobną zależność obserwowali i inni autorzy, a mianowicie Wertheim [14] mierząc dla krzemu oporność ρ i stałą Halla R_H , oraz Bemski [15] i Watkins [16] badając elektronowy rezonans spinowy w krzemie typu n bombardowanym elektronami o energii 500 keV i 1,5 MeV.

Możliwym sposobem wytłumaczenia takiej zależności temperaturowej, zgodnie z ideą Loferskiego i Rappaporta, byłoby uwzględnienie oddziaływania fononowego.

Atom sieci może posiadać pęd, który przypadkowo pokrywa się co do kierunku z pędem otrzymanym przez atom w wyniku zderzenia z szybkim elektronem. Ponieważ pędy dodają się, wypadkowa energia atomu jest większa niż suma jego energii drgań (fononowej) i energii zderzenia dla przypadku atomu w spoczynku. Dla temperatury $T \sim 263^\circ\text{K}$ otrzymamy w przybliżeniu

$$E_t \simeq E_c + 2(E_c E_v)^{1/2} = 16,5 \text{ eV},$$

gdzie E_t — energia wypadkowa, $E_c \simeq 15 \text{ eV}$ — energia uzyskana w zderzeniu dla atomu w spoczynku (założona) i $E_v \simeq 0,03 \text{ eV}$ — energia drgań. Dodatkowy przyrost energii wynosi więc aż około $1,5 \text{ eV}$. Należy oczywiście pamiętać, że sytuacja taka zachodzić może tylko dla drgań optycznych, dla których sąsiednie atomy drgają w przeciwfazie.

Otrzymana nadwyżka energii jest w przybliżeniu wystarczająca, aby uzyskać otrzymane przesunięcie krzywych $\sigma(E_B)$ dla $T = 263^\circ\text{K}$ i $T = 78^\circ\text{K}$.

Sytuację komplikują jednak dwie dalsze krzywe wydajności otrzymane dla $T = 21^\circ\text{K}$. Jak widać, również i tu otrzymujemy podobne przesunięcia temperaturowe krzywych $\sigma(E_B)$.

Tym razem przesunięcia tego nie można wytłumaczyć współdziałaniem fononowym, bowiem nie tylko w temperaturze $T = 21^\circ\text{K}$, lecz również i w temperaturze $T = 78^\circ\text{K}$ nie otrzymujemy pobudzenia optycznego (jedynie drgania zerowe) wobec wysokiej temperatury Debye'a w germanie $\theta \sim 400^\circ\text{K}$.

4. Model defektu metastabilnego

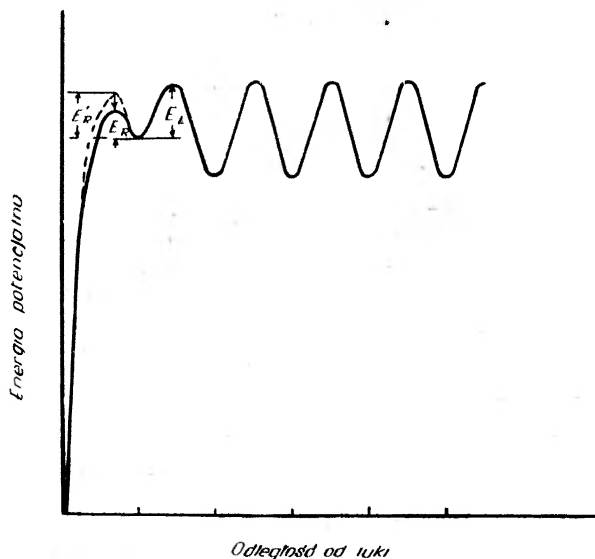
Celem wytłumaczenia wzrostu wydajności wytwarzania defektów ze wzrostem temperatury Wertheim [17], a następnie MacKay i Klontz [18], zaproponowali model metastabilnego stanu pary frenklowskiej luka — atom międzywęzłowy. Model energetyczny takiego metastabilnego defektu punktowego pokazano na rys. 3.

Atom wybity do pozycji międzywęzłowej w odległości $l = a$ (stała sieci) może na skutek pobudzenia termicznego rekombinować z własną luką poprzez barierę E_R , lub może być uwolniony przechodząc przez barierę E_L , tworząc wtedy defekt stabilny.

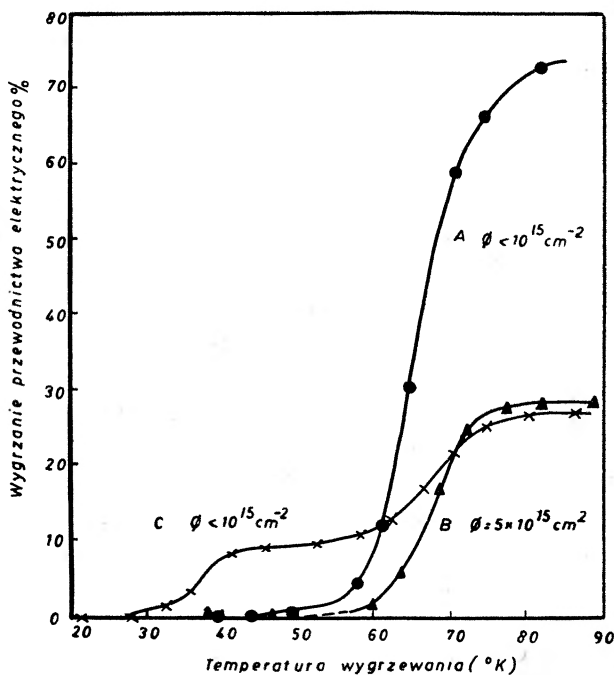
Ponadto, uprzedzając ściślejszą motywację powyższego modelu zakłada się, że bariera na rekombinację E_R zależy od obsadzenia stanu elektronowego atomu znajdującego się w stanie metastabilnym.

Gdy defekt chwyci elektron z kryształu, bariera na rekombinację wzrośnie do wartości $E'_R > E_R$. Wtedy zaś względne prawdopodobieństwo termicznego uwolnienia atomu wybitego wzrośnie.

Przeprowadzone wcześniej prace MacKaya i Klontza [18], [19] na germanie naświetlanym elektronami w niskich temperaturach, a zwłaszcza obserwacje progów wygrzewania termicznego wprowadzonych defektów, pozwoliły na sklasyfikowanie 3 typów defektów — A , B i C .



Rys. 3. Model energetyczny defektu metastabilnego



Rys. 4. Wygrzewanie defektów w germanie napromieniowanym elektronami o energii 1,1 MeV w temperaturze 10°K

Na rys. 4 podano krzywe wygrzewania otrzymane przez powyższych autorów.

Defekty klasy A odpowiadają za wygrzewanie w temperaturze $T \sim 65^\circ\text{K}$.

Defekty klasy C odpowiadają za wygrzewanie w temperaturze $T \sim 35^\circ\text{K}$,

zaś defekty klasy B są pozostałymi defektami stabilnymi, aż do temperatury $T \sim 130^\circ\text{K}$.

Z analizy powyższych krzywych można wyciągnąć następujące wnioski. Defekty A są wynikiem pierwotnego procesu wybicia i są defektami metatrwałymi. Defekty B i C stanowią defekty wtórne będące modyfikacjami defektów A . W dalszym ciągu zajmujemy się tylko defektami klasy A .

Defekt klasy A będący w stanie metastabilnym wprowadza do przerwy energetycznej w modelu pasmowym poziom lokalny akceptorowy w odległości około $E_{el} = 0,1$ eV od dna pasma przewodnictwa. Ocena położenia tego poziomu opiera się na analizie krzywej wygrzewania w otoczeniu progu $T = 65^\circ\text{K}$. Wygrzewanie zachodzi tu z prędkością ograniczoną stałą czasu pobudzenia elektronu z tego poziomu do pasma przewodnictwa. Częstość przejść jest tu $\nu_0 \simeq 10^3 \text{ sek}^{-1}$, podczas gdy dla przejść sieciowych $\nu_g \simeq 10^{13} \text{ sek}^{-1}$.

Wracając do modelu energetycznego defektu metastabilnego z rys. 3 postaramy się ocenić odpowiednie wysokości barier.

Wysokość bariery na rekombinację defektu A w przypadku elektronu wychwyconego na poziom E_{el} musi być $E'_R \geq 0,2$ eV, bowiem mniejsza wartość oznaczałaby, że w temperaturze $T \sim 65^\circ\text{K}$ może zachodzić rekombinacja defektu A nie ograniczona termicznym pobudzeniem elektronu.

Wysokość bariery na rekombinację defektu A bez elektronu na poziomie E_{el} powinna zawierać się w granicach $E'_R - E_{el} < E_R < E_L$. Dolną granicą na E_R byłaby wysokość bariery równa energii wiązania elektronu E_{el} .

Różnica wysokości barier $E_L - E'_R$ powinna być równa w przybliżeniu kT dla $T \sim 80^\circ\text{K}$, aby można było otrzymać dostateczną liczbę defektów metastabilnych przechodzących następnie w stan stabilny w tej temperaturze.

Dla powyższego modelu temperatura w sposób oczywisty będzie grała zasadniczą rolę.

W niskich temperaturach, dla których $E_L - E'_R$ oraz $E_L - E_R$ są dużo mniejsze od kT , jedynie mała część defektów A będzie termicznie rozdzielona przechodząc w defekty trwałe.

W temperaturach pośrednich, dla których $kT \ll E_L - E_R$, lecz $kT \approx E_L - E'_R$, produkcja defektów stabilnych zależeć będzie od temperatury i koncentracji elektronów swobodnych.

W temperaturach wysokich, kiedy $E_L - E_R \ll kT$ i $E_L - E'_R \ll kT$, produkcja defektów nie będzie zależeć od koncentracji elektronów swobodnych.

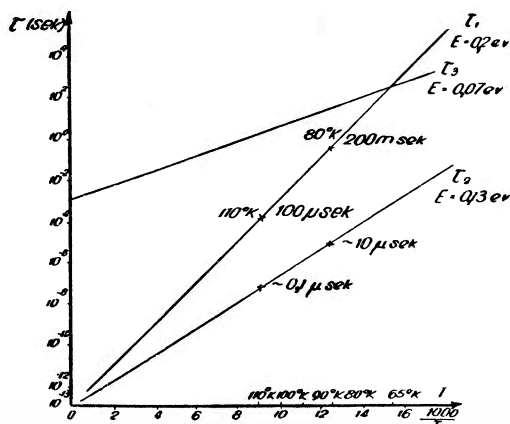
Mamy więc w ogólnym przypadku trzy możliwe procesy związane z wytworzeniem defektu A :

1. proces elektronowy pułapkowania elektronu na poziom o energii $E_{el} \approx 0,1$ eV,
2. proces sieciowy rekombinacji defektu przez barierę E_R lub E'_R ,
3. proces sieciowy rozdzielania defektu przez barierę E_L .

Możemy teraz przedstawić przebiegi stałych czasu dla odpowiednich procesów w zależności od temperatury [20], jak na rys. 5.

Dotychczas prowadzone badania wygrzewania defektów sieci pozwoliły

jedynie na oszacowanie energii aktywacji $E_{el} \simeq 0,1$ eV i ustalenie, że jest to proces elektronowy. Wyznaczenie energii E'_R byłoby już bardzo trudne ze względu na udział obu procesów o energiach aktywacji E_R i E_{el} oraz krótkie czasy występujące dla obszaru, gdzie przeważa proces o energii E'_R .



Rys. 5. Stałe czasu dla procesów τ_1 , τ_2 i τ_3 w funkcji temperatury. τ_1 — stała czasu rekombinacji atomu międzywęzłowego przez barierę $E'_R = 0,2$ eV, τ_2 — stała czasu rekombinacji atomu międzywęzłowego przez barierę $E_R = 0,13$ eV, τ_3 — stała czasu pułapkowania elektronu na poziomie energetycznym $E_{el} = 0,07$ eV

Dobre rozdzielenie obu energii aktywacji E_{el} i E'_R przy pomiarze E'_R otrzymać można jedynie w wyższym zakresie temperatur, np. dla $T > 80^\circ\text{K}$. W tym zaś obszarze stałe czasu procesu zachodzącego z energią aktywacji $E'_R \simeq 0,2$ eV są rzędu 100 msek i mniej. Dotychczasowe metody pomiaru mierzonych parametrów, jak np. metodą wygrzewania izotermicznego lub wygrzewania izochronicznego, na taki pomiar nie pozwalają.

5. Impulsowa metoda badania kinetyki wygrzewania defektów metastabilnych

Zaproponowana metoda [21] pomiaru szybkich procesów wygrzewania defektów punktowych polega w ogólności na impulsowym wytwarzaniu defektów w Ge w odpowiedniej temperaturze wyższej od progu wygrzewania, a następnie ciągłej obserwacji w czasie metodami oscyloskopowymi zmian spowodowanych ich wygrzewaniem.

Nie dyskutując bliżej technicznych możliwości zastosowania tej metody, należy zwrócić uwagę na pewne szczególne cechy modelu metatrwałego z punktu widzenia możliwości pomiaru zachodzących zmian, związanych z wygrzewaniem defektów punktowych.

Jeżeli przyjmiemy, że naszym parametrem mierzonym jest koncentracja swobodnych nośników prądu (elektronów w przypadku germanu typu n), wtedy muszą być spełnione pewne warunki, aby móc zaobserwować proces E'_R .

Przed wszystkim widać od razu, że proces E_R anihilacji bliskiej pary bez

uprzedniego wychwycenia elektronu przez atom międzywęzłowy nie może być zaobserwowany, ponieważ nie zmienia koncentracji początkowej n_0 . Proces E'_R może być zaobserwowany z chwilą wychwycenia elektronu z pasma przewodnictwa i tym samym zmiany początkowej koncentracji n_0 . Wygrzaniu takiego defektu będzie towarzyszył wzrost koncentracji n do uprzedniej wartości n_0 .

Jak więc widać, możliwość zaobserwowania procesu E'_R zależy od prawdopodobieństwa wychwycenia elektronu z pasma przewodnictwa przez defekt w stanie o energii E_R . Prawdopodobieństwo to jest oczywiście funkcją wielu czynników, takich jak: koncentracja nośników n , temperatura T , przekrój czynny defektu na wychwyt elektronu, energia elektronu w pasmie przewodnictwa.

Metoda [21] zapoczątkowana w pracy [22] polega na wytworzeniu defektów sieci w germanie typu n , bombardując próbkę przy pomocy krótkotrwałego impulsu elektronów o energii $E \sim 1$ MeV z akceleratora Van de Graaffa.

Impulsowe wytworzenie defektów sieci w próbce powoduje skokową zmianę koncentracji nośników prądu na skutek obsadzenia elektronami z pasma przewodnictwa poziomów akceptorowych związanych z wytworzonymi defektami.

W temperaturze $T > 60^\circ\text{K}$, powyżej progu wygrzewania, część defektów zrekombinuje poprzez barierę E'_R uwalniając spulapkowane elektrony (patrz rys. 3). Zmiana koncentracji nośników prądu zachodzić więc będzie z pewną stałą czasu τ odpowiadającą wysokości bariery E'_R .

Proces ten można zaobserwować mierząc zmiany koncentracji nośników prądu metodą pomiaru stałej Halla R_H .

W czasie naświetlania próbki impulsem elektronowym występuje jednak szereg efektów utrudniających pomiar napięcia Halla V_H .

Elektrony o energii $E \sim 1$ MeV bombardujące próbkę wywołują następujące efekty:

1. silny efekt jonizacyjny,
2. wytworzenie pewnej liczby defektów sieci i związaną z tym zmianę koncentracji nośników większościowych,
3. zanik wywołanej zmiany koncentracji (par. 2) na skutek anihilacji defektów — jest to efekt obserwowany,
4. podwyższenie temperatury próbki w czasie przyłożonego impulsu elektronowego, która to zmiana zanika z pewną stałą czasu zależną od materiału, wymiarów i warunków chłodzenia próbki,
5. zmiana oporności właściwej próbki na skutek zmiany temperatury (par. 4),
6. dalsze efekty temperaturowe: *a.* siła termoelektryczna na kontaktach Halla wynikająca z niejednorodnego rozkładu temperatury, *b.* zakłócenie symetrii oporowej w pomiarze Halla wynikające z zakłócenia równowagi termicznej wewnątrz bombardowanego obszaru (niejednorodne naświetlanie lub niejednorodna próbka).

W ogólności mamy do czynienia z trzema rodzajami przebiegów czasowych o pewnej amplitudzie i widmie częstotliwości.

A. Efekt jonizacyjny. Stała czasu zaniku tego efektu zależy od objętościowego czasu życia nośników prądu w danej temperaturze i może być rzędu $\tau \sim 1 \mu\text{sek}$. W warunkach eksperymentu poziom pobudzenia jonizacyjnego jest o 3 do 4 rzędów wielkości większy od początkowej koncentracji nośników prądu w próbce. Zanik tego pobudzenia do wartości zaniedbywalnej następuje po czasie całkowitym $t \approx (40-50) \mu\text{sek}$. Po tym więc czasie od chwili zadziałania impulsu elektronowego można prowadzić obserwacje właściwego efektu.

B. Procesy rekombinacyjne defektów metastabilnych. Czasy te określone są temperaturą próbki i mogą być dobrane w szerokim zakresie wartości. Są one ograniczone: od góry (długie stałe czasu) przez przejścia elektronowe pomiędzy poziomem akceptorowym defektu i pasmem przewodnictwa, zaś od dołu (krótkie stałe czasu) — przez całkowity czas zaniku jonizacji lub stałe czasu pułapkowania elektronów przez stany metastabilne [22]. Orientacyjne wartości najkrótszych stałych czasu wynoszą $\tau \approx 100 \mu\text{sek}$.

C. Efekty temperaturowe. W ogólności stałe czasu tych efektów mogą zawierać się w zakresie wartości czasów rekombinacyjnych defektów metastabilnych, zaś co do wartości amplitudy — przewyższać te ostatnie o rząd wielkości i więcej.

Celem omawianej metody jest wyeliminowanie wpływu efektów temperaturowych stanowiących główną trudność eksperymentu. Metoda polega na pomiarze impulsowej zmiany napięcia Halla w zmiennym polu magnetycznym o częstotliwości akustycznej, jaka zachodzi przy naświetleniu próbki germanu krótkotrwałym, pojedynczym impulsem elektronów z Van de Graaffa.

Wyrażając odpowiednio stały prąd płynący przez próbkę J , pole magnetyczne H oraz stałą Halla R_H przez wyrażenie:

$$J = J_0 \quad (3)$$

$$H = H_0 \cos \omega_1 t \quad (4)$$

$$R_H = R_0(1 + m \cos \omega_2 t) \quad (5)$$

otrzymamy wyrażenie na napięcie Halla V_H w postaci

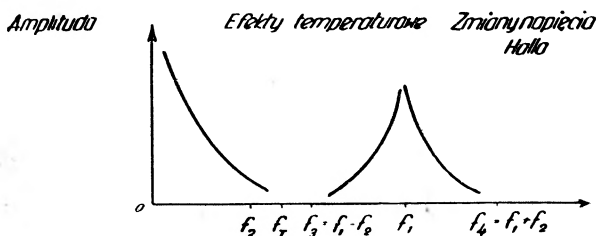
$$V_H = V_{H0} \left\{ \cos \omega_1 t + \frac{1}{2} m [\cos(\omega_1 - \omega_2)t + \cos(\omega_1 + \omega_2)t] \right\}. \quad (6)$$

Wyrażenie na stałą Halla R_H jest przybliżone i słuszne dla $m \ll 1$. $m = \frac{\Delta R}{R_0}$ oznacza względną zmianę współczynnika Halla wywołaną wytworzonymi defektami. Dla uproszczenia przyjęto również górną częstotliwość ω_2 zmian stałej Halla zamiast, w ogólności, rozkładu widmowego częstotliwości.

