

FIZYKA

w Szkole z Astronomią

CZASOPISMO DLA NAUCZYCIELI

368 (LXIV) indeks 35810X Nr 3 maj/czerwiec 2020 CENA 30,00 zł (w tym 8% VAT)

Fizyka w przetwarzaniu
danych biometrycznych

Sztuczna grawitacja
w układzie obracającym się

Wielkie pęki z kosmosu

Epidemia koronawirusa
a reakcja łańcuchowa

Diagramy
Aleksandra Jabłońskiego

Metoda potencjałów węzłowych
w obliczaniu obwodów elektrycznych

ASTRONOMIA

30 lat Teleskopu Hubble'a



Aerodynamika
nadwozi samochodowych
Zjawiska fizyczne przy dużych prędościach

ISSN 0426-3383



9 770426 338001

e-wydanie
specjalne

CSAROWISKO DLA NAUCZYCIELI
Chemia
w Szkole

ISSN 0411-8524 Nr 1/2018
CENA 15,00 zł (w tym 23% VAT)



Laboratorium chemiczne

Wybór artykułów z najciekawszymi eksperymentami opisywanymi w „Chemii w Szkole” w latach 2015–2017

Wydanie specjalne
w wersji elektronicznej
(plik PDF)

Tylko 15 zł!

Każdy może eksperymentować!

Dowiedz się:

- Jak otrzymać metaliczne, a jak chemiczne rośliny?
- Co i jak można zrobić z piasku?
- Jak wyhodować chemicznego węża?
- Czy zielona herbata może świecić?
- Do czego chemikowi przydaje się aspiryna?

Formularz zamówienia na: www.aspress.com.pl/specjalne/

Drodzy Czytelnicy!

Zwielką radością zapraszam Państwa do lektury najnowszego, trzeciego tegorocznego wydania *Fizyki w Szkole*. W tym numerze kontynuujemy temat grawitacji. O ile w poprzednim rozważaliśmy wpływ zmian wprowadzanych przez człowieka na pole grawitacyjne, to w tym poszliśmy o krok dalej. Otóż w artykule „Sztuczna grawitacja w układzie obracającym się” dyskutowana jest szczegółowo analogia pomiędzy siłą odśrodkową a grawitacją. Jak twierdzi autor, problem ten nie jest tylko teoretyczny. Artykuł stanowi pewną propozycję rozwiązania ważnego problemu inżynierskiego jakim, w przypadku budowy pozaziemskich baz kosmicznych, jest brak grawitacji. Problem jest o tyle istotny, ponieważ jesteśmy istotami wychodowanymi w warunkach wszechobecności grawitacji, więc jej brak na dłuższą metę prowadzi do degeneracji naszych organizmów.

Rozważania dotyczące budowy kosmicznych miast na obręczach koła mają mimo wszystko pewien posmak fantastyki naukowej. Współczesnym i realnie działającym obiektem jest niewątpliwie teleskop Hubble’a. Jego skróconą historię znajdziemy w artykule dr Marcina Wesołowskiego. Kosmiczną tematykę znajdziemy też w artykule Tadeusza Wibiga – „Zaskakująco wielkie pęki”.

Fizyka to nie tylko kosmos. Zdecydowanie ziemską tematykę porusza bardzo interesujący, choć dość trudny, artykuł prof. Jana Piechny traktujący o dynamice pojazdów.

Zapraszam do lektury zarówno wspomnianych, jak i pozostałych artykułów.

W imieniu redakcji
Redaktor prowadzący
Zbigniew Wiśniewski

46 Promieniowanie kosmiczne (1938-2000). Zaskakująco wielkie pęki

■ Tadeusz Wibig



Fizyka wczoraj, dziś, jutro

4 Sygnały i obrazy, czyli fizyka w przetwarzaniu i analizie danych biomedycznych

■ Tomasz Kubiak

Współcześnie trudno już nawet wyobrazić sobie możliwość wykrycia i prawidłowego rozpoznania wielu poważnych schorzeń bez udziału aparatury, bazującej na najnowszych osiągnięciach fizyki.

13 2020 - Rok Fizyki. Diagramy Aleksandra Jabłońskiego

■ Kazimierz Mikulski

16 Sztuczna grawitacja w układzie obracającym się. Część I - spadek swobodny bez prędkości początkowej

■ Jan Kurzyk

Określenie *sztuczna grawitacja* w odniesieniu do układów nieinercjalnych brzmi lekceważąco. Przymiotnik „sztuczna” sugeruje, że mamy do czynienia z czymś niepełnowartościowym, substytutem czegoś lepszego, prawdziwego.



22 Wpływ własności aerodynamicznych nadwozi samochodów na ich właściwości jezdne

■ Janusz Piechna

33 Kartka z kalendarza. 200 lat elektromagnetyzmu

■ Zbigniew Wiśniewski

34 Epidemie i fizyka jądrowa

■ Jerzy Kuczyński

Epidemia to niemal identyczne zjawisko jak to co znamy z opisu reaktora jądrowego czy bomby atomowej – zarażanie to to samo co, doskonale nam znana, reakcja łańcuchowa.



Z naszych lekcji

36 Siły bezwładności - problem dydaktyczny

■ Waldemar Reńda

40 Zastosowanie metody potencjałów węzłowych do obliczania obwodów elektrycznych

■ Czesław Surowiec

Astronomia dla każdego

36 30-ta rocznica wyniesienia

w przestrzeń kosmiczną.

Trzy dekady Teleskopu Hubble’a

■ Marcin Wesołowski

Kosmiczny Teleskop Hubble’a (ang. HST – Hubble Space Telescope) to podstawowe narzędzie wykorzystywane w dzisiejszej astronomii obserwacyjnej. Taki stan rzeczy trwa 30 lat.



FIZYKA

w Szkole z Astronomią

NUMER 3 MAJ/CZERWIEC 2020 Nakład 3000 egz. CENA 30,00 zł
368 (LXIII) indeks 35810X ISSN 0426-3383 (w tym 8% VAT)

Komitet redakcyjny Krystyna Jabłońska-Ławniczak, Jerzy Kreiner, Andrzej Majhofer (Przewodniczący Komitetu), Zygmunt Mazur, Andrzej Szymacha, Mirosław Trociuk
Redakcja Zbigniew Wiśniewski (redaktor prowadzący – fizykc@wp.pl) **Adres redakcji** ul. Warchałowskiego 2/58, 02-776 Warszawa **Wydawnictwo** Agencja AS Józef Szewczyk, ul. Warchałowskiego 2/58, 02-776 Warszawa, e-mail: szewczyk24@gmail.com, tel. 606 201 244, www.aspress.com.pl, NIP: 951-134-91-51 **Wydawca i redaktor naczelny** Józef Szewczyk, szewczyk24@gmail.com **Prenumerata** www.aspress.com.pl/prenumerata-2020/, e-mail: szewczyk24@gmail.com, tel. 606 201 244 **Reklama** Jędrzej Chodakowski, jchodakowski1953@gmail.com **Skład i tamania** Aleksandra Milewska **Druk i oprawa** Paper & Tinta, ul. Ceglana 34, 05-270 Nadma
Zdjęcie na okładce: Mercedes.

Redakcja nie zwraca nadesłanych materiałów, zastrzega sobie prawo formalnych zmian w treści artykułów i nie odpowiada za treść płatnych reklam.

Współcześnie trudno już nawet wyobrazić sobie możliwość wykrycia i prawidłowego rozpoznania wielu poważnych schorzeń bez udziału aparatury, bazującej na najnowszych osiągnięciach fizyki.

Sygnaty i obrazy,

czyli fizyka w przetwarzaniu i analizie danych biomedycznych



Tomasz Kubiak

W medycynie wykorzystuje się zarówno bardzo skomplikowane maszyny, np. hybrydowe skanery PET-CT, stanowiące połączenie pozytonowej tomografii emisyjnej (PET od ang. *Positron Emission Tomography*) z rentgenowską tomografią komputerową (CT – *Computed Tomography*) jak również zdecydowanie prostsze urządzenia, czego przykładem mogą być elektrokardiografy czy stetoskopy elektroniczne. Rezultatem procesu diagnostycznego nie jest już zatem wyłącznie informacja alfanumeryczna (zapis wywiadu chorobowego czy wyniki badań laboratoryjnych), ale przede wszystkim dane w postaci obrazów bądź sygnałów. Wymagają one przetworzenia, analizy oraz bezpiecznej archiwizacji.

W przeszłości lekarze opierali się wyłącznie na wzrokowej ocenie zdjęć rentgenowskich podświetlonych z wykorzystaniem negatostopu czy wykresów EKG wydrukowanych na papierze milimetrowym. Takie postępowanie nie pozwala na pogłębioną interpretację wyników a dodatkowo stwarza ryzyko popełnienia błędu, chociażby na skutek przemęczenia czy braku doświadczenia diagnosty.

Jeszcze gorzej sytuacja przedstawia się w przypadku badań osłuchowych, przeprowadzanych klasycznym stetoskopem.¹ Ich wyniki cechuje bowiem duży subiektywizm, chociażby ze względu na różnice w percepcji i klasyfikacji dźwięków docierających z wnętrza ciała pacjenta. Jeśli bowiem sygnały akustyczne nie są w żaden sposób rejestrowane i archiwizowane, nie istnieje możliwość późniejszej weryfikacji diagnozy bądź przeprowadzenia jakiegokolwiek analizy porównawczej otrzymanych rezultatów.

Na szczęście postęp technologiczny, rozwój mikroelektroniki oraz wszechobecna cyfryzacja zaowocowa-

ły wprowadzeniem komputerowych systemów przetwarzania i analizy danych biomedycznych, co zdecydowanie zrewolucjonizowało diagnostykę medyczną zarówno w ujęciu jakościowym jak i ilościowym. Możliwe jest już zatem nie tylko zestawianie ze sobą informacji zebranych w kolejnych badaniach a także zgromadzonych w bazach danych, ale przede wszystkim obliczanie konkretnych parametrów charakteryzujących sygnały biologiczne, które cechują się przecież dużą złożonością.

Zastosowanie komputerów i odpowiednich algorytmów obliczeniowych umożliwia również dogłębną analizę informacji obrazowej. Przyjrzyjmy się nieco dokładniej, w jaki sposób osiągnięcia fizyków i inżynierów wspomagają lekarzy w prowadzeniu procesu diagnostycznego.

Wstępne przetwarzanie sygnałów biomedycznych

Badania medyczne dostarczają nam szeregu informacji, które często mają formę biosygnałów, czyli zmieniających się w czasie przebiegów wartości określonych wielkości fizycznych. Stali czytelnicy „Fizyki w Szkole” zapewne uświadamiają sobie, z jak wielką różnorodnością fizycznych metod diagnostycznych spotykamy się w praktyce klinicznej. Stąd rejestrowane sygnały mogą mieć charakter elektryczny, magnetyczny, akustyczny, cieplny, chemiczny itp. Wybrane przykłady zamieszczone są w tabeli.

W przypadku, gdy mierzoną wielkością fizyczną jest napięcie, co ma miejsce chociażby w badaniu elektrokardiograficznym (EKG), do rejestracji sygnału wystarczy dobór odpowiednich elektrod pomiarowych, charakteryzujących się możliwie najmniejszym oporem elektrycznym.²

¹ O stetoskopach klasycznych i elektronicznych można przeczytać w artykule: T. Kubiak, Fizyka u internisty, Fizyka w Szkole, nr 2 (2018), s. 4-7.

² Szczegóły dotyczące badania EKG w artykule: T. Kubiak, Fizyczne podstawy badania czynności bioelektrycznej serca, Fizyka w Szkole, nr 3 (2015), s. 47-50.

Metoda diagnostyczna	Natura sygnału	Znaczenie rejestrowanego sygnału
elektrokardiografia (EKG)	elektryczny	odzwierciedla elektryczną aktywność serca
elektroencefalografia (EEG)	elektryczny	odzwierciedla elektryczną aktywność mózgu
elektromiografia (EMG)	elektryczny	odzwierciedla elektryczną aktywność mięśni
elektroencefalografia (ENG)	elektryczny	pokazuje przewodzenie impulsów elektrycznych w nerwach obwodowych
magnetokardiografia (MKG)	magnetyczny	zmiany pola magnetycznego generowanego przez serce
tokografia	mechaniczny	zapis skurczów mięśni macicy
fonokardiografia	akustyczny	zapis dźwięków generowanych podczas pracy serca
manometria	mechaniczny	pomiar zmian ciśnienia (np. w obrębie zwieracza czy mięśniówki przełyku)

Jeśli jednak sygnał posiada inną naturę, musi on zostać w procesie przetwarzania wstępnego przekonwertowywany na sygnał elektryczny za pomocą odpowiednich przetworników. Przykładowo w stetoskopach elektronicznych fale dźwiękowe docierające z wnętrza ciała pacjenta są zamieniane na sygnał elektryczny z wykorzystaniem wbudowanego w głowicę czujnika piezoelektrycznego albo przetwornika akustyczno-elektrycznego o charakterze pojemnościowym.

W kolejnych etapach przetwarzania wstępnego sygnał może być poddawany linearyzacji (jeśli pochodzi z przetwornika o nieliniowej charakterystyce), wzmocnieniu i filtrowaniu (usunięciu szumów periodycznych i nieperiodycznych).

W przypadku badań elektrokardiograficznych (EKG) bądź elektroencefalograficznych (EEG)³ na wejściu układu pomiarowego oprócz zmiennego sygnału użytkowego pojawia się również składowa stała, czyli tzw. sygnał współbieżny. Występuje on na wszystkich elektrodach i posiada wartość około 220 – 500 mV, czyli znacznie większą niż właściwy sygnał biologiczny o amplitudzie rzędu kilku miliwoltów (EKG) bądź mikrowoltów (EEG). Dlatego zarówno w elektrokardiografach jak i elektroencefalografach stosuje się wzmacniacze pomiarowe, które wzmacniają różnicę napięć pomiędzy elektrodami oraz tłumią sygnał współbieżny. Impedancja wejściowa takich urządzeń musi być duża (> 20-100 MΩ) a tym samym znacznie przewyższać opór elektrod, aby sygnał użyteczny był stabilny.

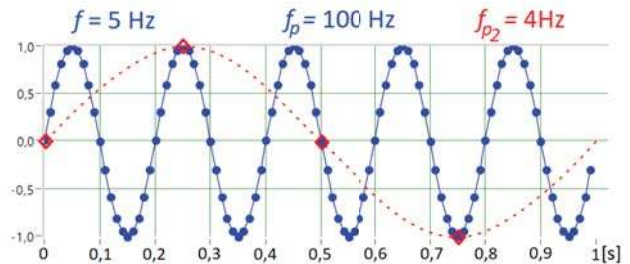
Wspomniane już przebiegi EKG i EEG są ciągłe, czyli określone dla każdej chwili. Ponadto mogą przyjmować dowolne wartości napięć w charakterystycznym dla nich zakresie, posiadają więc naturę analogową. Oczywiście, aby umożliwić komputerową analizę tego typu sygnałów biomedycznych konieczne trzeba zamienić je na postać cyfrową, w której zarówno dziedzina jak i zbiór wartości są dyskretne.

Przetwarzanie analogowo – cyfrowe

W procedurze konwersji sygnału analogowego na cyfrowy wyróżnia się zazwyczaj trzy etapy: próbkowanie, kwantyzację i kodowanie. Przyjrzyjmy się dokładnie każdemu z nich i zwróćmy uwagę na niebezpieczeństwa wiążące się z niewłaściwym przeprowadzeniem tych procedur.

Zacznijmy od próbkowania sygnału analogowego, które zazwyczaj jest równomierne i odbywa się z określoną częstotliwością f_p :

$$f_p = \frac{1}{\Delta t}$$



Rys. 1. Sinusoidalny sygnał analogowy o częstotliwości $f = 5 \text{ Hz}$ próbkowany z dwiema częstotliwościami próbkowania: $f_p = 100 \text{ Hz}$ (niebieskie kropki) oraz $f_{p2} = 4 \text{ Hz}$ (czerwone romby). W przypadku próbkowania ze zbyt niską częstotliwością ($f_{p2} = 4 \text{ Hz}$) wyraźnie widoczny jest efekt aliasingu.

Amplituda mierzona jest zatem tylko w ściśle określonych odstępach czasowych Δt , które nazywane są okresem próbkowania. Ciągły w czasie sygnał $V(t)$ zostaje zamieniony więc na szereg próbek:

$$V(t) \rightarrow \{V(0), V(\Delta t), V(2\Delta t), V(3\Delta t), \dots, V(n\Delta t)\}$$

W tym miejscu w głowach większości czytelników zapali się zapewne przysłowiowa lampka ostrzegawcza. I słusznie, bo niewłaściwie przeprowadzone próbkowanie może prowadzić do istotnych zniekształceń sygnału bądź utraty jego istotnych szczegółów. Klasycznym przykładem jest tzw. efekt aliasingu. Jeśli sygnał analogowy próbkujemy zbyt rzadko, czyli, mówiąc bardziej fachowo, zastosujemy zbyt małą częstotliwość próbkowania, składowe o wysokich częstotliwościach zostają nieprawidłowo zinterpretowane jako niskoczęstotliwościowe. Najłatwiej zobrazować to na przykładzie próbkowania przebiegu napięciowego mającego postać pojedynczej sinusoidy (rys. 1).

Jak zatem ustrzec się błędnej reprezentacji sygnału okresowego przez ciąg jego wartości chwilowych? Przede wszystkim trzeba odpowiednio dobrać częstotliwość próbkowania. Podstawą jest kryterium Shannona, które mówi, że sygnał analogowy zawierający częstotliwość f trzeba próbkować z częstotliwością $f_p \geq 2f$. Wówczas maksymalna częstotliwość składowa, która może być prawidłowo odtworzona w sygnale cyfrowym to tzw. częstotliwość Nyquista f_N :

$$f_N = \frac{f_p}{2}$$

Jeśli zatem sygnał analogowy, zawierający częstotliwości: $f_1 = 25 \text{ Hz}$, $f_2 = 70 \text{ Hz}$ i $f_3 = 165 \text{ Hz}$, będziemy próbkować z częstotliwością $f_p = 80 \text{ Hz}$, tylko f_1 zostanie odtworzona prawidłowo. Pozostałe składowe zostaną

niewłaściwie zinterpretowane, tzn. f_2 jako 10 Hz, natomiast f_3 jako 5 Hz. Poszczególne wartości f_{alias} , czyli mylnie odtworzonych częstotliwości, można wyznaczyć, posługując się wzorem:

$$f_{alias} = |n \cdot f_p - f_i|$$

gdzie: $n \cdot f_p$ – całkowita wielokrotność częstotliwości próbkowania, która leży najbliższej składowej o częstotliwości f_i .

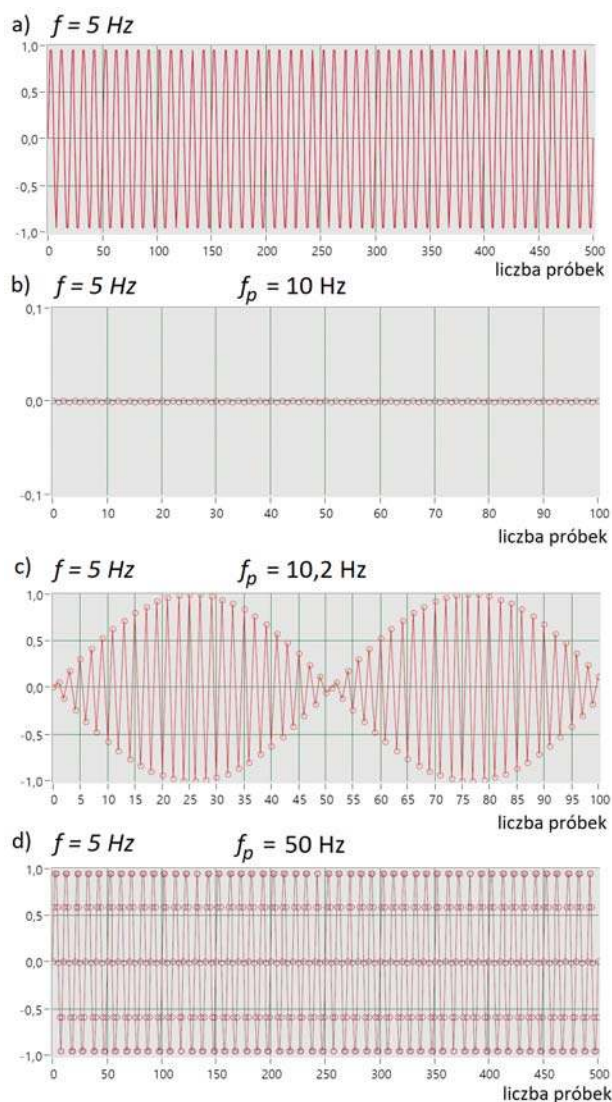
Warto wspomnieć, że w celu zminimalizowania ryzyka wystąpienia efektu aliasingu przed próbkowaniem sygnału można zastosować analogowy filtr dolnoprzepustowy, który usunie wszelkie składowe o częstotliwości przekraczającej zadaną wartość, w tym zakłócenia wysokoczęstotliwościowe, mogące negatywnie wpłynąć na strukturę widma.

Należy jednak mieć na uwadze, że nawet w przypadku spełnienia kryterium Shannona zrekonstruowany sygnał może zupełnie nie przypominać wyjściowego. Stanie się tak, gdy np. przebieg sinusoidalny o częstotliwości f spróbkujemy z $f_p = 2f$, zaczynając w jego fazie początkowej $\varphi = 0$ (otrzymamy wówczas sygnał o zerowej amplitudzie) albo gdy f_p tylko nieco zwiększymy ponad $2f$ (uzyskamy sygnał modulowany), co obrazuje rys. 2. Dla bezpieczeństwa warto zatem stosować częstotliwość próbkowania kilkukrotnie przewyższającą najwyższą częstotliwość występującą w sygnale analogowym.

Innym problemem, który może pojawić się podczas próbkowania sygnału, jest tzw. „efekt sztachetowy”. Wyobraźmy sobie, że podczas badania medycznego wystąpi nagle krótkotrwała zmiana amplitudy sygnału, czego przykładem może być iglica w zapisie EKG, pochodząca od kardiostymulatora. Jeśli taki gwałtowny skok napięcia pojawi się pomiędzy dwiema próbkami, nie zostanie zarejestrowany a tym samym informacja o nim zostanie utracona. Badania wskazują, że do wiarygodnej detekcji pochodzących od rozrusznika serca impulsów o amplitudzie 2-5 mV i czasie trwania 0,5-2 ms zazwyczaj potrzebne są częstotliwości próbkowania w zakresie 4-16 kHz.⁴ Dodajmy jeszcze, że w standardowych aparatach EKG częstotliwość próbkowania to zazwyczaj około 2000 Hz.

Warto również wspomnieć o sytuacji, w jakiej sygnał biomedyczny posiadałby składową zmieniającą się bardzo powoli w czasie. Wówczas informacja o niej zostanie utracona na skutek faktu, że próbkowaniu podlega tylko wycinek sygnału analogowego, posiadający ograniczoną długość.

Kolejnym zagadnieniem związanym z przetwarzaniem analogowo - cyfrowym jest kwantyzacja, która dotyczy amplitudy sygnału. Do jej odtwarzania wykorzystuje się bowiem określoną liczbę poziomów, wynikającą z rozdzielczości zastosowanego przetwornika. W ogólności,



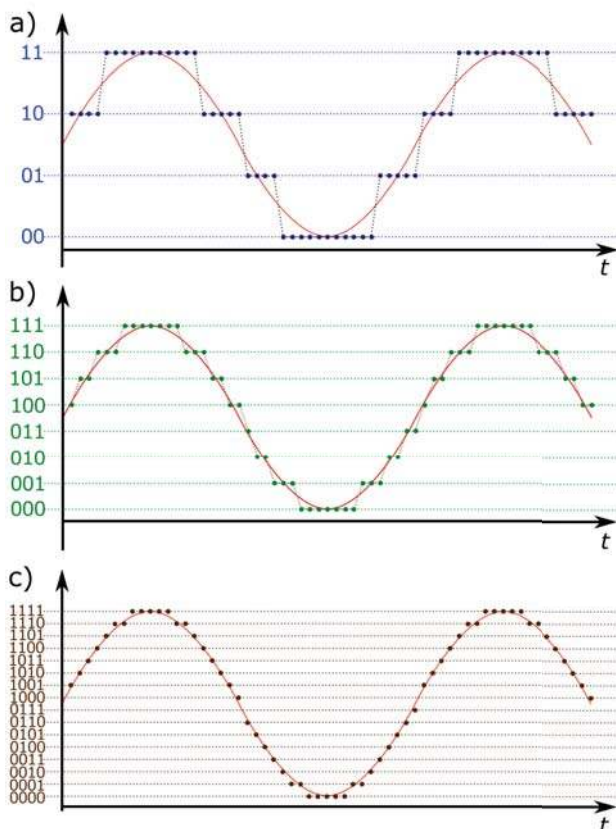
Rys. 2. Gdy przebieg sinusoidalny o częstotliwości $f = 5 \text{ Hz}$ (a) spróbkujemy z $f_p = 2f = 10 \text{ Hz}$, rozpoczynając w jego fazie początkowej $\varphi = 0$, otrzymamy sygnał o zerowej amplitudzie (b). Gdy zastosujemy $f_p = 10,2 \text{ Hz}$, uzyskamy sygnał modulowany (c). Użycie $f_p = 50 \text{ Hz}$ pozwala w sposób prawidłowy zrekonstruować sygnał (d).

gdy posiada on rejestr n bitowy, amplituda sygnału będzie reprezentowana przez 2^n dyskretnych stanów.

Już pierwszy rzut oka na rys. 3 uświadamia nam, że zbyt mała rozdzielczość przetwornika sprawi, że sygnał będzie przypominał schodki. Przetwornik 2 bitowy wyróżniałby bowiem tylko $2^2 = 4$ poziomy, 3 bitowy: $2^3 = 8$ poziomów, 4 bitowy: $2^4 = 16$ poziomów itd. Im większą zatem długością rejestru dysponujemy, tym dokładniejsze jest cyfrowe odwzorowanie sygnału i mniej widoczna jego nieciągłość. Stąd w urządzeniach medycznych, np. cyfrowych elektrokardiografach, stosuje się przetworniki 12, 16 a nawet 24 bitowe. Pozwala to ograniczyć tzw. błąd kwantyzacji, związany z utratą informacji o niewielkich zmianach wartości sygnału analogowego.

³ Badanie EEG opisano w artykule: T. Kubiak, Elektroencefalografia, czyli jak fizyka pomaga odkrywać tajemnice mózgu, Fizyka w Szkole, nr 6 (2016), s. 4-6.

⁴ Dane na podstawie pracy: M. Shoaib, H. Garudadri, Digital Pacer Detection in Diagnostic Grade ECG, IEEE 13th International Conference on e-Health Networking, Applications and Services, 2011, 326 - 331.



Rys. 3. Sygnał analogowy (czerwona krzywa) poddany kwantyzacji z wykorzystaniem przetworników o rozdzielczości odpowiednio: 2 bitów (a), 3 bitów (b) oraz 4 bitów (c).

Warto również przypomnieć, że wzmocnienie słabych sygnałów analogowych już na etapie przetwarzania wstępnego ułatwia dopasowanie przedziału zmienności ich amplitud do zakresu napięć wejściowych stosowanych przetworników analogowo - cyfrowych.