Jak więc widać, rozpatrywane przebiegi V_H zachodzą tak jak w przypadku przebiegów o modulowanej amplitudzie. Fala nośną jest tu zmienne napięcie Halla V_H o częstotliwości ω_1 , sygnałem modulującym jest stała Halla R_H (poprzez zmiany koncentracji nośników prądu) o częstotliwości ω_2 , zaś głębokością modulacji jest wielkość $m = \frac{\Delta R}{R_0}$.

Rozkład widmowy mierzonego napięcia Halla V_H podano na rys. 6.

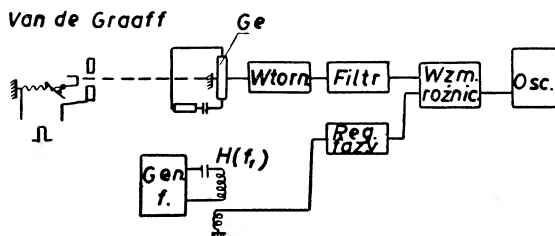
Tak więc, w wyniku zastosowania zmiennego pola magnetycznego o częstotliwości f_1 , badany efekt zmian stałej Halla o widmie częstotliwości zawierającym się od $0 \div f_2$ zostaje przesunięty do pasma częstotliwości zawierającego się pomiędzy $f_3 \div f_4$.



Rys. 6. Widmo częstości mierzonych efektów

Ponieważ efekty temperaturowe, nie modulują częstotliwości f_1 pola magnetycznego, ich widmo rozciągające się od $0 \div f_T$ pozostanie nieprzesunięte. Jeśli warunki doświadczalne zostaną tak dobrane, że częstotliwość f_T będzie leżała poniżej pasma $f_3 \div f_4$, wtedy można będzie wyeliminować szkodliwe efekty temperaturowe stosując filtr pasmowy o pasmie przepuszczania od $f_3 \div f_4$.

Schemat blokowy układu pomiarowego przedstawia rys. 7.



Rys. 7. Blokowy schemat układu do pomiaru impulsowych zmian współczynnika Halla w zmiennym polu magnetycznym

Zastosowanie układu przesuwania fazy pozwala na skompensowanie częstotliwości nośnej f_1 , będącej sumą napięcia Halla V_H oraz napięcia indukującego się w układzie od pola magnetycznego.

Pomiarowi podlega wtedy jedynie impulsowa zmiana, co w rezultacie daje większą czułość pomiaru.

6. Uwagi końcowe

Reasumując można powiedzieć:

1. Stan metatrwały związany jest z bliską parą Frenkla luka—atom międzywęzłowy,

2. Wszystkie inne defekty wytwarzane bombardowaniem elektronowym są wynikiem modyfikacji defektu frenklowskiego,
3. Wydajność wytwarzania defektów jest mała, jeśli proces ten idzie przez defekty metatrwałe,
4. Energia progowa E_{L0} jest być może niższa niż obecnie mierzona.

Metastable Defects in Germanium

Abstract: The model of metastable defect in germanium proposed by Wertheim [15] and MacKay and Klontz [16] is discussed.

The efficiency of defect production is strongly dependent on the temperature during irradiation and the charge state of the defect.

The new method of annealing kinetics investigation of metastable defects in *n*-type germanium is described.

Literatura

- [1] F. Seitz, Disc. Faraday Soc. (GB), **5**, 271 (1949).
- [2] H. B. Huntington, Phys. Rev., **93**, 1414 (1954).
- [3] W. Kohn, Phys. Rev., **94**, 1409 (1954) (A).
- [4] P. Manca, J. Phys. Chem. Solids (GB), **20**, 268 (1961).
- [5] L. Pauling, *The Nature of the Chemical Bond* (3rd Ed) Cornell University Press, New York (1960).
- [6] E. E. Klontz, Phys. Rev., **82**, 763 (1951) (A); **86**, 643 (1952) (A).
- [7] J. J. Loferski, P. Rappaport, Phys. Rev., **98**, 1861 (1955); **100**, 1261 (1955) (A); **111**, 432 (1958).
- [8] J. J. Harwood and al., *The Effects of Radiation on Materials*, Reinhold Publishing Corporation, New York (1958), s. 10.
- [9] D. T. Eggen, praca nie publikowana. Patrz G. R. Hennig, J. E. Hove, publikacja nr 751 zgłoszona na Konferencję Genewską Pokojowego Wykorzystania Energii Atomowej (1955).
- [10] D. T. Eggen, M. J. Laubenstein, Phys. Rev., **91**, 238 (1955) (A).
- [11] J. M. Denney, Phys. Rev., **92**, 531 (1953) (A).
- [12] R. A. Dugdale, *Report on the Conference on Defects in Crystalline Solids*, Bristol, 1954 (London, Physical Society) s. 246.
- [13] W. L. Brown, M. Augustyniak, J. appl. Phys. **30**, 1300 (1959).
- [14] G. K. Wertheim, odsyłacz w pracy [13].
- [15] G. Bemski, G. Fehar, Gere., Bull. Amer. Phys. Soc. Ser. II, **3**, 135 (1958), G. Bemski, J. appl. Phys., **30**, 1195 (1959).
- [16] G. D. Watkins, J. W. Corbett, R. M. Walker, Bull. Amer. Phys. Soc. Ser. II, **4**, 159 (1959) oraz J. appl. Phys., **30**, 1198 (1959).
- [17] G. K. Wertheim, Phys. Rev., **115**, 568 (1959).
- [18] J. W. Mackay, E. E. Klontz, J. appl. Phys., **30**, 1269 (1959); *Radiation Damage in Solids*, **3**, 27 (1962), Proceedings, Venice.
- [19] J. W. Mackay, E. E. Klontz, G. W. Gobeli, Phys. Rev. Letters, **2**, 146 (1959).
- [20] J. W. MacKay pryw. informacja.
- [21] H. Rzewuski, *Radiation Damage in Semiconductors*, 7th International Conference on the Physics of Semiconductors, **3**, 35 (1964), Paryż—Royaumont.
- [22] H. Rzewuski, Raport IBJ No 473/II (1963), Warszawa.

Jerzy Dera

Stacja Morska ZG-PAN

Jan Kalinowski

II Katedra Fizyki Politechniki Gdańskiej

Wybrane zagadnienia fizyki morza**Część I. Przenoszenie energii promienistej w morzu****1. Uwagi wstępne**

W związku z realnymi przewidywaniami wszechstronnej eksploatacji oceanów, w ostatnich kilku latach obserwuje się coraz większe zaangażowanie nauk ścisłych do badań morza. Przedmiotem badań fizyki morza jest przebieg zjawisk fizycznych w morzu. Prawa rządzące tymi zjawiskami mają niejednokrotnie specyficzny charakter, wynikający z makroskali ośrodka. W olbrzymich masach wodnych oceanu światowego, zasadniczą rolę odgrywają często inne zjawiska fizyczne niż w małych objętościach wody. Na przykład zmiana objętości wody pod wpływem ciśnienia hydrostatycznego, praktycznie niezauważalna w warunkach laboratoryjnych (ściślność wody $\approx 5 \cdot 10^{-5}$ cm²/KG), powoduje średnie obniżenie poziomu powierzchni oceanów o ok. 30 m.

Wśród wielu działów fizyki morza znajdujemy dynamikę, optykę, akustykę, termodynamikę i inne. Treścią tego artykułu jest omówienie jednego z głównych zagadnień optyki morza, jakie stanowi transport i rozkład energii promienistej w morzu.

Energia promienista w morzu, głównie pochodzenia słonecznego, podtrzymuje wszelkie zachodzące w nim procesy biologiczne i większość dynamicznych.

Syntezę współczesnego stanu wiedzy z zakresu przenoszenia energii promienistej w morzu zawdzięczamy głównie pracom Duntley'a [1], [2], Jerlova [3], [45], Josepha [24], Kozlaninowa [4], [5], Preisendorfera [6], [7], Schellenbergera [8], [9] i Tylera [7], [10].

2. Woda morska jako ośrodek optyczny

Zjawiskiem powszechnie spotykanym w morzach i oceanach jest przestrzenne zróżnicowanie wartości parametrów opisujących stan mas wodnych (np. temperatury, gęstości, składu chemicznego). Parametry te nie są niezależne. Szczególnie istotną jest zależność gęstości ρ od zasolenia S , temperatury t i ciśnienia p . Określa ona uwarstwienie gęstościowe mas wodnych wraz ze wszystkimi jego konsekwencjami [11], [12], [13], [14], [15], [16], [17]. Jakkolwiek obserwowane różnice $\rho(S, t, p)$ występują na ogół dopiero na trzecim miejscu po przecinku, to należy podkreślić, że w olbrzymiej masie wodnej oceanów stanowią one bezpośrednią przyczynę wielu morskich procesów dynamicznych [11], [15], [18]. Uwarstwienie gęstościowe mas wodnych ma wprawdzie mały wpływ na warunki propagacji fal świetlnych poprzez zróżnicowanie ich współczynnika załamania, powoduje ono jednak zróżnicowanie wód pod innymi względami charakteryzującymi absorpcyjno-rozpraszający ośrodek optyczny jakim jest woda morska. Uwarstwienie warunkuje np. gromadzenie się zawieszin, rozkład koncentracji substancji organicznych, rozwój mikroplanktonu itp. [17], [19], [20], [22], [23].

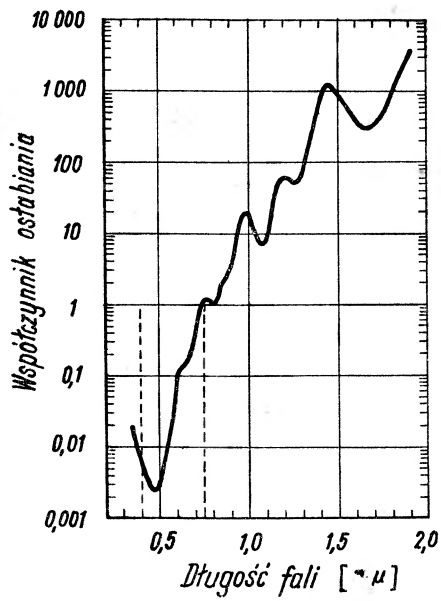
a) Podstawowy skład chemiczny

Woda morska stanowi roztwór soli niemal wszystkich spotykanych w przyrodzie pierwiastków chemicznych. Stężenie tych soli nie przekracza jednak 4%, a ich podstawowy skład jest na ogół stały (tablica I).

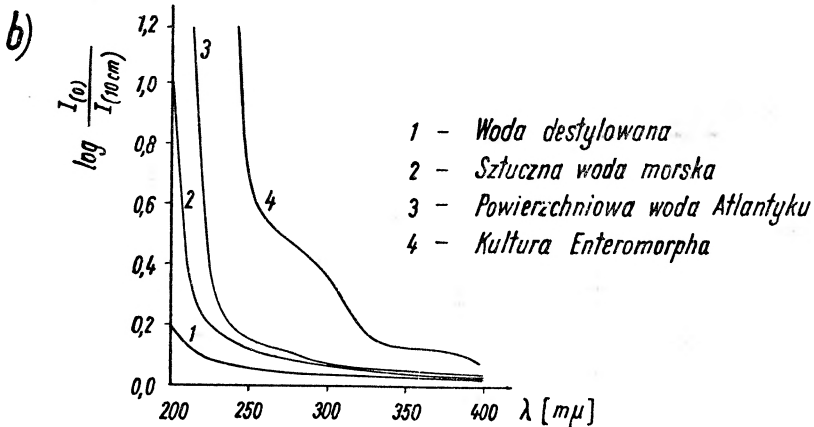
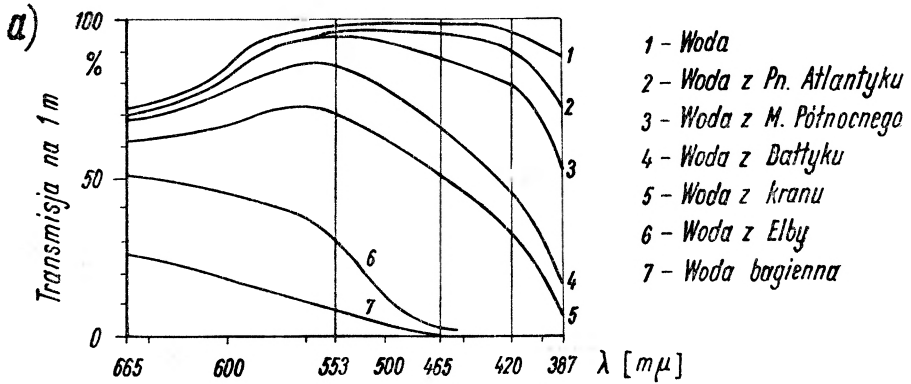
Tablica I
Główny skład jonów soli morskiej ($S = 34,4\%$ wg Defanta [11].)

	Składnik	Zawartość w g na 1 kg wody	Zawartość procentowa w całej soli
kationy	Sód	10,47	30,4
	Potas	0,38	1,1
	Magnez	1,28	3,7
	Wapń	0,41	1,2
	Stront	0,013	0,05
aniony	Chlorki	18,97	55,2
	Bromki	0,065	0,2
	Siarczany	2,65	7,7
	Kwaśne węglany	0,14	0,4
	Borany	0,027	0,08

Poza głównymi składnikami soli morskiej, podanymi w tablicy I, występuje w niej liczna grupa tzw. pierwiastków śladowych, np. rubid 200 mg/m^3 wody, uran 2 mg/m^3 , złoto $0,004 \text{ mg/m}^3$ i wiele innych [11]. Sól morska rozumiana jako zespół określonych wyżej składników nie wpływa praktycznie na widmo optyczne absorpcji wody morskiej [11], [13], [24], wobec czego krzywa absorpcji na rys. 1 z niewielkim błędem odnosi się zarówno do wody strukturalnie czystej,



Rys. 1. Widmo absorpcyjne wody [13]

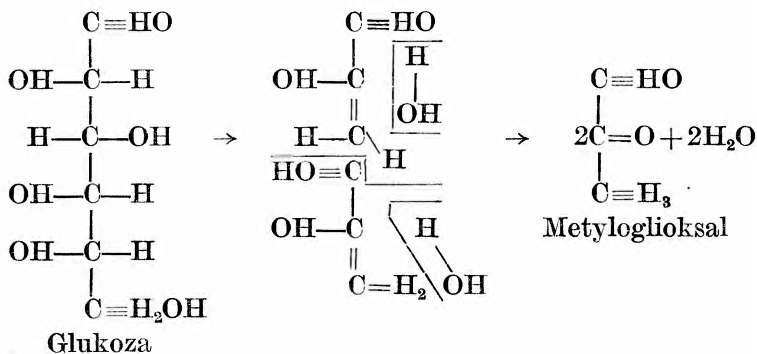


Rys. 2. Spektralne krzywe absorpcji optycznej w wodach naturalnych a) Kalle [20], b) Armstrong i Boalch [52]

jak i do optycznie czystej wody oceanicznej. Widoczne silne pochłanianie energii w pasmie długofalowym spowodowane jest, jak wiadomo, przejściami rotacyjnymi i oscylacyjnymi cząsteczek wody [72]. Wynika stąd, że na zróżnicowanie widm pochodzących z różnych obszarów wodnych (rys. 2), wpływają inne substancje rozpuszczone bądź zawieszony w wodzie.

b) Inne substancje

Oprócz rozpuszczonych gazów [13], oddzielną grupę składników wody morskiej stanowią substancje organiczne. Są to przede wszystkim produkty przemiany i rozpadu licznych organizmów morskich. Niektóre z nich mają selektywne właściwości absorpcyjne w pasmie widzialnym i nadfiolecie. Absorbując światło krótkofalowej części widma widzialnego powodują żółte zabarwienie własne, co nadaje im miano substancji żółtych (*Gelbstoff*) [20]. Substancje żółte dzielą się na dwie główne grupy: 1 — kwasy humusowe natury fenolowej (nadające wodzie zabarwienie brunatne), 2 — kwasy humusowe natury węglowodanowej lub melanoidy (nadające wodzie zabarwienie żółte). Te ostatnie są względnie trwalsze i powstają łatwo z niezwiązanych węglowodanów i aminokwasów jako produktów metabolizmu w biosferze morskiej. Na przykład w wyniku rozkładu glukozy powstaje metylogliksal, który kondensuje się tworząc wielocząsteczkową żółtą żywicę (substancja żółta):



Zróżnicowanie biosfery morskiej powoduje złożony rozkład koncentracji substancji żółtych w morzach, a stąd zróżnicowanie absorpcyjnych właściwości wód różnego pochodzenia (rys. 2). Stosunek współczynnika osłabienia światła we fioletcie i w czerwieni stanowić może wskaźnik zabarwienia wody [20].

Obok substancji rozpuszczonych w wodzie morskiej, znajduje się w niej cały szereg substancji zawieszonych w postaci drobnych zawiesin lub hydrozoli. Jeśli za kryterium podziału przyjąć pochodzenie zawiesin, to można wśród nich wyróżnić cząstki mineralne pochodzenia lądowego, cząstki organiczne i nieorganiczne, będące wynikiem rozkładu organizmów morskich i mikroorganizmy morskie zwane powszechnie planktonem.

Różnorodność zawiesin pociąga za sobą duże różnice w ich kształtach i rozmiarach (od ułamków mikrona do wielu milimetrów) [13]. Fakt ten od-

grywa istotną rolę w procesie rozpraszania energii promienistej w morzu [17], [21], [22], [25], [26], [27], [28], [29].

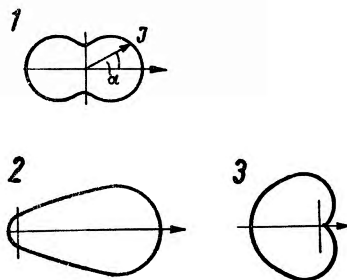
c) Elementarne zjawiska optyczne

Czynnikami warunkującymi przenoszenie energii i rozkład pola promieniowania w morzu są elementarne procesy absorpcji i rozpraszania. Procesy te spowodowane są głównie cząsteczkami wody i rozpuszczonymi w niej substancjami, jeśli chodzi o absorpcję, oraz zawiesinami i fluktuacjami gęstości, jeśli chodzi o rozpraszanie. Elementarne procesy absorpcji polegające na wzbudzeniu dyskretnych poziomów energetycznych molekuł prowadzą najczęściej do zamiany energii promienistej w inny rodzaj energii.

Przyjęty w b) podział cząstek rozpraszających w morzu daje z grubsza dwa różne przypadki rozpraszania. Pierwszy przypadek to rozpraszanie rayleighowskie na centrach o rozmiarach porównywalnych z długością fali λ , drugi przypadek, to rozpraszanie na centrach $\gg \lambda$ [14], [30], [31], [32], [33]. Ogólnie, dla dielektrycznych centrów rozpraszających (jakie najczęściej występują w morzu), strumień rozproszonej energii ΔF jest odwrotnie proporcjonalny do p -tej potęgi λ [14], [33]

$$\Delta F \sim \frac{1}{\lambda^p}, \quad (1)$$

gdzie $0 \leq p < 4$ odpowiednio dla przedziału średnic zawiesin $1,2\mu \geq d > 0,42\mu$; dla $d \geq 1,2\mu$ $p = 0$; oznacza to stałość składu widmowego strumienia padającego i rozproszonego; dla $d < 0,42\mu$ $p = 4$ [33]. Widmowa zależność strumienia rozpraszanego związana jest z mechanizmem rozpraszania. W przypadku małych średnic cząstek d w stosunku do λ , centrum rozpraszające znajduje się w jednorodnym polu wektora świetlnego. Można je zatem traktować jako pojedynczy oscylator, podczas gdy duże centra rozpraszające zachowują się jak zespół oscylatorów pojedynczych, drgających w różnych fazach. Ze względu na różny charakter rozpraszania w tych dwóch przypadkach należy spodziewać się zróżnicowania w kształtach kątowych rozkładów światła rozproszonego — funkcji (indykatrys) rozpraszania (rys. 3). Ze wzrostem rozmiarów centrów



Rys. 3. Indykatrysy rozpraszania światła [33]. 1 — wg prawa Rayleigha (małe centrum rozpraszania), 2 — dla dużych centrów przezroczystych, załamujących (wg Szulejkina), 3 — dla dużych centrów nieprzezroczystych (wg Szulejkina)

rozpraszających pojawia się coraz silniejsza asymetria indykatrysy. Fakt ten tłumaczy w/w złożony charakter rozpraszania, wynikający głównie z efektów interferencyjnych i dyfrakcyjnych ciągów falowych emitowanych przez poszczególne oscylatory dużego centrum rozpraszającego.