W wyniku opisanej powyżej operacji kwantyzacji zakres zmian wartości sygnału zostaje podzielony na określoną liczbę rozłącznych przedziałów. Każdemu z nich przypisuje się jedną wartość wyjściową (pochodzącą najczęściej z środka przedziału), nazywaną poziomem reprezentacji. Chwilowa wartość próbki sygnału analogowego przyporządkowywana jest zatem do najbliższego jej poziomu reprezentacji.

Z kolei poszczególnym poziomom kwantowania przydziela się w procesie kodowania tzw. słowa kodowe w postaci ciągu znaków, aby umożliwić komunikację przetwornika z dalszymi elementami systemu cyfrowego. Jeśli liczba przedziałów kwantowania określona jest jako 2^n , to naturalnym rozwiązaniem wydaje się zastosowanie systemu binarnego.⁵ Wówczas słowo kodowe to n-elementowy ciąg zer i jedynek (bitów). Te dwie cyfry mogą być reprezentowane fizycznie na kilka sposobów, np. jako różniące się poziomy napięcia.

⁵ Przypomnijmy, że zamiana liczby zapisanej w systemie dwójkowym na odpowiadającą jej wartość w systemie dziesiętnym polega na mnożeniu cyfr 0 i 1 przez kolejne potęgi liczby 2. Przykładowo: $1011_{(2)} = 1 \cdot 2^3 + 0 \cdot 2^2 + 1 \cdot 2^1 + 1 \cdot 2^0 = 11_{(10)}$.

⁶ W systemie uzupełnień do dwóch (w skrócie U2) najbardziej znaczący bit (czyli ten znajdujący się w zapisie słowa jako pierwszy z lewej strony) to tzw. bit znaku. Gdy jest on jedynką, to kodowana liczba jest ujemna, natomiast jeśli zerem, to liczba ta jest dodatnia albo równa 0. Przykładowo: $1011_{(U2)} = 1 \cdot (-2^3) + 0 \cdot 2^2 + 1 \cdot 2^1 + 1 \cdot 2^0 = -5_{(10)}$

Należy jednak pamiętać, że naturalny kod binarny pozwala przedstawić tylko liczby nieujemne. Zamiast takiego unipolarnego kodu w praktyce stosuje się zatem kod bipolarny, gdzie jeden z bitów słowa kodowego mówi o znaku. Przykładem może być tzw. kod uzupełnień do dwóch, który umożliwia zapis liczb całkowitych⁶.

Sygnały z urządzeń diagnostycznych trafiają do komputera z użyciem odpowiednich interfejsów w standardzie RS 232, IEEE-488, USB, bądź są przesyłane przez wi-fi czy bluetooth. Ten ostatni sposób wykorzystywany jest m.in. do transferu zapisów fali tętna z pulsoksymetrów.⁷ Wprowadzone do komputera dane mogą być podane dalszym etapem przetwarzania.

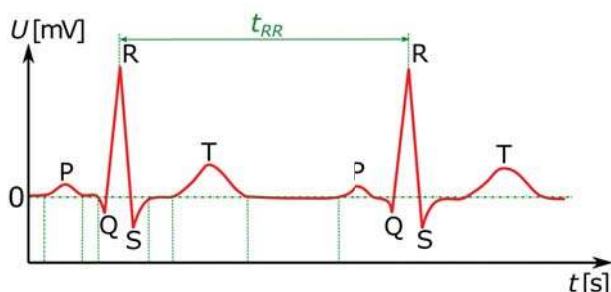
Cyfrowa analiza sygnałów

Jak już wcześniej wspomniano powierzchownie, wzrokowa ocena sygnałów biomedycznych pozwala wydobyć z nich tylko niewielką część informacji, co może prowadzić do błędów w interpretacji wyników badań a w konsekwencji do postawienia złej diagnozy. Dlatego, szczególnie w przypadku złożonych sygnałów, zdecydowanie przydatne jest wspomaganie się metodami ich cyfrowego przetwarzania. Sygnały można wówczas analizować nie tylko w dziedzinie czasu, ale również w dziedzinie częstotliwości a nawet w połączonej dziedzinie czasu i częstotliwości.

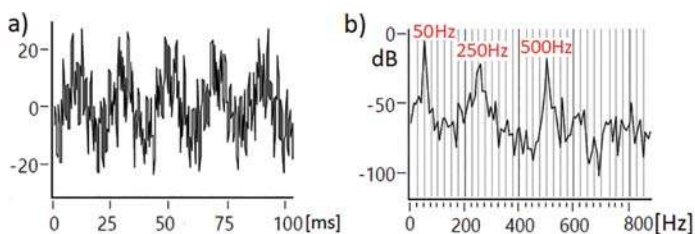
Oczywiście cyfrowe przetwarzanie sygnału może obejmować również jego filtrację. Pozwala ona usunąć zarówno wąskopasmowe zakłócenia cykliczne (np. 50 Hz pochodzące od zasilacza sieciowego) jak i te niecykliczne o charakterze szerokopasmowym, np. związane z drżeniem mięśni pacjenta podczas badania. Współczesne urządzenia diagnostyczne standardowo posiadają zaimplementowane filtry cyfrowe, czego przykładem są aparaty EKG wyposażone w filtry: sieciowe 50 i 60 Hz, zakłóceń mięśniowych 25, 35, 45 Hz czy tzw. pływania izolinii (np. 0,15 Hz, 0,45 Hz, 0,75 Hz, 1,5 Hz). Usunięcie zakłóceń znacznie ułatwia późniejszą analizę sygnału.

Podczas stawiania diagnozy przedstawiciele środowiska medycznego najchętniej posługują się wynikami badań, które mają postać przebiegów czasowych. Sposób oceny takich sygnałów zależy oczywiście od ich struktury, niemniej jednak zazwyczaj uwzględniana jest informacja o amplitudzie, zmienności oraz występowaniu charakterystycznych elementów.

W celach egzemplifikacyjnych odwołajmy się do wspomnianego już wielokrotnie elektrokardiogramu. Wzorcową krzywą EKG zobaczyć można na rys. 4. Jak widać jej przebieg z fizycznego punktu widzenia odzwierciedla zmiany wartości napięcia w czasie. Sygnał nie jest skomplikowany, dlatego w stosunkowo prosty sposób można wyznaczyć charakteryzujące go kluczowe parametry. Należą do nich m.in. amplitudy załamek (wychyleń



Rys. 4. Standardowa krzywa EKG, która ma postać przebiegu zmian napięcia w czasie, odzwierciedla aktywność elektryczną mięśnia sercowego.



Rys. 5. Złożony sygnał okresowy zakłócony szumem Gaussowskim (a). Widmo tego sygnału ujawnia, że składa się on z trzech składowych sinusoidalnych o częstotliwościach: 50, 250 i 500 Hz (b).

od poziomej linii podstawowej, czyli izoelektrycznej), czasy trwania poszczególnych odcinków (fragmentów krzywej między załamkami) czy odstępów (obejmujących załamek i sąsiedni odcinek).

Co ciekawe, podczas analizy kilkudziesięciosiekundowych zapisów standardowego badania EKG niektórzy lekarze wciąż posługują się jedynie wydrukami badania wykonanymi na papierze termoczułym z nadrukiem siatki milimetrowej. Wartości zmierzone z wykorzystaniem wspomnianej siatki lub linijki porównują następnie z wzorcowymi.

Tego typu analiza nie jest jednak możliwa w przypadku 24 bądź 48 godzinnych zapisów badania Holterowskiego, gdzie metody komputerowe okazują się wręcz niezbędne. Dlatego na rynku dostępne są programy, które wykonują automatyczną analizę badania, nie wymagając od lekarza wiedzy z zakresu statystyki czy cyfrowych metod przetwarzania sygnału.

Należy jednak mieć świadomość, że algorytmy diagnostyczne, aczkolwiek wciąż rozwijane, nie są jeszcze doskonale i pełne zaufaniem generowanym przez nie raportom może być ryzykowne.⁸ Przykładem może być komputerowa ocena zmienności rytmu zatokowego HRV (od ang. *Heart rate variability*), gdzie kluczową kwestią jest wiarygodne wyznaczenie kolejnych odstępów RR (odległości czasowych t_{RR} między szczytami załameków R, patrz rys.4).

Widoczne w zapisie holterowskim artefakty czy skurcze pozazatokowe mogą zostać bowiem błędnie zaklasyfikowane przez algorytm jako pobudzenia dominujące a tym samym uwzględnione w analizie. Doświadczony

diagnosta powinien zatem zawsze krytycznie ocenić automatycznie generowane wyniki, weryfikując np. skrajne wartości odstępów RR przy wykorzystaniu stosownych histogramów.

W kontekście analizy w dziedzinie czasu warto również nadmienić, że oprócz metod bazujących na prostej statystyce używa się też nieco bardziej skomplikowanych narzędzi matematycznych, np. funkcji konwolucji, autokorelacji czy korelacji wzajemnej. Zainteresowanym szczegółami polecam książki uwzględnione w bibliografii.

W przypadku słabych, złożonych sygnałów zakłóconych dodatkowo szumem typu Gaussowskiego (rys. 5a) przeprowadzenie analizy w dziedzinie czasu może być znacznie utrudnione a ponadto mało efektywne. Na szczęście w tego typu przypadkach z pomocą przychodzą metody przetwarzania sygnału w dziedzinie częstotliwości. Spróbkowany przebieg czasowy jest wówczas przekształcany, aby wydobyć z niego informacje o właściwościach periodycznych. Przykładowo na rys. 5b widać już wyraźnie, że w sygnale trzy częstotliwości dominują a pozostałe najprawdopodobniej charakteryzują szum.

W tym miejscu czytelnicy zapewne zaczną zastanawiać się, w jaki sposób realizowane jest przejście z dziedziny czasu w dziedzinę częstotliwości? Wstępem do wyjaśnienia może być przywołanie tzw. twierdzenia Fouriera. W uproszczeniu mówi ono, że dowolny, złożony sygnał okresowy można przedstawić w postaci szeregu składowych sinusoidalnych o określonych amplitudach, częstotliwościach i fazach.

Co ciekawe, niezwykle wagę tego twierdzenia dostrzegł już w 1867 r. Lord Kelvin, podkreślając, iż jest „nie tylko jednym z najpiękniejszych wyników współczesnej analizy, ale można o nim powiedzieć, że dostarcza niezastąpionego instrumentu przy rozważaniu niemal każdego abstrakcyjnego pytania w fizyce współczesnej”.⁹

Nie będziemy tu jednak przytaczać klasycznego przekształcenia Fouriera dotyczącego ciągłego sygnału analogowego. Interesują nas bowiem sygnały cyfrowe. Aby przenieść je z dziedziny czasu w dziedzinę częstotliwości, wyznacza się zatem tzw. dyskretną transformatę Fouriera $X[k]$ określoną wzorem:

$$X[k] = \sum_{n=0}^{N-1} x[n] \exp\left(\frac{-2i\pi kn}{N}\right)$$

gdzie: $x[n]$ - wartość próbki sygnału w dziedzinie czasu; N - liczba próbek, i - jednostka urojona ($i^2 = -1$), n - numer próbki, k - numer harmonicznej; $k = 0, 1, \dots, N - 1$.

Na podstawie transformaty Fouriera otrzymujemy widmo złożone z odpowiednich składowych częstotliwościowych. Teoretyczna rozdzielczość częstotliwościowa Δf takiej analizy to:

⁷ O podstawie działania pulsoksymetrów można przeczytać w artykule: T. Kubiak, Fizyka w karetce pogotowia, Fizyka w Szkole nr 6 (2014), s. 4-7.

⁸ Zainteresowanych problemami z automatyczną analizą EKG odsyłam do pracy: J. Schläpfer, H.J Wellens, Computer-Interpreted Electrocardiograms: Benefits and Limitations, J. Am. Coll. Cardiol. 70(9) 2017, 1183-1192.

⁹ Cytat przełożony z angielskiego na podstawie książki: R. Bracewell, Fourier Analysis and Imaging, New York, 2003.

$$\Delta f = \frac{f_p}{N} = \frac{1}{N\Delta t}$$

Już na pierwszy rzut oka widać zatem, że najlepszym sposobem na jej poprawę jest pobranie większej liczby próbek przy niezmienionej częstotliwości próbkowania f_p . Takie rozwiązanie przekłada się na wydłużony czas akwizycji sygnału i nie zawsze jest możliwe, chociażby ze względu na możliwość wystąpienia zmian w charakterystyce sygnału w związku z procesami zachodzącymi w żywym organizmie.

Z drugiej strony nasuwa się pomysł, by zachować liczbę próbek N , ale zmniejszyć częstotliwość próbkowania f_p . Trzeba jednak być ostrożnym i przypomnieć sobie o kryterium Shannona, które może wymusić ograniczenie pasma analizy, aby uniknąć zjawiska aliasingu.

Widać zatem, iż optymalizacja procesu cyfrowego przetwarzania sygnału wymaga od eksperymentatora pewnej wiedzy i doświadczenia. Nadmienimy jeszcze, że dzisiejsze obliczenia komputerowe opierają się o tzw. szybką transformację Fouriera (FFT - ang. *Fast Fourier Transform*). Wykorzystuje ona bardzo efektywne algorytmy, z zastrzeżeniem, że uwzględniana w obliczeniach liczba próbek jest całkowitą potęgą liczby 2. Jeśli warunek ten nie jest spełniony, możemy albo zmniejszyć liczbę próbek (co oczywiście wpłynie negatywnie na rozdzielczość analizy, o czym była już mowa wcześniej), albo na końcu przebiegu czasowego dodać określoną liczbę zer (*zero padding*).

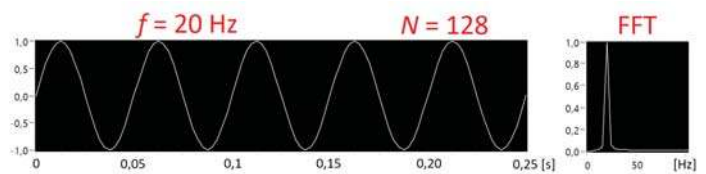
Przykładowe widmo uzyskane w wykorzystaniu metody FFT dla sygnału sinusoidalnego $N = 2^7 = 128$ próbek pokazano na rys.6. Co ciekawe, moce sygnału w reprezentacji czasowej i częstotliwościowej są sobie równe, o czym informuje nas twierdzenie Parsewala, przywołujące na myśl znaną każdemu z lekcji fizyki zasadę zachowania energii.

$$\sum_{n=0}^{N-1} |x[n]|^2 = \frac{1}{N} \sum_{k=0}^{N-1} |X[k]|^2$$

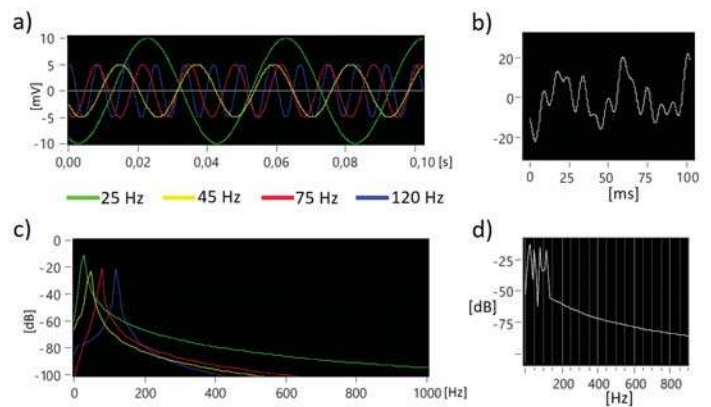
gdzie: $X[k]$ to dyskretna transformata Fouriera z $x[n]$.

Spójrzmy jeszcze na rys. 7, gdzie przedstawiono przykładowy, złożony sygnał okresowy, przebiegi sinusoidalne wchodzące w jego skład oraz obliczone widmo mocy. Niektórzy czytelnicy zastanawiają się zapewne, dlaczego piki odpowiadające poszczególnym częstotliwościom (rys. 7 c, d) są poszerzone? Spodziewali się bowiem ujrzeć wąskie linie, ściśle odpowiadające częstotliwościom występującym w analizowanym zapisie. Otóż byłoby tak, gdyby badany sygnał miał nieskończoną długość.

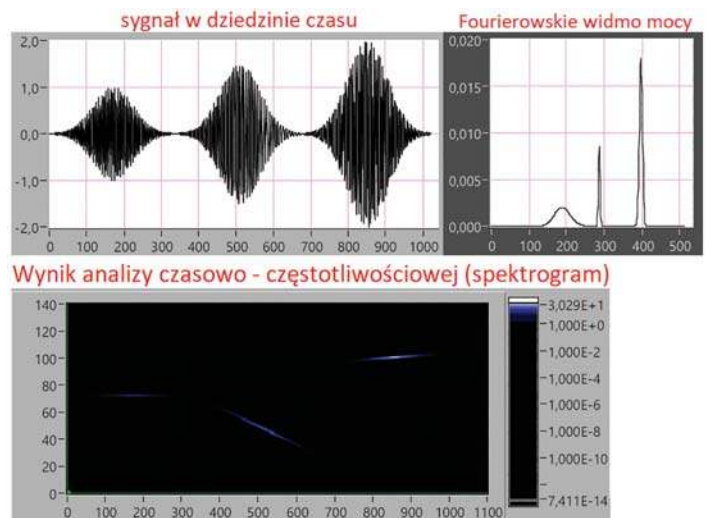
W praktyce jednak posługujemy się tzw. oknem pomiarowym, czyli z sygnału wycinamy fragment zawierający określoną liczbę cykli. Jeśli będzie ona niecałkowita, to dla okna prostokątnego zaobserwujemy spadek amplitudy widma i jego poszerzenie, czyli tzw. efekt przecieku. Jest on szczególnie uciążliwy w przypadku konieczności rozdzielania składowych częstotliwościowych położonych blisko siebie lub wyodrębnienia niskoamplitudowej skła-



Rys. 6. Sygnał sinusoidalny o częstotliwości 20 Hz i jego jednostronne widmo częstotliwościowe uzyskane z wykorzystaniem szybkiej transformacji Fouriera (FFT).



Rys. 7. Sinusoidalne składowe o częstotliwościach 25, 45, 75 i 120 Hz (a), tworzące złożony sygnał okresowy (b). Widmo mocy poszczególnych sygnałów składowych (c) oraz widmo całkowite sygnału złożonego (d).



Rys. 8. Wynik analizy złożonego sygnału w połączonej dziedzinie czasu-częstotliwości (metoda adaptacyjna).

dowej zlokalizowanej w bliskim sąsiedztwie wysokiego piku.¹⁰

Próbą zminimalizowania efektu przecieku jest zastosowanie okien czasowych o innym kształcie niż prostokątne. Mamy wiele ich typów, dlatego zainteresowani szczegółami powinni sięgnąć do polecanej literatury. Warto tylko wspomnieć, że nazwy okien nieprostokątnych biorą się zazwyczaj od nazwisk ich twórców, np. Juliusa von Hanna (okno Hanna/Hanninga), Richarda Hamminga (okno Hamminga), czy Maurice Stevensona Bartletta (okno Bartletta). Okna takie zawierają funkcje wagowe powodujące sprowadzenie sygnału do zera na jego końcach w celu zmniejszenia (wygładzenia) występujących tam nieciągłości. Oczywiście eksperymentator musi wie-

¹⁰ Od rozdzielczości częstotliwościowej analizy zależy, czy i w jakim stopniu piki są od siebie odseparowane.

dzieć, jakie okno najlepiej zastosować w danej sytuacji. Z reguły jednak te nieprostokątne są używane w przypadku długich sygnałów, w których analizie poddajemy tylko fragment zawierający niecałkowitą liczbę cykli.

Analizę widmową w medycynie wykorzystuje się m.in. do oceny wyników badań elektromiograficznych (EMG), elektroencefalograficznych (EEG) czy fonokardiograficznych (FKG). Potrzebne do tego jest jednak spore doświadczenie. Sygnały biomedyczne mogą wykazywać bowiem niestacjonarność, czego przykładem jest zapis częstości akcji serca, która ulega przyspieszeniu bądź spowolnieniu w zależności od aktywności pacjenta oraz jego stanu emocjonalnego. Widmo uśrednione nie uwzględni takich zmian, dlatego jednym z rozwiązań jest przeprowadzenie analizy w połączonej dziedzinie czasu i częstotliwości. Umożliwi to m.in. dokładne zbadanie wieloskładnikowych sygnałów, w których sukcesywnie pojawiają się różne częstotliwości.

Tego typu przebiegi można spróbować podzielić na stacjonarne fragmenty i dla każdego z nich wyznaczyć transformatę Fouriera a na jej podstawie widmo. Gdy otrzymane widma uszeregujemy kolejno wzdłuż osi czasu, zbudujemy tzw. spektrogram krótkoczasowej transformaty Fouriera (STFT - *ang. short-time Fourier transform*). Takie podejście ma jednak istotną wadę - nie można uzyskać jednocześnie dobrej rozdzielczości czasowej i częstotliwościowej (dłuższe okna wycinające poprawiają Δf , ale pogarszają umiejscowienie poszczególnych składowych w czasie).

Opracowano zatem kilka innych metod analizy sygnału w połączonej dziedzinie czasu i częstotliwości. Istotnym mankamentem większości z nich są jednak zaburzające spektrogram człony interferencyjne, generowane przez pary właściwych składowych częstotliwościowych. Trzeba zatem zdecydować, jaki rodzaj informacji jest najistotniejszy i na tej podstawie dobrać odpowiednią technikę. Przykładowo na rys. 8 zobrazowano wynik analizy sygnału metodą adaptacyjną, którą cechuje bardzo dobra jakość dekompozycji oraz brak członów interferencyjnych.

Dodajmy jeszcze, że w diagnostyce medycznej analiza czasowo-częstotliwościowa stosowana jest m.in. przy ocenie sygnałów otoemisji akustycznych¹¹ czy zapisów z dopplerowskich badań ultrasonograficznych¹² a także w badaniu desynchronizacji i synchronizacji rytmów EEG oraz w celu wykrycia późnych potencjałów komorowych w zapisie EKG.

Analiza obrazów

Mówi się, że obrazy bardzo dobrze przemawiają do wyobraźni, ale w medycynie są przede wszystkim źródłem niezwykle cennych informacji diagnostycznych. Współcześnie trudno nawet wyobrazić sobie możliwość prawidłowego rozpoznania wielu urazów czy schorzeń bez wykorzystania badań obrazowych. Powiedzmy śmia-

ło, że ich stworzenie to wielki triumf fizyki stosowanej. Przypomnijmy tylko, że wszystko zaczęło się od odkrycia promieni X przez Wilhelma Conrada Röntgena w 1895 r. Promienie te zostały później wykorzystane w tomografii komputerowej (CT od *ang. computed tomography*), której podstawy stworzyli Allan McLeod Cormack i Godfrey Newbold Hounsfield w latach 70-tych XX w.

Drugi wielki przełom to odkrycie przez dwa niezależne zespoły badawcze (kierowane przez Feliksa Blocha i Edwarda Purcella) zjawiska magnetycznego rezonansu jądrowego, co miało miejsce w 1945 roku. Dzięki temu w 1973 r. Paul Charles Lauterbur oraz Peter Mansfield mogli stworzyć jedną z najlepszych znanych obecnie technik diagnostycznych, czyli obrazowanie metodą rezonansu magnetycznego (MRI od *ang. Magnetic Resonance Imaging*). Wagę tych osiągnięć potwierdza fakt, że wszyscy wzmiankowani powyżej badacze zostali uhonorowani nagrodami Nobla.

Oczywiście we współczesnej medycynie, oprócz tomografii rentgenowskiej i magnetycznorezonansowej stosuje się jeszcze wiele innych rodzajów badań obrazowych, m.in. ultrasonografię, pozytonową tomografię emisyjną (PET), tomografię emisyjną pojedynczych fotonów (SPECT), czy termografię. Warto spojrzeć na rys. 9, gdzie przedstawione są przykładowe obrazy anatomiczne uzyskane z wykorzystaniem różnych technik. Można zaryzykować twierdzenie, że nawet laik dostrzeże ogromny potencjał diagnostyczny, jaki w sobie zawierają. Zadaniem ekspertów, np. radiologów, jest przede wszystkim poprawna interpretacja uzyskanych obrazów, ale również wydobycie z nich jak największej ilości informacji przydatnych do oceny stanu pacjenta.

Zastosowanie komputerowych metod przetwarzania obrazów pozwala poprawić ich jakość, uwidocznic wybrane elementy czy stworzyć prezentacje trójwymiarowe na podstawie dwuwymiarowych przekrojów anatomicznych. Ułatwia też przeprowadzenie dokładnej analizy zarejestrowanych obrazów a także porównanie ich z innymi, które zgromadzono w specjalistycznych bazach danych.

Rozwój teleradiologii sprawił, że badania wykonywane w poszczególnych pracowniach, z reguły nie są opisywane na miejscu, tylko trafiają do wyspecjalizowanych centrów opisowych zlokalizowanych niekiedy setki kilometrów dalej. Komunikację pomiędzy aparaturą medyczną, stacjami diagnostycznymi i serwerami zdecydowanie ułatwia wspólny, ujednoczony standard zapisu, transferu, przechowywania oraz prezentacji danych pozyskiwanych w badaniach obrazowych, jakim jest DICOM (*ang. Digital Imaging and Communications in Medicine*). Należy zwrócić uwagę, że nie mówimy tu o plikach graficznych, tylko wieloatrybutowych obiektach zawierających również dane pacjenta, nastawy urządzenia diagnostycznego, sposób prezentacji pojedynczych obrazów lub ich sekwencji itp.¹³

¹¹ Otoemisja akustyczna to badanie pozwalające wykryć niedosłuch odbiorczy pochodzenia ślimakowego.

¹² O badaniach dopplerowskich można przeczytać w artykule: T. Kubiak, Ultrasonografia, czyli fale akustyczne w służbie medycynie, Fizyka w Szkole, nr 5 (2017), s. 4-9.

¹³ Szczegółów można dowiedzieć się na stronie: <https://www.dicomstandard.org/> (dostęp z 22.03.2020).

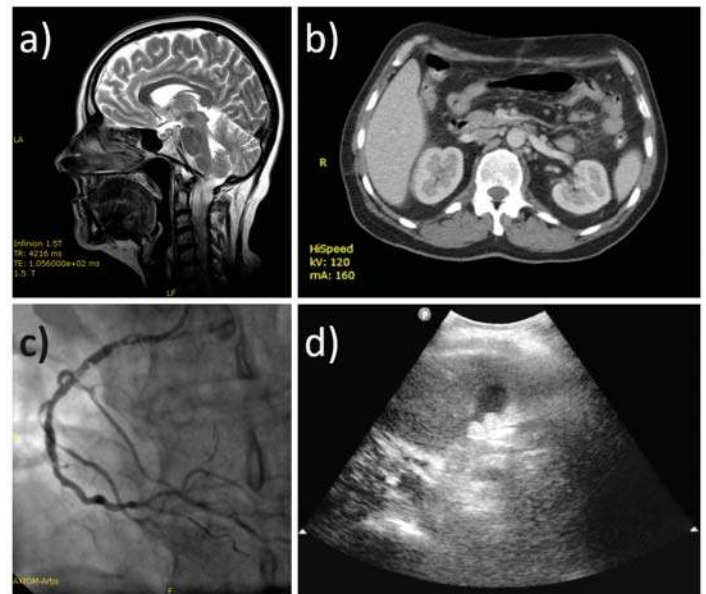
Analogowe monochromatyczne obrazy medyczne mogą być reprezentowane przez dwuwymiarową funkcję. Jej argumentami są współrzędne x, y określające położenie poszczególnych punktów, natomiast zbiór wartości stanowią odpowiadające im poziomy jasności. Cyfrowy obraz rastrowy składa się natomiast ze skończonej liczby pikseli (nazwa od ang. *picture element*), tworzących regularną, np. prostokątną siatkę. Reprezentuje go zatem macierz, której wymiar (liczba wierszy i kolumn) definiuje rozdzielczość obrazu w pionie i w poziomie. Elementy tej macierzy to skwantowane wartości intensywności. Standardowo liczba dyskretnych wartości, reprezentujących jasność poszczególnych punktów jest potęgą dwójki. Przykładowo dla obrazów 8 bitowych mamy $2^8 = 256$ poziomów szarości. Dla zainteresowanych wspomnijmy również, iż obrazy w systemie RGB¹⁴ mogą być definiowane przez trzy macierze dwuwymiarowe, z których każda reprezentuje inny kanał barwny bądź poprzez macierz trójwymiarową. Jeśli posłużymy się 24-bitowym zapisem kolorów, to na każdą z barw składowych przypadnie wówczas 8 bitów, zatem przyjmie ona wartość z zakresu 0-255.