Rozpraszanie w ośrodkach rozciąglonych, jaki stanowi morze, ma charakter rozpraszania wielokrotnego. Sprawia to, że promieniowanie w morzu rozchodzi się w sposób przypominający dyfuzję; skąd nazwa „dyfuzja światła“ i wynikający z niej specyficzny rozkład pola promieniowania w morzu [34], [35], [36], [38], [39], [40], [41], [42], [43], [44], [45], [46].

Podobnie jak w innych ośrodkach rozpraszających [30], [31], [33], również w morzu rozpraszaniu towarzyszy częściowa polaryzacja światła [47], [48], [49], [50], [51].

3. Właściwości optyczne morza i ich klasyfikacja

Pojęcie właściwości optycznych morza obejmuje zbiór funkcji, które porządkowują danemu punktowi ośrodka liczby określające warunki przenoszenia energii promienistej.

Współczesny [6], [12], najbardziej operatywny sposób wyrażania tych funkcji oparty jest na trzech zasadniczych wielkościach fotometrycznych. Są to: funkcja radiacyjna lub krótko radiacja L (*radiance*), oświetlenie skalarne E_0 i oświetlenie E nazywane czasem oświetleniem wektorowym.

Radiacja $L(\vec{r}_p, \vec{\xi})$, (wymiar: W/m^2 sr) jest strumieniem energii promienistej padającym z kierunku $\vec{\xi}$ w jednostkowy kąt bryłowy na jednostkową powierzchnię prostopadle do niej w punkcie \vec{r}_p i chwili t .

Radiację L ze strumieniem F i natężeniem promieniowania J wiążą zależności: $L = \frac{dJ}{dA_n} = \frac{d^2F}{dA_n d\omega}$, gdzie dA_n — element powierzchni prostopadły do kierunku obserwowanego strumienia, $d\omega$ — element kąta bryłowego. Oświetlenie skalarne E_0 (wymiar: W/m^2) definiuje się wówczas jako całkę radiacji po wszystkich kierunkach:

$$E_0 = \int_G L(\vec{r}_p, \vec{\xi}) d\omega(\vec{\xi}), \quad (2)$$

gdzie G jest sferą wektorów jednostkowych $\vec{\xi}$, a $d\omega(\vec{\xi})$ — infinytezymalnym kątem bryłowym wokół kierunku $\vec{\xi}$.

Wielkość ta jest miarą energii promienistej w punkcie \vec{r}_p przychodzącej ze wszystkich kierunków wokół tego punktu, a więc miarą objętościowej gęstości energii promienistej w tym punkcie.

Oświetlenie E (wymiar: W/m^2) określa powierzchniową gęstość energii promienistej i zgodnie z definicją Preisendorfera [6] wyraża się wzorem:

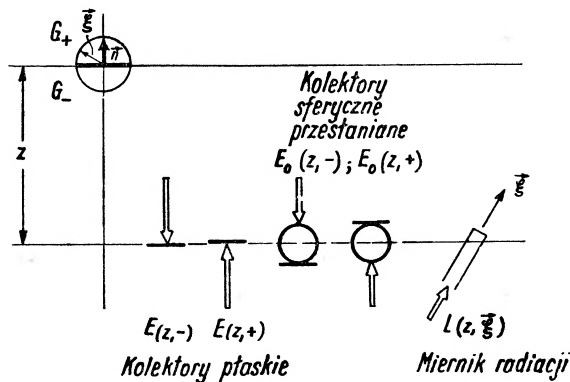
$$\vec{E} = \int_G \vec{\xi} L(\vec{r}_p, \vec{\xi}) d\omega(\vec{\xi}). \quad (3)$$

Jest to sumaryczna wartość składowych radiacji przypadająca na skierowaną w dowolnym kierunku $\vec{\xi}$ płaszczyznę jednostkową.

Ponieważ naturalne pole promieniowania w morzu ma dwa wyróżnione kierunki (zenit, nadir), w definicjach oświetleń, zarówno skalarnego, jak i wektorowego, wyróżnia się ponadto oświetlenia odgórne $E_0(z, -)$ i $E(z, -)$ oraz oświetlenia oddolne $E_0(z, +)$ i $E(z, +)$ na dowolnej głębokości z . Wielkości te dane są odpowiednio wzorami:

$$E_0(z, \pm) = \int_{G_{\pm}} L(\vec{r}_p, \vec{\xi}) d\omega(\vec{\xi}), \quad (4)$$

$$E(z, \pm) = \int_{G_{\pm}} (\vec{\xi} \cdot \vec{n}) L(\vec{r}_p, \vec{\xi}) d\omega(\vec{\xi}). \quad (5)$$



Rys. 4. Dwukierunkowa analiza pola świetlnego w morzu (Preisendorfer [6])

Ilustrację tych wielkości podaje rys. 4. W wyrażeniach (4) i (5) oraz na rysunku, \vec{n} oznacza wektor jednostkowy prostopadły do płaszczyzny poziomej, a G_- i G_+ dolną i górną półsferę wektorów jednostkowych.

Podstawą definicji operacyjnych, wskazujących na sposób bezpośredniego pomiaru właściwości optycznych morza¹, jest równanie przenoszenia energii promienistej [6]:

$$\frac{n^2}{v} \frac{D[L/n^2]}{Dt} = -cL + L_* + L_n, \quad (6)$$

gdzie

$$L_*(\vec{r}_p, \vec{\xi}, t) = \int_G L(\vec{r}_p, \vec{\xi}', t) \beta(\vec{r}_p, \vec{\xi}', \vec{\xi}, t) d\omega(\vec{\xi}'), \quad (7)$$

a pozostałe oznaczenia są następujące: n — współczynnik załamania światła, v — prędkość światła w środowisku, c — objętościowy współczynnik osłabiania, β — objętościowa funkcja rozpraszania.

¹ Bliższa interpretacja tych właściwości i metody pomiarowe podane będą w późniejszej pracy autorów [56].

Dla dowolnego kierunku przenoszenia energii promienistej, $L_*(\vec{r}_p, \vec{\xi})$ jest radiacją uzyskaną w kierunku $\vec{\xi}$ w wyniku rozpraszania strumieni promieniowania z wszystkich kierunków $\vec{\xi}'$ na jednostkowej długości rozpraszającego elementu objętości wokół punktu \vec{r}_p . Funkcję L_* , podobnie jak funkcję L_n odpowiedzialną za radiację pochodzącą ze źródeł wewnętrznych, wyraża się w jednostkach radiacji na jednostkę długości w kierunku obserwacji $\vec{\xi}$.

W stanie quasi-stacjonarnym przy założeniu bezźródłowości pola promieniowania ($L_n = 0$), równanie (6) przyjmuje postać:

$$\frac{dL}{dr} = -cL + L_*, \quad (8)$$

przy czym z założenia $n = \text{const}$. Równanie (8) daje szybkość zmian radiacji L na drodze dr w kierunku $\vec{\xi}$ w punkcie ośrodka \vec{r}_p .

Równanie przenoszenia energii promienistej z punktu widzenia eksperymentatora jest symbolem makrozjawisk optycznych w morzu. Dzięki temu, skądinąd wyłaniająca się potrzeba podziału właściwości optycznych opisujących te zjawiska, prowadzi do ich klasyfikacji o charakterze fenomenologicznym. Jak się okazuje, najbardziej naturalny jest podział właściwości optycznych morza na dwie klasy, to jest na tzw. właściwości rzeczywiste lub spójne (*inherent optical properties*) i właściwości pozorne (*apparent optical properties*) [6], [7].

Do pierwszej z tych klas zalicza się z definicji wszystkie te właściwości, których operacyjne wartości wykazują niezmienniczość wobec dowolnych zmian rozkładu pola promieniowania w morzu; do drugiej — wszystkie pozostałe właściwości. Definiowane dalej funkcje będą odnosiły się do światła monochromatycznego i niespolaryzowanego. Uogólnienia ich na przypadek światła spolaryzowanego zaniechano celem uniknięcia nieistotnych komplikacji rachunkowych.

a) Właściwości rzeczywiste

Przyjęte wyżej kryterium podziału pozwala wyróżnić cztery zasadnicze właściwości rzeczywiste: całkowity współczynnik osłabiania radiacji $c = a + b$, a — całkowity współczynnik absorpcji, $\beta(\vec{r}_p, \vec{\xi}, \vec{\xi}')$ — objętościową funkcję rozpraszania i całkowity współczynnik rozpraszania $b = \int_G \beta(\vec{r}_p, \vec{\xi}, \vec{\xi}') d\omega(\vec{\xi}')$.

W wyrażeniu na b całkowanie rozciąga się na pełny kąt bryłowy odpowiadający sferze G utworzonej przez wektory jednostkowe $\vec{\xi}'$. Przy analizie pola promieniowania w morzu wygodnie jest czasem posługiwać się tzw. współczynnikiem rozpraszania do przodu b_f i wstecz b_b (całkowanie po półsferach) w punkcie \vec{r}_p w kierunku \vec{n} :

$$b_f(\vec{r}_p, \vec{n}) = \int_{\vec{\xi} \cdot \vec{n} \geq 0} \beta(\vec{r}_p, \vec{\xi}, \vec{\xi}') d\omega(\vec{\xi}'), \quad (9)$$

$$b_b(\vec{r}_p, \vec{n}) = \int_{\vec{\xi} \cdot \vec{n} < 0} \beta(\vec{r}_p, \vec{\xi}, \vec{\xi}') d\omega(\vec{\xi}'). \quad (10)$$

Sumując (9) i (10) otrzymujemy:

$$b(\vec{r}_p) = \int_G \beta(\vec{r}_p, \vec{\xi}, \vec{\xi}') d\omega(\vec{\xi}') = b_f(\vec{r}_p) + b_o(\vec{r}_p). \quad (11)$$

Poniżej podamy definicje operacyjne wymienionych właściwości rzeczywistych. Dla współczynnika osłabienia radiacji c definicja ta wynika wprost z równania (8):

$$c = \frac{L_*}{L} - \frac{1}{L} \frac{dL}{dr}. \quad (12)$$

Jeżeli w ośrodku biegnie jedynie wąska wiązka promieniowania w kierunku r , to zgodnie z wyżej podaną definicją (7), $L_* \approx 0$ i wówczas:

$$c = -\frac{1}{L} \frac{dL}{dr} \quad (13)$$

skąd

$$c = -\frac{1}{r} \ln \left(\frac{L_r}{L_0} \right), \quad (14)$$

gdzie L_r/L_0 jest transmisją na drodze r . Równanie (14) jest dobrze znanym wyrażeniem służącym za podstawę transmisyjnej metody pomiaru współczynnika osłabiania światła.

W naturalnym polu promieniowania w morzu na niedużych obszarach w chwili t , radiacja L w poziomie jest stała względem r i wobec tego z równania (12) wynika alternatywna definicja współczynnika osłabiania, tj.:

$$c = \frac{L_*}{L} \quad (15)$$

W oparciu o tę definicję można dokonać również pomiaru współczynnika osłabiania w naturalnym polu świetlnym [74].

Operacyjna definicja objętościowej funkcji rozpraszania wynika z równania (7):

$$\beta(\vec{r}_p, \vec{\xi}_0, \vec{\xi}) = \frac{\Delta L_*(\vec{r}_p, \vec{\xi})}{L_0(\vec{r}_p, \vec{\xi}_0) \Delta \omega_0}. \quad (16)$$

W wyrażeniu tym $L_0(\vec{r}_p, \vec{\xi}_0)$ oznacza stałą wartość radiacji w małym kącie bryłowym $\Delta \omega_0$ wokół kierunku $\vec{\xi}_0$ w punkcie \vec{r}_p .

Ponadto z założenia poza $\Delta \omega_0$, $L = 0$ (wąska wiązka promieniowania). $\Delta L_*(\vec{r}_p, \vec{\xi})$ jest radiacją powstałą jako wynik rozpraszania radiacji L_0 w kierunku $\vec{\xi}$.

Korzystając z (7) i z (11) znajdujemy operacyjną definicję całkowitego współczynnika rozpraszania b :

$$b(\vec{r}_p) = \frac{\int_G L_*(\vec{r}_p, \vec{\xi}) d\omega(\vec{\xi})}{\int_G L(\vec{r}_p, \vec{\xi}) d\omega(\vec{\xi})}. \quad (17)$$

Pomiaru współczynnika $b(\vec{r}_p)$ można zatem dokonać przez sumowanie wartości radiacji powstałych w wyniku rozpraszania (L_*) w punkcie ośrodka \vec{r}_p oraz niezależne sumowanie wartości radiacji wchodzących do tego punktu (L), a następnie wzięcie stosunku otrzymanych sum.

Definicję operacyjną współczynnika absorpcji w postaci różniczkowej znajdujemy na podstawie wyrażen (8) i (17) oraz definicji oświetleń:

$$a(\vec{r}_p) = - \frac{\vec{V} \cdot \vec{E}(\vec{r}_p)}{E_0(\vec{r}_p)}. \quad (18)$$

Zastosowanie wyrażenia (18) do eksperymentu wymaga jego pewnej modyfikacji. Modyfikację tę ułatwia powszechnie przyjęty fakt, że pole świetlne w morzu można uważać za uwarstwione poziomo, tj., że w pierwszym przybliżeniu zależy ono jedynie od głębokości z . Załóżmy, że z przyjmuje wartości dodatnie od powierzchni morza w dół i że \vec{n} stanowi wektor normalny do powierzchni morza skierowany ku górze (rys. 4).

W tej sytuacji dywergencja wektora oświetlenia \vec{E} przyjmuje prostą postać:

$$\vec{V} \cdot \vec{E}(\vec{r}_p) = \frac{d}{dz} [E(z, -) - E(z, +)]. \quad (19)$$

Korzystając z wyrażenia (19) można teraz operacyjną definicję współczynnika absorpcji (18) przedstawić za pomocą wielkości bezpośrednio mierzalnych w naturalnym polu świetlnym w morzu:

$$a(z) = - \frac{1}{E_0(z)} \frac{d}{dz} [E(z, -) - E(z, +)]. \quad (20)$$

W wyrażeniu tym oświetlenie skalarne na głębokości z (p. wzór 4) $E_0(z) = E_0(z, +) + E_0(z, -)$.

b) Właściwości pozorne

Zgodnie z definicją, tzw. pozorne właściwości optyczne morza charakteryzują ośrodek w określonych warunkach oświetleniowych. Mają one szczególnie ważne znaczenie przy pomiarach prowadzonych *in situ* w naturalnym, aktualnie panującym polu promieniowania w morzu. Pełny zbiór pozornych właściwości optycznych morza wyraża się liczbą 20. Podamy tutaj kilka z nich, które dadzą się opisać przez wprowadzony wyżej kwartet oświetleniowy $E(z, +)$, $E(z, -)$, $E_0(z, +)$, $E_0(z, -)$ i wystarczają do opisu wielu zjawisk optycznych w morzu o znaczeniu praktycznym.

W wyniku procesów rozpraszania promieniowania w morzu (p. § 2) powstaje między innymi oddolny strumień energii promienistej, uwarunkowany rzeczywistymi właściwościami optycznymi danego obszaru ośrodka i oświetleniem zewnętrznym. Zatem na dowolnej głębokości z istnieje określony stosunek oświetlenia oddolnego do górnego i odwrotnie:

$$\frac{E(z, \pm)}{E(z, \mp)} = R(z, \mp). \quad (21)$$

Pomiaru współczynnika $b(\vec{r}_p)$ można zatem dokonać przez sumowanie wartości radiacji powstałych w wyniku rozpraszania (L_*) w punkcie ośrodka \vec{r}_p oraz niezależne sumowanie wartości radiacji wchodzących do tego punktu (L), a następnie wzięcie stosunku otrzymanych sum.

Definicję operacyjną współczynnika absorpcji w postaci różniczkowej znajdujemy na podstawie wyrażen (8) i (17) oraz definicji oświetleń:

$$a(\vec{r}_p) = - \frac{\vec{\nabla} \cdot \vec{E}(\vec{r}_p)}{E_0(\vec{r}_p)}. \quad (18)$$

Zastosowanie wyrażenia (18) do eksperymentu wymaga jego pewnej modyfikacji. Modyfikację tę ułatwia powszechnie przyjęty fakt, że pole świetlne w morzu można uważać za uwarstwione poziomo, tj., że w pierwszym przybliżeniu zależy ono jedynie od głębokości z . Załóżmy, że z przyjmuje wartości dodatnie od powierzchni morza w dół i że \vec{n} stanowi wektor normalny do powierzchni morza skierowany ku górze (rys. 4).

W tej sytuacji dywergencja wektora oświetlenia \vec{E} przyjmuje prostą postać:

$$\vec{\nabla} \cdot \vec{E}(\vec{r}_p) = \frac{d}{dz} [E(z, -) - E(z, +)]. \quad (19)$$

Korzystając z wyrażenia (19) można teraz operacyjną definicję współczynnika absorpcji (18) przedstawić za pomocą wielkości bezpośrednio mierzalnych w naturalnym polu świetlnym w morzu:

$$a(z) = - \frac{1}{E_0(z)} \frac{d}{dz} [E(z, -) - E(z, +)]. \quad (20)$$

W wyrażeniu tym oświetlenie skalarne na głębokości z (p. wzór 4) $E_0(z) = E_0(z, +) + E_0(z, -)$.

b) Właściwości pozorne

Zgodnie z definicją, tzw. pozorne właściwości optyczne morza charakteryzują ośrodek w określonych warunkach oświetleniowych. Mają one szczególnie ważne znaczenie przy pomiarach prowadzonych *in situ* w naturalnym, aktualnie panującym polu promieniowania w morzu. Pełny zbiór pozornych właściwości optycznych morza wyraża się liczbą 20. Podamy tutaj kilka z nich, które dadzą się opisać przez wprowadzony wyżej kwartet oświetleniowy $E(z, +)$, $E(z, -)$, $E_0(z, +)$, $E_0(z, -)$ i wystarczają do opisu wielu zjawisk optycznych w morzu o znaczeniu praktycznym.

W wyniku procesów rozpraszania promieniowania w morzu (p. § 2) powstaje między innymi oddolny strumień energii promienistej, uwarunkowany rzeczywistymi właściwościami optycznymi danego obszaru ośrodka i oświetleniem zewnętrznym. Zatem na dowolnej głębokości z istnieje określony stosunek oświetlenia oddolnego do górnego i odwrotnie:

$$\frac{E(z, \pm)}{E(z, \mp)} = R(z, \mp). \quad (21)$$

$R(z, \mp)$ stanowi jedną z właściwości pozornych i nosi nazwę odbicia na głębokości z . Oczywiście $R(z, +) = \frac{1}{R(z, -)}$. Dwa wyrażenia na $R(z)$ odnoszą się do odbicia ze względu na kierunki odgórny i oddolny.