W niniejszym artykule nie będą omawiane poszczególne techniki akwizycji obrazów cyfrowych wykorzystywane w różnych urządzeniach diagnostycznych. Wypada tylko wspomnieć, że w procesach próbkowania i kwantowania obrazów trzeba mieć na uwadze efekty analogiczne do tych, obserwowanych w przypadku digitalizacji sygnałów, czyli np. możliwość wystąpienia aliasingu. Zarejestrowany obraz poddaje się wstępnemu przetwarzaniu, wykonując filtrację, poprawę kontrastu wyostrenie czy redukcję kolorów.

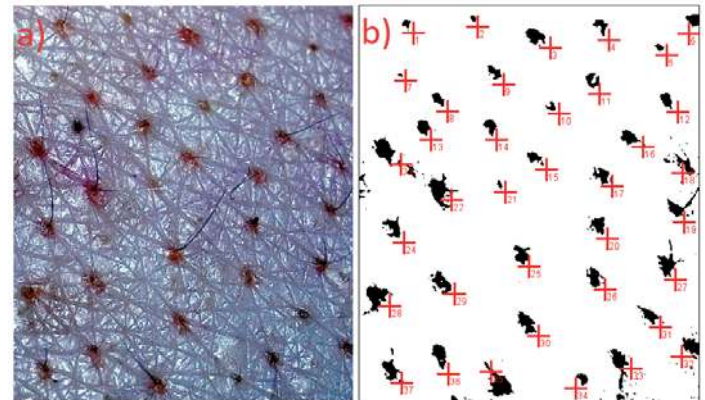
Zabiegi te mają na celu pozbycie się zakłóceń, poprawę jakości obrazu, uwypuklenie pewnych szczegółów, czyli przygotowanie do właściwej analizy. Korektę tonalną fotografii mikroskopowych ułatwia wyznaczenie ich histogramów, czyli wykresów przedstawiających liczbę pikseli posiadających określony poziom jasności (w przypadku zdjęć RGB histogramy są wyświetlane osobno dla poszczególnych kanałów). Procedura wyrównania histogramu pozwala poprawić globalny kontrast obrazu, co często przekłada się na możliwość wydobywania z niego większej ilości informacji. Właściwie przygotowane obrazy poddaje się już ocenie szczegółowej.

Dedykowane oprogramowanie, np. popularny darmowy program ImageJ, stworzony przez naukowców z Narodowego Instytutu Zdrowia USA, umożliwia wykonywanie pomiarów bezpośrednio na obrazie z uwzględnieniem ustawionej uprzednio skali (kalibrację wykonuje się na podstawie wzorca o znanych wymiarach). Oczywiście podczas procedur metrologicznych trzeba mieć na uwadze błędy dyskretyzacji i posługiwać się obrazem o jak najlepszej rozdzielczości przestrzennej.

W wielu przypadkach, np. podczas oceny zdjęć preparatów mikroskopowych, bardzo przydatną procedurą jest



Rys. 9. Obrazy anatomiczne uzyskane w badaniu: (a) MRI; (b) CT; (c) koronarograficznym; (d) ultrasonograficznym.



Rys. 10. Fotografia skóry ludzkiej (a) wraz z wynikiem segmentacji przez progowanie (b).

segmentacja, czyli podział obrazu na rozłączne obszary o interesujących badacza właściwościach oraz nieistotne tło. Można wówczas z większą łatwością wyznaczyć zarówno parametry lokalne, np. rozmiar pojedynczej komórki jak i globalne, np. liczbę struktur znajdujących się w obszarze zainteresowania (tzw. ROI - ang. *region of interest*).

Segmentację obrazu monochromatycznego wykonuje się na wiele sposobów, a jednym z najpopularniejszych jest tzw. progowanie. Jak sama nazwa wskazuje sprawą kluczową jest tu ustalenie progu, czyli wartości intensywności, która rozgraniczy nam poszukiwane obiekty od tła (można ustalić też podwójny próg definiujący zakres poziomów jasności, charakteryzujących interesujące nas regiony). Próg ten jest określany ręcznie bądź automatycznie na podstawie histogramu obrazu bazowego. Piksele w zależności od swojej jasności są następnie zaliczane do jednej z dwóch kategorii. W rezultacie obraz monochromatyczny zamieniany jest na binarny, posiada-

¹⁴ O widzeniu barwnym można przeczytać w artykule: T. Kubiak, Od biofizyki układu wzrokowego do złudzeń optycznych, Fizyka w Szkole z Astronomią, nr 6 (2019), s. 4-10.

jący tylko dwa kolory: czarny (oznaczone objekty) i biały (tło).

Warto spojrzeć na rys. 10, gdzie przedstawiono makrofotografię skóry ludzkiej wraz z wynikiem przeprowadzonej segmentacji przez progowanie. Wyodrębnienie ze zdjęcia poszczególnych porów pozwala na ich łatwe policzenie. Dla zadanego ROI program komputerowy może zrobić to automatycznie, uwzględniając zawarte wewnątrz objekty i przyjmując poprawkę na te, które zostały obcięte przez brzeg zaznaczenia (wklęsła struktura przecięta przez ramkę mogłaby być zliczana wielokrotnie).

Warto jeszcze wspomnieć o istnieniu innych technik segmentacji. Przykładowo spełniające zadany warunek piksele mogą być dołączane do grupy (rozrost obszaru), jeśli znajdują się w bezpośrednim sąsiedztwie innych posiadających tę samą własność.

Można również, stosując określone kryterium, rekurencyjnie dzielić cały obraz na coraz mniejsze regiony, tak, aby były w przybliżeniu jednolite, a następnie łączyć je, jeśli są przyległe i wykazują podobieństwo (podział i łączenie). Inną strategię wykorzystuje tzw. metoda działów wodnych, która opiera się na poszukiwaniu lokalnych minimów jasności. Piksele zaliczane są do jednego obszaru („zlewiska”), gdy wyliczona droga wiedzie je do wspólnego minimum (analogia do wody, spływającej po zboczach gór w określone miejsce). Z drugiej strony, zamiast definiować całe obszary, można skupić się na wykrywaniu krawędzi, czyli miejsc gwałtownych zmian jasności.

W zależności od zawartości obrazu i celu, jaki przyświeca diagnoście, musi on zdecydować o wyborze najbardziej optymalnej metody segmentacji. Trzeba jednak mieć na uwadze, że wraz z rozwojem techniki wszystkie procedury będą na pewno coraz bardziej zautomatyzowane. Przykładowo zdalne wykonanie obrysu zmiany nowotworowej na kolejnych etapach leczenia powinno zapewnić lepszą powtarzalność analizy niż każdorazowa ręczna segmentacja a tym samym dostarczyć bardziej miarodajnej informacji o postępach terapii.

Przyszłością są algorytmy bazujące na sztucznej inteligencji oraz systemy uczące się w oparciu o duże ilości danych testowych zgromadzonych w dedykowanych bazach medycznych. Będą one automatycznie wyszukiwać określone objekty na obrazach, klasyfikować je i identyfikować ich kluczowe własności. Naukowcy zwracają jednak uwagę na konieczność rozróżnienia dwóch pojęć: komputerowego wspomaganie diagnostyki obrazowej (CAD od ang. *computer-aided diagnosis*) oraz komputerowej diagnozy (ACD od ang. *automated computer diagnosis*), dodając, że na razie metody komputerowe mają za zadanie wspierać radiologów, a nie autonomicznie podejmować za nich decyzje.

Obrazy uzyskiwane w procedurach diagnostycznych, np. rentgenografii, ultrasonografii czy MRI, są monochromatyczne. Można jednak nadać im umowne barwy, aby uwidocznic pewne cechy. Przykładowo podczas

wizualizacji wyników badania echokardiograficznego różnym szybkościom oraz kierunkom przepływu krwi zmierzonym metodą dopplerowską przyporządkowuje się odmienne kolory, co poprawia czytelność i ułatwia postawienie diagnozy.

Zastosowanie barw umownych przydaje się również podczas wspólnej prezentacji danych pozyskanych dwóch różnych systemów akwizycji, np. pozytonowej tomografii emisyjnej i tomografii rentgenowskiej. Takie złożone obrazy multimodalne dostarczają wiedzy zarówno o budowie organizmu jak i procesach metabolicznych zachodzących w wybranych regionach ciała. Informacje anatomiczne przedstawiane są wówczas w skali szarości, natomiast funkcjonalne właśnie za pomocą kolorów, co pozwala na ich wizualne oddzielenie.

Dodatkowo z ciągów kolejno zarejestrowanych obrazów można tworzyć filmy, ukazujące działanie określonego narządu (np. pracę mięśnia sercowego) bądź przepływ środka kontrastującego przez naczynie krwionośne.

Innym ułatwieniem jest opcja tworzenia wirtualnych rekonstrukcji trójwymiarowych. Ma to szczególne znaczenie podczas wizualizacji obiektów słabo widocznych na obrazach dwuwymiarowych. Przykładowo możemy próbować wyobrazić sobie przebieg tętnicy na podstawie zestawu jej kolejnych tomograficznych przekrojów poprzecznych, ale zdecydowanie pewniejszym rozwiązaniem jest ukazanie jej w prezentacji 3D.

Dogłębna analiza sygnałów i obrazów biomedycznych pozwala ocenić stan zdrowia pacjenta a w rezultacie ułatwia postawianie właściwej diagnozy. Zrozumienie natury sygnałów oraz mechanizmów ich generowania stwarza możliwość zgłębiania tajników funkcjonowania ludzkiego organizmu, rozwija naszą wiedzę a przede wszystkim jest wspaniałą przygodą dla miłośników fizyki i nauk przyrodniczych.

dr Tomasz Kubiak

**Państwowa Wyższa Szkoła Zawodowa im. Hipolita Cegielskiego
w Gnieźnie
oraz Poznański Park Naukowo-Technologiczny Fundacji UAM**

LITERATURA

- [1] J. Moczko, L. Kramer, Cyfrowe metody przetwarzania sygnałów biomedycznych, Wydawnictwo Naukowe UAM, Poznań 2001.
- [2] A. M. Bianchi, L. T. Mainardi, S. Cerutti, Time frequency analysis of biomedical signals, Transactions of the Institute of Measurement and Control 22 (2000), 215 - 230.
- [3] J. Szabatin, Przetwarzanie sygnałów, Politechnika Warszawska 2003.
- [4] R. J. Rak, A. Majkowski, Czasowo-częstotliwościowa analiza sygnałów, Przegląd Elektrotechniczny 5 (2004), 515-520.
- [5] J. Woźnicki, Podstawowe techniki przetwarzania obrazu, Wydawnictwa Komunikacji i Łączności, Warszawa, 1996.
- [6] P. Horowitz, W. Hill, Sztuka elektroniki, t2, WKŁ, Warszawa 2006, s. 161-190; 570-629.
- [7] A. Materka, P. Strumillo, Wstęp do komputerowej analizy obrazów, Politechnika Łódzka, Łódź 2009.
- [8] W. Bieniecki, S. Grabowski, Wybrane zagadnienia przetwarzania i analizy obrazów mikroskopowych w diagnostyce medycznej. Zeszyt Naukowy X lat KIS, Politechnika Łódzka, Łódź 2005.
- [9] B. Halallı, A. Makandar, Computer Aided Diagnosis - Medical Image Analysis Techniques, Intechopen 2018, 85-109.
- [10] K. Doi, Current status and future potential of computer-aided diagnosis in medical imaging, British Journal of Radiology, 78 (2005), S3 - S19.

100 LAT
1920-2020



POLSKIEGO
TOWARZYSTWA
FIZYCZNEGO



2020 – Rok Fizyki

Diagramy Aleksandra Jabłońskiego

Kazimierz Mikulski

Rok 2020 jest szczególnym rokiem w nauce polskiej. W Monitorze Polskim jako Dzienniku Urzędowym Rzeczypospolitej Polskiej w dniu 4 listopada 2019 r. poz. 1043, umieszczono uchwaloną na ostatnim posiedzeniu Senatu IX kadencji – 8 października 2019 r., Uchwałę Senatu Rzeczypospolitej Polskiej z dnia 18 października 2019 r. ustanawiającą rok 2020 Rokiem Fizyki.^{1, 2}

W uchwale zapisano, że Senat Rzeczypospolitej Polskiej ustanawia rok 2020 Rokiem Fizyki, w 100-lecie powstania Polskiego Towarzystwa Fizycznego. Czytamy także: „*Jednocześnie Senat pragnie w ten sposób podkreślić zasługi polskich fizyków dla światowej nauki oraz przypomnieć niezwykły rozwój fizyki polskiej, jaki nastąpił po odzyskaniu przez Polskę niepodległości w roku 1918*”.³ Wskazano też na cel jakim jest „*uhonorowanie polskich fizyków i ich osiągnięć, a także docenienie cywilizacyjnego znaczenia tej dziedziny nauki*”.⁴

Jak wskazano w uchwale, mimo kolosalnych zniszczeń osobowych i materialnych podczas II wojny światowej i okupacji, fizyka w Polsce szybko się odrodziła, w efekcie: „*W 1952 r. Marian Danysz i Jerzy Pniewski z Zakładu Fizyki Doświadczalnej Uniwersytetu Warszawskiego odkryli nowy rodzaj materii – hiperjądra, zawierające obok protonów i neutronów także hiperony. Powstała w ten sposób nowa gałąź fizyki jądrowej – fizyka hiperjądrowa. W 1962 r. w Warszawie odkryto także pierwsze hiperjądro podwójne, zawierające dwa hiperony lambda. Za to wielkie osiągnięcie Danysz i Pniewski byli kilkakrotnie nominowani do Nagrody Nobla z fizyki*” – napisano w treści uchwały.⁵

W dalszej części tejże uchwały czytamy: „*Chcemy, aby Rok Fizyki przyniósł jeszcze większe uznanie w naszym społeczeństwie dla roli nauki, w szczególności fizyki oraz jej wkładu w rozwój gospodarki i wpływu na nasze życie codzienne. Zainteresowanie fizyką, niezwykle ciekawą nauką eksperymentalną, przekłada się na wzrost umiejętności w tym zakresie kolejnych pokoleń. Podejmowana działalność naukowa i zawodowa w tym obszarze jest szczególnie ważna w czasie budowy społeczeństwa informacyjnego i gospodarki opartej na wiedzy*”. Tak głosi uchwała.⁶

Na stronie internetowej Polskiego Towarzystwa Fizycznego zawarto informację <100 lat Fizyki w Polsce>. Na tle dat i miejsc zjazdów PTF umieszczono najważniejsze wydarzenia w Polsce oraz jedno z ważniejszych faktów w rozwoju fizyki w XX i XXI wieku.⁷

Nazwiska polskich fizyków okresu dwudziestolecia międzywojennego, którzy osiągnęli naukowe wyniki wielkiej wagi, tworzą długą listę. Na jej czele niejednokrotnie wymieniani są:

- Wojciech Rubinowicz (1889-1974) – uzupełnił i rozwinął teorię dyfrakcji światła – obecnie powszechnie jest ona znana jako teoria Younga-Rubinowicza. Szczegóły znajdzie czytelnik na opracowaniu umieszczonym w Internecie pod adresem <http://rubinowicz.canonia.pl/>.
- Aleksander Jabłoński podał schemat poziomów energii, który jest obecnie podstawowym narzędziem badania cząsteczek (powszechnie znany dziś jako diagram Jabłońskiego).
- Marian Mięśowicz (1907-1995) – odkrył anizotropię lepkości ciekłych kryształów i wyznaczył odpowiednie współczynniki, znane dziś pod nazwą współczynników Mięśowicza. Patrz także na treść strony o adresie internetowym https://pl.wikipedia.org/wiki/Marian_Mięśowicz
- Mieczysław Wolfke (1883-1947), który wraz z Holendrem Willemem Keesomem odkrył hel-2, pierwszą ciecz kwantową o niezwykłych właściwościach. Polski fizyk, prekursor telewizji i holografii, odkrywca metody zestalenia helu oraz dwóch odmian ciekłego helu. Wielki Mistrz Wielkiej Łoży Narodowej Polski w latach 1931–1934. Więcej pod adresem https://pl.wikipedia.org/wiki/Mieczysław_Wolfke
- Henryk Niewodniczański (1900-1968) i Jan Blaton (1907-1948) odkryli w 1933 r. bardzo rzadki rodzaj promieniowania atomów: dipolowe promieniowanie magnetyczne. Więcej na stronie internetowej pod adresem https://pl.wikipedia.org/wiki/Henryk_Niewodniczański oraz https://pl.wikipedia.org/wiki/Jan_Blaton.

¹ <http://prawo.sejm.gov.pl/isap.nsf/download.xsp/WMP20190001043/O/M20191043.pdf>

² <http://www.ptf.net.pl/>

³ Treść http://g.ekspert.infor.pl/pl/_dane/akty_pdf/MPO/2019/194/1043.pdf#zoom=90

⁴ <http://naukawpolsce.pap.pl/aktualnosci/news%2C79083%2Csenat-ustanowil-rok-2020-rokiem-fizyki.html?fbclid=IwAR35dky3yX-g-QtViuSSH4z4cgbOZuKclEgd09E6RwIPg-VT2S3kSqQpY>

⁵ <http://naukawpolsce.pap.pl/aktualnosci/news%2C79083%2Csenat-ustanowil-rok-2020-rokiem-fizyki.html?fbclid=IwAR35dky3yX-g-QtViuSSH4z4cgbOZuKclEgd09E6RwIPg-VT2S3kSqQpY>

⁶ Ibidem

⁷ <https://100lat.ptf.net.pl/?page=zjazd-100-lat>

- Jan Czochralski (1885-1953) – ojciec epoki krzemowej. Bez niego i jego metody nie byłoby współczesnej elektroniki półprzewodnikowej, technologii i techniki cyfrowej. Codziennie na świecie otrzymuje się kilkadziesiąt ton monokryształów hodowanych jego metodą. Jest autorem szeroko cytowanych publikacji oraz wielu patentów.⁸ Przez wielu uważany za „ojca elektroniki”.⁹

W wielu publikacjach naukowych podkreśla się, że polscy badacze w fizyce cząstek elementarnych i wysokich energii byli współautorami ważnych odkryć, takich jak: oscylacji neutrin oraz bozonu Higgsa. Wiadomo także, że największy światowy kongres odbył się w 1996 roku w Warszawie. Tym samym podkreślono wkład polskiej nauki w dokonania w tej dziedzinie fizyki.

Niejednokrotnie wskazuje się na osiągnięcia polskich fizyków-teoretyków, którzy uzyskali w swoich opracowaniach naukowych wyniki na światowym poziomie. Między innymi Jan Weysenhoff z Uniwersytetu Jagiellońskiego wprowadził do fizyki pojęcie cieczy spinowej. Natomiast uczeń Leopolda Infelda, prof. Andrzej Trautmann z Uniwersytetu Warszawskiego podał przekonujący dowód istnienia fal grawitacyjnych. Te wskazania w znaczący sposób przyczyniły się do wzmożenia badań w tej dziedzinie i w konsekwencji odkryto fale (grawitacyjne) w dniu 14 września 2015 r., przy współudziale polskich eksperymentatorów i teoretyków, a 3 października 2017 r. Rainer Weiss, Barry C. Barish i Kip S. Thorne otrzymali Nagrodę Nobla z dziedziny fizyki „za decydujący wkład w stworzenie detektora LIGO¹⁰ i obserwację fal grawitacyjnych”^{11, 12}. Zwrócono uwagę, że w ostatnich latach polscy fizycy znaleźli się trzykrotnie wśród laureatów niedawno ustanowionej prestiżowej Nagrody „Breakthrough Prize”: za badania neutrin, za odkrycie bozonu Higgsa i za odkrycie fal grawitacyjnych.

Na XXXV Zjeździe Fizyków Polskich w Białymstoku prof. Andrzej Kajetan Wróblewski (Instytut Fizyki Doświadczalnej UW) przedstawił referat nt. „Fizyka w Polsce wczoraj, dziś i jutro”, w którym zaprezentował krótką charakterystykę fizyki w Polsce XX w. Przedstawił wskaźniki bibliometryczne, liczby dotyczące stopni i tytułów naukowych oraz liczby studentów fizyki. W podsumowaniu ocenił wkład Polaków do światowej fizyki XX w., wyodrębniając przede wszystkim – poza Marią Skłodowską-Curie i czterech fizyków, którzy dokonali



Fotografia 1. Profesor Aleksander Jabłoński (1898–1980) i karykatura z 1978 r. autorstwa matematyka toruńskiego prof. dr Leona Jeśmianowicza (1904-1989)
Źródło: https://upload.wikimedia.org/wikipedia/commons/cc/Aleksander_Jab%C5%82o%C5%84ski.jpg
http://www.faj.org.pl/files/home_faj/A_Jablonski_biografia.pdf
<https://www.ifiz.umk.pl/nauka/wyklady-im-aleksandra-jablonskiego/aleksander-jablonski-1898-1980/>
https://upload.wikimedia.org/wikipedia/commons/1/1c/Diagram_jablonskiego.svg
https://pl.wikipedia.org/wiki/Aleksander_Jab%C5%82o%C5%84ski

odkryć na miarę Nagrody Nobla: Mariana Smoluchowskiego¹³, Mariana Danysza, Jerzego Pniewskiego i Karola Olszewskiego.

Wśród osiemnastu mniej zasłużonych, którzy jednak wnieśli bardzo poważny wkład do rozwoju fizyki i w pewnym okresie należeli do liderów światowej fizyki (których nazwiska są wymieniane w syntetycznych obcojęzycznych historycznych opracowaniach) znalazł się prof. Aleksander Jabłoński.^{14, 15}

W 1938 roku w „Nature” ukazała się praca A. Jabłońskiego, przedstawiająca wyniki pomiarów zależności szerokości linii widmowych od temperatury, wykonanych wspólnie z Henrykiem Horodniczym. Ogólną teorię ciśnieniowego poszerzenia linii widmowych opisał w latach 1944-1945 i opublikował w 1945 roku w „Physical Review”.

Jest uważany za twórcę „toruńskiej szkoły fizyki”. Był promotorem doktoratów m.in. prof. Kazimierza Antonowicza (1914-2003) i prof. Stanisława Łęgowskiego (1931-2015). W jego zespole pracowali także Marian Kryszewski, Wanda Hanusowa, Stanisław Dembiński, a także jego córka Danuta Jabłońska-Frąckowiak i jej mąż Mieczysław Frąckowiak.

⁸ Polacy w fizyce współczesnej – Justyna Chojnacka http://dydaktyka.fizyka.umk.pl/FWs_fragmenty/Polacy.pdf

⁹ Szanowany jako wynalazca, zapomniany jako człowiek. Kim był Jan Czochralski? Co pozostawił w nauce? „Jan Czochralski i jego rok 2013”. „Fizyka w Szkole” nr 4, lipiec/sierpień 2013, s. 4-7.

¹⁰ LIGO (Laser Interferometer Gravitational Wave Observatory) – detektor fal grawitacyjnych, bazujący na zasadzie interferometru Michelsona. LIGO jest wspólnym przedsięwzięciem naukowców z MIT, Caltech i wielu innych szkół wyższych. Więcej na stronie o adresie <https://pl.wikipedia.org/wiki/LIGO>

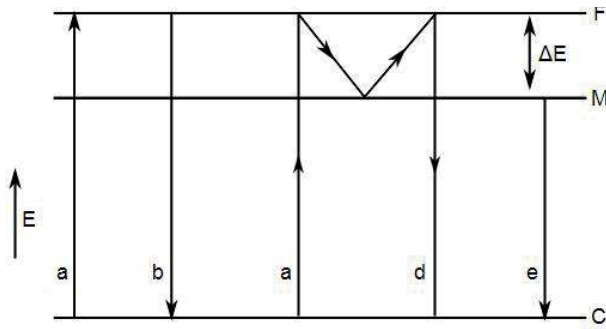
¹¹ *The 2017 Nobel Prize in Physics – Press Release, Nobel Prize*

¹² <https://www.nobelprize.org/prizes/physics/2017/press-release/>

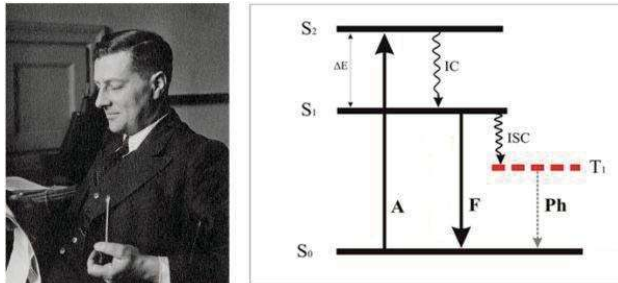
¹³ „Rok 2017 rokiem Mariana Smoluchowskiego w 100-lecie śmierci, w 190-lecie odkrycia ruchów Browna”, „Fizyka w Szkole z Astronomią” nr 3, maj/czerwiec 2017, s.4-8

¹⁴ https://pl.wikipedia.org/wiki/Aleksander_Jab%C5%82o%C5%84ski Andrzej Kajetan Wróblewski *Fizyka w Polsce wczoraj, dziś i jutro*, <http://labfiz.uwb.edu.pl/ptf/echa/html/akw.html>

¹⁵ http://www.faj.org.pl/files/home_faj/A_Jablonski_biografia.pdf



Rysunek 1. Pierwsze Diagramy Jabłońskiego z 1933 r.



Rysunek 2. Uproszczony schemat poziomów energetycznych Jabłońskiego przedstawiający wzbudzone stany elektronowe o różnych wartościach całkowitego spinu: 0 – stan singletowy (S) i 1 – stan trypletowy (T). Przejścia między stanami S a T są zabronione. Strzałki proste ilustrują procesy promieniste, a strzałki faliste procesy bezpromieniste. Literami oznaczono poszczególne procesy fizyczne. A – Absorpcja, F – Fluorescencja (fotoluminescencja natychmiastowa), Ph – Fosforescencja (fotoluminescencja opóźniona), IC – konwersja wewnętrzna, ISC – konwersja interkombinacyjna (międzysystemowa).