Jako odpowiednik rzeczywistego współczynnika osłabiania radiacji c , w dyfuzyjnym polu promieniowania *in situ* wprowadza się pojęcie funkcji dyfuzyjnego osłabiania radiacji na głębokości z :

$$K(z, \vec{\xi}) = \frac{-1}{L(z, \vec{\xi})} \frac{dL(z, \vec{\xi})}{dz}, \quad (22)$$

pojęcie współczynnika dyfuzyjnego osłabiania oświetlenia na głębokości z ze względu na kierunek oddolny i odgórny

$$K(z, \pm) = \frac{-1}{E(z, \pm)} \frac{dE(z, \pm)}{dz} \quad (23)$$

i współczynnika dyfuzyjnego osłabiania oświetlenia skalarnego na głębokości z również ze względu na dwa kierunki:

$$k(z, \pm) = \frac{-1}{E_0(z, \pm)} \frac{dE_0(z, \pm)}{dz}. \quad (24)$$

Ponadto analogicznym wzorem wyraża się całkowity współczynnik dyfuzyjnego osłabiania (całkowitego) oświetlenia skalarnego na głębokości z :

$$k(z) = \frac{-1}{E_0(z)} \frac{dE_0(z)}{dz}. \quad (25)$$

Różnice interpretacyjne $K(z, \pm)$ i $k(z, \pm)$ wynikają z różnic w interpretacji $E(z)$ i $E_0(z)$ (p. str. 542). Kierunkowy charakter oświetlenia $E(z)$ pozwala wprowadzić następną właściwość pozorną, nazywaną funkcją rozkładu strumienia na głębokości z , zdefiniowaną dla oddolnego i odgórnego strumienia energii promienistej:

$$D(z, \pm) = \frac{E_0(z, \pm)}{E(z, \pm)}. \quad (26)$$

Wartości $D(z, \pm)$ określają kierunkową strukturę rozkładu promieniowania na głębokości z . Przypuśćmy przykładowo, że L jest radiacją o kierunku $\vec{\xi}$ pochodzącą z obszaru α_- , wtedy na mocy definicji (26) oraz (4) i (5) $D(z, -) = \frac{1}{|\vec{\xi} \cdot \vec{n}|}$. Jeżeli w obszarze α_- , $L(\vec{\xi}) = \text{const.}$, to $D(z, -) = 2$.

Warto tu jeszcze raz podkreślić, że chociaż pozorne właściwości optyczne morza silnie zależą od zewnętrznych warunków oświetleniowych, to jednak wykazują one pewne prawidłowości, zarówno jako funkcje zmiennych przestrzennych, jak i czasu. Zależność kilku z nich od głębokości z przedstawia Tab. II.

Tablica II

Z [m]	$D(z, -)$	$K(z, -)$	$K(z, +)$	$a^2(z)$	$E(z, -)$
10,42	1,288	0,153	0,150	0,115	0,0184
16,58	1,291	0,174	0,172	0,118	0,0204
28,96	1,313	0,169	0,169	0,117	0,0227
41,30	1,315	0,165	0,165	0,117	0,0235
53,71	1,307	0,158	0,158	0,112	0,0234

Wartości podane w tej tablicy zaczerpnięto z pracy [6]. Były one obliczone na podstawie pomiarów radiacji i oświetleń w obszarze długości fali 480 m μ . Dla porównania przytoczone są wartości współczynnika absorpcji.

Szerszą interpretację zdefiniowanych tu funkcji znaleźć można w pracach [6], [7], [57], [58].

c) Właściwości mieszane (*Hybrid Optical Properties*)

Mieszane właściwości optyczne morza stanowią podgrupę właściwości pozornych i ich wyróżnienie wynika jedynie z trudności w znalezieniu prostych definicji operacyjnych. Mogą one być natomiast łatwo zdefiniowane przez pewne analityczne związki pomiędzy wcześniej podanymi rzeczywistymi i pozornymi właściwościami.

Objętościowy współczynnik osłabiania, ze względu na dwa wyróżnione kierunki strumienia energii promienistej w morzu, zdefiniowany jako iloczyn rzeczywistego współczynnika osłabiania przez odpowiednią funkcję rozkładu $D(z, \pm)$

$$c(z, \pm) = c(z)D(z, \pm) \quad (27)$$

oddaje kierunkowy rozkład współczynnika osłabiania wynikający z kierunkowego zróżnicowania E_0/E .

W analogiczny sposób wprowadza się pojęcia pozostałych właściwości mieszanych. Całkowite współczynniki absorpcji:

$$a(z, \pm) = a(z)D(z, \pm) \quad (28)$$

i całkowite współczynniki rozpraszania:

$$b(z, \pm) = b(z)D(z, \pm). \quad (29)$$

Ponadto, jako odpowiedniki rzeczywistych współczynników rozpraszania do przodu i wstecz, z wyróżnieniem kierunków oddolnego i odgórnego, definiuje się mieszane współczynniki rozpraszania do przodu

$$b_f(z, \pm) = \frac{1}{E(z, \pm)} \int_{\sigma_{\pm}} \left[\int_{\sigma_{\pm}} L(z, \vec{\xi}') \beta(z, \vec{\xi}', \vec{\xi}) d\omega(\vec{\xi}') \right] d\omega(\vec{\xi}) \quad (30)$$

² Wartości $a(z)$ obliczono na podstawie wzoru (20).

i mieszane współczynniki rozpraszania wstecz

$$b_b(z, \pm) = \frac{1}{E(z, \pm)} \int_{\Omega^\mp} \left[\int_{\Omega^\pm} L(z, \vec{\xi}') \beta(z, \vec{\xi}', \vec{\xi}) d\omega(\vec{\xi}') \right] d\omega(\vec{\xi}). \quad (31)$$

Oczywiście podobnie jak w (11)

$$b(z, \pm) = b_f(z, \pm) + b_b(z, \pm) \quad (32)$$

oraz podobnie jak przy rozumieniu $c(z)$

$$c(z, \pm) = a(z, \pm) + b(z, \pm). \quad (33)$$

Niezależnie od podanego tutaj schematu klasyfikacji optycznych właściwości morza zilustrowanego ich przykładami, zbiór definicji standardowych pojęć i jednostek stosowanych w optyce morza znaleźć można w raportach UNESCO [53] i IAPO³ [73].

4. Elementy ogólnej teorii właściwości optycznych morza

Teoria właściwości optycznych morza rozumianych jak w § 3 wywodzi się z ogólnej teorii przenoszenia energii promienistej Chandrasekhara [54]. Teorię tę dla środowiska morskiego opracował Preisendorfer [6], przyjmując za punkt wyjścia równanie przenoszenia energii promienistej (6). Z równania (6) zapisanego w postaci:

$$\vec{\xi} \cdot \vec{\nabla} L(z, \vec{\xi}) = -c(z)L(z, \vec{\xi}) + L_*(z, \vec{\xi}) \quad (34)$$

otrzymał on bezczasowe równania zmienności oświetleń $E(z, \pm)$:

$$\frac{dE(z, -)}{dz} = -[a(z, -) + b_b(z, -)]E(z, -) + b_b(z, +)E(z, +), \quad (35a)$$

$$-\frac{dE(z, +)}{dz} = -[a(z, +) + b_b(z, +)]E(z, +) + b_b(z, -)E(z, -). \quad (35b)$$

W sposób ogólny pokazują one jak zmienia się oświetlenie w morzu ze zmianą głębokości z , gdy obserwuje się je od dołu i od góry w środowisku o określonych właściwościach optycznych [57], [58].

Z równań (35a, b), które są bezpośrednią konsekwencją równania przenoszenia energii promienistej, wyprowadza się cały szereg związków pomiędzy poszczególnymi właściwościami optycznymi morza. Związki te mogą być wykorzystane podczas wykonywania pomiarów optycznych, interpretacji wyników i wreszcie przy konstrukcji aparatury pomiarowej.

Podzielenie równania (35a) przez $E(z, -)$ i wykorzystanie operacyjnych definicji $K(z, -)$ i $R(z, -)$ pozwala otrzymać równanie

$$K(z, -) = a(z, -) + b_b(z, -) - b_b(z, +)R(z, -) \quad (36)$$

³ IAPO — International Association of Physical Oceanography.

i podobnie w oparciu o (35b) — równanie:

$$-K(z, +) = a(z, +) + b_b(z, +) - b_b(z, -)R(z, +). \quad (37)$$

Oba te równania wiążą bezpośrednio wyznaczane na drodze doświadczalnej współczynniki dyfuzyjnego osłabiania oświetlenia $K(z, \pm)$ z pozornymi, a pośrednio z rzeczywistymi, współczynnikami absorpcji i rozpraszania dla danego ośrodka, tkwiącymi w $a(z, \pm)$ i $b(z, \pm)$.

Na podstawie następnej zależności:

$$k(z, \pm) = K(z, \pm) - \frac{1}{D(z, \pm)} \frac{dD(z, \pm)}{dz} \quad (38)$$

zauważmy, że k może być identyczne z K . Nastąpi to wtedy, gdy w okolicy pewnego poziomu z $\frac{dD(z, \pm)}{dz} = 0$, tzn. gdy kierunkowa struktura funkcji rozkładu w pionowym sąsiedztwie tego poziomu nie ulega zmianie.

Funkcje $K(z, \pm)$, $k(z, \pm)$ i $k(z)$, które zgodnie z definicjami można wyznaczyć przez pomiar odpowiednich oświetleń, pozostają w ścisłym związku z funkcją dyfuzyjnego osłabiania radiacji $K(z, \vec{\xi})$. Ponieważ poszczególne wartości tej ostatniej wyróżniają ścisłe chociaż dowolne kierunki w ośrodku, to funkcje $K(z, \pm)$ i $k(z, \pm)$ wyróżniające jedynie generalny kierunek oddolny lub odgórny dadzą się wyrazić średniowaniem radiacji po $K(z, \vec{\xi})$, np.:

$$K(z, \pm) = \frac{\int_{\sigma_{\pm}} L(z, \vec{\xi}) K(z, \vec{\xi}) \vec{\xi} \cdot \vec{n} d\omega(\vec{\xi})}{\int_{\sigma_{\pm}} L(z, \vec{\xi}) \vec{\xi} \cdot \vec{n} d\omega(\vec{\xi})}. \quad (39)$$

Jeżeli na podstawie definicji mierzonych właściwości optycznych oraz równania (23) przepiszemy (35a) w innej postaci:

$$K(z, -) = c(z, -) - \frac{1}{E(z, -)} \int_{\sigma_-} L_*(z, \vec{\xi}) d\omega(\vec{\xi}), \quad (40)$$

to możemy wnioskować jak wpływa na wartość $K(z, -)$ zmiana zewnętrznych warunków oświetlenia i jak $K(z, -)$ zależy od funkcji rozkładu $D(z, -)$. Zauważmy w tym celu, że wyrażenie:

$$\int_{\sigma_-} L_*(z, \vec{\xi}) d\omega(\vec{\xi}) = E_{0*}(z, -) \quad (41)$$

jest odgórnym oświetleniem skalarnym wywołanym rozpraszaniem strumienia energii promienistej w jednostce objętości ośrodka na głębokości z . Wtedy po wykorzystaniu (26) i (27) wyrażenie (40) przyjmie postać:

$$K(z, -) = \left[c(z) - \frac{E_{0*}(z, -)}{E_0(z, -)} \right] D(z, -). \quad (42)$$

Wynika stąd, że współczynnik dyfuzyjnego osłabiania $K(z, -)$ rośnie ze wzrostem, a maleje z maleniem $D(z, -)$, lecz w ogólności nieliniowo. Zmiana

$D(z, -)$ oznacza bowiem również zmianę drugiego składnika w nawiasie wyrażenia (42) (porównaj wzór 26).

Pokażemy teraz za Preisendorferem, że oświetlenie odgórne w morzu zmienia się wolniej z głębokością niż oświetlenie oddolne. W tym celu, po odpowiedniej kombinacji równań (36) i (37) można zauważyć, że ze względu na prawą stronę otrzymanego po tej operacji równania $K(z, -) - K(z, +)R(z, -) = a(z, -) + a(z, +)R(z, -)$, lewa strona równania jest

$$K(z, -) - K(z, +)R(z, -) \geq 0. \quad (43)$$

Z nierówności (43) na mocy definicji $K(z, \pm)$ i $R(z, -)$ znajdujemy wykazywaną nierówność:

$$\frac{dE(z, -)}{dz} \leq \frac{dE(z, +)}{dz}. \quad (44)$$

W nomenklaturze pozornych i mieszanych właściwości optycznych morza można również dowieść, że dla strumienia odgórnego

$$a(z, -) \leq K(z, -) \leq c(z, -) \quad (45a)$$

i strumienia oddolnego

$$a(z, +) \leq -K(z, +) \leq c(z, +). \quad (45b)$$

Wreszcie na zakończenie tych krótkich rozważań sformułujemy twierdzenie o tzw. polu granicznym w morzu. Chodzi tu o charakter zmian pola promieniowania na dużych głębokościach, gdzie skierowany strumień bezpośredniego promieniowania słonecznego praktycznie zanika i można mówić właściwie o czystym polu promieniowania dyfuzyjnego.

W takich warunkach, jak wykazał to doświadczalnie po raz pierwszy Whitney [68], a teoretycznie dowiódł Preisendorfer [46], radiacja z głębokością zmienia się eksponencjalnie, podobnie jak to ma miejsce dla skolimowanej wąskiej wiązki promieniowania w ośrodku jednorodnym.

Granicznym polem promieniowania w morzu nazywa się zatem pole na dużych głębokościach ($z \rightarrow \infty$) — pole, którego właściwości nie zależą od zewnętrznych (na powierzchni) warunków oświetlenia. Wtedy

$$L(z, \vec{\xi}) = L(z_0, \vec{\xi}) e^{-k_\infty(z-z_0)}, \quad (46)$$

gdzie z_0 jest głębokością poniżej której $K(z, \vec{\xi})$ przyjmuje stałą wartość k_∞ , zależną jedynie od rzeczywistych właściwości optycznych morza. Stąd również charakter funkcji (46) określony jest właściwościami spójnymi, a więc np. funkcjami e i β . Ponieważ warunkiem koniecznym i dostatecznym istnienia pola granicznego jest stałość $K(z, \vec{\xi})$, twierdzenie o polu granicznym można wyrazić następująco:

jeżeli istnieje $\lim_{z \rightarrow \infty} \beta(z, \vec{\xi}, \vec{\xi}') = \beta(\vec{\xi}, \vec{\xi}')$ i $\lim_{z \rightarrow \infty} c(z) = c$, to dla każdego $\vec{\xi}$ w obszarze G istnieje

$$\lim_{z \rightarrow \infty} K(z, \vec{\xi}) = \lim_{z \rightarrow \infty} k(z) = k_\infty. \quad (47)$$

Wyrażenia (46) i (47) są równoważnymi wersjami twierdzenia o polu granicznym w morzu. Z twierdzenia tego wynika szereg konsekwencji dla różnych innych właściwości optycznych pozornych i mieszanych, z których przykładowo podamy dwie. W polu granicznym współczynniki dyfuzyjnego osłabiania dla oświeleń wektorowych i skalarnych są sobie równe

$$\lim_{z \rightarrow \infty} K(z, \pm) = \lim_{z \rightarrow \infty} k(z, \pm) = k_{\infty}. \quad (48)$$

Funckje rozkładu strumienia $D(z, \pm)$ stają się niezależne od z

$$\lim_{z \rightarrow \infty} D(z, \pm) \equiv D(\pm). \quad (49)$$

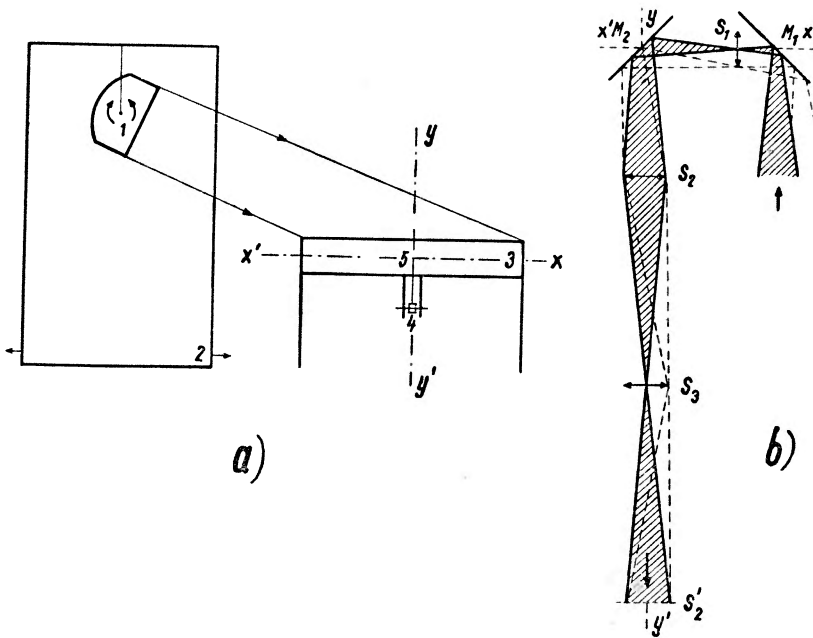
Podane w tym paragrafie fragmenty teoretycznych podstaw przenoszenia energii promienistej w morzu potwierdzają przydatność wprowadzonej w § 3 klasyfikacji właściwości optycznych i dowodzą istnienia pewnych prawidłowości w ich przestrzennym zróżnicowaniu.

Szczegółową matematyczną analizę podanych wyżej rozważań, jak i szersze ujęcie teorii, znaleźć można w pracach Preisendorfera [55], [57], [58], [59], [61], [62], [6] i innych pracach zawartych w materiałach *Scripps Institution of Oceanography — La Jolla, California*.