Na zdjęciu na lewo – profesor Jabłoński z nieodłączną fajką.
 Źródło: na podstawie opracowania prof. dr hab. Józef Szudy pt. ALEKSANDER JABŁOŃSKI (1898–1980) TWÓRCA TORUŃSKIEJ SZKOŁY FIZYKI
http://www.faj.org.pl/files/home_faj/A_Jablonski_biografia.pdf
http://dydaktyka.fizyka.umk.pl/Pokazy_2012/33.html

Diagramy Jabłońskiego

W tym miejscu warto przypomnieć diagramy Jabłońskiego, ewaluowanie których następowało na przełomie kilku dziesięciu lat. Rysunek 1 przedstawia pierwszy diagram.

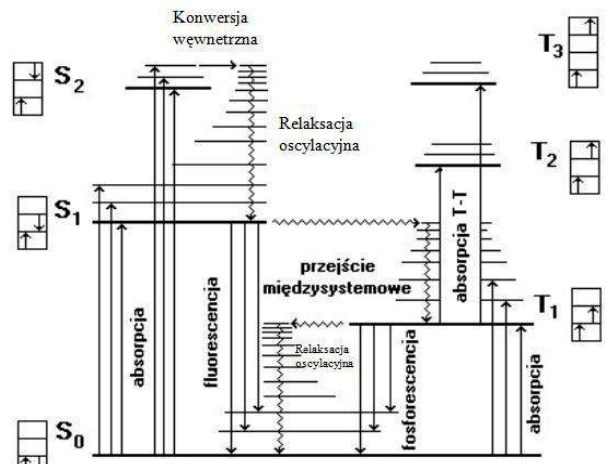
Natomiast poniżej wskazuję na kolejne, przedstawiające wzbudzone stany elektronowe o różnych wartościach całkowitego spinu. Po raz pierwszy termin ten (the Jabłoński diagram) wprowadził do literatury w 1941 roku wybitny amerykański fizykochemik Gilbert N. Lewis i jego współpracownicy D. Lipkin i T. T. Magel w ich słynnej pracy [1], w której w oparciu o ten diagram przeprowadzili analizę wyników swoich doświadczeń dotyczących fluorescencji i fosforescencji cząsteczek organicznych¹⁶.

Obok (rys. 3) kolejna wersja diagramów Jabłońskiego, umieszczana w literaturze przedmiotu.

W ciekawym opracowaniu pt. „Spektroskopia absorpcyjna i emisyjna cząsteczek organicznych” zapisano: „Każda cząsteczka ma charakterystyczny dla siebie układ poziomów energetycznych – elektronowych, oscylacyjnych i rotacyjnych, przy czym tych ostatnich nie obserwuje się w cieczach. W wyniku absorpcji promieniowania ultrafioletowego lub widzialnego cząsteczka przechodzi do jednego ze stanów wzbudzonych, a następnie na różne sposoby wytraca nadmiar energii. Procesy pochłaniania i deekscytacji przedstawia się często schematycznie na tzw. diagramie Jabłońskiego”.

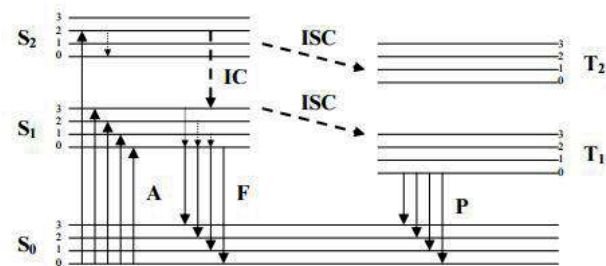
W związku z Rokiem Fizyki warto byłoby zaprezentować „wspomnienia lub doświadczenia” nauczycieli fizyki. A może ogłosić konkurs w tym obszarze zainteresowań.

**dr Kazimierz Mikulski
 Maksymilianowo**

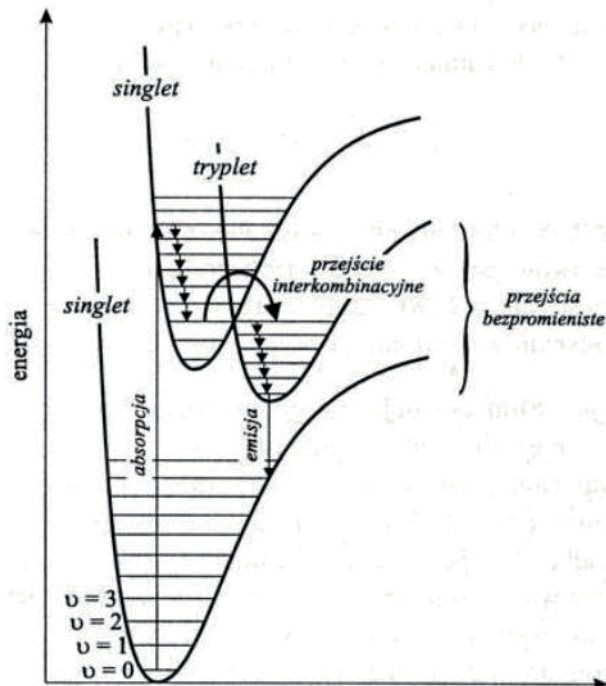


Rysunek 3. Współczesny Diagram Jabłońskiego. Jeżeli cząsteczka znajdującej się w elektronowym stanie podstawowym dostarczymy odpowiednią energię to cząsteczka ta ulegnie wzbudzeniu do wyższego stanu elektronowego opisanego inną krzywą energii potencjalnej. Źródło: Absorpcja światła Daniel T. Gryko Instytut Chemii Organicznej PAN

Omówienie na stronie WWW pod adresem <https://notatek.pl/diagram-jablonskiego-omowienie>



Rysunek 4. Diagram Jabłońskiego. Dla przypomnienia i utrwalenia oznaczenia są następujące: A – absorpcja, F – fluorescencja, P – fosforescencja, IC – konwersja wewnętrzna, ISC – przejścia międzysystemowe (interkombinacyjne), S₀, S₁, S₂, stany elektronowe singletowe podstawowy i dwa wzbudzone, T₁, T₂, stany wzbudzone trypletowe, liczby 0,1,2,3 numerują poziomy oscylacyjne, strzałki z liniami ciągłymi oznaczają przejścia promieniste, strzałki z liniami przerywanymi - przejścia bezpromieniste. Dla przejrzystości rysunku poziomy rotacyjne pominięto. Źródło: Spektroskopia absorpcyjna i emisyjna cząsteczek organicznych <https://www.fuw.edu.pl/IIPRACOWNIA/home/Opisy-cwiczen/SF5.pdf>



Rysunek 5. Diagramy Jabłońskiego i wyjaśnienie np. Mechanizm fosforescencji T₁→S₀

Źródło: więcej szczegółów - https://chemia.uwb.edu.pl/media/uploads/2016/01/19/w11_b.pdf



Sztuczna grawitacja w układzie obracającym się

Część I – spadek swobodny bez prędkości początkowej

Jan Kurzyk

Określenie *sztuczna grawitacja* w odniesieniu do układów nieinercjalnych brzmi lekceważąco. Przymiotnik „sztuczna” sugeruje, że mamy do czynienia z czymś niepełnowartościowym, substytutem czegoś lepszego, prawdziwego.

Przychodzi nam na myśl sztuczny miód lub wyrób czekoladopodobny (pojęcie znane osobom, które pamiętają „czasy słusznie minione”). Tymczasem grawitacja jakiej doświadczamy w układach nieinercjalnych nie jest gorsza od tej „naturalnej”. Zauważył to Albert Einstein, który swoją teorię grawitacji oparł na postulatcie równoważności, w którym stwierdza, że lokalnie nie da się odróżnić spadku swobodnego w polu grawitacyjnym od ruchu swobodnego w układzie nieinercjalnym.

Oba typy sił – siłę grawitacyjną i siłę bezwładności – cechuje ta sama własność: ruch pod wpływem każdej z tych sił prowadzi do przyspieszenia niezależnego od masy poruszającego się ciała. Jedynie źródła tych sił są różne. W przypadku grawitacji tradycyjnie rozumianej jej źródłem jest masa, a w przypadku tzw. sztucznej grawitacji nieinercjalność układu. Zgodnie z koncepcją Einsteina w obu przypadkach czasoprzestrzeń ulega zakrzywieniu, a jedną z konsekwencji tego zakrzywienia są zjawiska nazywane przez nas grawitacją.

W artykule chciałbym pokazać podobieństwa obu typów grawitacji bez uciekania się do Ogólnej Teorii Względności. Wszystkie rozważania będą prowadziły na gruncie zwykłej mechaniki newtonowskiej. Nie będę zajmował się układem poruszającym się ze stałym przyspieszeniem, gdyż równoważność takiego układu z układem inercjalnym w jednorodnym polu grawitacyjnym jest widoczna od razu. Zajmę się układem obracającym się ze stałą prędkością kątową. Ruchy swobodne w tym ukła-

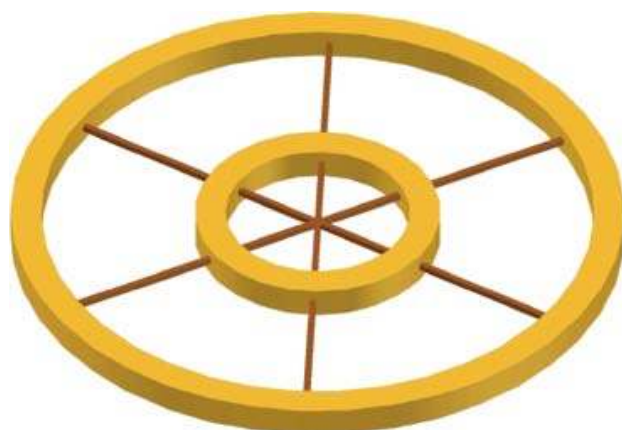
dzie będę porównywał z ruchami swobodnymi w polu grawitacyjnym planety obracającej się wokół własnej osi. W obu przypadkach będę pomijał opór powietrza.

Definicje pojęć

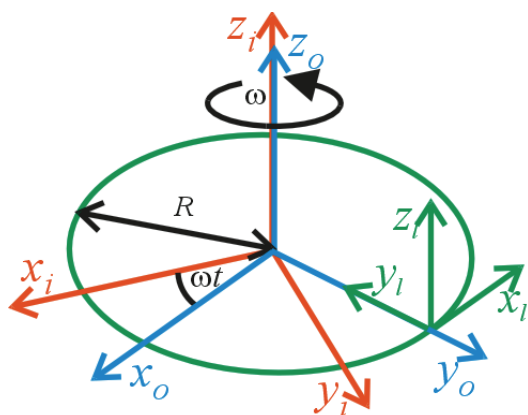
O sztucznej grawitacji mówimy często w kontekście budowy miast kosmicznych. Wyobraźmy sobie obiekt o kształcie przedstawionym na rysunku 1. Część mieszkalną stanowi zewnętrzny pierścień. Obiekt obraca się ruchem jednostajnym wokół głównej osi symetrii konstrukcji.

Wprowadźmy oznaczenia i definicje pojęć, którymi będę się posługiwał. Określenie *powierzchnia miasta* będzie oznaczało skrajną powierzchnię zewnętrznego pierścienia równoległą do osi obrotu. Na niej będzie toczyć się życie mieszkańców.

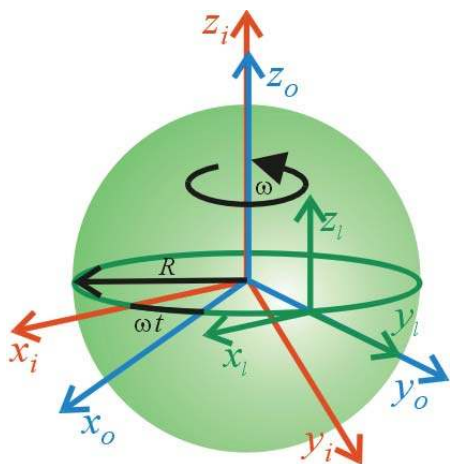
Odpowiednikiem tego pojęcia w przypadku planety będzie *powierzchnia planety*. Niech R oznacza promień *powierzchni miasta*. Jeśli miasto wiruje z prędkością kątową ω , to na *powierzchni miasta* wszystkie ciała doświadczają przyspieszenia odśrodkowego o wartości $\omega^2 R$. Przyspieszenie



Rysunek 1. Uproszczona konstrukcja miasta kosmicznego.



(a)



(b)

Rysunek 2. (a) Trzy układy współrzędnych powiązanych z miastem kosmicznym. Układ inercyjalny x_i, y_i, z_i , układ nieinercyjny obracający się wraz z miastem x_o, y_o, z_o i układ lokalny związany z jakimś punktem na powierzchni miasta x_l, y_l, z_l . Osie z_i, z_o pokrywają się z osią obrotu miasta, a osie x -owe i y -owe są prostopadłe do powierzchni miasta. (b) Analogiczne układy współrzędnych powiązanych z obracającą się planetą.

to będą oznaczał symbolem g_o . Dla przykładu, jeśli $R = 10 \text{ km}$, i $\omega \approx 0,03132 \text{ s}^{-1}$ (okres $T \approx 200,6 \text{ s}$), to $g_o \approx 9,81 \text{ m/s}^2$. Czyli przyspieszenie odśrodkowe ma wartość przyspieszenia grawitacyjnego na powierzchni Ziemi na szerokości geograficznej ok. 50° .

Z naszym miastem przyszłości zwiążmy trzy układy współrzędnych (patrz rysunek 2 (a)). Układ inercyjalny x_i, y_i, z_i , główny układ nieinercyjny obracający się wraz z miastem x_o, y_o, z_o i lokalny układ nieinercyjny x_l, y_l, z_l . Osie z_i, z_o pokrywają się z osią obrotu miasta, a ich początek związany jest ze środkiem konstrukcji miasta. Oś z_l jest równoległa do osi obrotu i styczna do *powierzchni miasta*. Wszystkie osie x -owe i y -owe są prostopadłe do *powierzchni miasta*.

Podobne układy zwiążemy z planetą przy czym osie x -owe i y -owe będą leżały w płaszczyźnie równika. W obu przypadkach oś y_l jest prostopadła do osi obrotu, a jej zwrot jest przeciwny do zwrotu *siły ciężkości*, czyli w przypadku kosmicznego miasta jest skierowana w kierunku środka konstrukcji miasta, a w przypadku planety od środka planety (patrz rysunek 2). Kierunek osi y_l wyznacza lokalny pion, a jej zwrot definiuje lokalne pojęcia *góra*.

Transformacja współrzędnych z układu inercyjnego do głównego układu nieinercyjnego ma postać

$$\begin{aligned}x_o(t) &= x_i(t) \cos(\omega t), \\y_o(t) &= -x_i(t) \sin(\omega t) + y_i(t) \cos(\omega t), \\z_o(t) &= z_i(t).\end{aligned}$$

Z kolei transformacja współrzędnych z głównego układu nieinercyjnego do lokalnego opisują równania: w układzie lokalnym związanym z punktem na *powierzchni miasta*

$$x_l = -x_o, \quad y_l = R - y_o, \quad z_l = z_o,$$

w układzie lokalnym związanym z punktem na równiku planety

$$x_l = x_o, \quad y_l = y_o - R, \quad z_l = z_o.$$

Gdybyśmy ograniczyli się do odpowiednio małych obszarów w pobliżu *powierzchni miasta* i *powierzchni planety* oraz odpowiednio małych prędkości nadawanych ciałom, to problem pola siły odśrodkowej w naszym kosmicznym mieście i problem siły grawitacyjnej na planecie sprowadziłyby się do problemu pola jednorodnego, co jak wspominałem wyżej jest mało interesujące. Dlatego w artykule chcę wyjść poza to przybliżenie.

W układzie obracającym się otrzymanie dokładnych rozwiązań problemu ruchu swobodnego wymagałoby uwzględnienia zależności siły odśrodkowej od odległości od osi obrotu oraz uwzględnienie siły Coriolisa. Dokładne rozwiązanie takich równań ruchu byłoby problemem samym w sobie, o ile w ogóle byłoby możliwe analitycznie. Ale możliwe jest podejście znacznie prostsze. Możemy rozwiązać problem ruchu w układzie inercyjnym, a następnie przetransformować te wyniki do układu obracającego się.

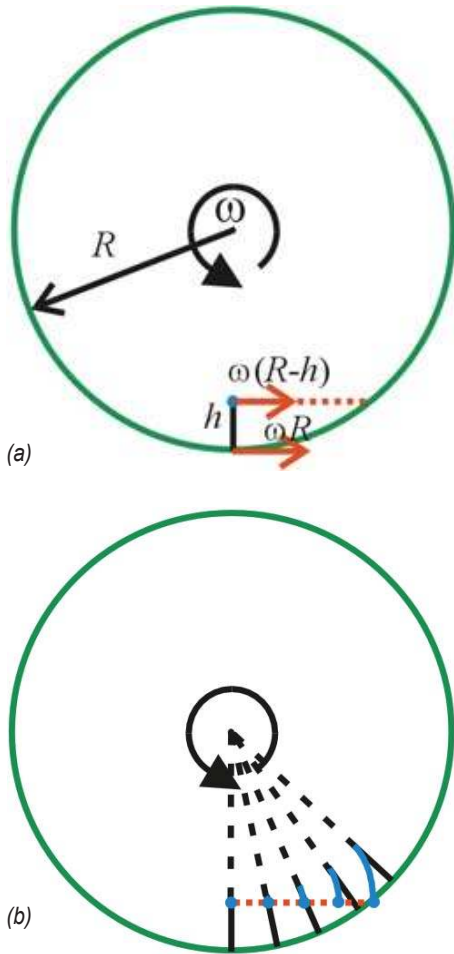
W tym momencie proponuję czytelnikowi zabawę. Zadam pytanie i proszę o zastanowienie się nad odpowiedzią przed przejściem do dalszej części artykułu. Załóżmy, że stoimy na powierzchni naszego kosmicznego miasta i wypuszczamy z rąk kamień. Wiemy, że w układzie obracającym się działa siła odśrodkowa, pod wpływem której kamień upadnie na podłogę (*powierzchnię miasta*). Ale moje pytanie brzmi: jak wytłumaczyć upadek tego kamienia na podłogę w układzie inercyjnym, w którym przecież nie ma żadnej siły odśrodkowej?

Spadek swobodny bez prędkości początkowej w układzie obracającym się

- Czas spadku i równania ruchu

Na początku odpowiem na pytanie, jakie zadałem w poprzednim punkcie. Kamień, który trzymaliśmy w dłoni przed jego wypuszczeniem porusza się względem układu inercyjnego z prędkością liniową o wartości $\omega(R-h)$ gdzie h jest wysokością kamienia nad *powierzchnią miasta*.

Zgodnie z pierwszą zasadą dynamiki, po wypuszczeniu kamienia z dłoni, będzie on się poruszał ruchem jednostajnym prostoliniowym z taką prędkością względem układu inercyjnego, jaką miał w momencie wypuszczenia z dłoni. Czyli po pewnym czasie uderzy w *powierzchnię miasta* (patrz rysunek 3 (a)). Jednocześnie podczas jego ruchu miasto obraca się, więc jeśli $h \ll R$, czyli $\omega(R-h)t \approx \omega R t$, to kamień upadnie obok nas. Kolejne fazy tego ruchu po-



Rysunek 3. (a) Po upuszczeniu ciała z wysokości h porusza się ono względem układu inercyjnego ruchem jednostajnym prostoliniowym z szybkością $\omega(R-h)$. Czerwona kropkowana linia pokazuje tor ruchu względem układu inercyjnego. (b) Kolejne fazy ruchu kamienia widziane z układu związanego z miastem kosmicznym. Niebieskie linie pokazują tor ruchu ciała w tym układzie.

kazuje rysunek 3 (b). Widzimy, że kamień nie spada po linii prostej, lecz jego tor odchylił się w kierunku przeciwnym do kierunku obrotu miasta. W układzie obracającym się efekt ten tłumaczymy siłą Coriolisa [1,2]. Podobny efekt obserwujemy na Ziemi. Tory spadających swobodnie ciał na Ziemi są odchylane w kierunku wschodnim.

Nasze rozważania ograniczmy do stosunkowo małych wysokości. Jeśli wysokość h będzie zbliżała się do R , to tory ciał staną się bardzo złożone. Zresztą w zwykłym polu grawitacyjnym obracającej się planety tory ciał spadających z bardzo dużych wysokości również będą bardzo skomplikowane. Takimi ekstremalnymi przypadkami być może zajmę się kiedyś w osobnym artykule.

W momencie puszczenia ciała w wybranym przez nas układzie inercyjnym ma ono prędkość

$$v_{ix} = -\omega(R-h), \quad v_{iy} = 0, \quad v_{iz} = 0.$$

Czyli w tym układzie równania ruchu ciała są następujące

$$x_i(t) = -\omega(R-h)t, \quad y_i(t) = R-h, \quad z_i(t) = 0.$$

Czas t_k ruchu ciała do momentu upadku możemy znaleźć z równania

$$\omega^2(R-h)^2 t_k^2 + (R-h)^2 = R^2.$$

Stąd po odpowiednich przekształceniach znajdujemy, że ciało upadnie po czasie

$$t_k = \sqrt{\frac{2h}{g_0} \frac{\sqrt{1 - \frac{1}{2} \frac{h}{R}}}{1 - \frac{h}{R}}}.$$

Jeśli $h \ll R$, to korzystając z rozwinięcia w szereg Maclaurina [3] do wyrazów rzędu h/R możemy napisać następujący przybliżony wzór na czas spadku ciała

$$t_k \approx \sqrt{\frac{2h}{g_0}} \left(1 + \frac{3h}{4R} \right).$$

Jeśli ograniczymy się do pierwszego wyrazu rozwinięcia dostaniemy typowy wzór na czas spadku swobodnego w jednorodnym polu

$$t_k = \sqrt{2h/g_0}.$$

Zapisane wyżej równania ruchu w układzie inercyjnym możemy przetransformować do lokalnego układu korzystając z podanych w pierwszym punkcie transformacji. Otrzymamy wówczas następujące równania

$$\begin{aligned} x_l(t) &= \omega(R-h)t \cos(\omega t) - (R-h) \sin(\omega t), \\ y_l(t) &= R - \omega(R-h)t \sin(\omega t) - (R-h) \cos(\omega t), \\ z_l(t) &= 0. \end{aligned}$$

Można sprawdzić, że dla niewielkich wysokości ($h \ll R$) czas spadku swobodnego t_k będzie dużo mniejszy od okresu obrotu miasta ($t_k \ll T$), wobec tego $\omega t \ll 1$. Stosując rozwinięcie funkcji \sin i \cos w powyższych wyrażeniach w szereg Maclaurina do pierwszych dwóch niezerowych wyrazów dostaniemy po odpowiednich przekształceniach następujące przybliżenia

$$\begin{aligned} x_l(t) &\approx -\frac{1}{3} g_0 \omega t^3 \left(1 - \frac{h}{R} \right), \\ y_l(t) &\approx h - \frac{1}{2} g_0 t^2 \left(1 - \frac{h}{R} \right) \left(1 - \frac{1}{3} \omega t \right). \end{aligned}$$

W najbardziej zgrubnym przybliżeniu, gdy zaniedbamy wyrazy $\frac{1}{3} \omega t$ i h/R dostajemy typowe równania opisujące spadek w jednorodnym polu obracającej się planety.

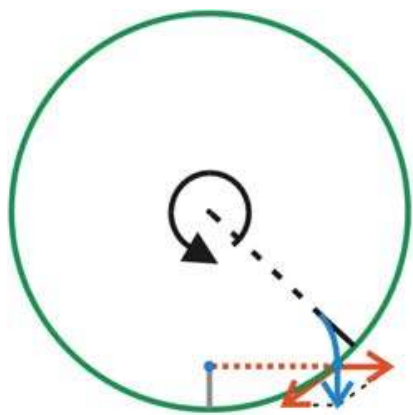
• Zasada zachowania energii mechanicznej

Można sprawdzić, że rotacja siły odśrodkowej jest zerowa. W związku z tym siła odśrodkowa jest siłą zachowawczą. A zatem ciału w polu siły odśrodkowej możemy przypisać energię potencjalną. Jeżeli energię potencjalną będziemy liczyć względem powierzchni miasta, to ciało o masie m na wysokości h będzie miało energię równą

$$E_p(h) = -\int_R^{R-h} m \omega^2 r dr = m g_0 h \left(1 - \frac{1}{2} \frac{h}{R} \right).$$

W momencie upadku energia potencjalna będzie zera i w całości zamieni się w energię kinetyczną ciała

$$\frac{mv^2}{2} = m g_0 h \left(1 - \frac{1}{2} \frac{h}{R} \right).$$



Rysunek 4. Konstrukcja geometryczna znajdująca prędkość ciała względem punktu upadku w momencie upadku.

Stąd znajdujemy, że szybkość ciała spadającego z wysokości h w momencie upadku jest równa

$$v = \sqrt{2g_0 h \left(1 - \frac{h}{2R}\right)}.$$

Sprawdźmy, czy jest to zgodne z równaniami podanymi wcześniej. Dokładne rozwiązania na prędkość spadającego ciała w układzie lokalnym mają postać

$$\begin{aligned} \frac{dx_t(t)}{dt} &= -g_0 t \left(1 - \frac{h}{R}\right) \sin(\omega t), \\ \frac{dy_t(t)}{dt} &= -g_0 t \left(1 - \frac{h}{2R}\right) \cos(\omega t). \end{aligned}$$

Pierwiastek z sumy kwadratów tych składowych prędkości w chwili t_k daje nam dokładnie ten sam wynik, jaki dostaliśmy z zasady zachowania energii. Jest to wynik dokładny dla dowolnej wysokości $h < R$. Przybliżony wynik dla $h \ll R$ z dokładnością do wyrazów rzędu h/R ma postać

$$v = \sqrt{2g_0 h \left(1 - \frac{h}{2R}\right)}.$$

Gdy pominiemy wyrazy rzędu h/R , dostaniemy typowy wzór na szybkość osiąganą w momencie upadku ciała spadającego bez prędkości początkowej w polu jednorodnym.

Prędkość ciała w momencie upadku względem punktu upadku możemy znaleźć odejmując od wektora prędkości ciała względem układu inercyjnego wektor prędkości chwilowej punktu upadku. Geometryczną konstrukcję tej operacji pokazuje rysunek 4.

Spadek swobodny bez prędkości początkowej w polu grawitacyjnym planety

- Czas spadku i równania ruchu

Zajmijmy się teraz spadkiem swobodnym w polu grawitacyjnym planety z uwzględnieniem jej ruchu obrotowego. Ograniczymy się do eksperymentów odbywających się w płaszczyźnie równika. Tak jak w poprzednim punk-

cie pomijamy opory powietrza. Rozwiązując poprzedni problem skorzystaliśmy z tego, że w układzie inercyjnym upuszczone ciało porusza się ruchem jednostajnym prostoliniowym. To pozwoliło nam łatwo znaleźć dokładne rozwiązania. Teraz sytuacja jest znacznie bardziej skomplikowana, gdyż siła grawitacji działająca na ciało działa zarówno w układzie inercyjnym jak i nieinercyjnym związanym z obracającą się planetą.

Spójrzmy najpierw na ruch ciała w układzie inercyjnym. Upuszczamy ciało z wysokości h . W tym momencie ciało jest nieruchome względem planety, ale względem układu inercyjnego ma prędkość o wartości $\omega(R+h)$ (patrz rysunek 5 (a)). Załóżmy, że ta prędkość jest mniejsza od pierwszej prędkości kosmicznej na tej wysokości. Wobec tego torem ruchu planety w układzie inercyjnym będzie elipsa¹, a dokładniej fragment elipsy, której jedno z ognisk znajduje się w środku planety (patrz rysunek 5 (a)).

Ponieważ podczas lotu ciała planeta cały czas obraca się, tor ciała widziany z planety będzie inny. Jego kształt pokazano na rysunkach 5 (b) i 5 (d) (niebieskie linie). Warto porównać ten tor z torem ciała pokazanym na rysunku 3 (b). Jak widzimy obie krzywe mają podobny charakter. Obie są „wklęsłe” i wygięte w kierunku przeciwnym do kierunku obrotu (w kierunku wschodu).

Ekstremalnie wysoka „wieża” na rysunku 5 ma wysokość równą połowie promienia planety. Gdyby masa i promień planety były takie jak w przypadku Ziemi, to tor o kształcie przedstawionym na rysunkach 5 (a) i (b) dostalibyśmy, w przypadku prędkości kątowej planety czterokrotnie większej od prędkości kątowej Ziemi (doba musiałaby trwać 6 godzin). W przypadku prędkości kątowej takiej jaką ma Ziemia, mała pół elipsy byłaby ponad trzy razy mniejsza od poprzedniej (patrz rysunek 5 (c)). Tym samym tor ciała widziany z Ziemi byłby znacznie mniej wygięty (ciało spadłoby bliżej podstawy „wieży”, porównaj rysunki 5 (b) i (d)).

Stosunkowo prosto możemy znaleźć czas spadku ciała z dużej wysokości w przypadku planety nieobracającej się. Z zasady zachowania energii dostajemy następujące równanie

$$v_y^2 = 2GM \left(\frac{1}{R+y} - \frac{1}{R+h} \right),$$

gdzie G jest stałą grawitacji i M masą planety.

Stąd mamy

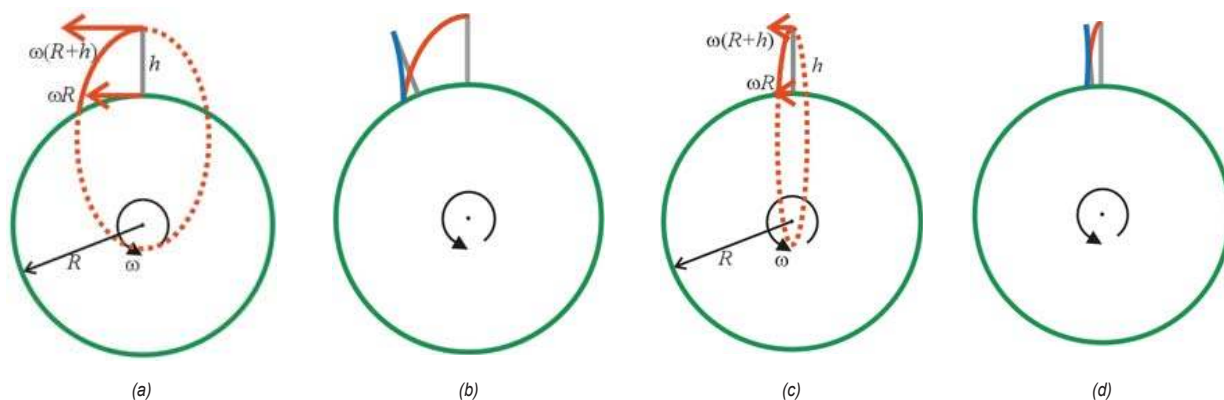
$$\frac{dy}{dt} = -\sqrt{2GM \left(\frac{1}{R+y} - \frac{1}{R+h} \right)}.$$

Czyli

$$t_k = -\frac{1}{\sqrt{2GM}} \int_h^0 \frac{dy}{\sqrt{\frac{1}{R+y} - \frac{1}{R+h}}}.$$

¹ Bardzo ładny dowód tego, że torami ruchów swobodnych w polu grawitacyjnym jednorodnej kuli są krzywe stożkowe można znaleźć w [4].

² Przybliżone rozwiązanie problemu odchylenia toru od pionu podczas spadku swobodnego w płaszczyźnie równika można znaleźć w [4], a w przypadku dowolnej szerokości geograficznej w [5].



Rysunek 5. Spadek ciała z wysokości $h = \frac{1}{2} R$ w płaszczyźnie równika. (a) W momencie upuszczenia ciała z wysokości h nad powierzchnią planety jej prędkość względem układu inercyjnego ma wartość $\omega(R+h)$. W tym układzie ciało będzie poruszać się po fragmencie elipsy (czerwona linia) z ogniskiem w środku planety. (b) Niebieska linia pokazuje tor ciała widziany z planety. Rysunki (a) i (b) dotyczą planety o masie i promieniu takich jak ma Ziemia, ale o prędkości kątowej czterokrotnie większej niż w przypadku Ziemi. Rysunki (c) i (d) dotyczą tej samej planety, ale obracającej się z prędkością kątową taką jak Ziemia. W obu przypadkach zaniedbano opór powietrza.

I ostatecznie

$$t_k = \frac{R}{v_H} \sqrt{1+h/R} \left((1+h/R) \operatorname{atan} \sqrt{h/R} + \sqrt{h/R} \right),$$

gdzie

$$v_H = \sqrt{2GM/R}$$

jest drugą prędkością kosmiczną planety. Dla $h = R/2$ czas spadania na Ziemi (bez oporu powietrza) wyniósłby ok. 19 minut ($\approx 1139,18$ s).

Przy odpowiednio małej prędkości kątowej planety i odpowiednio małych wysokości otrzymany wynik możemy z dobrym przybliżeniem zastosować do przypadku spadku swobodnego w polu grawitacyjnym obracającej się planety. Ponadto możemy wówczas otrzymać przybliżony wzór na czas spadku z wysokości h . Rozwijając powyższe wyrażenie w szereg Maclaurina do wyrazów rzędu h/R dostajemy

$$t_k \approx \sqrt{\frac{2h}{g}} \left(1 + \frac{5h}{6R} \right).$$

Czas spadania na Ziemi (bez oporu powietrza) z wysokości $h = R/2$ wyliczony z tego przybliżonego wzoru byłby o niecałe 4 sekundy dłuższy od dokładnego czasu podanego wyżej. Błąd procentowy wyniósłby ok. 0,34%. Ale zastosowanie przybliżenia

$$t_k \approx \sqrt{2h/g}$$

zaniżyłoby czas aż o ok. 332 sekundy (błąd procentowy wyniósłby ok. 29%).

Warto porównać otrzymane wyniki przybliżonego czasu spadania ciał w polu grawitacyjnym planety i w układzie obracającym się. Oba wzory różnią się jedynie współczynnikiem przed wyrazem h/R . W polu grawitacyjnym dostaliśmy $5/6$, a w układzie obracającym się $3/4$.

W układzie obracającym się mogliśmy znaleźć dokładną postać równań ruchu. Gdy planeta obraca się wokół własnej osi, raczej nie możliwe jest znalezienie dokładnych analitycznych wyrażeń na zależności od czasu współrzędnych ciała. Nie da się tego zrobić ani w układzie inercyjnym, ani w układzie obracającym się wraz z planetą.

W przypadku planety obracającej się i spadku wzdłuż osi y mogliśmy skorzystać z podanego wyżej przybliżonego wzoru na czas spadku swobodnego z wysokości h . Po odpowiednich przekształceniach i zastosowaniu rozwinięcia Maclaurina można podać przybliżoną postać funkcji $y_i(t)$

$$y_i(t) \approx h - \frac{1}{2} g t^2 \left(1 - \frac{5}{3} \frac{h}{R} \right).$$

Dodanie ruchu obrotowego planety komplikuje zagadnienie jeszcze bardziej. Jak już pokazywałem w inercyjnym układzie odniesienia ruch będzie się odbywał po fragmencie elipsy (rysunek 5 (a) i (c)). W układzie obracającym się wraz z planetą torem jest krzywa pokazana na rysunkach 5 (b) i (d). W tym układzie musielibyśmy poza siłą grawitacji uwzględnić siłę odśrodkową i siłę Coriolisa. Równania ruchu odbywającego się w płaszczyźnie równika miałyby postać

$$\begin{aligned} \frac{d^2 x_0}{dt^2} &= 2\omega \frac{dy_0}{dt} + \omega^2 x_0, \\ \frac{d^2 y_0}{dt^2} &= -\frac{GM}{x_0^2 + y_0^2} - 2\omega \frac{dx_0}{dt} = \omega^2 y_0. \end{aligned}$$

Jak widzimy jest to układ zależnych równań różniczkowych. Nie mamy możliwości rozwiązania tego układu w sposób analityczny. Wobec tego w przypadku zmiennej y ograniczmy się do podanego wyżej przybliżenia otrzymanego dla planety nieobracającej się.

To rozwiązanie siłą rzeczy nie zawiera wyrazów zależnych od prędkości kątowej planety. Taką zależność widzieliśmy w funkcji, jaką dostaliśmy rozwiązując problem spadku ciała w układzie obracającym się. Ale gdybyśmy mogli tę zależność zaniedbać (dla bardzo małej prędkości kątowej), to otrzymalibyśmy przybliżenie podobne do tego w polu grawitacyjnym planety. Przybliżenie to różniłoby się tylko czynnikiem stojącym przed ułamkiem h/R . W polu grawitacyjnym mamy $5/3$, a w układzie obracającym się 1.

Dla małych wysokości, w pierwszym z powyższego układu równań, wyraz opisujący x -ową składową przyspieszenia odśrodkowego $\omega^2 x_0$ jest bardzo mały, gdyż $x_0 \ll R$. Jeśli opuścimy ten wyraz i wykorzystamy przybliże-

nie funkcji $y_i(t)$ otrzymane dla nieobracającej się planety, to dostaniemy równanie

$$\frac{d^2 x_0}{dt^2} = -2\omega g t \left(1 - \frac{5}{6} \frac{h}{R}\right).$$

A stąd znajdujemy², że

$$x_0(t) \approx -\frac{1}{3} g \omega t^3 \left(1 - \frac{5}{6} \frac{h}{R}\right).$$

Przypomnijmy, że przybliżenie analogicznej funkcji w układzie obracającym się miało bardzo podobną postać

$$x_0(t) \approx -\frac{1}{3} g_0 \omega t^3 \left(1 - \frac{h}{R}\right).$$

Jak widzimy, cechy spadku swobodnego bez prędkości początkowej w układzie obracającym się i w polu grawitacyjnym obracającej się planety wykazują bardzo duże podobieństwo. Jest tak nawet wtedy, gdy mamy do czynienia ze spadkiem z bardzo dużej, w porównaniu z R , wysoko-

ści. Dla bardzo małych wysokości, dla których możemy zaniedbać wyrazy rzędu h/R oba zjawiska są wręcz identyczne.

Aby zamknąć temat, w następnym artykule opiszę jeszcze krótko problem rzutu pionowego i rzutu ukośnego w układzie obracającym się.

Jan Kurzyk

Instytut Fizyki Politechniki Krakowskiej

LITERATURA

- [1] T. Wibig, „Zagubiona przyszłość”, czyli o realności sił pozornych. *Fizyka w Szkole* nr 6, 2019.
- [2] J. Kurzyk, Ruch jednostajny po okręgu pod wpływem sił bezwładności. *Fizyka w Szkole* nr 2, 2020.
- [3] https://pl.wikipedia.org/wiki/Wz%C3%B3r_Taylora. (dostęp 11.03.2020).
- [4] C. Kittel, W.D. Knight, M.A. Ruderman, *Mechanika*, PWN 1975.
- [5] A. Januszajtis, *Fizyka dla politechnik*, tom I. PWN 1977.

Co w fizyce piszczy

O egzoplanecie, która zniknęła

Odkąd Aleksander Wolszczan metodą zaćmieniową odkrył pierwszą egzoplanetę zaroilo się w literaturze naukowej od planet okrążających inne gwiazdy. Zazwyczaj były to obiekty bardzo blisko okrążające macierzystą gwiazdę. Do wykrycia planety można stosować dwie grupy metod. Pierwsza polega na obserwacji kołysania się głównej gwiazdy, druga polega na obserwacji zaćmień wspomnianej gwiazdy. W takiej sytuacji planetę można czasem, używając teleskopów kosmicznych, zaobserwować jako punkt.

Do pomiarów astronomicznych dotyczących egzoplanet należy jednak podchodzić z ostrożnością. Przekonali się o tym ostatnio naukowcy badający obiekt o nazwie Fomalhaut b. Obiekt ten został odkryty w 2008 roku i był, oczywiście nie w sposób ciągły, badany do 2014 roku. Wyznaczono wiele jego parametrów i postanowiono niedawno wrócić do jego badań. Niestety obiekt zniknął. Pojawiają się dwa wyjaśnienia tego faktu. Wyjaśnienie pierwsze zakłada, że obiekt opuścił swój układ planetarny. Wyjaśnienie drugie jest takie, że go tam nigdy nie było. W najlepszym razie była to kula gwiazdowego pyłu. Oczywiście nie można wykluczyć interwencji Lorda Wadera. Wszystkie skojarzenia z historią badań kanałów na Marsie są oczywiście nieuzasadnione i przypadkowe.

Źródło: András Gáspár, George H. Rieke. New HST data and modeling reveal a massive planetesimal collision around Fomalhaut. *Proceedings of the National Academy of Sciences*, April 20, 2020; DOI: 10.1073/pnas.1912506117

Wszechświat nie do końca jednorodny

Jednym z podstawowych założeń współczesnej kosmologii jest stwierdzenie, że Wszechświat jest w dużej skali jednorodny. Możemy go sobie wyobrazić jako jednorodnie pęczniejącą kulę wypełnioną gazem składającym się z gwiazd lub galaktyk. Założenie to wydawało się na tyle naturalne, że nikt go nie sprawdzał. Tymczasem ostatnie badania emisji promieniowania rentgenowskiego z odległych gromad ga-

laktyk zdają się poddawać w wątpliwość to założenie. Otóż istnieją kierunki, w których Wszechświat rozszerza się szybciej i kierunki, w których jego ekspansja jest wolniejsza. Jeśli powyższe rezultaty się potwierdzą wiele niewzruszalnych z pozoru teorii kosmologicznych będzie musiało zostać zrewidowanych.

Źródło: https://www.technology.org/2020/04/09/universes-expansion-may-not-be-the-same-in-all-directions/#offer_link

Jak tanio pozyskać wodór?

Jednym z problemów przed jakimi stoi ludzkość to zastąpienie silników wykorzystujących spalanie ropy silnikami elektrycznymi zasilanymi... No właśnie czym? Jedną z odpowiedzi jest wodór, który podczas spalania w tlenie produkuje wodę a tak zwany ślad węglowy tego procesu wynosi dokładnie zero. Jest jeden mały problem, otóż na ziemi praktycznie nie ma wolnego wodoru. Trzeba go najpierw pozyskać rozkładając wodę. Czyli zgodnie z II zasadą dynamiki do takiego procesu trzeba by wpompowywać energię. Rośliny też w swoich reakcjach wykorzystują wodór, który uwalniają z wody. Aby przeprowadzić ten proces wykorzystują światło i cały proces nosi nazwę fotosyntezy. Naukowcy z sukcesem starają się przeprowadzić podobny proces zmierzający do rozkładu wody na tlen i wodór. Przeprowadzane w laboratoriach reakcje noszą nazwę sztucznej fotosyntezy. Niestety, jak do tej pory nie znalazły one większego zastosowania na skalę przemysłową. Jedną z przyczyn jest używanie w tych urządzeniach platyny, która to znacznie podraża koszty. Naukowcy z uniwersytetu w Waszyngtonie wykazali jednak, że platynowe elektrody można z powodzeniem zastąpić żelaznymi albo niklowymi bez straty wydajności. Ten z pozoru mały krok znacznie przybliżył nas do nowej ery wodorowej energetyki.

Źródło: Dongguo Li, Eun Joo Park, Wenlei Zhu, Qiurong Shi, Yang Zhou, Hangyu Tian, Yuehe Lin, Alexey Serov, Barr Zulevi, Ehren Donel Baca, Cy Fujimoto, Hoon T. Chung, Yu Seung Kim. Highly quaternized polystyrene ionomers for high performance anion exchange membrane water electrolyzers. *Nature Energy*, 2020; DOI: 10.1038/s41560-020-0577-x.



Wpływ własności aerodynamicznych nadwozi samochodów na ich własności jezdne

Samochodami jeździmy wszyscy. Jako pasażerowie lub kierowcy. Nie wszyscy zastanawiają się nad zjawiskami towarzyszącymi jeździe samochodem, szczególnie z dużą prędkością.

Janusz Piechna

Typowy samochód ma nadwozie zawieszane elastycznie na czterech ogumionych kołach. Silnik samochodu przenosi moment obrotowy na koła wytwarzając na styku opony i jezdni siłę napędową równoważącą opory ruchu samochodu, głównie jego opory aerodynamiczne.

Opony kół muszą przenosić siły pionowe będące sumą ciężaru samochodu i pionowej siły aerodynamicznej nazywanej siłą nośną. Gdy siła nośna ma znak ujemny jest uważana za docisk aerodynamiczny. Na rys. 1 pokazano samochód wyposażony w poziomą płytkę nazywaną spojlerem umieszczoną z przodu i dociskającą przednie koła do jezdni oraz płat tylny dociskający tylne koła.

Podczas hamowania powstają siły bezwładności równoważone przez siły tarcia opon o jezdnię.

Podczas pokonywania łuków dróg pojawiają się siły bezwładności usiłujące wyrzucić samochód na zewnątrz łuku, czemu przeciwstawiają się siły tarcia występujące na styku opon i jezdni.

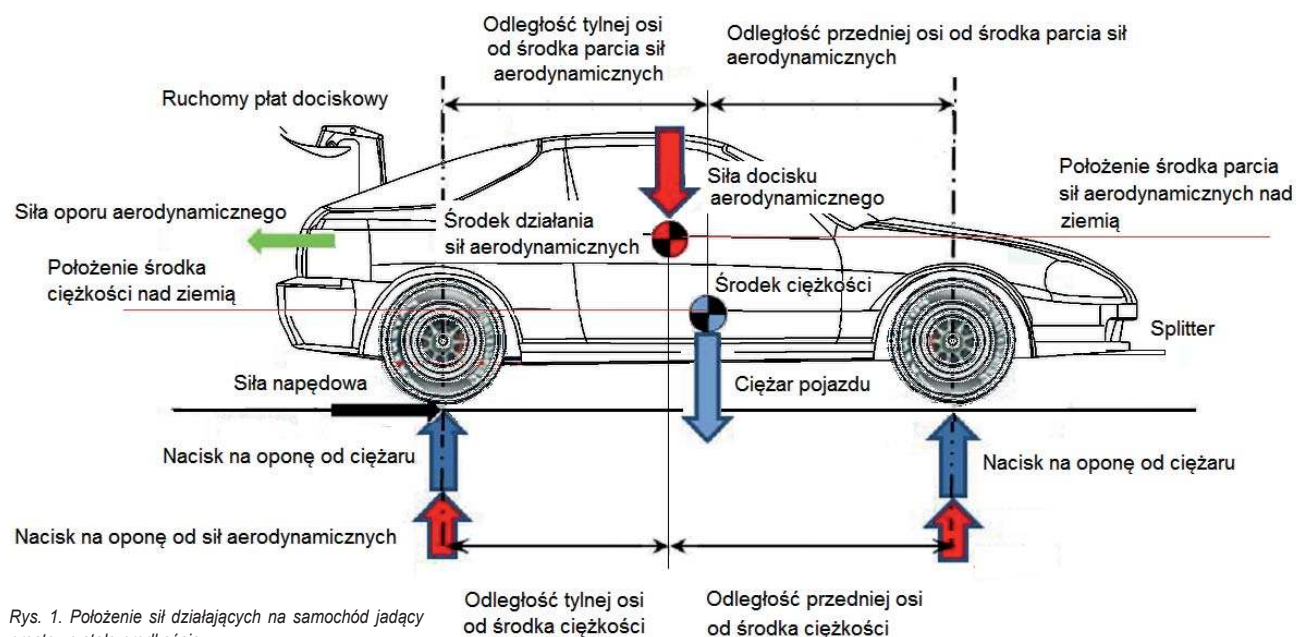
Zatem styk opony z jezdnią jest miejscem, gdzie powstają siły tarcia pozwalające na utrzymanie samochodu na drodze i jego sterowanie.

Opony samochodu muszą przenosić siły podłużne (hamowania i napędu) oraz siły poprzeczne.

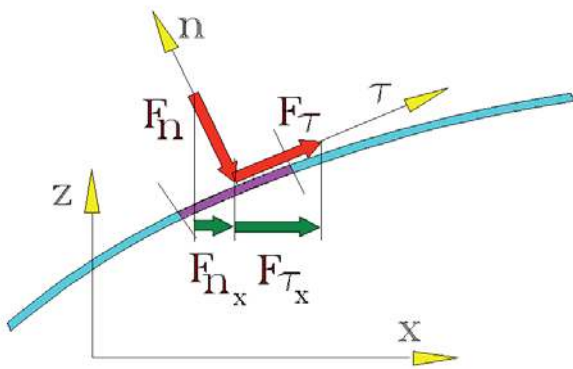
Siły jakie działają na nadwozie to siły bezwładności i siły aerodynamiczne. Zmiany prędkości samochodu lub kierunku jego ruchu powodują pojawienie się sił bezwładności.

Siły aerodynamiczne zależą od kształtu samochodu, prędkości jazdy i gęstości powietrza.

Powietrze opływające powierzchnię nadwozia wytwarza dwa rodzaje sił. Siły pochodzące od ciśnienia



Rys. 1. Położenie sił działających na samochód jadący prosto ze stałą prędkością.



Rys. 2. Siły pochodzące od opływu powietrza wokół nadwozia samochodu, pochodzące od działania ciśnienia i naprężeń stycznych na powierzchni nadwozia, mają składowe w kierunku ruchu tworząc opór aerodynamiczny.

powietrza działają zawsze prostopadłe do powierzchni. Siły pochodzące od tarcia powietrza przepływającego wzdłuż powierzchni wytwarzają siły równoległe do tej powierzchni.

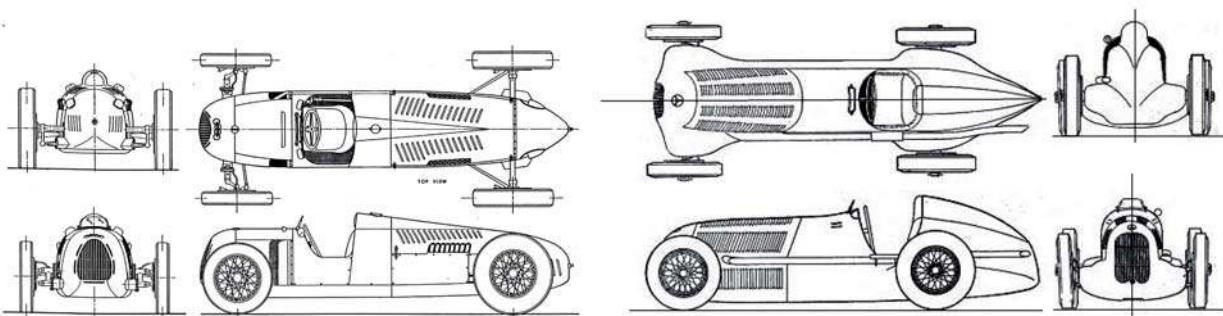
Zwykle interesują nas siły: działająca wzdłuż samochodu – siła oporu aerodynamicznego, siła pionowa – aerodynamiczna siła nośna i siła pozioma działająca poprzecznie do płaszczyzny symetrii samochodu – aerodynamiczna siła boczna.

Nadwozie samochodu opływane przez powietrze wymusza lokalne zmiany prędkości w pobliżu nadwozia samochodu, co wiąże się z lokalnymi zmianami ciśnienia powietrza na powierzchni nadwozia. Ponieważ siły aerodynamiczne zależą od kwadratu prędkości, własności aerodynamiczne samochodu opisuje się bezwymiarowymi współczynnikami, w których siły aerodynamiczne oporu – F_x , poprzeczna – F_y i nośna – F_z odnosi się do ciśnienia dynamicznego q (zależnego od gęstości powietrza ρ i prędkości U) i powierzchni czołowej A .

$$q = \frac{\rho U^2}{2}; \quad C_x = \frac{P_x}{qA}; \quad C_y = \frac{P_y}{qA}; \quad C_z = \frac{P_z}{qA};$$

Przyjmuje się, iż współczynniki te są stałe, nie zależą od prędkości opływu, jeśli struktura przepływu nie ulega zmianie.

Dla zrozumienia zjawisk fizycznych występujących podczas ruchu samochodu istotne jest zrozumienie i zapamiętanie związków występujących pomiędzy siłami aerodynamicznymi, siłami bezwładności i siłami tarcia pomiędzy oponami a jezdnią.



Rys. 3. Auto Union typ C (po lewej) i Mercedes W25 (po prawej) z lat 30-tych ubiegłego stulecia.

Aerodynamicy mają istotny wpływ na siły dociskające oponę do jezdni. Kształt nadwozia może przy dużych prędkościach zwiększać lub zmniejszać ten docisk.

Pomimo długiego rozwoju aerodynamiki lotniczej i samochodowej, dopiero w latach 60-tych poprzedniego stulecia konstruktorzy dostrzegli możliwości poprawy własności jezdnych samochodu przez wykorzystanie sił aerodynamicznych. Słynna jest odpowiedź Enzo Ferrari, w roku 1960, na zastrzeżenia jego kierowcy Paula Frere, co do kształtu przedniej szyby Ferrari 250 TR ograniczającej jego prędkość maksymalną – „Aerodynamika jest dla ludzi, którzy nie potrafią budować silników”.

Dzisiaj samochody firmy Ferrari charakteryzują się jednymi z najlepiej opracowanych kształtów pod względem aerodynamiki.

Do lat 60. poprzedniego stulecia szybkie samochody budowane były tak, by ich kształty generowały małe opory aerodynamiczne.

Takie samochody jak rewolucyjne Auto Union typ C z kierowcą siedzącym przed silnikiem i Mercedes W25 zbudowany w 1934 roku z silnikiem przed kierowcą (nazywanymi w owym czasie pojazdami Silver Arrows) o masie ograniczonej do 750 kg, budowano tak, by charakteryzowały się małymi oporami aerodynamicznymi.

Współczesne analizy numeryczne własności aerodynamicznych tych samochodów pokazują, iż Auto Union typ C uzyskał współczynnik oporu $C_x = 0.6$, współczynnik siły nośnej $C_z = -0.01$ (co oznacza niewielki docisk aerodynamiczny do jezdni) przy czym $C_{zp} = 0.13$ (co oznacza unoszenie przodu samochodu) i $C_{zt} = -0.14$ (co oznacza docisk tyłu).

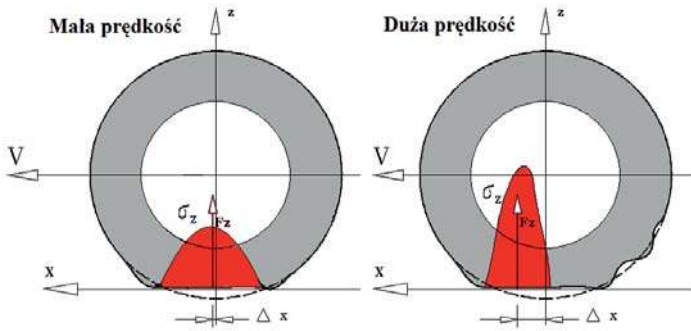
Mercedes charakteryzował się współczynnikiem oporu $C_x = 0.42$, współczynnikiem siły nośnej $C_z = -0.16$ (co oznacza docisk) a rozkład na osie wynosił $C_{zp} = -0.046$ i $C_{zt} = -0.114$.