5. Prace doświadczalne

a) Pomiary na modelach

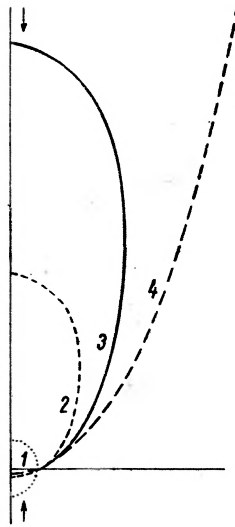
Autorzy prac przeprowadzanych na modelach [63], [64], wychodzili z założenia, że laboratoryjne modele pozwalają uniknąć wysokich kosztów pomiarów w warunkach naturalnych i oddalić niepożądane zmiany naturalnego oświetlenia zewnętrznego. Celem pierwszej z cytowanych prac było wykonanie pomiarów rozkładu promieniowania w ośrodku imitującym nieskończenie grubą optycznie, nieskończenie rozciąglą, jednorodną i równomiernie oświetloną warstwę morza. Ośrodek stanowiła dobrze zdefiniowana zawiesina (cząstki w przybliżeniu kuliste o rozmiarach 0,2 do 75 μ) octanu poliwinylu w wodzie, umieszczona w prostopadłościennym zbiorniku o kwadratowej podstawie (100 \times 100 cm^2) i wysokości 20 cm. Schemat urządzenia pomiarowego pokazany jest na rys. 5. Wzrost absorpcji w ośrodku uzyskiwano przez rozpuszczenie czarnej farby. Obrót zwierciadła M_1 wokół osi o kierunku XX' i obrót tej osi wraz z M_1 i M_2 wokół osi o kierunku YY' umożliwiał detekcję oświetlenia w różnych kierunkach poziomych (kąt azymutalny φ) i pionowych (kąt θ). Do opisu ośrodka przyjęła Lenoble [63] współczynnik osłabienia c , parametr $\omega_0 = \frac{b}{c}$ i funckję rozpraszania. Jeżeli głębokość optyczna $\tau = (c \cdot z) \rightarrow \infty$, wtedy zgodnie z (47) radiacja przyjmuje postać eksponencjalną, a kątowy rozkład natężenia promieniowania przy danym z jest symetryczny względem poziomu (oczywista jest symetria względem pionu). Należy zwrócić uwagę, że symetria



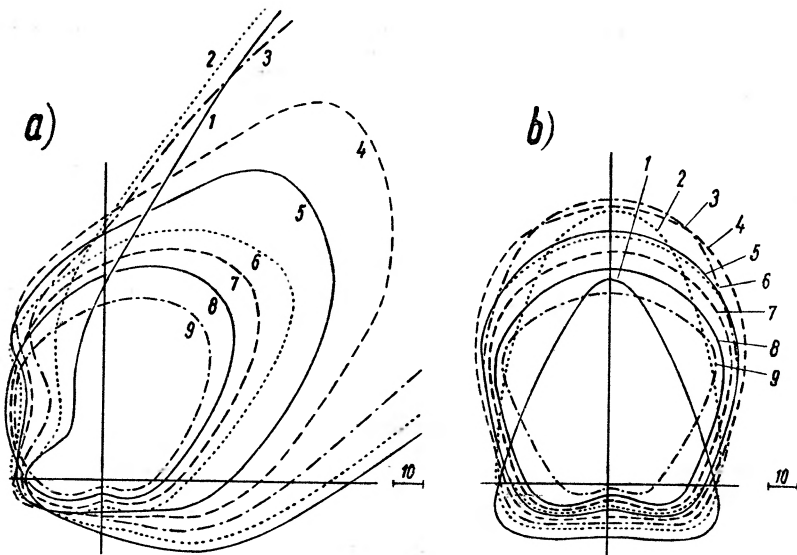
Rys. 5. a) Schemat urządzenia do pomiaru rozkładu oświetlenia w zbiorniku doświadczalnym. 1 — źródło światła (projektor), 2 — konstrukcja umożliwiającą zmianę położenia projektora, 3 — zbiornik, 4 — fotopowielacz (RCA typ 1P28) z niebiesko-zielonym filtrem, 5 — układ optyczny. b) Zasada działania układu optycznego a) (Lenoble [63])

względem poziomu wystąpi jedynie przy $\bar{\omega}_0 = 1$, tzn. gdy osłabienie radiacji w ośrodku następuje jedynie na skutek rozpraszania. Z chwilą pojawienia się i wzrostu absorpcji ($\bar{\omega}_0$ maleje) symetria zanika ustępując miejsca rosnącej asymetrii rozkładu. Teoretyczny rozkład $J(\theta)$ dla różnych wartości $\bar{\omega}_0$ pokazany jest na rys. 6. Doświadczalne krzywe rozkładu dla różnych τ przy $\bar{\omega}_0 = 1$ w układzie biegunowym przedstawia rys. 7. Rozkłady te posiadają zdecydowaną asymetrię, co widać szczególnie dobrze na przykładzie a); silne maksimum w kierunku oświetlenia zewnętrznego i minimum w kierunku do niego prostopadłym. Maksima są silniejsze na mniejszych głębokościach. Tłumaczy się to większym wkładem do natężenia wiązki światła skierowanego w stosunku do dyfuzyjnego. W innej wersji rozkład pola świetlnego w medium podany jest na rys. 8. Widać, że począwszy od pewnego τ krzywe dla różnych kątów przechodzą w odcinki prostych równoległych. Wyjaśnienia tych zależności należy szukać w warunkach stanu stacjonarnego przewidywanego przez teorię (§ 4). Osłabianie jest tu jednak znacznie mniejsze niż przewiduje teoria.

Stosowanie sztucznie wytworzonego medium imitującego środowisko morskie może wyeliminować błędy pomiarów związane z niezdefiniowanym składem centrów rozpraszających w warunkach naturalnych. Jest to szczególnie ważne gdy chodzi o określenie przydatności metody pomiarowej i ekstrapolacji wyników. Zagadnienie to podjął Tyler [64] w oparciu o skonstruowany przez

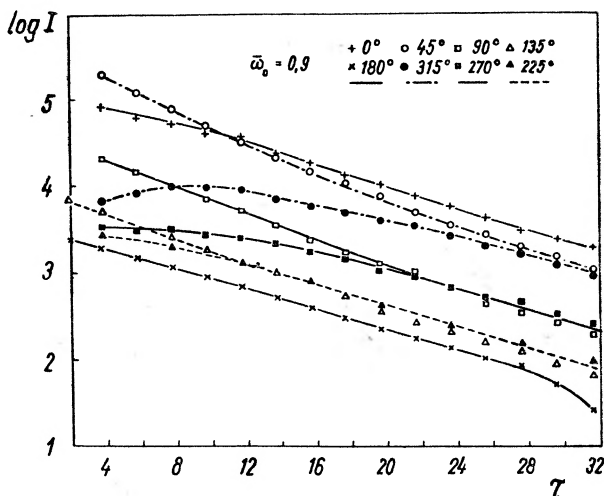


Rys. 6. Teoretyczny biegunowy rozkład $J(\theta)$ w stanie stacjonarnym dla różnych $\bar{\omega}_0$. 1 — $\bar{\omega}_0 = 1$; 2 — $\bar{\omega}_0 = 0,66$; 3 — $\bar{\omega}_0 = 0,5$; 4 — $\bar{\omega}_0 = 0,33$ (Lenoble [63])

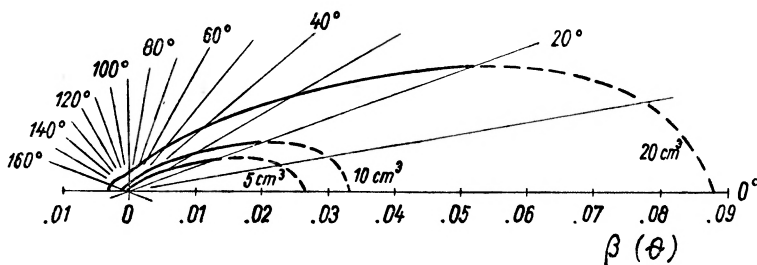


Rys. 7. Biegunowe rozkłady $J(\theta)$ dla różnych τ i $\bar{\omega}_0 = 1$. a) w płaszczyźnie rysunku 5a, b) w płaszczyźnie prostopadłej do rysunku 5a. Krzywe: 1 — $\tau = 3,18$; 2 — $\tau = 5,7$; 3 — $\tau = 8,2$; 4 — $\tau = 10,7$; 5 — $\tau = 13,2$; 6 — $\tau = 16,1$; 7 — $\tau = 18,2$; 8 — $\tau = 20,7$; 9 — $\tau = 24,0$ (Lenoble [63])

siebie przyrząd do pomiaru funkcji rozpraszania β . Biegunowe wykresy funkcji $\beta(\theta)$ przedstawione są na rys. 9. Ze względu na konstrukcję urządzenia wartości β dla $\theta < 20^\circ$ nie mogły być zmierzone i zostały teoretycznie ekstrapolowane. Trzy krzywe wykreślone są dla trzech ośrodków różniących się kon-



Rys. 8. Zmiana natężenia światła w funkcji głębokości optycznej dla różnych kierunków obserwacji ($\bar{\omega}_0 = 0,9$) (Lenoble [63])

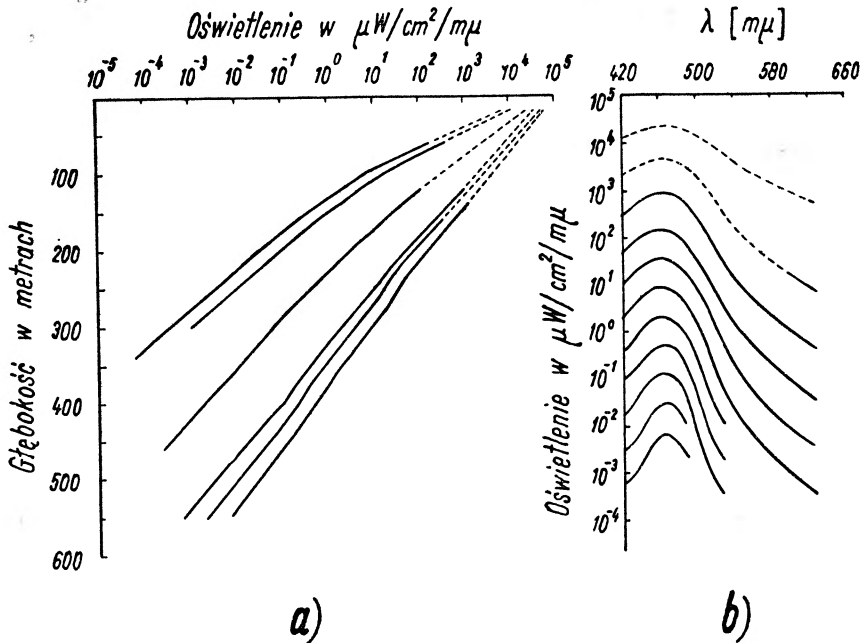


Rys. 9. Przebieg funkcji rozpraszania $\beta(\theta)$ w układzie biegunowym dla trzech próbek o różnych stężeniach hydrozoli (Tyler [64])

centracją określonych hydrozoli. Obliczony w oparciu o te wykresy całkowity współczynnik rozpraszania mógł być porównany z niezależnie zmierzoną przez tego autora wartością doświadczalną. Różnica wartości współczynnika b wyznaczonego tymi dwoma różnymi metodami nie przekroczyła 10%. Świadczy to na rzecz dopuszczalności ekstrapolacji kształtu funkcji rozpraszania β , przynajmniej dla stosowanych przez Tylera hydrozoli.

b) Pomiar w wodach naturalnych

Z chwilą gdy światło słoneczne przeniknie powierzchnię morza [66], [67], poddane zostaje omawianym w § 2 procesom absorpcji i rozpraszania charakterystycznym dla wody morskiej *in situ*. W wyniku tego zostaje ono osłabione, przy czym wraz ze wzrostem głębokości ulega zmianie jego względny skład widmowy. Oba zjawiska przedstawione są na rys. 10. Obie części (*a* i *b*) tego rysunku pokazują, że maksimum przenikania osiąga światło o długości fali ok. 461 m μ .



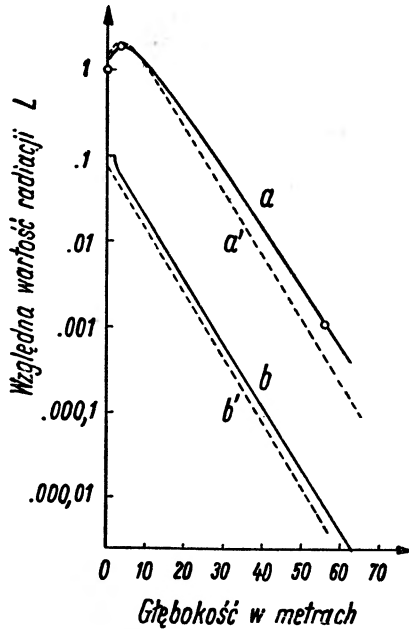
Rys. 10. Osłabienie a) i rozkład widmowy b) światła naturalnego wnikaącego w głąb morza (pomiaru wykonane w okolicy Bermudów (Kampa [38] krzywe na rys. a) od góry — dla: 630 $\text{m}\mu$, 599 $\text{m}\mu$, 530 $\text{m}\mu$, 418 $\text{m}\mu$, 491 $\text{m}\mu$ i 461 $\text{m}\mu$).

Klasycznym przykładem eksperymentu weryfikującego stosunkowo młodą teorię Preisendorfera (§ 4) jest praca Tylera [36], dotycząca osłabienia radiacji w morzu. Za pomocą przyrządu umożliwiającego pomiar wartości radiacji L badał on jej zmianę ze wzrostem głębokości w morzu. Typowe krzywe doświadczalne pokazuje rys. 11. Krzywe a , b różnią się między sobą szczególnie dla małych głębokości. Znikanie w przypadku a pierwszej pochodnej radiacji w okolicy $z = 3,7$ m tłumaczy się zjawiskiem wzrostu wraz z głębokością dyfuzji światła i stosunkowo jeszcze małą absorpcją bezpośredniego strumienia światła słonecznego. Dalej przebieg krzywych zbliżony jest do siebie, a przesunięcie pionowe świadczy o ogólnie niższej wartości radiacji oddolnej. Średnie nachylenie każdej z krzywych daje średni współczynnik osłabiania dla radiacji. Współczynnik osłabiania można obliczyć również z nachylenia krzywych teoretycznych radiacji jako funkcji głębokości (rys. 11, krzywe a' , b').

Z równania przenoszenia energii (6) Preisendorfer [57] wyprowadził ogólne wyrażenie na radiację w kierunku θ , φ na głębokości z :

$$L_r(z, \theta, \varphi) = L_0(z_t, \theta, \varphi) e^{-cr} + \frac{L_*(z, \theta, \varphi)}{c + K \cos \theta} [1 - e^{-(c - K \cos \theta)z}]. \quad (50)$$

W równaniu tym pierwszy składnik prawej strony jest radiacją na głębokości z osłabioną na drodze r w kierunku θ , φ , pochodzącą od bezpośredniego stru-



Rys. 11. Osłabianie radiacji w morzu w warunkach dnia słonecznego. a , b — krzywe doświadczalne (a — dla radiacji odgórnej, b — dla radiacji oddolnej); a' , b' — krzywe teoretyczne (a' — dla radiacji odgórnej, b' — dla radiacji oddolnej (Tyler [36]))

mienia światła słonecznego. Drugi składnik reprezentuje radiację światła dyfuzyjnego, które w wyniku wielokrotnego rozpraszania znalazło się na drodze r o kierunku θ , φ . Poszczególne symbole są następujące: $L_0(z_i, \theta, \varphi)$ — radiacja źródła na głębokości z_i obserwowana w kierunku θ , φ , K — funkcja dyfuzyjnego osłabiania światła, θ — kąt między pionem i kierunkiem drogi, φ — kąt azymutalny. Wyrażenie (50) znacznie się upraszcza w przypadku obserwacji pionowych, tj. odgórnej i oddolnej. Wtedy dla światła odgórnego $\theta = \varphi = 0$, $r = z$ i wobec tego

$$L(z) = L(0)e^{-cz} + \frac{L_*(z)}{c-K} [1 - e^{-(c-K)z}], \quad (51)$$

dla światła oddolnego $\theta = 180^\circ$, $\varphi = 0^\circ$, $r = z$ i

$$L(z) = \frac{L_*(z)}{c+K}. \quad (52)$$

Równania (51) i (52) są teoretycznymi wyrażeniami na osłabianie radiacji odgórnej i oddolnej. Celem przedstawienia graficznego tych wyrażeń jako funkcji z należy znać wartości c , K , $L(0)$ oraz L_* . Na ogół dość duże trudności nastęrcza znalezienie wartości funkcji L_* . Celem jej oszacowania dla warstwy powierzchniowej Tyler posłużył się równaniem (6) i skorzystał z zerowania się pochodnej $\frac{dL(z)}{dz}$ na krzywej a (rys. 11). Do określenia wartości L_* dla więk-

szych z wykorzystał on twierdzenie (47) o polu granicznym. Obliczając dalej na podstawie (51) wartości L_* dla danego z z pomiaru $L(z)$, mógł znaleźć $L_*(O)$. Teraz przy stałym i znanym $L_*(O)$ wyrażenie na $L_*(z)$ można podstawić do (51)

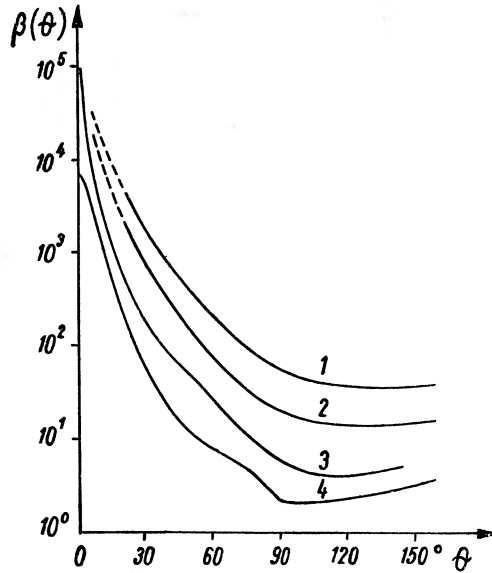
$$L(z) = L(O)e^{-cz} + \frac{L_*(O)e^{-Kz}}{c-K} [1 - e^{-(c-K)z}]. \quad (53)$$

Stąd dla każdego z można znaleźć wartość $L(z)$ radiacji odgórnej i w sposób podobny wartość radiacji oddolnej. Wykreślone na podstawie tych danych wykresy funkcji (51) i (52) przedstawione są na rys. 11 (krzywe a' , b'). Rozbieżność w stosunku do krzywych doświadczalnych wynika ze zbyt grubego przybliżenia asymptotycznej wartości współczynnika rozpraszania dyfuzyjnego k_∞ wartością średnią K . Jeżeli zamiast K średniego użyć asymptotycznej (granicznej) wartości k_∞ , to jak pokazano na kilku przykładach (kółeczka na krzywej a) zgodność wartości teoretycznych z doświadczalnymi jest bardzo dobra. Podobne pomiary wykonane w dzień pochmurny [36], [68] (możliwość zaniedbania skierowanej wiązki bezpośredniego promieniowania słonecznego), w granicach błędów pomiarów współczynnika c , dały również dobrą zgodność z przewidywaniami teoretycznymi.

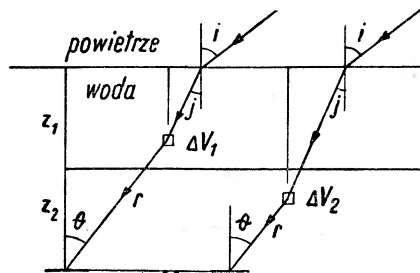
O rozkładzie pola radiacyjnego w morzu decyduje kształtąt katowej funkcji rozpraszania $\beta(\theta)$ (p. rys. 3). Wyznaczenie funkcji rozpraszania możliwe jest zarówno na bezpośredniej drodze doświadczalnej [70], [28], jak i pośrednio w oparciu o znany rozkład radiacji [9]. Matematyczną postać $\beta(\theta)$ uzyskuje się przez aproksymację szeregiem wielomianów Legendre'a, przy czym współczynniki rozwinięcia oblicza się z odpowiednich wzorów rekurencyjnych. W pierwszym przypadku na podstawie wartości $\beta(\theta)$ z krzywych doświadczalnych, w drugim na podstawie doświadczalnego rozkładu radiacji i współczynnika osłabiania.

Szczegółową dyskusję tych metod przeprowadził Schellenberger [9]. Między innymi dokonał on próby oceny poprawności ekstrapolacji doświadczalnych funkcji rozpraszania na małe kąty rozpraszania. Do analizy posłużyły wyniki, nielicznych jak dotychczas, prac doświadczalnych pokazanych na rys. 12. Dokładna analiza i wyprowadzenie ścisłej postaci funkcji $\beta(\theta)$ w przedziale małych kątów dla warunków środowiska morskiego jest szczególnie istotne, gdyż na przedział ten przypada ok. 76% rozpraszanej energii promienistej.