W tym czasie Bugatti typ 51 (1933) charakteryzował się wartością $C_x = 0.74$. W owym czasie nikt nie interesował się siłami aerodynamicznymi działającymi w kierunku pionowym stąd docisk aerodynamiczny tych samochodów był raczej dziełem przypadku.

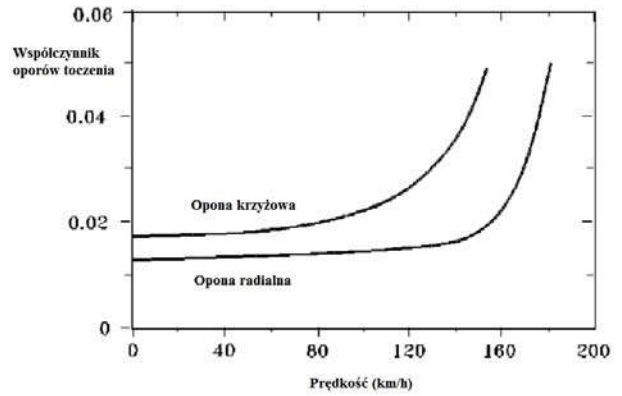
Z powodu odsłoniętych kół współczynniki oporu nie były rewelacyjne, ale mniejsze niż innych samochodów z tego okresu.

Dopiero później zorientowano się, iż dla szybkiego pokonania toru, ważniejsze jest szybkie pokonywanie zakrętów niż maksymalne prędkości osiągnięte na prostych.

Opona jest wykonana z włókien układanych w sposób



Rys. 4. Rozkład ciśnienia na styku opony z jezdnią przy małej prędkości ruchu i przy prędkości dużej.



Rys. 5. Zmiany współczynnika toczenia opon w zależności od typu opony i prędkości jazdy.

krzyżowy lub radialny zatopionych w masie gumowej tworząc jednocześnie elastyczną, ale wytrzymałą strukturę.

Podczas toczenia się koła, w okolicy styku opony z jezdnią opona odkształca się zmieniając swój łukowy kształt w płaski. Po przetoczeniu się koła płaski element powinien powrócić do łukowego kształtu. Dzieje się tak na skutek działania wysokiego ciśnienia powietrza w oponie i sprężystości materiału opony.

Ze wzrostem prędkości uginanie i prostowanie opony zachodzi w coraz krótszym czasie i po przekroczeniu pewnej krytycznej prędkości opona nie nadąża za zmieniającą się geometrią kontaktu z jezdnią i naciski na jezdnię w części uginanej opony są większe od części występującej z tyłu w obszarze powrotu opony do pierwotnego kształtu. Wypadkowa siła pionowa przenoszona przez oponę przesuwa się spod osi obrotu w przód powodując powstawanie dodatkowego oporu toczenia (rys. 4). Opony posiadają parametr określający maksymalną prędkość użytkowania. Nie jest to prędkość, powyżej której opona się rozleci, a prędkość, powyżej której jej opory toczenia znacząco rosną, bo odkształcenia opony nie nadążają za wymaganymi od niej jej kształtami. Zatem podczas jazdy z dużą prędkością pojawia się istotny wzrost oporów toczenia (rys. 5).

Na styku z jezdnią, opona musi przenosić nie tylko siły pionowe, ale także siły podłużne, jak i siły poprzeczne.

Siły tarcia pomiędzy oponą a jezdnią zależą głównie od współczynnika tarcia gumy z materiałem, z którego wykonana jest jezdnia i siły docisku opony do jezdni. W przypadku opon zamiast nazwy współczynnik tarcia, używa się nazwy współczynnik przyczepności opony.

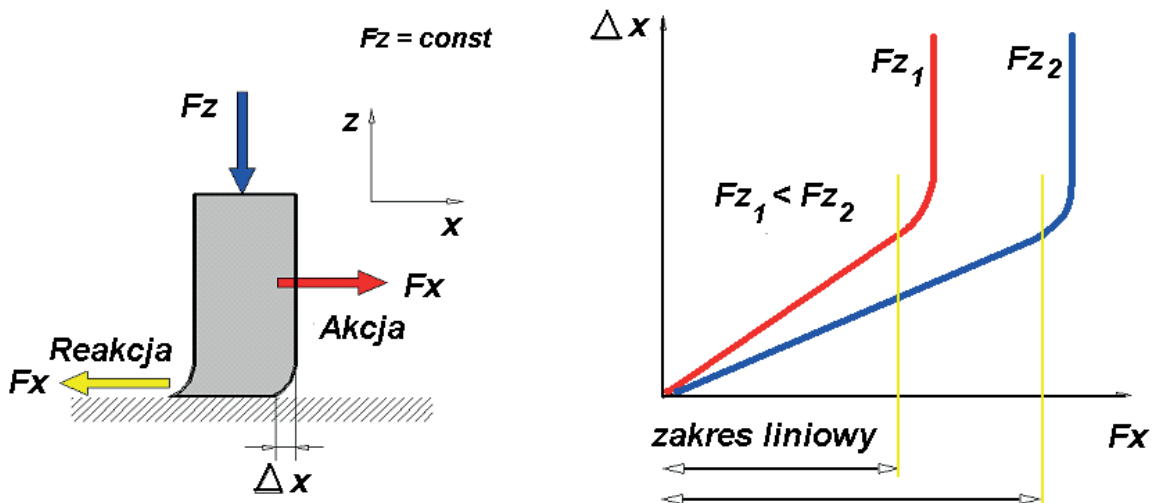
Współczynnik przyczepności opony zależy głównie od rodzaju nawierzchni jezdni, jej stanu i dodatkowych substancji znajdujących się na niej, takich jak woda, śnieg, lód.

Rodzaj gumy, kształt bieżnika, temperatura opony, ciśnienie powietrza w jej wnętrzu, mają także wpływ na wartość współczynnika tarcia opony o jezdnię.

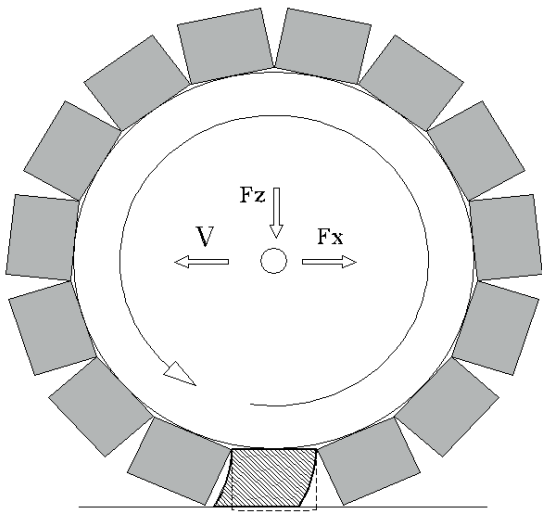
Konstruktorzy samochodów nie mają wpływu na jezdnię. Mogą wybierać opony wykonane z określonego gatunku gumy, dobrać kształt bieżnika opony i ciśnienie w niej panujące.

Charakter zjawisk fizycznych występujący podczas przyspieszania i hamowania jest odmienny od tych jakie występują podczas przenoszenia sił bocznych.

Zwiększenia docisku opony do jezdni pozwala na przenoszenie większych sił podłużnych (rys. 6). Zewnętrzna część elastycznej opony, podczas przyspieszania, jazdy ze stałą prędkością, w czasie hamowania, odkształca się i prędkość styku opony z jezdnią jest nieco inna niż prędkość sztywnego koła. Ta różnica prędkości odniesiona do prędkości obwodowej sztywnego koła nazywana jest

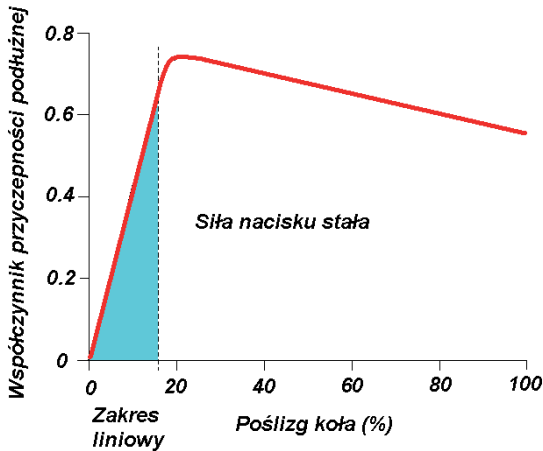


Rys. 6 Zależność wielkości przenoszonej siły na styku opony z jezdnią w zależności od docisku opony do jezdni.



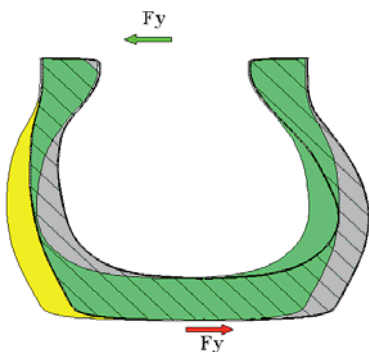
Rys. 7. Elastyczna część opony wygina się co powoduje pozorny poślizg koła.

poślizgiem koła. W pewnym małym zakresie poślizgu, opona się odkształca i powraca do pierwotnego stanu przenosząc siłę podłużną, bez fizycznego ślizgania się opony po jezdni. Przenoszona siła podłużna rośnie liniowo ze wzrostem poślizgu. Gdy różnica prędkości – poślizg staje się większy, opona zaczyna rzeczywiście ślizgać się po powierzchni jezdni. Gdy opona się ślizga, przenoszona siła podłużna jest stała lub lekko spada.

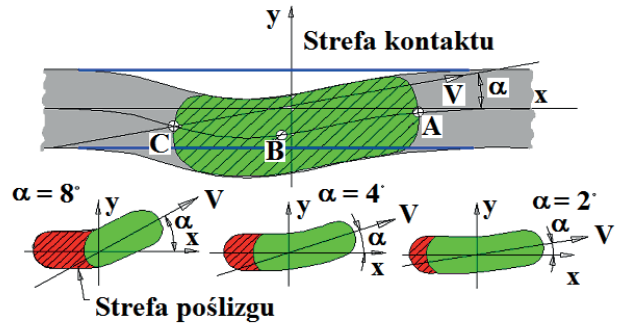


Rys. 8. Zależność współczynnika przyczepności podłużnej od wielkości poślizgu koła.

Opona widziana w przekroju ma kształt Ω i pozwala na elastyczne jej uginanie podczas toczenia.



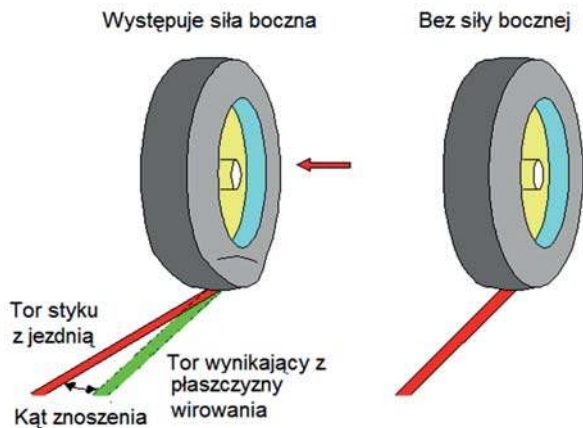
Rys. 9. Poprzeczny przekrój opony nie odkształconej i odkształconej



Rys. 10. Kształt obszaru kontaktu opony z jezdnią.

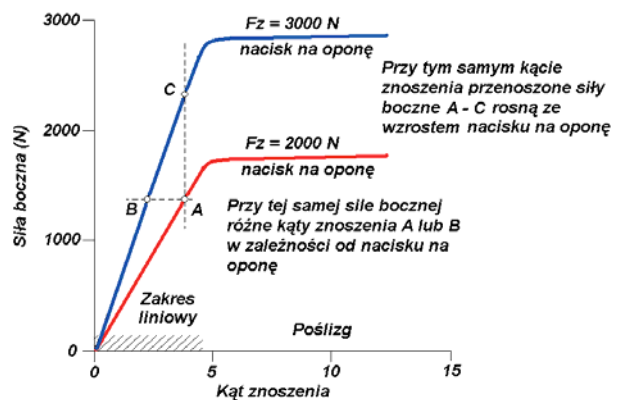
Gdy koło poddane jest działaniu siły bocznej, opona odkształca się poprzecznie. Pamiętać jednak należy, iż koło jednocześnie się toczy. Przenosząc siłę boczną opona musi się odkształcić i wrócić do poprzedniego kształtu.

Opona tocząc się i odkształcając zachowuje płaszczyzną wirowania koła, ale styk opony z jezdnią przemieszcza się w bok. Koło zachowując stałą płaszczyznę wirowania przesuwa się jednocześnie w bok. Linia śladu koła na jezdni odchyła się od płaszczyzny wirowania koła o kąt nazywany kątem znoszenia.



Rys. 11. Kąty znoszenia opony przenoszącej siłę poprzeczną.

Kąt znoszenia jest bardzo istotnym parametrem pracy opony. Im większa jest siła boczna, tym większy jest kąt znoszenia opony. Jednocześnie, im większa jest siła docisku opony do jezdni, tym kąt znoszenia jest mniejszy. Przy



Rys. 12. Zależność siły bocznej przenoszonej przez oponę od kąta znoszenia opony i siły docisku opony do jezdni.

stałej sile docisku opony do jezdni i wzroście siły bocznej, kąt znoszenia rośnie liniowo ze wzrostem siły bocznej do pewnej granicznej wartości. Po jej przekroczeniu opona się ślizga i nie przenosi większej siły bocznej.

Czy kierowca może wyczuć, iż zbliżył się do granic przyczepności opon?

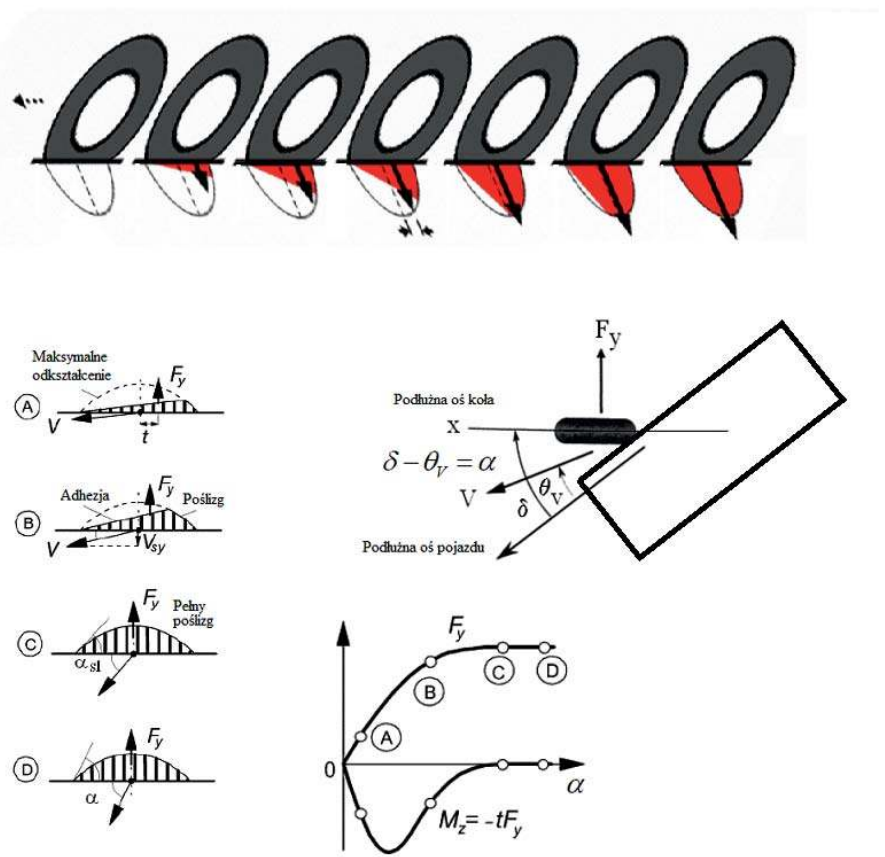
Okazuje się, że tak. Dotyczy to jednak tylko kół sterowanych, a zatem zwykle kół osi przedniej.

Gdy opona zaczyna przenosić siłę boczną to początkowo siła ta nie jest umiejscowiona dokładnie pod osią koła a znajduje się nieco z tyłu za osią (rys. 13). Powoduje to powstanie na kole momentu skręcającego koło. Gdy chcemy skręcić w lewo, skręcamy kierownicą i koło w lewo. Powstająca siła boczna działa w kierunku wyprostowania koła i kierownicy. Ten wzrost oporu koła przekazywany na kierownicę jest sygnałem wyczuwanym przez doświadczonego kierowcę. Ze wzrostem siły bocznej przenoszonej przez oponę moment prostujący koło wzrasta a w miarę zwiększania kąta znoszenia i siły bocznej, potem zaczyna spadać i zanika, gdy opona wyczerpie możliwości przenoszenia siły bocznej i zaczyna się ślizgać. Typowi kierowcy zwykle tak prowadzą samochód, by nie przekraczać maksimum momentu prostującego koła. Kierowcy rajdowi przekraczają to maksimum (rys. 14).

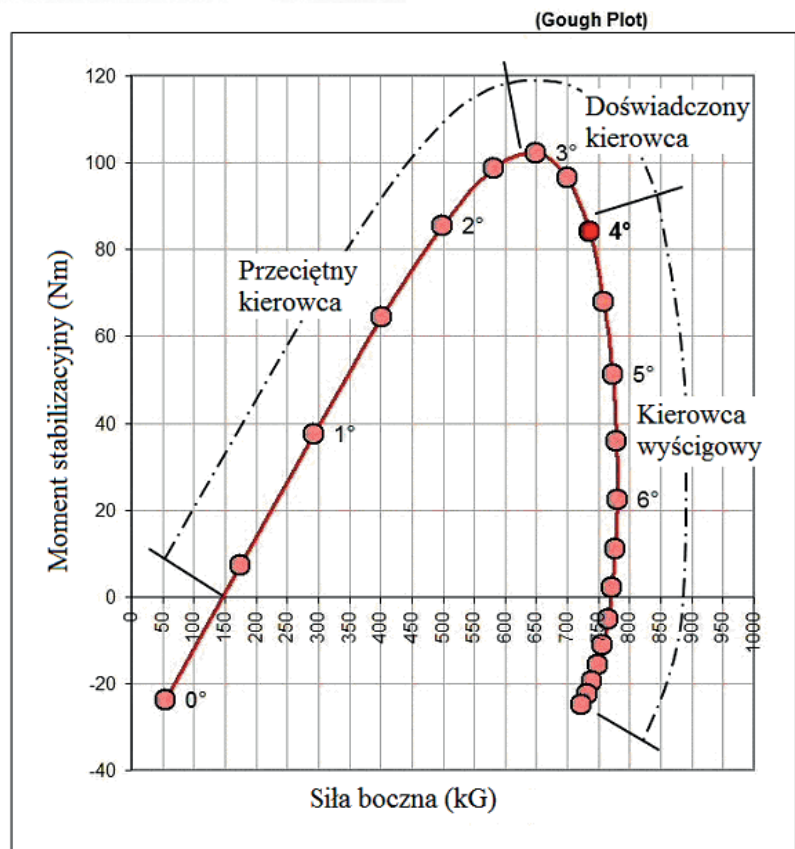
Jeśli wyobrazimy sobie samochód poddany działaniu bocznego wiatru to zdajemy sobie sprawę, iż opony samochodu muszą się przeciwstawić działaniu dodatkowej siły aerodynamicznej o kierunku poprzecznym do płaszczyzny symetrii samochodu (rys. 15). Opony muszą się odkształcać poprzecznie i samochód wolno spychany jest z drogi. Kąty znoszenia opon są na ogół bardzo małe.

Jeśli kąty znoszenia opon kół przednich są takie same jak tylnych, samochód nie zmienia kierunku natomiast jest lekko spychany przez wiatr na bok. I o takim samochodzie mówimy, że ma charakterystykę neutralną.

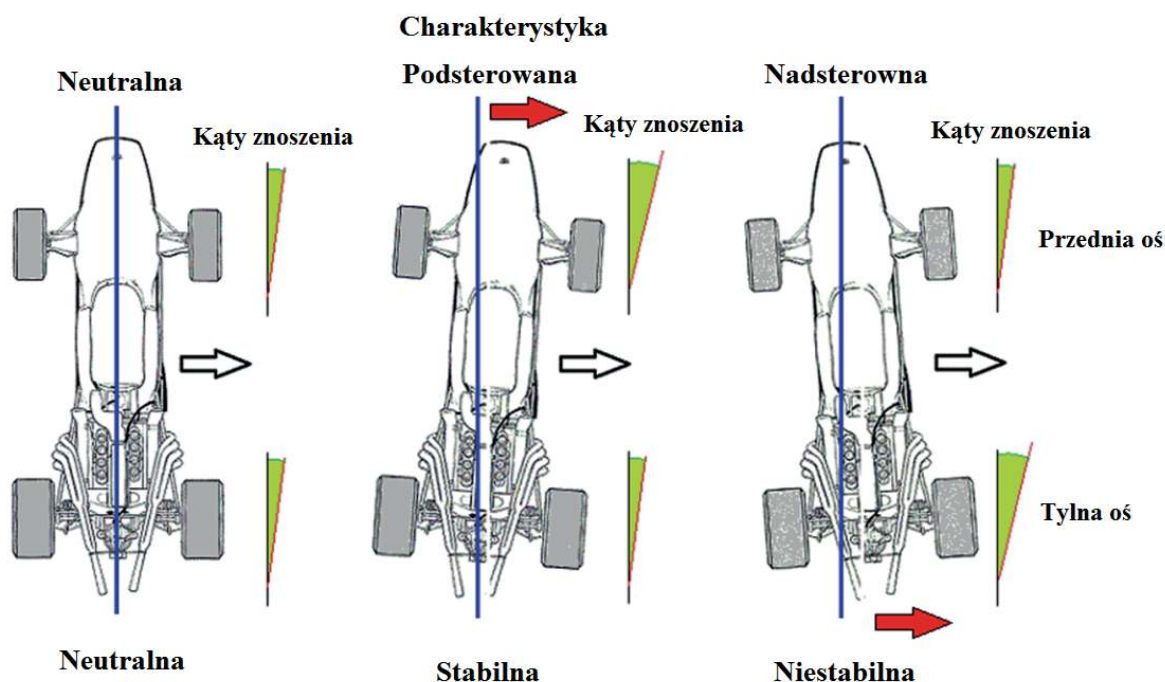
Gdy kąty znoszenia opon tylnej osi są większe niż przedniej, samochód zacznie skręcać pod wiatr. I o takim samochodzie mówimy, że ma charakterystykę nadsterowną.



Rys. 13 Zmiany położenia wypadkowej siły poprzecznej przy jej wzroście i wzroście kąta znoszenia.



Rys. 14. Wykres zależności momentu stabilizacyjnego od wartości siły bocznej przenoszonej przez oponę. Wykres Gough'a.



Rys. 15. Wpływ kątów znoszenia kół osi przedniej i tylnej na zachowanie się samochodu obciążonego siłą boczną.

Gdy kąty znoszenia opon osi tylnej są mniejsze niż przedniej, samochód zacznie skręcać z wiatrem. I o takim samochodzie mówimy, że ma charakterystykę podsterowną.

Konstruktorzy starają się by charakterystyki samochodów były neutralne.

Zwykle samochody przednio napędowe mają charakterystykę podsterowną, tylna napędowe odwrotnie, nadsterowną.

Podczas jazdy po łuku siłą działającą poprzecznie do płaszczyzny symetrii pojazdu jest siła odśrodkowa.

Kąty znoszenia poszczególnych kół są ze sobą powiązane w taki sposób, iż linie prostopadłe do przemieszczania się kół przecinają się w jednym punkcie nazywanym chwilowym punktem obrotu (rys. 16). Punkt obrotu samochodu, który porusza się wolno jest inny od punktu obrotu, gdy samochód porusza się szybko, gdy występują siły boczne i pojawiają się kąty znoszenia.

Samochód obraca się inaczej, gdy porusza się wolno, a inaczej, gdy porusza się szybko.

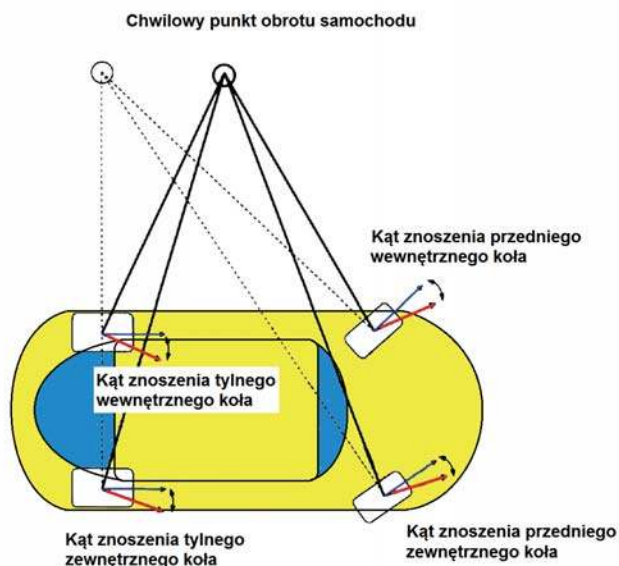
Zatem samochód nadsterowny z większymi kątami znoszenia na oponach tylnej osi będzie zacieśniał zakręt. Jego tył będzie uciekał na zewnątrz łuku.

Na tym samym łuku samochód podsterowny z większymi kątami znoszenia na oponach osi przedniej będzie poszerzał łuk. Jego przód będzie uciekał na zewnątrz łuku (rys. 17).

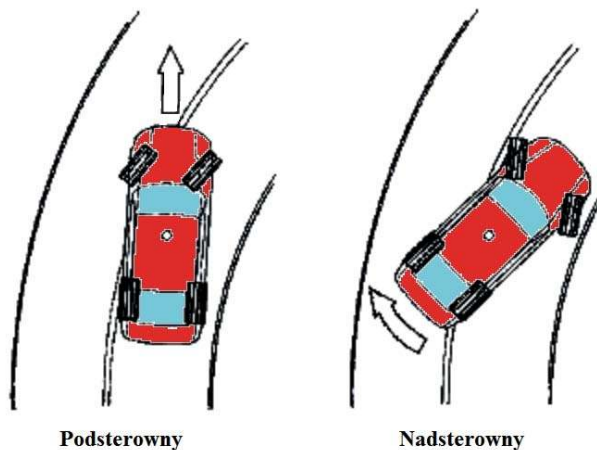
Co jest lepsze? Z punktu widzenia stateczności ruchu charakterystyka podsterowna charakteryzuje samochód stabilny, nadsterowna niestabilny.

Doświadczeni kierowcy twierdzą, iż łatwiej panować nad samochodem nadsterownym niż podsterownym.

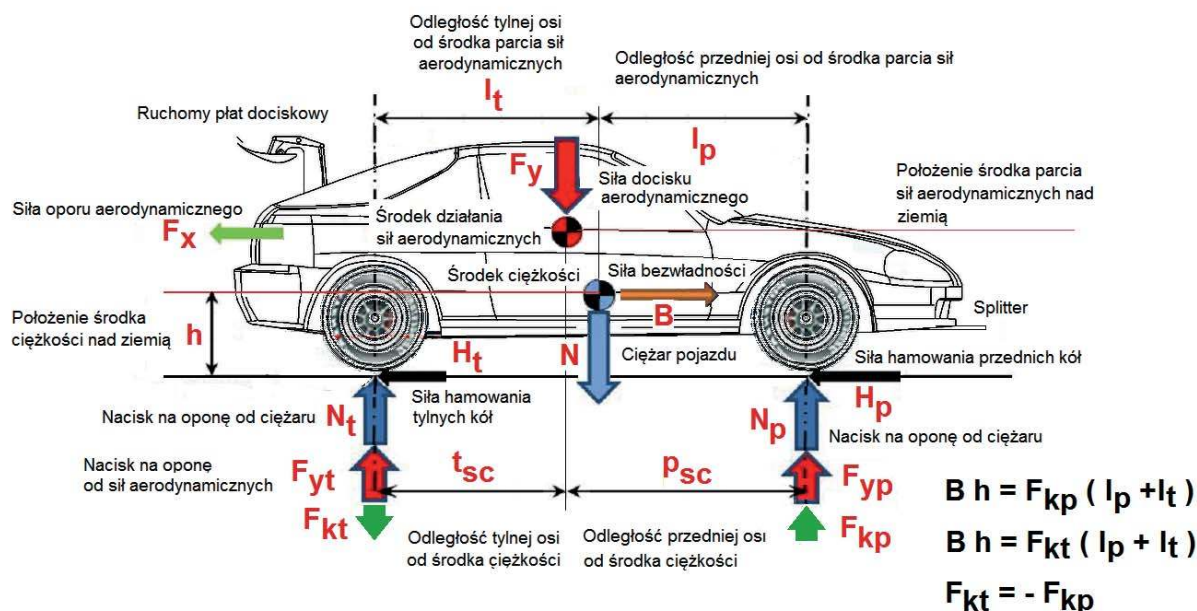
Czy mamy wpływ na charakterystyki samochodu? Mamy, w sposób świadomy lub nieświadomy.



Rys. 16. Chwilowy punkt obrotu samochodu zależy od wartości kątów znoszenia kół.



Rys. 17. Zachowanie się samochodu podsterownego i nadsterownego na zakręcie.



Rys. 18. Siły działające na samochód podczas hamowania.

Podczas hamowania, z powodu położenia środka ciężkości samochodu na pewnej wysokości nad jezdnią, siła bezwładności przyłożona w środku ciężkości samochodu wytwarza moment powodujący, iż przednie koła są dociskane do jezdni większą siłą niż tylne (rys. 18). A to oznacza, że kąty znoszenia opon tylnej osi są większe niż przedniej i podczas hamowania na łuku samochód będzie miał charakterystykę nadsterowną.

Hamując na łuku nieświadomie zmieniamy charakterystykę samochodu w kierunku wzrostu nadsterowności.

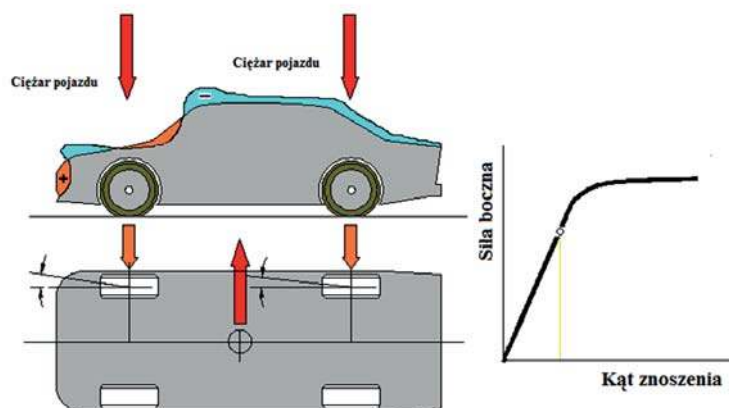
Przyspieszając na łuku, zmieniamy charakterystykę samochodu w kierunku wzrostu podsterowności. To tylko skutki uboczne naszych działań.

Gdy jedziemy wolno na samochód działa tylko siła ciężkości i koła są obciążone tylko ciężarem samochodu. Zwykle oznacza to równe obciążenie osi i takie same kąty znoszenia opon (rys. 19). Samochód ma charakterystykę neutralną.

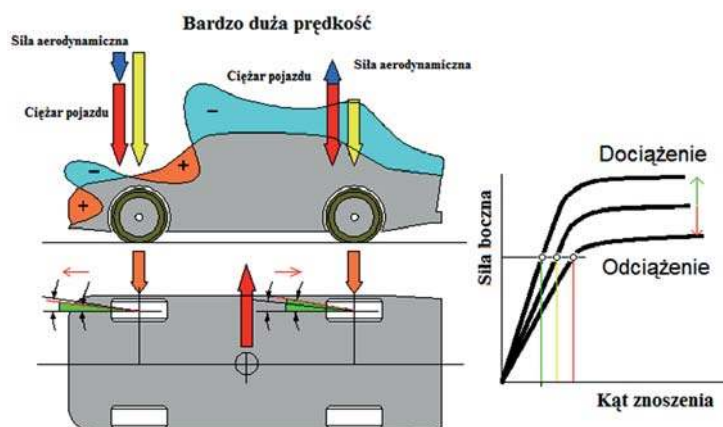
Niestety, gdy jedziemy szybciej pojawiają się siły aerodynamiczne. Zwykle nie bardzo wiemy, jak się zmieniają z prędkością i jaki wpływ mają na zachowanie się samochodu. Typowo, wytwórcy samochodów skupiają się na ograniczeniu zużycia paliwa, co oznacza walkę z oporami aerodynamicznymi. Tworzą kształty nadwozi, których opory aerodynamiczne są co prawda małe ale konsekwencją tego jest zwykle pojawienie się pionowej siły aerodynamicznej unoszącej tył samochodu.

Gdy jedziemy szybciej, rosną siły aerodynamiczne unoszące tył samochodu. A to oznacza, iż kąty znoszenia tylnych kół stają się większe od występujących na osi przedniej i charakterystyka samochodu zmienia się na nadsterowną (rys. 20). Samochód staje się niestabilny. I niestabilność ta powiększa się ze wzrostem prędkości.

Niestety, siły aerodynamiczne zależą od kwadratu prędkości jazdy, zatem początkowo ich działanie jest niewidoczne. Rozpędzając się w obszar prędkości wyższych

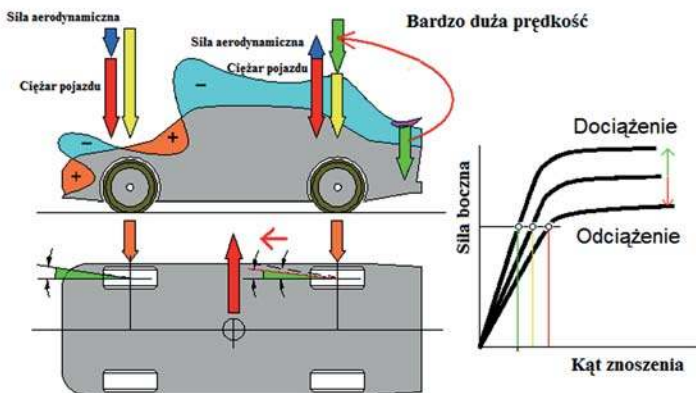


Rys. 19. Docisk kół obu osi tylko przez ciężar samochodu – kąty znoszenia kół obu osi identyczne.



Rys. 20. Siły aerodynamiczne (niebieskie wektory) pojawiające się ze wzrostem prędkości zmieniają siły docisku opon do jezdni (żółte wektory) zwiększając kąty znoszenia kół tylnej osi i zmniejszając przedniej

od 160 km/h możemy natrafić na nieoczekiwane zachowanie samochodu. Należy tu zauważyć, iż istotna jest masa samochodu a właściwie stosunek sił aerodynamicz-



Rys. 21. Kompensacja redukcji docisku tylnej osi przez dodatkowy płat dociskowy (wektor zielony)

nych do ciężaru samochodu. W przypadku ciężkich samochodów problemy zaczynają się pojawiać przy wyższych prędkościach.

Wytwórcy samochodów częściowo walczą z tym zjawiskiem stosując dodatkowe płyty aerodynamiczne stałe lub wysuwane przy dużej prędkości, by przy ich pomocy zredukować niekorzystne siły aerodynamiczne powstające na nadwoziu a powodujące niestabilność pojazdu (rys. 21).

Przed wieloma laty firma Porsche wykonała interesujący test. Zadanie polegało na hamowaniu pojazdu po wjeździe w łuk drogi. Używano samochodu bez elementów aerodynamicznych i tego samego samochodu z dodatkowymi stałymi elementami aerodynamicznymi.

Czym się różniły charakterystyki aerodynamiczne tych wersji samochodu? Wersja bazowa, bez dodatkowych elementów aerodynamicznych, z powodu swojego kształtu generowała siłę aerodynamiczną unoszącą przód nadwozia i w znacznie większym stopniu tył nadwozia (rys. 22).

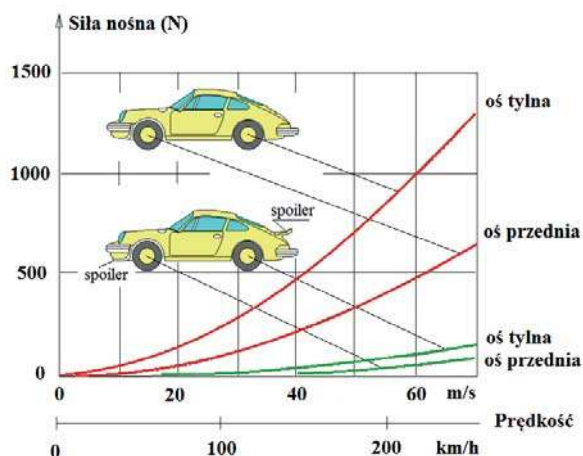
Wersja wyposażona w dodatkowe elementy aerodynamiczne redukujące powstające na nadwoziu pionowe siły aerodynamiczne ze wzrostem prędkości praktycznie nie zmieniała swoich charakterystyk (rys. 22).

Podczas testów zarejestrowano w czasie zmiany położenia kierownicy i opóźnienia występujące podczas hamowania. Zmiany tych wielkości w czasie pokazano na rys. 23.

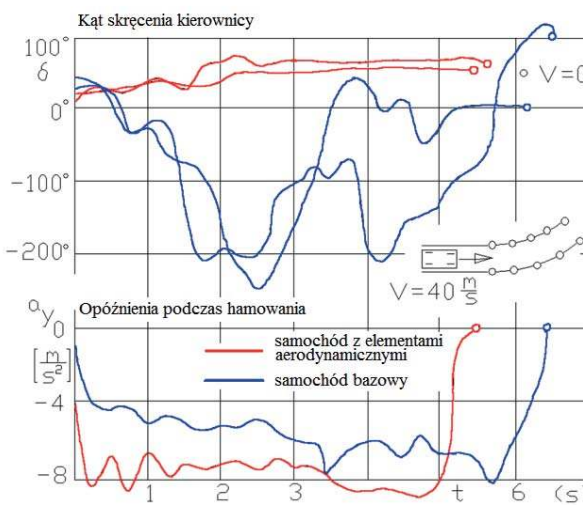
Co z tych przebiegów można wywnioskować?

Samochody wchodzą w lewy zakręt, gdy kierowcy lekko skręcają kierownicę w lewą stronę, naciskając jednocześnie na hamulec. Jak pamiętamy hamowanie powoduje odciążenie kół tylnej osi i dociążenie przedniej.

Samochód bazowy podlega silnemu działaniu sił aerodynamicznych. W momencie rozpoczęcia manewru tył samochodu unoszony jest siłą prawie 50 kG a przód 20 kG. Po dodatkowym odciążeniu tyłu na skutek hamowania, opony tylnej osi tracą przyczepność. Tył samochodu wysuwa się w prawo. Kierowca skręca silnie kierownicę w prawo by opanować obrót samochodu w lewo. Jednocześnie nie naciska zbyt silnie na hamulec. Konieczny do opanowania samochodu skręt kierownicy w prawo wymaga jej skręcenia o ponad 200 stopni. Kierowca



Rys. 22. Wielkości siły nośnej generowanej przez nadwozie bazowe i wyposażone w kompensujące elementy aerodynamiczne.



Rys. 23. Zmiany położenia kierownicy i opóźnienia samochodu podczas hamowania na łuku.

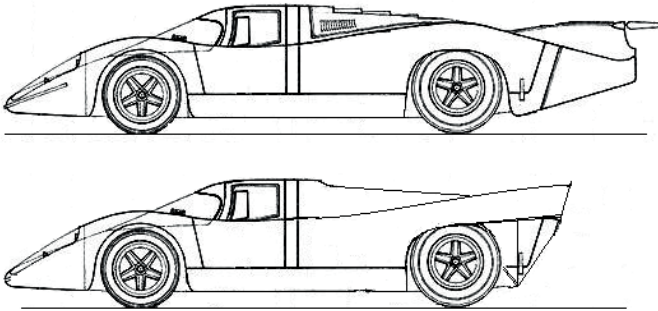
ma skrzyżowane ręce. Za moment musi wykonać kontrę kierownicą w drugą stronę. W końcu zatrzymuje samochód po ponad 6-ciu sekundach.

Druga próba wygląda podobnie, ale nie całkiem. Konieczne ruchy kierownicy są bardzo obszerne i do tego różne od wykonanych podczas wcześniejszego przejazdu. Samochód zachowuje się nieprzewidywalnie. Kierowca po pierwszym przejeździe nie nauczył się, jak panować nad samochodem.

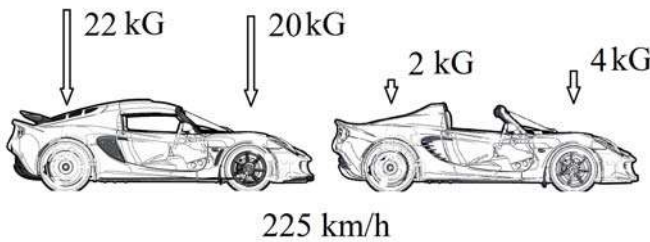
Przejazdy samochodem z elementami aerodynamicznymi prawie całkowicie kompensującymi aerodynamiczną siłę unoszącą samochód do góry wyglądają zupełnie inaczej.

Kierowca tylko lekko skręca kierownicę w lewo i małymi ruchami koryguje tor jazdy jednocześnie bardzo silnie hamując. Zatrzymuje się o sekundę wcześniej. Drugi przejazd jest podobny. Samochód zachowuje się przewidywalnie.

W tej chwili na ulicach można spotkać bardzo wiele samochodów mających stałe lub ruchome elementy aerodynamiczne kompensujące złe charakterystyki aerodynamiczne samego nadwozia. Ze względów estetycznych konstruktorzy ukrywają je w obrysach nadwozia wysuwając przy prędkościach przekraczających 100 km/h.



Rys. 24. Sylwetki dwóch wersji Porsche 917, wersja L długa na górze, wersja K krótka na dole.



Rys. 25. Podobne samochody z innym typem nadwozia generują inne wartości sił docisku aerodynamicznego.

Skrajnym przypadkiem wpływu sił aerodynamicznych jest zachowanie się samochodu Porsche 917 w dwóch wersjach (K – krótkie nadwozie i L – nadwozie długie) przy dużej prędkości jazdy. Prędkość maksymalna, w 600-konnej wersji, wynosiła 354 km/h, przy masie własnej poniżej 1000 kg. Samochód był bardzo szybki. Jednak z jakiegoś powodu 917 prowadziły się bardzo źle, a moc silnika była zdecydowanie zbyt duża dla lekkiego nadwozia. Podczas debiutu w 1000-kilometrowym wyścigu na Nürburgringu w czerwcu 1969 kierowcy fabryczni odmówili startu 917 L, twierdząc, że jest zbyt niebezpieczny w prowadzeniu. Testy fabryczne kończyły się na prędkościach 250 km/h i inżynierowie nie widzieli żadnego problemu. Kierowcy rozwijający podczas wyścigu prędkości 350 km/h uważali wersję L za skrajnie niestabilną.

Dopiero skrócenie tyłu nadwozia i zmiana jego kształtu, tak by generować siłę docisku kosztem wzrostu oporów, doprowadziło do uzyskania stabilnego samochodu, który zaczął wygrywać wyścigi (rys. 24).

Współczynnik oporu aerodynamicznego wersji długiej wynosił 0.36 (przy powierzchni czołowej 1.57 m²) a wersji krótkiej 0.44 (powierzchnia czołowa 1.55 m²). Współczynniki siły nośnej nie są znane.

Długie opływowe tyły zmniejszały siłę oporu, ale generowały siłę unoszącą tył.

Prędkości w czasie wyścigu były tylko 40% większe od prędkości podczas testów fabrycznych, ale siły aerodynamiczne już 2-krotnie większe. Aerodynamiczne siły unoszące długi tył wersji L przy 250 km/h nie stanowiły problemu, ale dwa razy większe przy 350 km/h były już problemem dla lekkiego samochodu.

Współczesne szybkie samochody mają zwykle neutralne charakterystyki niezależne od prędkości jazdy (rys. 25).

Do tego momentu rozważaliśmy tylko skutki działania sił aerodynamicznych na stabilność samochodu. Szybkie pokonywanie zakrętów jest równie ważne, a czasem ważniejsze, od szybkiej jazdy po prostej drodze.

Maksymalne prędkości pokonywania zakrętów wynikają z równowagi pomiędzy wartością siły odśrodkowej i siły tarcia poprzecznej jakie mogą przenieść opony.

W pokazanej na rys. 26 nierówności, siła poprzeczna możliwa do przeniesienia przez opony musi być większa od siły odśrodkowej.

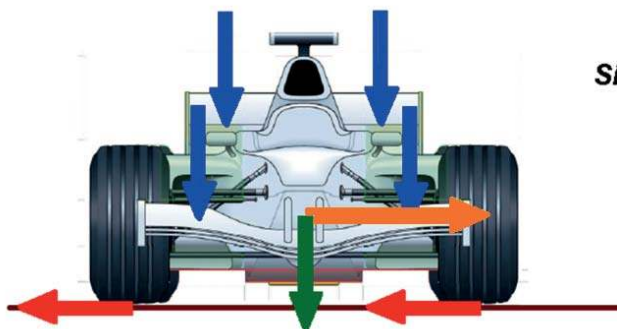
Siła odśrodkowa usiłująca wyrzucić samochód na zewnątrz zakrętu zależy od masy samochodu, kwadratu prędkości jazdy i promienia zakrętu.

Siła możliwa do przeniesienia przez opony zależy od współczynnika przyczepności opon i siły docisku opon do jezdni. Ta zaś zależy od ciężaru samochodu i aerodynamicznej siły docisku.

Każda ze stron nierówności zawiera masę samochodu.

Na konstrukcję samochodu mają wpływ nie tylko konstruktorzy i aerodynamicy, ale przede wszystkim księgowi.

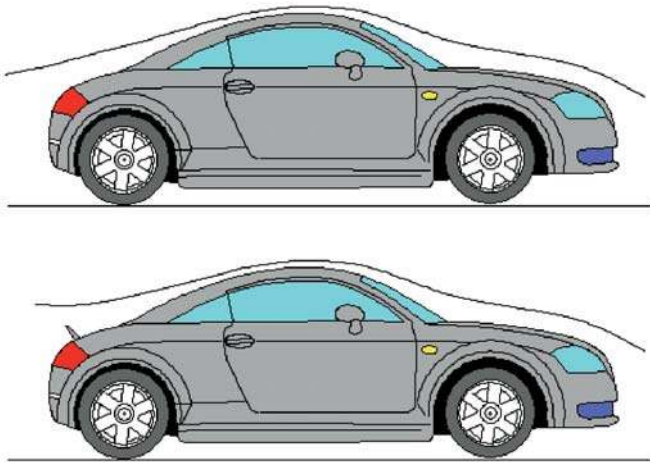
Przed laty, gdy na drogach pojawił się rewelacyjny samochód Audi TT, zarówno aerodynamicy jak i kierowcy testowi uprzedzali, iż samochód o takim kształcie będzie niestabilny przy dużych prędkościach, gdyż przepływ będzie długo przyklejony do obłych części tyłu generując siłę unoszącą tył. Nie chciano się zgodzić na żaden element aerodynamiczny z tyłu by skompensować to zjawisko. Zepsułby design samochodu. Decydenci uznali, że dołożenie 15 kg masy z tyłu pojazdu rozwiąże pro-



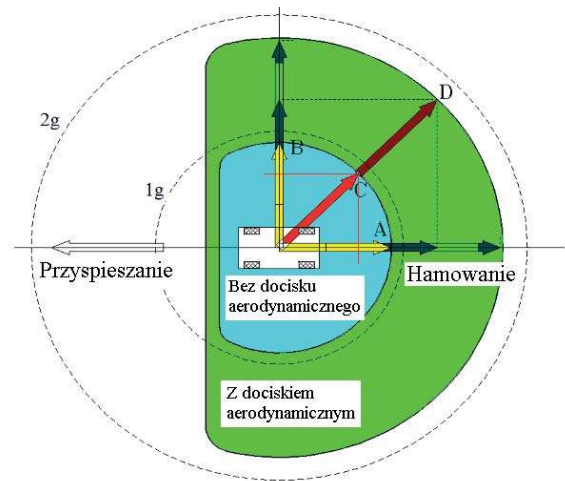
Rys. 26. Związki pomiędzy siłami tarcia przenoszonymi przez opony a siłą odśrodkową.

Sily :

$$\underbrace{\left[mg + \rho C_z S \frac{V^2}{2} \right]}_{\text{Ciężar Aerodynamika}} \mu \geq \underbrace{m \frac{V^2}{R}}_{\text{Odśrodkowa}}$$



Rys. 27. Schemat opływu dwóch wersji nadwozia samochodu Audi TT.



Rys. 28. Wykres G-G. Porównanie wielkości przyspieszeń poprzecznych i podłużnych możliwych do osiągnięcia przez samochody wykorzystujące lub nie docisk aerodynamiczny.

blem. Wykorzystując zawartość rys. 26 wiemy, że nie rozwiązało.

Użytkownicy stwierdzili, że tył samochodu jest unoszony i tracą nad nim kontrolę podczas zmiany pasa przy dużych prędkościach (180 km/h i więcej) lub wykonywania ostrych zakrętów.

Po serii wypadków na tyłach tego samochodu pojawiły się listwy (spoilery) korygujące opływ nadwozia (rys. 27). Użytkowników dziwiło to, że subtelny spojler stanowił różnicę między wymykaniem się spod kontroli a utrzymaniem stabilności samochodu. W tym przypadku spoiler spiętrza powietrze podnosząc ciśnienie na dużej powierzchni tylnej części nadwozia. Suma iloczynów lokalnego wzrostu ciśnienia i powierzchni, na której nastąpił generuje aerodynamiczną siłę dociskającą nadwozie.

Konstruktorzy szybkich samochodów zdają sobie sprawę z korzyści jakie wynikają z uwzględnienia docisku aerodynamicznego. Jeśli uda się uzyskać docisk aerody-

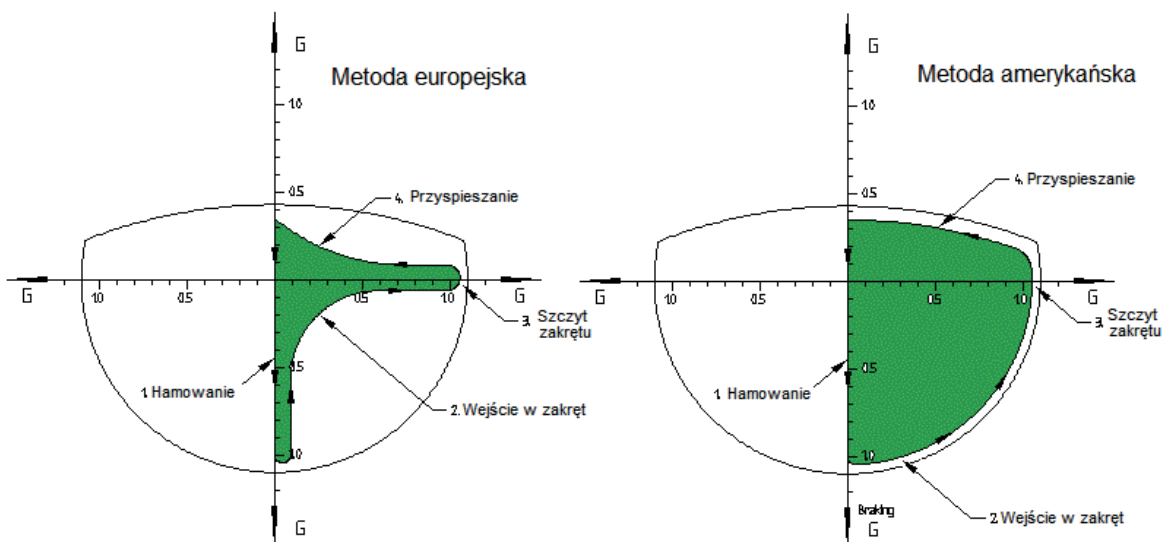
namiczny na nadwoziu samochodu, jego własności pokonywania zakrętów wzrosną w znaczący sposób.

Najwyraźniej jest to widoczne na rys. 28, na wykresach (G-G) pokazujących obwiednie dopuszczalnych przyspieszeń zarówno bocznych jak i powstających podczas hamowania i przyspieszania a wynikających z maksymalnej przyczepności opon.

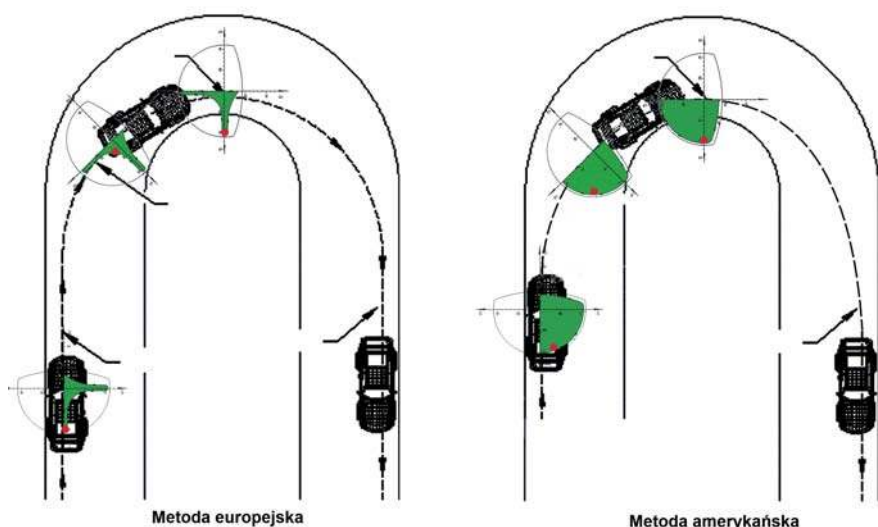
Niebieski obszar to obszar występowania przyspieszeń występujących podczas jazdy samochodem bez docisku aerodynamicznego a zielony z dociskiem aerodynamicznym. Wykorzystanie docisku aerodynamicznego w istotny sposób poprawia własności jezdne samochodu pozwalając na szybsze pokonywanie zakrętów i efektywniejsze hamowanie.