Jerlov [71] zwrócił uwagę, że w górnych niejednorodnych warstwach morza, rozpraszanie można opisać stosunkowo prostym modelem pojedynczego rozpraszania (*single scattering*) w dwóch kolejnych warstwach powierzchniowych (rys. 13). Wiązka promieniowania słonecznego, załamana pod kątem j zostaje bezpośrednio rozproszona przez elementy objętości wody ΔV_1 i ΔV_2 w dwóch kolejnych warstwach o grubościach z_1 i z_2 . Podział na dwie warstwy jest pierwszym przybliżeniem pionowej niejednorodności (różnica w koncentracji zawiesin) całej warstwy ($z_1 + z_2$). W oparciu o ten model wyprowadził



Rys. 12. Doświadczalne funkcje rozpraszania w układzie współrzędnych kartezjańskich. 1 — jezioro Pend Oreille (Idaho — USA), z pomiarów Tylera (27. IV. 1960), 2 — jezioro Pend Oreille, z pomiarów Tylera (26. IV. 1960), 3 — morze Wschodnio-Chińskie, z pomiarów Kozłaninowa, 4 — Północny Atlantyk z pomiarów Jerlova (1961)



Rys. 13. Ilustracja dwuwarstwowego modelu pojedynczego rozpraszania (Jerlov [71])

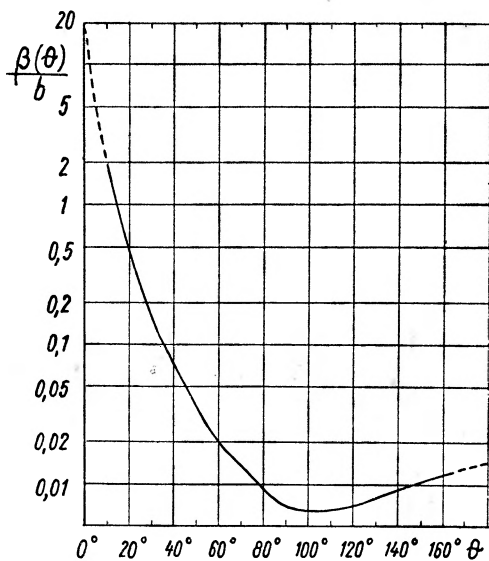
Jerlov wzór na oświetlenie w różnych przekrojach poziomych jako funkcję współczynników osłabiania i rozpraszania:

$$E = E(O)e^{-c_1 z_1 \sec j - c_2 z_2 \sec j} \left[1 + \frac{\sec j}{c_1} \int_0^{2\pi} \int_0^{\pi/2} \frac{\beta^I(\alpha) \sin \theta}{\sec j - \sec \theta} (1 - e^{-c_1 z_1 (\sec \theta - \sec j)}) \times \right. \\ \left. \times (e^{-c_2 z_2 (\sec \theta - \sec j)}) d\varphi d\theta + \frac{\sec j}{c_2} \int_0^{2\pi} \int_0^{\pi/2} \frac{\beta^{II}(\alpha) \sin \theta}{\sec j - \sec \theta} (1 - e^{-c_2 z_2 (\sec \theta - \sec j)}) d\varphi d\theta \right]. \quad (54)$$

W wyrażeniu tym $\beta^I(\alpha)$ i $\beta^{II}(\alpha)$ oznaczają funkcje rozpraszania względem kąta α tworzonego przez kierunki rozpraszania z kierunkiem wiązki załamanej. Łatwo wykazać, że

$$\cos \alpha = \cos j \cdot \cos \theta + \sin j \cdot \sin \theta \cdot \cos \varphi, \quad (55)$$

gdzie φ — kąt azymutalny. Numeryczne całkowanie (54), które przeprowadzone zostało na podstawie doświadczalnie wyznaczonych w morzu *in situ*, względnej funkcji rozpraszania $\beta(\theta)/b$ (rys. 14), b i c pozwoliło uzyskać wartości E dla różnych z . Porównanie tak obliczonych E z wartościami doświadczalnymi podaje tablica III.



Rys. 14. Przebieg stosunku $\beta(\theta)/b$ dla powierzchniowej warstwy wód w pobliżu Gibraltaru (Jerlov [71])

Tablica III

Głębokość z (m)	Współczynnik osłabiania c (m^{-1})	Współczynnik rozpraszania b (m^{-1})	Oświetlenie E %	
			dośw.	teoret.
0			100	100
10	0,051	0,035	76	77
25	0,050	0,034	47	47
50	0,055	0,038	16	16
75	0,066	0,050	3,9	3,8
90	0,063	0,047	1,73	1,68

Z tabeli widać dobrą zgodność dla E w zakresie z od 0 do 50 m. Zmieniając się w tym zakresie z , b pozwala wyróżnić postulowane przez Jerlova warstwy: górną 0—25 m (średnie $b = 0,034 m^{-1}$) i kolejną 25—50 m (średnie $b = 0,038 m^{-1}$). Z tablicy 3 widać ponadto, że różnica ($c - b$) dająca z definicji współczynnik absorpcji jest w podanym zakresie z stała i wynosi ok. $0,016 m^{-1}$.

Przedstawione w artykule wyniki stanowią pewną ilustrację zjawisk optycznych w morzu i ze względu na ciągły jeszcze brak materiału doświadczalnego

nie mogą stanowić pełnego potwierdzenia ogólnej teorii przenoszenia energii promienistej w morzu.

Panu Profesorowi Dr I. Adamczewskiemu dziękujemy za cenne uwagi merytoryczne podczas pisania tego artykułu.

Some Problems of Sea Physics
Part I. Radiant energy transfer in the sea

Abstract: The basic problem of contemporary sea optics i.e. propagation of radiant energy in the sea has been worked out. The topics considered were: a brief description of micro-phenomena connected with the transmission of light through the sea water, phenomenological classification of optical properties of sea based on the equation of radiant energy transfer as well as some elements of the theory of Preisendorfer with particular stress on consistent (coherent) and apparent optical properties of the sea. Furthermore an attempt of verifying the theory by experiments both on laboratory models and in natural water conditions was done.

Literatura

- [1] S. Q. Duntley, *J. Opt. Soc. Amer.*, **53**, 214 (1963).
- [2] S. Q. Duntley, *The Sea*, vol. I (*Physical Oceanography*) 452, New York, London 1962.
- [3] N. G. Jerlov, *Oceanography and Marine Biology*, **1**, 89 (1963).
- [4] M. W. Kozlaninow, *Trudy Inst. Okean. A. N. SSSR*, **35**, 3 (1959).
- [5] M. W. Kozlaninow, *Geofizika, Itogi Nauki AN SSSR*, 259 (1963).
- [6] R. W. Preisendorfer, *Symposium on Radiant Energy in the Sea*, IUGG Monogr. No 10 p. 83 (Helsinki 1961).
- [7] J. E. Tyler, R. W. Preisendorfer, *The Sea Vol. I (Physical Oceanography)* 397 (New York, London 1962).
- [8] G. Schellenberger, *Gerlands Beitr. Geophys.*, **72**, 315 (1963).
- [9] G. Schellenberger, *Acta Hydrophys.*, **10**, 79 (1965).
- [10] J. E. Tyler, *Proc. 10th Pacif. Sci. Congr.*, Bishop Museum Press s. 1, 344, Honolulu 1961
- [11] A. Defant, *Physical Oceanography*, Pergamon Press Oxford, London, New York, Paris 1961.
- [12] M. N. Hill (General Editor), *The Sea vol. 1 Physical Oceanography*, New York, London 1962.
- [13] G. Dietrich, K. Kalle, *Obszczeje Moriewiedienije*, Leningrad 1961 (tłum. z niem.).
- [14] W. W. Szulejkin, *Kratkij kurs fizyki moria*, Leningrad 1962.
- [15] W. T. Timofiejew, W. W. L. Panow, *Koswiennyje metody analizy i wydzielania wodnych mas; Gidrometeoizdat* 1962.
- [16] J. Dera, *Acta Geophys. Pol.*, **11**, 179 (1963).
- [17] J. Dera, *Acta Geophys. Pol.*, **13**, 15 (1965).
- [18] Dz. Proudman, *Dinamiczeskaja Okeanografija*, Moskwa 1957 (tłum. z ang.).
- [19] K. Wyrтки, *Symposium on Radiant Energy in the Sea*, IUGG Monogr. No 10, 51, Helsinki 1961.
- [20] K. Kalle, *Symposium on Radiant Energy in the Sea*, IUGG Monogr. nr 10, 59, Helsinki 1961.
- [21] A. N. Paramonow, *Okeanologia* **5**, 89 (1965).
- [22] B. P. Boden, E. M. Kampa, B. C. Abbott, *Proceedings 3rd. International Congress Photobiology*, Editors: Christensen B. Chr., Buchmann B. p. 189, 1960.
- [23] J. Dera, R. Bojanowski, *Acta Geophys. Pol.*, **14**, 23 (1966).

- [24] *Tablice Landolt-Börnstein*, B. III., 441 (1952).
- [25] J. Joseph, *Unesco Symposium on Physical Oceanography*, p. 53, Tokio 1955.
- [26] M. Dwing, E. Thorndike, *Science*, **147**, 1921 (1965).
- [27] D. W. Radakow, D. S. Nikolajew, *Okeanologia*, **3**, 517 (1963).
- [28] J. E. Tyler, *Limnology and Oceanography*, **6**, 451 (1961).
- [29] E. I. Gordiejew, *Trudy Inst. Okean. AN SSSR*, **64**, 202 (1964).
- [30] H. C. Van de Hulst, *Rassiejanije swieta małymi czasticami*, Moskwa 1961 (tłum. z ang.).
- [31] *Spektroskopia swietorassiejawajuszczich sried*, zbiór artykułów pod red. B. I. Stiepanowa, Mińsk, 1963.
- [32] K. J. Mysels, *J. Amer. Chem. Soc.*, **86**, 3503 (1964).
- [33] A. A. Sziszłowski, *Prikladnaja fiziczeskaja optika*, Fizmatgiz., Moskwa 1961.
- [34] Le Grand, *Symposium on Radiant Energy in the Sea* IUGG Monogr. No 10, 9, Helsinki 1961.
- [35] G. Schellenberger, *Pure and Appl. Geophys.* **58**, 129 (1964).
- [36] J. E. Tyler, *J. Mar. Res.*, **18**, 157 (1960).
- [37] J. E. Tyler, *Limnology and Oceanography*, **6**, 24 (1961).
- [38] E. M. Kampa, *Symposium on Radiant Energy in the Sea*, IUGG. Mon. No 10, 91, Helsinki 1961.
- [39] B. P. Boden, *Symposium on Radiant Energy in the Sea*, IUGG Monogr. No 10, 96, Helsinki 1961.
- [40] E. M. Kampa, B. P. Boden, *Nature* **174**, 869 (1954).
- [41] A. H. Ivanoff, *Symposium on Radiant Energy in the Sea*, IUGG Monogr. No 10, 45, Helsinki 1961.
- [42] F. S. Replogle, J. R., I. B. Steiner, *J. Opt. Soc. Amer.*, **55**, 1149 (1965).
- [43] R. N. Wolfe, J. J. De Palma, S. B. Saunders, *J. Opt. Soc. Amer.*, **55**, 956 (1965).
- [44] W. N. Pielewin, *Izw. AN SSSR, fiz. atmosf. i okeana*, **1**, 539 (1965).
- [45] N. Jerlov, *Rep. Swedish Deep — Sea Exp., 1947—48*, **3**, Fasc. 1 (1951).
- [46] R. W. Preisendorfer, *J. Mar. Res.*, **18**, 1 (1959).
- [47] W. A. Timofiejewa, *Dokl. AN SSSR* **147**, 604 (1962).
- [48] W. A. Timofiejewa, *Izw. AN SSSR ser. geofiz.* No 12, s. 1843 (1962).
- [49] J. Tyler, *Mar. Res.*, **21**, 102 (1963).
- [50] G. C. Holzworth, C. R. Nagaraja Rao, *J. Opt. Soc. Amer.*, **55**, 403 (1965).
- [51] T. H. Waterman, *Symposium on Radiant Energy in the Sea*, IUGG, Monogr. No 10, p. 103, Helsinki 1961.
- [52] F. A. J. Armstrong, G. T. Boalch, *Symposium on Radiant Energy in the Sea*, IUGG, Monogr. nr 10, 63, Helsinki 1961.
- [53] *Report First Meet on Photosynthetic Radiant Energy — Moscow 1964* Unesco Techn. Pap. Mar. Sci. No 2.
- [54] S. Chandrasekhar, *Radiative transfer*, Oxford 1950.
- [55] R. W. Preisendorfer, *J. Math. and Mech.*, **6**, 685 (1957).
- [56] J. Kalinowski, J. Dera, *Postępy Fizyki (w przygotowaniu do druku)*.
- [57] R. W. Preisendorfer, *Rep. Scripps Inst. Oceanography*, University of California, SIO Ref. 58—43 (1957).
- [58] R. W. Preisendorfer, *Rep. Scripps Inst. Oceanography University of California*, Ref. 58—46 (1958).
- [59] R. W. Preisendorfer, *J. Roy. Meteor. Soc.* **83**, 358 (1957).
- [60] R. W. Preisendorfer, *Proc. Nat. Acad. Sci. USA*, **44**, 320 (1958).
- [61] R. W. Preisendorfer, *Proc. Nat. Acad. Sci. USA*, **44**, 323 (1958).
- [62] R. W. Preisendorfer, *Proc. Nat. Acad. Sci. USA*, **44**, 328 (1958).
- [63] J. Lenoble, *Symposium on Radiant Energy in the Sea*, IUGG Monogr. No 10, 30, Helsinki 1961.
- [64] J. E. Tyler, *Symposium on Radiant Energy in the Sea* IUGG Monogr. No 10, 40, Helsinki 1961.
- [65] W. R. G. Atkins, and H. H. Poole, *Proc. Roy. Soc. B.*, **140**, 321 (1952).

- [66] Ju. A. R. Mullama, *Atlas optyczeskich charakteristik wzwolnowannoj powierzchni moria*, Tartu 1964.
- [67] Z. Sekera, *Symposium on Radiant Energy in the Sea*, IUGG Monogr. No 10, 66, Helsinki 1961.
- [68] L. V. Whitney, *J. Mar. Res.*, **4**, 122 (1941).
- [69] H. Hinzpeter, *Kielar Meeresforschungen* **18**, 36 (1962).
- [70] J. E. Tyler, W. H. Richardson, R. W. Holmes, *J. Geophys. Res.*, **64**, 667 (1959).
- [71] N. G. Jerlov, *Symposium on Radiant Energy in the Sea*, IUGG Monogr. No 10, 3, Helsinki 1961.
- [72] M. A. Eliazewicz, *Atomnaja i Molekularnaja Spiektroskopia*, Moskwa 1962.
- [73] *Standart Terminology on Optics of the Sea* — IAPSO Stock of Publications Oceanografiska Institutionen, Box 1038, Göteborg 1964.
- [74] J. E. Tyler, *Proc. Nat. Acad. Sci.*, **47**, 1726 (1961).

Bronisław Kuchowicz

Warszawa

Czy meteory mogą składać się z antymaterii?

W świetle teorii Kleina (patrz artykuł Alfvéna w tymże numerze Postępów Fizyki) na uwagę zasługuje rozpatrzenie możliwych konsekwencji wysunętej niedawno hipotezy o tym, jakoby komety zbudowane były z antymaterii i dochodziły do naszego Układu Słonecznego z takich obszarów wszechświata, które są wypełnione antymaterią. Jedną z takich konsekwencji, które można sprawdzić eksperymentalnie, jest powstawanie rojów meteorów w wyniku rozpadu komet. Roje meteorów składać się będą w tym wypadku z antymaterii i trafiając do wyższych warstw atmosfery ziemskiej dawać będą charakterystyczne promieniowanie anihilacyjne. W procesach anihilacji powstawać będzie twarde promieniowanie γ oraz szybkie neutrony i antyneutrony. Kwanty γ o energii powyżej 70 MeV, szczególnie nadające się do detekcji, wytwarzane będą w rozpadach mezonów π^0 z anihilacji. Wymienione rodzaje promieniowania, charakteryzujące się dużą wartością średniej drogi swobodnej w atmosferze, łatwo jest zarejestrować w dużych nawet odległościach od obszaru, w którym nastąpiła anihilacja. Według dostępnych danych natężenie strumienia promieniowania γ w wymienionym zakresie energii, na wysokości, na której pozostaje atmosfera ok. 10—100 g/cm³, wynosi mniej więcej 0,1 cm⁻² sek⁻¹ stereorad⁻¹ [1, 2].

Jeśli przyjmiemy, że cały strumień promieniowania γ pochodzi z procesów anihilacyjnych, można otrzymać górną granicę dla stężenia antymaterii w otaczającej nas przestrzeni. Wartość maksymalna wyniesie ok. 10⁻⁷ antynukleonów w 1 cm³.

Grupa uczonych radzieckich ogłosiła ostatnio [3] wyniki pomiarów średniego natężenia promieniowania γ na wysokości 25—30 km w okresach rozmaitej aktywności meteorów. Stwierdzono, że podczas występowania najsilniejszych rojów meteorów strumień kwantów γ wzrasta mniej więcej o 50%. Stąd można ocenić graniczną masę mikrometeoru (przy założeniu, że składać się on ma z antymaterii) na 10⁻⁹—10⁻¹⁰ g¹.

Oprócz pomiarów średniego strumienia kwantów γ na wysokości 25—30 km Konstantinow i współpracownicy [3] przeprowadzili pomiary na wysokości 13—18 km w celu wykrycia promieniowania, związanego z pojedynczym rejestrowanym meteorem, wchodzącym do atmosfery. Kwanty γ i neutrony rejestrowano za pomocą liczników scyntylacyjnych i borowych, a jednocześnie stosowano metody radiolokacyjne do śledzenia toru meteoru (długość fali 4 m). Odbite sygnały radarowe i impulsy z detektorów rejestrowano na tej samej taśmie, co pozwalało na śledzenie korelacji w czasie pomiędzy tymi dwoma zapisami do 0,1 sek.

Uczni radzieccy gromadzili informacje w ciągu trzech lat, a do opracowania ich stosowali technikę maszyn liczących. Jedną z metod obróbki danych polegała na tym, że sumowano liczby impulsów z detektorów dla pięciosekundowych interwałów, w ciągu których rejestrowano sygnał radiolokacyjny, odbity od meteoru, oddzielnie zaś sumowano analogiczne liczby dla

¹ Jeśli chodzi o wydzielenie energii, jest on równoważny meteorowi z materii o masie rzędu 0,1 g.

interwałów czasu, w ciągu których takiego odbicia nie było. Okazało się, że pierwsza z otrzymanych wartości systematycznie przewyższała drugą. Prawdopodobieństwo tego, że efekt ten był przypadkowy, wynosiło 0,0002. Tak więc z dużym prawdopodobieństwem statystycznym udało się stwierdzić, że podczas obserwacji śladu meteoru w atmosferze powiększa się rejestrowany strumień wysokoenergetycznych kwantów γ i neutronów. To podwyższenie wynosi około 2% w stosunku do tła, i równe jest liczbowo mniej więcej jednemu impulsowi na jeden meteor. Zainteresowanych wynikami Konstantinowa i współpracowników odsyłam do ich pracy [3], w której podane są szczegółowe tablice oraz kilka typowych diagramów danych eksperymentalnych.

Na zakończenie warto nadmienić, że w świetle zreferowanych tu wyników mniej fantastyczna wydaje się hipoteza o meteorycie tunguskim jako bryle antymaterii, na rzecz której w ubiegłym roku [4] przytoczono wyniki pomiarów zawartości ^{14}C w rocznych przyrostach drzew.

May the micrometeors consist of antimatter?

Abstract: If the comets consist of antimatter, then meteor showers originating in the decays of these comets may exhibit some annihilation properties in their interaction with the Earth's atmosphere. Results of Soviet investigations show that an increase in hard γ radiation and in neutrons (2% of the background) is observed during the time of observation (by radar techniques) of the meteor track. Observations have been performed during the years 1961—64. The estimated mass of a meteor from antimatter amounts to 10^{-9} — 10^{-10} g.

Literatura

- [1] T. L. Cline, Phys. Rev. Lett., **7**, 109 (1961).
- [2] W. L. Kraushaar, C. W. Clark, Phys. Rev. Letters, **8**, 106 (1962).
- [3] W. P. Konstantinow, M. M. Bredow, A. I. Bielajewskij, I. A. Sokołow, Kosmicheskije Issledowanija, **4**, Nr 1, 66 (1966).
- [4] C. Cowen, C. R. Atluri, W. F. Libby, Nature **206**, 861 (1965).

J. Z. Cypkin: *Teoria układów impulsowych*. Państwowe Wydawnictwo Naukowe, Warszawa 1965, str. 626.

Burzliwy i dynamiczny rozwój techniki impulsowej w ciągu ostatnich lat i olbrzymie zapotrzebowanie przemysłowe na efektywne, niezawodne i ekonomiczne rozwiązania automatyki przemysłowej sprawiło, że w dużym stopniu praktyczne rozwiązania inżynierskie wyprzedziły w tym zakresie badania podstawowe i teorię układów impulsowych. Zadaniem teorii układów tego rodzaju jest w pierwszym stadium analiza, w drugim zaś, ważniejszym, synteza układów o zadanych cechach konstrukcyjnych z uwzględnieniem możliwości ich realizacji technicznej. O ile w zakresie układów liniowych metody syntezy poczyniły już znaczne postępy, o tyle w zakresie układów nieliniowych, szczególnie interesujących dla praktyki, metody syntezy praktycznie nadal nie istnieją. Obok metod ciągłych pomiarów i sterowania rozwinęły się w ostatnich czasach także metody impulsowe, dające często poprawę stabilności układu w stosunku do regulacji ciągłej, szczególnie dla obiektów o dużym opóźnieniu lub inercji, dużą dokładność regulacji i możliwość sterowania bardzo małą mocą dużych organów wykonawczych, przy znacznej prostocie konstrukcyjnej. Kontynuacją i rozwinięciem metod regulacji impulsowej są układy regulacji cyfrowej, w których odpowiednie konwertyory analogowo-cyfrowe pozwalają na znaczne powiększenie dokładności regulacji i umożliwiają wykorzystanie do celów automatyki zalet współczesnych elektronowych maszyn cyfrowych. Wszystkie te nowe metody oparte są na technice impulsowej i wymagają zarówno teoretycznego opanowania analizy takich układów, jak i praktycznych metod ich syntezy.

Jedną z pionierskich prac z zakresu techniki impulsowej jest niezwykle obszerne dzieło znanego radzieckiego naukowca J. Z. Cypkina pt. *Teoria układów impulsowych*, wydane przez Państwowe Wydawnictwo Naukowe w r. 1965, na podstawie wydania rosyjskiego z r. 1958. Książka powyższa zainteresuje niewątpliwie szerokie grono inżynierów i studentów specjalizujących się w automatyce, telemechanice, maszynach liczących, urządzeniach radiolokacyjnych itp.

Układ rozdziałów jest przejrzysty i logiczny, umożliwiający szybkie odszukanie wybranego zagadnienia. Omawiając podstawowe pojęcia z zakresu układów impulsowych autor dokonuje ich podziału na układy otwarte (bez sprzężeń zwrotnych) i zamknięte (ze sprzężeniami zwrotnymi obejmującymi impulsator). Podstawy matematyczne teorii układów impulsowych przedstawia rozdział drugi. Zasadniczą metodą badawczą wprowadzoną przez autora jest tzw. dyskretne przekształcenie Laplace'a, które pozwala na sformalizowanie aparatu matematycznego analogicznie do przekształcenia ciągłego, szeroko znanego i stosowanego przy analizie układów liniowych. Niewątpliwą zaletą tej metody jest możliwość oparcia się przy jej przyswajaniu na znanych już i stosowanych pojęciach po odpowiedniej ich modyfikacji. Dalej wprowadza autor tzw. funkcje drabinkowe i rozpatruje równania różnicowe nie wspominając o innych stosowanych metodach analizy układów impulsowych, np. funkcjach schodkowych czy też tzw. transformacji Z. Na uwagę zasługuje ustęp o traktowaniu funkcji będących procesami stochastycznymi oraz ich wpływie na pracę układów impulsowych. Rozdział trzeci omawia teorię układów impulsowych otwartych, a czwarty badanie i projektowanie układów impulsowych otwartych na przykładzie szeregu konkretnych urządzeń. Wreszcie rozdział piąty daje podstawy teorii układów impulsowych zamkniętych, a siódmy przedstawia badanie i obliczanie takich układów, także w oparciu o szereg znanych urządzeń, przy czym jeden

z ustępów traktuje o maszynach liczących jako przetwornikach danych w układach automatyki. Dodatek zawiera ponadto szereg tablic i wykresów potrzebnych przy obliczeniach układów impulsowych. Bogaty zestaw literatury, mający 354 pozycje, podzielono na cztery działy i umieszczono na końcu monografii. Skorowidz umożliwia szybkie odszukanie potrzebnych zagadnień.

Zdaniem recenzenta monografia o tak dużej objętości jest źle dostosowana do potrzeb inżyniera-praktyka pragnącego rozszerzyć swoją wiedzę i może oddziaływać na czytelnika odstrasząco. Nie sposób czytać wybranych ustępów bez lektury zasadniczej części teoretycznej, ta zaś nie grzeszy jasnością i zwięzłością ujęcia i nie zawiera przykładowo zestawienia używanej symboliki. Układ rozdziałów ma materiał dobrany dość swobodnie, odpowiednio w stosunku do rozwoju konkretnych urządzeń technicznych. Tak np. ustępy 1.2 i 1.3 omawiają, wbrew tytułowi monografii, także konkretne urządzenia impulsowe zamiast ograniczyć się przede wszystkim do pojęć podstawowych. Wiele z wyprowadzeń można było umieścić na końcu w formie dodatków lub tablic, lub też wskazać ich źródła bibliograficzne. Zastrzeżenie budzi wyrażenie „kwantowanie w poziomie“ (chodzi o kwantowanie wielkości fizycznej sygnału) na str. 10, nie zwrócone uwagi na istotne cechy sygnałów rzeczywistych tak ciągłych, jak i dyskretnych, którymi są skończony czas ich trwania i ich skończona moc. Zapewne z winy tłumaczenia określa się wielkość ω raz jako częstość (str. 12), innym razem jako częstotliwość (str. 178), podczas gdy w innych miejscach (np. str. 96) mówi się poprawnie o pulsacji. Recenzent na próżno usiłował zastąpić często występujące wyrażenie „układ ciągły“ (str. 13) innym, lepiej oddającym istotę rzeczy dla układu niezawierającego impulsatora. Szkoda także, że redakcja wydawnictwa polskiego nie uzupełniła spisu literatury najnowszymi pozycjami z omawianego zakresu, gdyż siedem lat w dziedzinie nauk technicznych stanowi już okres, który znacznie dezaktualizuje dzieło.

Mimo to omawiana monografia stanowi oryginalną i cenną pozycję naszego technicznego rynku wydawniczego i będzie na pewno bardzo pomocna zarówno dla samouków, jak i inżynierów konstruktorów.

Roman Śledziewski

F. W. Name: *Fizyka współczesna*. Państwowe Wydawnictwo Naukowe, Warszawa 1965. str. 477. Tłumaczył Bronisław Średniawa.

Książka F. W. Van Name'a pt. *Fizyka współczesna*, która ukazała się ostatnio w polskim przekładzie, nie jest, jak tego można by się spodziewać, opracowaniem zagadnień będących aktualnie na warsztacie fizyków, ani nawet przedstawieniem osiągnięć fizyki ostatnich lat. Autor zajmuje się bowiem głównie problemami fizyki końca dziewiętnastego i początku dwudziestego stulecia, a mianowicie fizyką elektronu, atomu i jądra atomowego w takim zakresie, że tytuł podręcznika „Podstawy fizyki współczesnej“ lub „Wstęp do fizyki współczesnej“ byłby tu bardziej odpowiedni. Nazwę „fizyka współczesna“ trzeba tu zatem traktować jedynie jako przeciwstawienie do fizyki klasycznej.

Książka podzielona jest na trzy główne części, w których omówione są kolejno: a) odkrycie i własności elektronu oraz szczególna teoria względności, b) podstawowe prawidłowości w budowie atomów i drobin, głównie na podstawie ich widm optycznych i rentgenowskich, wreszcie c) fizyka jądra atomowego, cząstek jądrowych i promieni kosmicznych. Na końcu każdego rozdziału podane są zadania, dotyczące omówionych w tym rozdziale zagadnień.

Podręcznik utrzymany jest na ogół na takim poziomie, że może być z powodzeniem studiowany przez osoby stykające się z tymi zagadnieniami po raz pierwszy, może też być bardzo pomocny do prowadzenia niezbyt zaawansowanego wykładu z tej dziedziny.

Sposób wyłożenia materiału zawartego w tym podręczniku razić może miejscami swoją zwięzłością, chciałoby się czasem by autor nieco dokładniej objaśnił i uzasadnił omawiane prawo czy zjawisko. Jest to tym bardziej istotne, że podręcznik przeznaczony jest przeciw

dla początkujących. Niektóre rozdziały są wskutek tego mało zrozumiałe (np. rozdział o zasadzie nieokreśloności Heisenberga). Ta zwięzłość stylu pozwala jednak autorowi omówić szeroki wachlarz zagadnień w małej stosunkowo objętości.

Polskie wydanie omawianego podręcznika nie jest, niestety, bardzo staranne. Faktycznie istniejących błędów jest bowiem znacznie więcej niż to uwidoczniło na dwustronicowej wkładce *Errata*. Błędy te, zresztą łatwe do usunięcia przy staranniejszej korekcie, mogą utrudnić miejscami zrozumienie tekstu, zwłaszcza mało zaawansowanym czytelnikom. Dla ilustracji przytaczam przykładowo kilka z wielu dostrzeżonych usterek nie uwzględnionych w *Erracie*. I tak na str. 141 jest „wyjaśnić szczegółową strukturę subtelną“ zamiast chyba „wyjaśnić szczegółowo strukturę subtelną“, na str. 217 zamiast „kilowoltów“ jest „kilowatów“, na str. 227 jest „pionów“ zamiast „jonów“, na str. 263 jest promieniowanie „anihilacyjne“ zamiast „anihilacyjne“ itp.; częste są błędy w pisowni nazwisk np. jest „Doan“ zamiast „Duan“, a nazwisko van de Graaffa pisane jest prawie za każdym razem inaczej.

W tłumaczeniu polskim, na ogół bardzo starannym pod względem treści fizycznej, spotyka się miejscami niezbyt zręczne pod względem stylistycznym sformułowania, np. „przedozowanie zwykłej dawki promieni X“ lub nie używane w języku polskim terminy, np. „kwadruplet“ zamiast „kwartet“, „ciężar izotopowy“ i in. Wyraźne przejęzyczenia jak np. na str. 217, że „za pomocą betatronu przekroczono napięcie dwustu milionów woltów“, które trzeba chyba przypisać autorowi, należało opatrzyć objaśniającym komentarzem.

Mimo tych usterek *Fizyka współczesna* ze względu na tematykę i zakres wyłożonego materiału znajdzie na pewno licznych czytelników. Można ją polecić zwłaszcza tym, którzy pragną zaznajomić się z podstawowymi zagadnieniami fizyki współczesnej, bez specjalnego wnikania w szczegóły tych zagadnień.

Franciszek Leś

Maksymilian Pluta: *Mikroskopia fazowo-kontrastowa i interferencyjna*. Warszawa 1965
PWN, s. 370 + 14 wkł. cena 35.— zł.

Wydana nakładem Państwowego Wydawnictwa Naukowego książka wypełnia poważną lukę piśmiennictwa polskiego. W ostatnim dwudziestolecu nastąpił burzliwy rozwój techniki mikroskopii optycznej znacznie odbiegającej od „klasycznych“ metod mikroskopii. Jest to przede wszystkim mikroskopia fazowo-kontrastowa i interferencyjna. Stosowanie jej w różnych dziedzinach nauki i techniki jest olbrzymie, lecz wciąż jeszcze dalekie od pełnego wykorzystania. Stąd książka zawierająca podstawowe elementy teorii tych nowych technik, szereg zastosowań, a także i budowę najbardziej znanych urządzeń mikroskopowych, staje się bardzo ważną pomocą w rękach biologów, medyków, krytalografów i techników, którzy w swojej pracy stosują te techniki lub poszukują metod badawczych dla rozwiązania zagadnień w swych laboratoriach.

Książka składa się z dwu prawie równych części, z których pierwsza poświęcona jest mikroskopii fazowo-kontrastowej, a druga — mikroskopii interferencyjnej. Poza wstępem, bardzo prostym a niezbędnym dla zrozumienia zasady działania zwykłego mikroskopu i koniecznym dla wprowadzenia pojęć stosowanych w książce, autor przedyskutowuje właściwości i ograniczenia klasycznej mikroskopii amplitudowej. Wspomina nawet o współcześnie stosowanej metodzie optycznej funkcji przenoszenia, ogólniejszej i szerszej od stosowanych tradycyjn.

Dalej autor omawia przystępnie zasadę i teorię kontrastu fazowego. Kontrast fazowy pozwala na obserwację przedmiotów słabo pochłaniających światło i niedostrzegalnych bez techniki barwienia w zwykłym mikroskopie. Po omówieniu różnych wariantów mikroskopu typu Zernike następuje opis kontrastu fazowego typu anoptralnego oraz kontrastu fazowego zmiennego i barwnego. W dalszym ciągu autor opisuje użytkowanie poszczególnych rodzajów urządzeń kontrastowo-fazowych, ich regulację, a także właściwości i sposoby przygotowywania

preparatów. Przytacza także zastosowania w biologii i medycynie (mikrobiologia, hematologia, cytologia i in.), a także w chemii, krystalografii i mineralogii, włókiennictwie oraz metalografii.

W drugiej części książki autor zajmuje się mikroskopią interferencyjną. Podaje zasadę działania mikroskopii interferencyjnej dwupromieniowej i metodę interferencji dyferencjalnej oraz interferencji wielopromieniowej. Następnie daje przegląd znanych mikroskopów interferencyjnych dwupromieniowych, interferencyjno-polaryzacyjnych z płaskorównoległymi płytkami dwójłonnymi i pryzmatami dwójłonnymi.

Po omówieniu odmiennej specyfiki pomiarów za pomocą mikroskopu interferencyjnego, autor podaje szereg przykładów typowych pomiarów doskonale ilustrujących możliwości stosowania mikroskopu tego typu. Część tę zamyka wykaz zastosowań w naukach biologiczno-medycznych, chemii i fizyko-chemii, włókiennictwie, krystalografii i mineralogii oraz metalografii.

Książka charakteryzuje się wykładem prostym i przejrzystym. Aparat matematyczny jest sprowadzony do minimum. Tekst ilustrują dobrze dobrane rysunki i schematy. Książkę zamyka 25 tablic ze zdjęciami mikroskopowymi czarno-białymi i barwnymi. Dołączony wykaz literatury zawiera głównie prace źródłowe, w tym także i autora tej książki, konstruktora szeregu urządzeń mikroskopowych omawianych typów. Opisywane mikroskopy to urządzenia najczęściej spotykane i najbardziej charakterystyczne. Między innymi opisano także urządzenia produkcji krajowej.

Wyczuwa się także pewien niedostatek książki. Czytelnik pragnący nauczyć się posługiwania się mikroskopem dowiadyuje się, być może, zbyt wiele o metodach wytwarzania pierścieni fazowo-kontrastowych, których i tak nie będzie zapewne sam produkował, oraz może zbyt wiele o szczegółach konstrukcyjnych, choć przypuszczalnie książka nie jest napisana dla konstruktorów. Jest to przede wszystkim książka o mikroskopii i mikroskopach.

Książkę wydano na dobrym papierze. Strona graficzna jest dobra, redakcja dość staranna i błędów drukarskich jest niewiele. Niestety, mikrofotografie mimo papieru kredowego pod względem technicznym pozostawiają sporo do życzenia.

Mikroskopia fazowo-kontrastowa i interferencyjna znajduje niemal z każdym dniem coraz to dalsze zastosowania i pojawiają się coraz nowsze rozwiązania techniczne. Z książką powinni zapoznać się pracownicy naukowcy i technicy wielu dziedzin, szukający rozwiązania trudności w technice mikroskopowej i wszyscy ci, którzy poważnie zabierają się do pracy z szerokim stosowaniem mikroskopu fazowo-kontrastowego lub interferencyjnego. W sumie książka ta jest bardzo potrzebna i wypełnia, choć w części, ważną lukę w polskiej literaturze tego przedmiotu.

Zbigniew Lewandowski

N. W. Smirnow i I. W. Dunin-Barkowski: Krótki kurs statystyki matematycznej dla zastosowań technicznych. Państwowe Wydawnictwo Naukowe, Warszawa 1966, Biblioteka Naukowa Inżyniera, str. 520, cena zł 60.

Pisząc o tej książce w czasopiśmie, przeznaczonym dla fizyków, należy przede wszystkim zaznaczyć, że książka napisana jest dla techników. Z punktu widzenia fizyka, część pierwsza książki, obejmująca pięć rozdziałów, której przedmiotem jest wyłożenie elementów rachunku prawdopodobieństwa, napisana jest przy użyciu pojęć i języka, które są inne niż te, którymi posługują się fizycy. Z tego względu nie poleciłbym tych rozdziałów fizykowi, który chce zapoznać się z elementami rachunku prawdopodobieństwa, tym bardziej że mamy już dużo dobrych podręczników z tego zakresu (np. książki Fisz, Gliwienki czy też Arleya i Bucha, jeżeli zostanie przetłumaczona na język polski). Natomiast pozostała część książki zawiera sporo cennego dla fizyka materiału dotyczącego w szczególności zastosowań statystyki matematycznej. Wprawdzie zastosowania te dotyczą zagadnień technicznych, a nie fizycznych, ale wiele metod i część przykładów ma w fizyce odpowiednik. Duża ilość tablic, wykresów,

numerycznie przeliczonych przykładów ułatwia korzystanie z książki w praktyce. Co prawda, książka ta jest, jak gdyby skróconym wydaniem innej książki tych samych autorów, którą znam z oryginału rosyjskiego (Teorija wierojatnostiej i matematičeskaja statistika w tiechnike, GITTL, Moskwa 1955). Ta ostatnia książka jest znacznie lepsza pod względem bogactwa tablic, wykresów, przykładów liczbowych, czyniących z niej prawdziwą encyklopedię statystyki dla technika i również dla fizyka. Ze względu na rosnącą potrzebę znajomości metod statystycznych wydaje się, że przetłumaczenie tej drugiej książki, byłoby również bardzo pożądane. Wracając do książki, którą mamy już, chciałbym zwrócić w szczególności uwagę na rozdziały VI i VII, szczególnie użyteczne dla fizyka, zawierające metody statystyczne oceny parametrów rozkładu i testy hipotez statystycznych, a zatem zagadnienia, które w szczególności obejmują część zagadnień Rachunku Błędów. Rozdziały te napisane są bardzo nowocześnie i zawierają proste wyłożenie zasad i sposobu stosowania testu χ^2 , pojęcia przedziału ufności, testowania parametrów rozkładu i kształtu rozkładu i szereg innych ważnych dla fizyka metod stosowanych przy analizie pomiarów fizycznych. Krótki zarys teorii procesów stochastycznych zawarty w rozdziale XI jest dobrym wstępem, godnym przeczytania. Bardzo pożyteczny jest zbiór tablic zamieszczonych na końcu książki oraz szereg krótkich tablic w tekście.

Tłumaczenie książki jest bardzo ładne pod względem językowym, z uniknięciem narzucających się często przy tłumaczeniu z języka rosyjskiego podobieństw zwrotów i składni. Terminologia nie jest jeszcze dobrze ustalona i trudno ją oceniać. Książka wydana jest starannie, druk, wykresy i wzory przejrzyste, błędów druku nie znalazłem.

Wydanie tej książki w języku polskim jest cenne, ponieważ bardzo ważne staje się stosowanie statystyki w technice, a literatura w języku polskim w tej dziedzinie liczy dotychczas niewiele pozycji.

Przemysław Zieliński

Nagrody Leninowskie

Nagrody im. Lenina za rok 1966 w dziale fizyki otrzymali:

A. A. Abrikosow, L. P. Gorkow i W. L. Ginzburg za współudział w opracowaniu teorii nadprzewodnictwa stopów i właściwości nadprzewodników w silnych polach magnetycznych. Teoria ta, w której powstaniu wielką rolę odegrał Landau, nazywana jest często teorią G-L-A-G (Ginzburga, Landaua, Abrikosowa, Gorkowa).

Za badania teoretyczne i eksperymentalne ekscytonów w półprzewodnikach nagrodę otrzymał, o czym już informowaliśmy w poprzednim zeszycie Postępów, E. F. Gross wraz ze swymi współpracownikami B. P. Zacharcową i A. A. Kaplan-skim.

Również za badania ekscytonów nagrodę otrzymała grupa kijowska A. S. Dawydow, A. F. Prichotko, W. L. Broude, A. F. Lub-czenko, M. S. Brodin i E. I. Raszba.

Blackett członkiem honorowym edynburskiego Towarzystwa Królewskiego

Na zebraniu w dniu 4 lipca 1966 Towarzystwo Królewskie w Edynburgu nadało godność członka honorowego P. M. S. Blackettowi, profesorowi emerytowanemu fizyki w Uniwersytecie Londyńskim, laureatowi Nagrody Nobla z fizyki (1948).

Członkowie zagraniczni National Academy of Science

Hannes Alfvén (Królewski Instytut Technologiczny w Sztokholmie), P. M. S. Blackett (Uniwersytet Londyński) i Pol Swings (Uniwersytet w Liège) zostali wybrani na członków zagranicznych amerykańskiej Narodowej Akademii Nauk.

Nowi członkowie Akademii Nauk ZSRR

Akademia Nauk ZSRR dokonała wyboru nowych członków. Między innymi na członków rzeczywistych zostali powołani następujący fizycy: N. G. Basow, L. F. Wereszczagin, S. W. Wonsowski, W. L. Ginzburg, A. M. Prohorow, M. A. Markow i A. B. Migdał.

Medal dla Kapicy

Brytyjskie Towarzystwo Fizyczne i Instytut Fizyki nadały Medal i Nagrodę

Rutherforda za rok 1966 Piotrowi Kapicy (Instytut Problemów Fizycznych Akademii Nauk ZSRR, Moskwa) w uznaniu jego ogromnego wkładu w fizykę niskich temperatur. Jak wiadomo, Kapicę łączyły z Rutherfordem ścisłe kontakty naukowe w czasie, kiedy pracował w Cambridge.

Nowe władze Brytyjskiego Towarzystwa Fizycznego

Brytyjskie Towarzystwo Fizyczne i Instytut Fizyki na zebraniu ogólnym w dniu 5 lipca 1966 wybrały nowe władze w składzie: prezydent — sir James Taylor, wiceprezydent — J. V. Dunworth, skarbnik — P. T. Menzies, sekretarz — R. Press.

Nowi profesorowie

Rada Państwa na posiedzeniu w dniu 4 kwietnia 1966 r. nadała tytuł profesora nadzwyczajnego docentowi Marianowi Kryszewskiemu (Katedra Fizyki Wydziału Chemicznego Politechniki Łódzkiej) i docentowi Tadeuszowi Piechowi (Katedra Fizyki I Akademii Górniczo-Hutniczej).

Profesor zwyczajny

Rada Państwa na posiedzeniu w dniu 4 kwietnia 1966 r. nadała tytuł profesora zwyczajnego profesorowi Ludwikowi Kozłowskiemu (Katedra Fizyki I Akademii Górniczo-Hutniczej).

250 rocznica śmierci Leibniza

W roku bieżącym upływa 250 lat od śmierci wielkiego matematyka i filozofa niemieckiego Leibniza.

Gotfryd Wilhelm baron von Leibniz urodził się w 1646 r. w Lipsku. Korzystając z książek z biblioteki swego ojca w wieku 8 lat nauczył się łaciny, a mając lat 13 dobrze znał literaturę klasyczną. W 2 lata później wstąpił na Uniwersytet w Lipsku. Stopień naukowy uzyskał w Altdorf, gdyż Uniwersytet Lipski odmówił nadania mu stopnia ze względu na zbyt młody wiek. Następnie osiedlił się we Frankfurcie, gdzie został powołany na stanowisko doradcy Elektora do spraw prawnych, politycznych, teologicznych i filozoficznych. Wysłany do Paryża z misją do Ludwika XIV zetknął się tam z Huygensem, który skierował zainteresowania Leibniza ku matematyce. W r. 1676 został bibliotekarzem i radcą

stanu Księża Brunszwiku w Hanowerze i pozostał na tym stanowisku do końca życia. Zmarł w 1716 r.

Leibniz obdarzony był wszechstronnymi zdolnościami i opublikował prace z wielu rozmaitych dziedzin, jak meteorologia, paleontologia, etnografia, teologia. Między 1690 r. a 1716 r. wydał szereg dzieł filozoficznych, które czynią go jednym z największych filozofów epoki. W dziedzinie mechaniki wprowadził pojęcie „siły żywej“ (*vis viva*) — mv^2 . Uważał, że ta wielkość powinna być przyjęta jako „właściwa miara siły“ zamiast wielkości mv zalecanej przez Kartezjusza. Spór między zwolennikami obu poglądów trwał aż do 1743 roku, kiedy to d'Alembert pokazał, że nie ma sprzeczności między obu stanowiskami. Istotnie, wielkość mv^2 jest równa całce siły po czasie, a wielkość mv — całce siły po drodze. Zrozumienie tekstów z tego okresu utrudnione jest dla nowoczesnego czytelnika przez stałe używanie terminu „siła“ w znaczeniu energii, chociaż już Leonardo da Vinci dawał sobie sprawę z różnicy między tymi pojęciami.

Leibniz z powodzeniem budował maszyny do liczenia, przy czym doszedł do wniosku, że układ dwójkowy jest tu dogodniejszy od dziesiętnego. Dlatego można by go słusznie uważać za ojca współczesnych komputerów.

Nieśmiertelną zasługą Leibniza o podstawowym znaczeniu dla rozwoju nauki było odkrycie rachunku różniczkowego i całkowego. Formalnym śladem działalności matematycznej Leibniza są używane do dnia dzisiejszego symbole całki, różniczki, a także znaki mnożenia w postaci kropki i dzielenia w postaci dwukropka. Słynny spór o pierwszeństwo odkrycia rachunku różniczkowego i całkowego między Leibnizem a Newtonem jest już dziś przebrzmiały. Wydaje się, że obaj działali niezależnie i mniej więcej jednocześnie. Jednocześnie też prawie zostali wybrani przez Akademię Francuską na pierwszych jej członków zagranicznych (Leibniz 28 stycznia 1699, a Newton 21 lutego 1699).

Godne uwagi są również zasługi Leibniza w zakresie tworzenia instytucji poświęconych postępowi nauki. Głównie z jego inicjatywy powstała w 1700 r. w Berlinie Królewska Pruska Akademia Nauk. Przez 16 lat Leibniz kierował jej pracami. On również przyczynił się w znacznym stopniu do powstania w 1724 r. Akademii Nauk w Petersburgu.

Dwóchsetlecie urodzin Daltona

We wrześniu przypada dwóchsetlecie urodzin Johna Daltona, który odegrał wielką rolę w rozwoju fizyki i chemii przez wpro-

wadzenie do nauk ścisłych pojęcia atomu. Wprawdzie pojęcie to było znane już starożytnym filozofom greckim, ale przed Daltonem była to raczej abstrakcyjna koncepcja filozoficzna.

John Dalton urodził się 6 września 1766 r. w Eaglesfield w rodzinie kwakerskiej. Początkowo pracował jako robotnik rolny, później został nauczycielem. Mając lat 19 objął wspólnie z bratem Jonathanem kierownictwo szkoły w Kendal. W okresie następnych jedenastu lat pogłębiał swoją wiedzę przez samokształcenie, studiując grekę, łacinę i nauki ścisłe. W 1793 roku przeniósł się do Manchesteru, gdzie wykładał w New College — Oxford i Cambridge były wówczas niedostępne dla kwaków. Po sześciu latach zrezygnował ze swego stanowiska i zaczął uczyć prywatnie. W roku 1794 został członkiem Manchesterskiego Towarzystwa Literackiego i Filozoficznego, w r. 1822 został wybrany na członka Londyńskiego Towarzystwa Królewskiego, w r. 1826 otrzymał z rąk Davy'ego, jako prezydenta Royal Society, Medal Królewski, a od r. 1833 pobierał pensję królewską. Zmarł 20 maja 1844 r.

Rozważając występowanie ciał w trzech różnych stanach skupienia Dalton pisał: „Obserwacje te doprowadzają do wniosku, który wydaje się powszechnie przyjęty, że wszystkie (niezbyt małe) ciała składają się z wielkiej liczby nadzwyczaj drobnych cząstek, czyli atomów materii, związanych siłami przyciągania“, a w dziele swoim *New System of Chemical Philosophy* dodaje: „Każda dowolna (najdrobniejsza) cząstka wodoru nie różni się niczym od jakiegokolwiek innej (najdrobniejszej) jego cząstki. Procesy analizy i syntezy chemicznej polegają tylko na rozdzieleniu złączonych ze sobą cząstek lub na wytworzeniu takich połączeń. Czynniki chemiczne nie są jednak w stanie doprowadzić do stworzenia lub unicestwienia materii... Jedyne zmiany, jakie możemy wywołać, polegają na rozdzieleniu cząstek znajdujących się w stanie kohezji, czyli w postaci związku i na połączeniu ze sobą tych cząstek, które znajdują się w pewnej odległości od siebie“.

Charakterystyczne dla konkretności pojęcia atomu, tak jak rozumiał je Dalton, było jego zainteresowanie zagadnieniem ciężarów atomowych. W pracy przedstawionej Manchesterskiemu Towarzystwu Literackiemu i Filozoficznemu w październiku 1803 r. pisze: „Badanie względnych ciężarów elementarnych składników ciał (atomów) jest, o ile wiem, tematem zupełnie nowym. Prowadziłem te badania ze znacznym powodzeniem“. Aczkolwiek większość wyznaczonych przez Daltona ciężarów atomowych okazała się później błędna, to jednak próby jego mają niewątpliwie wielką wartość jako wskazanie właściwego kierunku poszukiwań. Słusznie wyraził się w 1925 r.

A. N. Whitehead w swoich odczytach „Nauka a świat współczesny“ wygłoszonych w Uniwersytecie Harvard: „W wieku XVIII każdy wykształcony człowiek czytał Lukrecjusza i idea atomu nie była mu obca, ale dopiero John Dalton uczynił atom pojęciem twórczym w nurcie nauki, a jako idea twórcza atomistyka była idea nową“.

Badania Daltona nad właściwościami gazów i par doprowadziły go do stwierdzenia, że ciśnienie mieszaniny gazów równa się sumie ciśnień oddzielnych składników. Prawo to znane jest obecnie jako prawo Daltona.

Przez wiele lat Dalton prowadził systematyczne obserwacje meteorologiczne w Manchesterze, przeważnie przy pomocy przyrządów własnej konstrukcji.

W 1798 r. przedstawił Manchesterskiemu Towarzystwu Literackiemu i Filozoficznemu pracę na temat „ślepoty barwnej“, której objawy zauważył u siebie samego. Mógł odróżnić błękit od fioletu, ale nie czerwien od zieleni. Widział tylko barwy błękitną, żółtą i fioletową, przedmioty pomarańczowe i zielone wydawały mu się żółte. Schorzenie to nazywane jest często daltonizmem.

Dwóchsetna rocznica urodzin Wollastona

W tym samym roku co Dalton, to jest w r. 1766, urodził się William Hyde Wollaston jako syn duchownego w East Dereham.

Wollaston ukończył studia medyczne w r. 1793, a w r. 1797 zaczął praktykować jako lekarz w Londynie. W trzy lata później zrezygnował z praktyki lekarskiej i rozpoczął pracę we własnym laboratorium. Wkrótce wynalazł metodę wytwarzania czystej kowalnej platyny. Wynalazek ten przyniósł mu 30 000 funtów, a później inne wynalazki przynosiły mu również znaczny dochód. Od roku 1793 był członkiem Londyńskiego Towarzystwa Królewskiego, a w latach 1804—1816 sekretarzem tego Towarzystwa. Jego badania w dziedzinie optyki i chemii zostały nagrodzone w 1802 r. Medalem Copleya. W 1823 r. został wybrany członkiem zagranicznym francuskiej Akademii Nauk.

Wollaston był znakomitym eksperymentatorem, wyróżniającym się nieprawdopodobną wprost zręcznością. Do najważniejszych jego osiągnięć naukowych należy odkrycie pierwiastków palladu i radu oraz szereg prac z dziedziny optyki i akustyki. Upamiętnił się przez wynalezienie pryzmatu polaryzującego, który dotychczas znany jest jako pryzmat Wollastona i przyrządu zwanego camera lucida oraz goniometru odbiciowego, pozwalającego na precyzyjne pomiary kątów w kryształach. Wydaje się, że jako pierwszy badał (w r. 1803) nad-

fiolkową dziedzinę widma, przy czym do detekcji użył chlorku srebra. Również jako pierwszy obserwował widmo pasmowe, a mianowicie w części niebieskiej płomienia świecy stwierdził występowanie widma CO, które dzisiaj nosi nazwę pasm Swana. W r. 1804 odkrył dichroizm badając przechodzenie światła przez sole palladu. W r. 1801 wykonał interesujący eksperyment mający na celu wykazanie identity elektryczności powstającej przez tarcie z elektrycznością galwaniczną. Doświadczenie polegało na wykorzystaniu wyładowań elektrostatycznych do wywołania elektrolizy wody. W związku z tymi doświadczeniami opracował metodę otrzymywania nici metalowych o średnicy zaledwie kilku mikronów.

Wollaston zmarł 22 grudnia 1828 roku.

Van Zandt Williams (1916—1966)

Dnia 13 maja 1966 zmarł nagle w Londynie Van Zandt Williams, dyrektor Amerykańskiego Instytutu Fizyki.

Williams urodził się w r. 1916 w Providence (Rhode Island), studiował w Brown University, doktoryzował się w Princeton w 1945 r. Zajmował się przede wszystkim spektroskopią podczerwieni. W okresie, kiedy uczeni pracujący w dziedzinie spektroskopii podczerwieni zmuszeni byli posługiwać się spektrografami, które sami budowali, Van Zandt Williams, wspólnie z R. B. Barnesem, zaprojektował dla firmy Perkin-Elmer pierwszy spektrograf na podczerwień produkowany seryjnie. W tym samym mniej więcej czasie R. R. Brattain opracował podobny przyrząd dla firmy Beckman Instruments. Był to okres II wojny światowej, kiedy analiza widmowa w podczerwieni wzbudzała wielkie zainteresowanie w związku z jej zastosowaniem w produkcji sztucznego kauczuku i lekkiej benzyny lotniczej. Od roku 1948 Williams był stałym pracownikiem firmy Perkin-Elmer, gdzie zajmował wysokie stanowisko w dziele badawczym.

Od 1960 r. Van Zandt Williams był członkiem zarządu Amerykańskiego Towarzystwa Optycznego, w roku 1965 został wybrany wiceprezydentem i prezydentem-elektem, a w styczniu 1966 objął urząd prezydenta tego Towarzystwa. Od kwietnia 1965 r. Williams był dyrektorem Amerykańskiego Instytutu Fizyki. Zmarł nagle w Londynie, dokąd pojechał z ramienia Amerykańskiego Instytutu Fizyki w sprawie ustalenia odpowiedniego programu międzynarodowej wymiany informacji fizycznych.

Williams był współautorem świetnej książki *Infrared Spectroscopy*.

ZARZĄD

Przewodniczący
Prof. dr WOJCIECH RUBINOWICZ
czł. rzecz. PAN

Wiceprzewodniczący
Prof. dr ALEKSANDER JABŁOŃSKI
czł. rzecz. PAN

Sekretarz
Doc. dr ANIELA WOLSKA

Skarbnik
Dr KAZIMIERZ ROSIŃSKI

Członek Zarządu
Doc. dr HALINA CHĘCIŃSKA
Dr TOMASZ HOFMOKL

PRZEWODNICZĄCY ODDZIAŁÓW TOWARZYSTWA

Gdańsk — Doc. dr A. KAWSKI
Głiwice — Prof. dr KAZIMIERZ GOSTKOWSKI
Kraków — Prof. dr HENRYK NIEWODNICZAŃSKI
Lublin — Doc. dr DANUTA STACHÓRSKA
Łódź — Doc. dr MARIAN KRYSZEWSKI
Opole — Doc. dr CZESŁAW JANKIEWICZ
Poznań — Doc. dr ZDZISŁAW PAJAK
Szczecin — Dr WIKTOR NOWAK
Toruń — Doc. dr MIECZYSLAW FRĄCKOWIAK
Warszawa — Doc. dr JULIAN AULEYTNER
Wrocław — Doc. dr ZYGMUNT BODNAR

TREŚĆ

H. Alfvén — Antymateria i ewolucja metagalaktyki (<i>tłum. J. Pochmurska</i>)	441
Laurie M. Brown — Poprzez kwarki do symetrii cząstek (<i>tłum. K. Rybicki</i>)	471
A. Kotański — Podstawowe pojęcia z teorii grup stosowane w fizyce cząstek elementarnych	485
J. Kryłow — Wpływ dyslokacji na własności elektryczne germanu II	497
H. Rzewuski — Defekty metatrwałe w germanie	525
J. Dera, J. Kalinowski — Wybrane zagadnienia fizyki morza	537
NOWOŚCI NAUKOWE	
B. Kuchowicz — Czy meteory mogą składać się z antymaterii?	565

RECENZJE

J. Z. Cypkin — Teoria układów impulsowych (<i>R. Śledziwski</i>)	567
F. W. Name — Fizyka współczesna (<i>F. Leś</i>)	568
M. Pluta — Mikroskopia fazowo-kontrastowa i interferencyjna (<i>Z. Lewandowski</i>)	569
N. W. Smirnow i I. W. Dunin-Barkowski — Krótki kurs statystyki matematycznej dla zastosowań technicznych (<i>P. Zieliński</i>)	570
KRONIKA	572

Warunki prenumeraty czasopisma

„POSTĘPY FIZYKI“ — dwumiesięcznik

Prenumeratę na kraj przyjmują urzędy pocztowe, listonosze oraz Oddziały i Delegatury „Ruch“.

Można również dokonywać wpłat na konto PKO nr 4-6-777, Przedsiębiorstwo Upowszechnienia Prasy i Książki „Ruch“ w Krakowie, ul. Worcella 6.

Prenumeraty przyjmowane są do 15 dnia miesiąca poprzedzającego okres prenumeraty.

Cena prenumeraty:

półrocznie zł 45.—, rocznie zł 90.—

Prenumeratę na zagranicę, która jest o 40% droższa, przyjmuje Biuro Kolportażu Wydawnictw Zagranicznych „Ruch“, Warszawa, ul. Wronia 23, tel. 20-46-88, konto PKO nr 1-6-100024.

Egzemplarze numerów zdezaktualizowanych można nabywać w Przedsiębiorstwie Upowszechnienia Prasy i Książki „Ruch“ w Krakowie, ul. Worcella 6, konto PKO nr 4-6-777.

Bieżące numery można nabywać lub zamówić w księgarniach „Domu Książki“ oraz w Ośrodku Rozpowszechniania Wydawnictw Naukowych Polskiej Akademii Nauk — Wzorcownia Wydawnictw Naukowych PAN — Ossolineum — PWN, Warszawa, Pałac Kultury i Nauki (wysoki parter).

Tylko prenumerata zapewnia regularne otrzymywanie czasopisma.