Ponieważ docisk aerodynamiczny zależy od kwadratu prędkości jazdy, jego przejawy i zalety uwidaczniają się dopiero przy większych prędkościach.

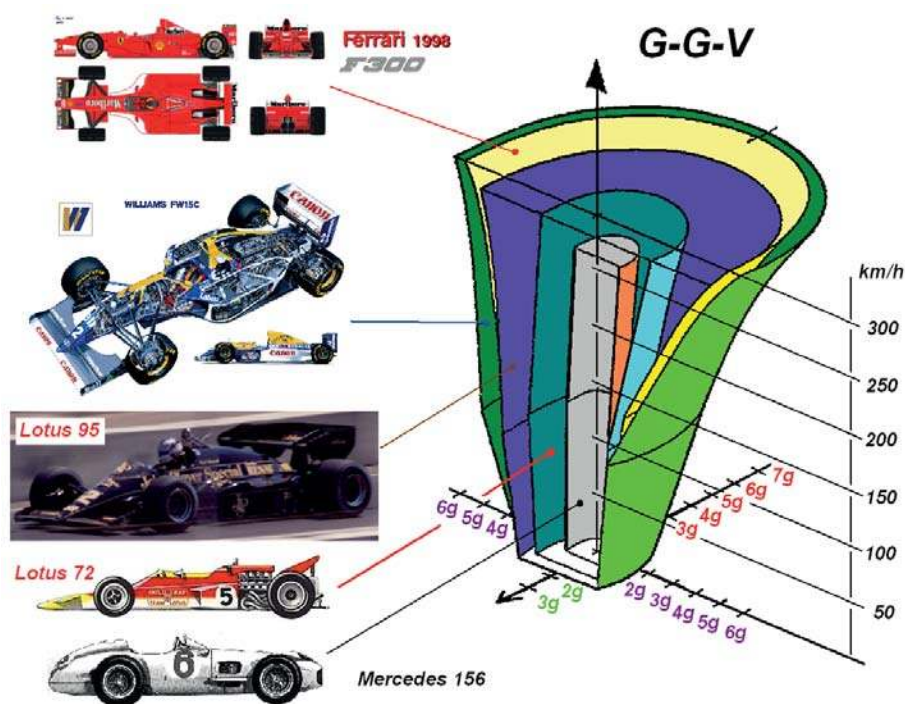
Jako ciekawostkę można na wykresie G-G pokazać różnice w technice pokonywania zakrętów stosowanej w Europie i w Ameryce (rys. 29).



Rys. 29. Wykres G-G ilustrujący zmiany przyspieszeń podczas pokonywania zakrętów techniką europejską i amerykańską.



Rys. 30. Trajektorie ruchu samochodu po łuku metodą europejską i amerykańską.



Rys. 31. Zależność maksymalnych przyspieszeń poprzecznych i podłużnych w zależności od prędkości jazdy i budowy samochodu.

LITERATURA

- [1] I.J. Piechna, L. Rudniak, A. Piechna, 2009, „CFD Analysis of the Central Engine Generic Sports Car Aerodynamics”, Proceedings Book -Simulation for Innovative Design, 4th European Automotive Simulation Conference Munich, pp. 95-106
- [2] J. Piechna, T. Janson, P. Sadowski, S. Tudruj, A. Piechna, L. Rudniak, Sports Car with Movable Flaps and Deformable Airbags, Automotive Simulation World Congress, Frankfurt/Main, Germany.
- [3] M. Gugala, J. Piechna, J.D. Mueller, The Design of Formula SAE Car Body Aerodynamics Using ANSYS Fluent Optimisation Tools, Automotive Simulation World Congress, in Frankfurt/Main, Germany
- [4] T. Janson, J. Piechna, Numerical analysis of aerodynamic characteristics of a high-speed car with movable bodywork elements, The Archive of Mechanical Engineering, Vol. LXII, 2015, No. 4, pp451-476
- [5] K. Kurec, M. Remer, T. Mayer, S. Tudruj, J. Piechna, 2018, Flow control for a car-mounted rear wing, International Journal of Mechanical Sciences tom 152, pp 384- 399, Pergamon, 2019/3/1,
- [6] K. Kurec, M. Remer, J. Broniszewski, P. Bibik, S. Tudruj, J. Piechna, 2019, Advanced Modeling and Simulation of Vehicle Active Aerodynamic Safety, Journal of Advanced Transportation,
- [7] J. Broniszewski, J. Piechna, A fully coupled analysis of unsteady aerodynamics impact on vehicle dynamics during braking, Engineering Applications of Computational Fluid Mechanics Volume 13, 2018 - Issue 1,
- [8] K. Kurec, M. Remer, J. Piechna, The influence of different aerodynamic setups on enhancing a sports car's braking, International Journal of Mechanical Sciences 164 (2019) 105140, <https://doi.org/10.1016/j.ijmecsci.2019.105140>,
- [9] K. Kurec, J. Piechna, „Influence of Side Spoilers on the Aerodynamic Properties of a Sports Car ...”, Energies 2019, 12(24), <https://doi.org/10.3390/en12244697>

W Europie lansuje się bezpieczną technikę hamowania na wprost przed zakrętem i przejście zakrętu łukiem o stałym promieniu i stałej prędkości, przyspieszając po wyjściu z zakrętu.

Technika amerykańska, mniej bezpieczna, zakłada hamowanie po szerokim wejściu w zakręt, zmniejszając się w miarę zacieśniania zakrętu, przy zmniejszającej się prędkości ruchu i zaprzestaniu hamowania dopiero tuż przed szczytem zakrętu, a następnie natychmiastowe płynne przyspieszenie.

Amerykańska technika jest efektywniejsza. Pozwala na szybsze pokonywanie zakrętów, ale nie pozostawia marginesu bezpieczeństwa. Zastosowanie na samochodzie dodatkowych elementów zwiększających docisk aerodynamiczny daje taki margines bezpieczeństwa.

Na rys. 30 pokazano różnice w trajektoriach ruchu samochodu na zakręcie pokonywanym oboma technikami. Czerwonymi kropkami oznaczono przyspieszenia występujące na poszczególnych fragmentach łuku.

Na rysunku 31. zostały pokazane możliwe do uzyskania przyspieszenia boczne i podczas hamowania dla różnych pojazdów i różnych prędkości jazdy.

Wszystkie pokazane samochody rozwijają prędkości przekraczające 250 km/h.

Pokazany na rys. 31 Mercedes 156 pomimo dużej mocy i małych oporów aerodynamicznych przy dużych prędkościach jazdy nie może przejeżdżać zakrętów z dużą prędkością.

Lotus 72 ma bardziej przyczepne opony i proste płaty z przodu i z tyłu, i jest wyraźnie lepszy na zakrętach.

Samochody specjalnie budowane tak by uzyskiwać bardzo duże siły docisku aerodynamicznego, takie jak Lotus 95, Williams FW15C czy Ferrari F300, wykazują swoją absolutną przewagę przy bardzo dużych prędkościach.

Tylko formalne ograniczenia regulaminowe ograniczają ich możliwości.

Sposoby generowania aerodynamicznej siły docisku to już inna historia.

<https://www.carthrottle.com/post/the-science-behind-somersault-crashes-and-dangerous-aero/>

Kartka z kalendarza

200 lat elektromagnetyzmu

Zbigniew Wiśniewski

Dziś każdy student, a nawet każdy uczeń wie, że zjawiska elektryczne i zjawiska magnetyczne są ze sobą powiązane. W fizyce współczesnej zjawiska magnetyczne i elektryczne opisuje się wspólnie za pomocą teorii elektromagnetyzmu. Pełny opis tych zjawisk zyskujemy dzięki równaniom Maxwella. Ale nie zawsze tak było. Tak naprawdę wiedza o zjawiskach magnetycznych jest starsza niż o zjawiskach elektrycznych.



Hans Christian Ørsted, źródło wikipedia

Najstarszym znanym urządzeniem magnetycznym jest oczywiście igła magnetyczna, znana podobno już w starożytnych Chinach. Dzisiaj urządzenia oparte na zasadzie elektromagnetyzmu są wszechobecne. Najważniejsze z nich to oczywiście silnik elektryczny. Mierniki natężenia i napięcia też wykorzystują fakt, że wokół płynącego przewodnika z prądem powstaje pole magnetyczne, które jest w stanie z kolei obracać cewki.

Oczywiście, ta wiedza nie zawsze była nam dana. Dla przykładu Coulomb, odkrywca prawa oddziaływania dwóch ładunków elektrycznych był zdania, że nie da się powiązać oddziaływań magnetycznych z elektrycznymi. Przełomowym pod tym względem był rok 1820. Wówczas miały miejsce przynajmniej dwa rewolucyjne eksperymenty. Pierwszy był wykonany przez duńskiego fizyka Hansa Christiana Ørsteda a drugi przez André-Marie Ampère.

Eksperyment Ørsteda polegał na umieszczeniu w pobliżu obwodu z prądem kompasu, czyli igły magnetycznej. Źródłem prądu był opisywany w poprzednim numerze stos Volty. W wyniku przepływu prądu nastąpiło odchylenie igły kompasu co było dowodem, że przepływowi ładunku elektrycznego towarzyszy powstanie pola magnetycznego.

Dzisiaj eksperyment ten wydaje się banalny. Powstaje więc pytanie, dlaczego należało czekać aż 20 lat, aby go wykonać? Czyżby ówczesni przyrodnicy nie byli zainteresowani taką tematyką? Otóż wręcz przeciwnie. Związku pomiędzy polem elektrycznym a magnetycznym zaczęto poszukiwać od momentu odkrycia stosu Volty albo mówiąc prościej – od odkrycia ogniwa elektrycznego, a jeszcze prościej od odkrycia baterii.

Tymczasem musiało minąć około 20 lat zanim wspomniane badania zakończyły się sukcesem. Rozstrzygają-

cych rezultatów nie dały przeprowadzone w 1802 roku eksperymenty w wykonaniu Włocha Giano Domenico Romagnostiego, choć był on zdecydowanie bliski celu. Jean-Nicholas Pierre Hachette i Bernard Charles Desormes usiłowali zaobserwować wytwarzanie pola magnetycznego przez baterię zbudowaną z 1480 ogniw. Ich wynik był negatywny.

Dziś możemy powiedzieć, że gdyby zwarli obwód zostaliby odkrywcami elektromagnetyzmu. Tę zmianę właśnie wykonał Oersted. Prąd płynący w przewodniku odchyłał igłą magnetyczną. Odkrycie Oersteda zaowocowało lawiną odkryć i wynalazków. Jeszcze w tym samym roku François Dominique Arago (1786-1853) zauważył, że płynący przez przewodnik prąd przyciąga opilki żelazne, które po wyłączeniu prądu odpadały od przewodnika. Wykazał on w ten sposób, że prąd w przewodniku magnetyzuje żelazo.

Ampère (1775-1836) zasugerował, że efekt namagnesowania będzie znacznie większy, jeśli przewodnik zwiniemy. I w ten sposób stworzono pierwowzór elektromagnesu. Za wynalazcę elektromagnesu uchodzi jednak angielski inżynier elektryk William Sturgeon (1783-1850). Opracowane przez niego urządzenie zawierało, oprócz pokrytego izolacją przewodnika, rdzeń żelazny. Obecnie elektromagnesy są obecne w naszym życiu, lecz praktyczne najczęściej niezauważane.



André-Marie Ampère, źródło wikipedia

Jak podaje Witold Jaszczuk, we współczesnym gospodarstwie domowym jest od kilku do kilkunastu sztuk elektromagnesów. Elektromagnesy są obecne w sprzęcie audiowizualnym, komputerach, pojazdach mechanicznych itp. Szacuje się, że rocznie wytwarza się kilka miliardów elektromagnesów, są to głównie magnesy zasilane prądem stałym. Dodatkowo należy podkreślić, że elektromagnesy odegrały niebagatelną rolę w badaniach nad strukturą materii jako elementy akceleratorów – urządzeń do przyspieszania cząstek naładowanych. Same akceleratory poza zastosowaniami naukowymi znalazły z kolei zastosowanie również w medycynie w tym w terapii i diagnostyce nowotworów.

Wspomniany tu Ampère wykonał w 1820 roku jeszcze jedno fundamentalne doświadczenie. Otóż zademonstrował on wzajemne przyciąganie się przewodników. Wykrył on opracowany przez siebie przyrząd noszący nazwę

dokończenie na str. 39

Epidemie i fizyka jądrowa

Jerzy Kuczyński

1 W czasie epidemii, gdy się ciągle siedzi w domu po głowie chodzą różne myśli. W tym oczywiście wspomnienia. Mnie co chwilę przypomina się jeden z Kolegów. Starszy człowiek, pewnie w moim obecnym wieku, bo było to w czasach, gdy jeszcze pływałem jako oficer a on był moim załogantem. Tak dawno, że nie pamiętam ani jego imienia, ani nazwiska. Na pewno nie był to sztormowy rejs, bo „godziny wachty wolno płynęły” i nudę zapępiało się rozmowami. Załogant z zawodu był leśnikiem i opowiadał jak to w lesie bywa. W szczególności zapamiętałem opowieść o wścieklicznie.

2 Wściekliczna jest w lesie stale. I nie ma problemu. Po prostu trzeba uważać jak się zwierzę zachowuje. Prosta reguła – jak zwierzę się ciebie nie boi to znaczy, że Ty masz się bać jego. W zdrowym, prawidłowo zarządzanym lesie ilość chorych zwierząt jest niewielka i lekko oscyluje w czasie. Raz chorych jest nieco więcej a drugi raz ich liczba maleje niemal do zera. Podobnie oscyluje populacja zwierząt choć z nieco inną amplitudą i co ważniejsze w innej fazie.

3 Kiedy zwierząt jest dużo to po pewnym czasie rośnie liczba chorych. Zwierzęta padają i populacja maleje. Po pewnym czasie liczba zarażonych też maleje, bo gdy maleje populacja, czyli gęstość osobników, „odległość” między nimi rośnie, zarażeń jest mniej i populacja się odradza. To nie są wielkie wahania. Jeżeli się nie jest na co dzień w lesie, to trudno to nawet zauważyć. Po prostu, gdy zarażone zwierzę, zanim padnie, zaraża średnio 0,9 zwierzęcia (10 chorych zaraża 9 następnych),

liczba padnięć maleje a, że zwierzęta się rozmnażają więc gęstość rośnie i jedno, dwa pokolenia później chore zwierzę zaczyna zarażać 1,1 osobnika. Bo zwierząt jest więcej i o spotkanie łatwiej. Liczba upadków jest większa i populacja znowu maleje.

I tak to sobie spokojnie oscyluje. Pod warunkiem, że nie zmienia się jakieś uwarunkowania. A widzisz, mówił mi ów Załogant, był zwyczaj, że gajowi mogą polować na drobne drapieżniki takie jak kuny, łaski czy tchórze. Drapieżnik uchodzi za szkodnika a upolować go trudno. Samo znalezienie nie jest łatwe. Ale gajowy jest dzień w dzień jest w lesie. Zauważy jakieś ślady. Przyjrzy się dokładniej. Wytropi, zlokalizuje i w końcu złapie. Dużo tego nie upoluje, ale minie rok, dwa i dla żony na futro będzie.

Myśliwi się oburzali, że to kłusownictwo. Prawda, to kłusownictwo, ale ktoś lub coś musi kontrolować liczebność drobnych drapieżników. Naturalnych wrogów brakuje więc myśliwi zobowiązali się, że będą na nie polować a gajowym tego zabroniono. Ale taki myśliwy, nawet jak bardzo napalony na polowanie to jest w lesie co drugi tydzień. A znaleźć tchórza czy łaskę nie łatwo.

Nawet jak się znajdzie to kto ma tyle cierpliwości, żeby to upolować. Trzeba na to stracić sporo czasu a trofeum marne. Nawet za dziesięć lat niedzielnego polowania na futro może nie starczyć. Więc upolowanych sztuk było niewiele i drapieżniki się namnożyły. Oczywiście „wtrąciła się” wściekliczna. Tym razem jeden osobnik zarażał już kilka i drapieżniki prawie wyginęły. Po pewnym czasie wszystko wróciło do normy. Ale to była już inna norma. Drapieżników bardzo mało, ale za to gryzoni dużo.

Wspomnienie powracało wielokrotnie i dość natrętne, bo podobnych zdarzeń było sporo. Już nie tych opowiadanych na wachcie, ale spotykanych w życiu. Paręnaście lat temu, gdy szedłem w marcu na pomiary meteorologiczne ploszyłem nawet kilkanaście zajęcy przy niewiele ponad stu metrowej drodze. Pojawiła się jakaś epidemia i teraz zajęcia widzę raz, dwa razy w roku! Kuropatwy już nie widziałem od lat. Bażanty kiedyś widziałem codziennie, teraz nawet nie raz w tygodniu. Nie sadzę by na to miały jakiś wpływ polowania (kłusownictwo) co kiedyś było prawdopodobne, bo czasy się zmieniły i mięso znajduje się w sklepie a nie w lesie. I chyba z tego ostatniego powodu populacje zwierzęce w pewnym momencie wzrosły. Niestety na krótko.

4 Oczywiście to wszystko przychodzi mi na myśl, gdy siedzę w domu z powodu zakazu wychodzenia. I zastanawiam się jak uniknąć choroby. W skali kraju (ludzkości) oczywiście przez zarażanie mniej niż jednej osoby przez przeciętnego chorego. I czasem ponoszą mnie nerwy. Dziś (Wielki Czwartek) ogłoszono, że mamy (od za tydzień!) używać maseczek, oraz że nie musi być to być profesjonalna maska, bo ma chronić nie noszącego ją a otoczenie. Tydzień temu wyczytałem, że taka najlichsza maska zmniejsza „zasięg rażenia” koronawirusa do 20 centymetrów (z dwu metrów a gdzieś czytałem, że nawet z ośmiu!). Tym samym jest bardzo skuteczna. Ale chroni nie noszącego a otoczenie zmniejszając „promień rażenia”!

Zastanówmy się chwilę nad tym. Epidemia to niemal identyczne zjawisko jak to co znamy z opisu reaktora jądrowego czy bomby atomowej – zarażanie to to samo co doskonale nam znana reakcja łańcuchowa. Ilość zakażeń, czyli tzw. współczynnik reprodukcji odpowiada temu, ile nowych rozszczepień produkuje jedno. Zastosowanie masek i izolacji to dokładny odpowiednik rozrzedzenia materiału rozszczepialnego. A więc, przy odpowiedniej skuteczności, odpowiada to przejściu z reżimu (rodzaju paliwa) typu bomba (wyraźnie większy od jednego współczynnik reprodukcji) do reżimu typu reaktor (współczynnik w okolicy jednego). Stosowanie testów i eliminacja (kwarantanna) zarażających nosicieli to odpowiednik stosowania prętów sterujących i awaryjnych pochłaniających neutrony. Z kolei odpowiednikiem zastosowania moderatora są różnego rodzaju obyczaje zwiększające prawdopodobieństwo zajścia reakcji, czyli w tym



COVID-19

przypadku zakażenia. Takie jak np. podawanie ręki czy wylewne przywitania z całowaniem.

Zauważmy, w całej powyższej opowieści nie było ani słowa o odporności na choroby (w reakcji jądrowej odpowiadające przekrojowi czynnemu), leczeniu, szczepionkach itd. Model (bo opowieść o zarażaniu zwierząt to oczywiście model i to z tych najprymitywniejszych) przewidywał prostą regułę – kontakt = śmierć. I nawet przy takiej regule epidemia samoistnie wygasa przy odpowiednim wzroście „odległości” między osobnikami. Jak już wiemy „odległość” rośnie w wyniku eliminacji z populacji zarażających, więc epidemia wcześniej czy później wygaśnie.

5 Samoistne wygaśnięcie jest jednak dosyć kosztowne. W pracy [1] podano szacunki skutków epidemii przy braku przeciwdziałania. Otóż przy współczynniku powielania między 1,4 a 4,0 i okresie zarażenia 3-7 dni, co odpowiada aktualnej (zapewne dość ułomnej!) wiedzy o COVID-19, wg użytego modelu zarazi się między 51 a 98 procent populacji. Dla Polski oznaczałoby to 200 do 380 tysięcy zgonów. Jak widać przepisy kwarantannowe mają sens!

5 Zarówno w fizyce jądrowej jak i w epidemii kluczowy jest „współczynnik powielenia”. Jego oszacowanie jest dosyć trudne i zwykle jest realistyczne dopiero post factum, czyli w momencie, gdy epidemia jest w końcowej fazie. W początkowej fazie, czyli w momencie pisania tego tekstu dostępne są, i to mało wiarygodne, dane jedynie o liczbie zachorowań oraz zmarłych [2].

W epidemiologii nieco doskonalszym odpowiednikiem tego co tu opisałem są modele typu SIR [3]. Modele SIR są opisane równaniami różniczkowymi (równania Kermacka-McKendricka sformułowane w 1927 roku) postaci pierwsza pochodna czasowa jest proporcjonalna do kombinacji liniowej pozostałych zmiennych. Zmienne te to S liczba osób podatnych (wielkość populacji), I liczba chorych (roznoszących chorobę) i R liczba ozdrowieńców (stąd SIR). Konkretnie równania te są postaci

$$\begin{aligned}\dot{S} &= -r \cdot S \cdot I, \\ \dot{I} &= r \cdot S \cdot I - a \cdot I, \\ \dot{R} &= a \cdot I\end{aligned}$$

gdzie r i a są stałymi współczynnikami. Omawianie bliżej tych równań i ich rozwiązań nie ma tu sensu, bo bez trudu można o tym przeczytać (omówienie własności rozwiązań w [1]). Do tego my fizycy sporo własności epidemii znamy z fizyki reakcji jądrowych. Może tylko warto dodać, że odpowiednikiem współczynnika powielenia jest tu

$$R_0 = \frac{r \cdot S_0}{a},$$

gdzie S_0 jest początkową wielkością populacji. Do uzyskania rozwiązania potrzebne są jeszcze warunki brzegowe i wartości stałych. Te ostatnie oczywiście trzeba zmierzyć. Warto zauważyć, że model ma wbudowany warunek

$$\dot{I} + \dot{R} + \dot{S} = 0$$

co oczywiście oznacza, że populacja w trakcie epidemii się nie zmienia (ignorujemy urodzenia, w reaktorze odpowiadające częściowo powielaniu paliwa, czyli produkcji plutonu), a wyleczonych i zmarłych traktuje się jednakowo, jako już nie zarażających!). Dlatego opisanych wcześniej zjawisk typu oscylacyjnego (druga i następne fale epidemii) tu nie ma.

6 Ekonomiczne podejście do problemu sprowadza się do zapewnienia możliwie tanio tej odległości. Nawet teraz maska kosztuje kilka złotych a godzina bezproduktywnego siedzenia domu to kilkadziesiąt złotych straty. I tak się zastanawiam, jakim tytanem intelektu trzeba być, by to zauważyć dopiero po trzech tygodniach.

Jerzy Kuczyński

Wyższa Szkoła Techniczna w Katowicach

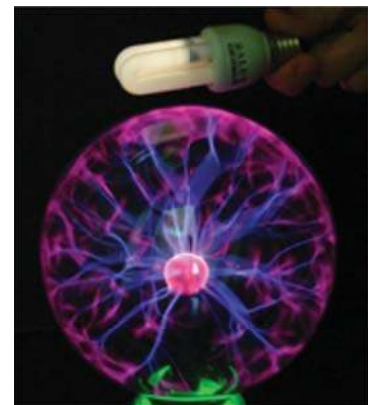
LITERATURA

- [1] A. Kleczkowski, Matematyka i wirusy, PAUza Akademicka 509-510 s. 4. W artykule wbrew tytułowi bardzo opisowo omówiono modele epidemii. Praktycznie cały numer tego czasopisma elektronicznego jest poświęcony epidemii. Czasopismo dostępne w Internecie i zdecydowanie wiarygodne.
- [2] Niestety w trakcie epidemii trudno o rzetelne i nie obciążone bieżącą publicystyką dane. W tym przypadku niemal jedynym źródłem informacji są ulotne teksty Internetu. Tu chciałbym zacytować tekst (nie potrafię go już odnaleźć), w którym podano, że ilość zmarłych w Lombardii (jej części?) wzrosła w stosunku do poprzednich lat o 370%, przy czym ilość stwierdzonych zgonów z powodu COVID-19 była zbliżona do standardowej całkowitej ilości zgonów w poprzednich latach. Sugeruje to oczywiście, że rzeczywista śmiertelność była mocno niedoceniona i wyniosła blisko trzykrotnie więcej od oficjalnie podanej.
- [3] Model dostępny w Internecie. Wystarczy w wyszukiwarce kliknąć „model SIR”.

W następnym numerze

O kuli plazmowej i wyładowaniach w gazach

Kula plazmowa od 20 lat pojawia się na stronach internetowych i okładkach czasopism. Cóż jest więc w niej jeszcze do opisania? Tak naprawdę to nie wiemy, jaki gaz jest w środku, dlaczego przepływa prąd między dwoma szklanymi ściankami i dlaczego pojawiają się takie kolory a nie inne. Z kulą plazmową można zrobić 10 ciekawych doświadczeń i jedno nierozsądne. Jakie? Przeczytacie o tym pod koniec sierpnia.



Siły bezwładności – problem dydaktyczny

Waldemar Reñda

W dziale II w punkcie 8) *Podstawy programowej – zakres podstawowy* – czytamy: „Uczeń wskazuje siłę dośrodkową jako przyczynę ruchu jednostajnego po okręgu, a w punkcie 9) „... rozróżnia układy inercjalne i nieinercjalne, posługuje się pojęciem siły bezwładności.” Natomiast zakres rozszerzony programu fizyki dodaje: „... omawia różnice między opisem ruchu ciał w układach inercjalnych i nieinercjalnych.”

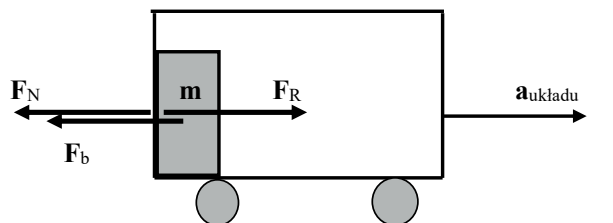
Tematyka ta jest uważana za trudną, zarówno w przekazie lekcyjnym, jak i w uczniowskiej percepcji. I choć z siłami bezwładności mamy na co dzień do czynienia, to jednak interpretacja tych zjawisk jest dość skomplikowana i niejednoznaczna. Zresztą cała kinematyka i dynamika to działy trudne w realizacji zarówno pojęciowo, jak i ze względu na konieczność stosowania dość zaawansowanej matematyki.¹ I to tym bardziej, że zarówno brak jest korelacji programów fizyki i matematyki, jak i konieczność wprowadzenia takich pojęć jak iloczyn skalarny czy iloczyn wektorowy, bez których nie można poprawnie omówić takich zagadnień jak praca różnych rodzajów sił, równowaga bryły sztywnej i wielu innych, jak chociażby występowanie sił Coriolisa.

W latach 80-tych w starszych klasach szkoły podstawowej badano poziom sympatii uczniów w stosunku do różnych przedmiotów. Okazało się, że o ile w klasie VI fizyka należała do przedmiotów na ogół lubianych,² to już w klasie VII znalazła się znacznie niżej w tym rankingu, a przyczyną była właśnie kinematyka i dynamika. Niestety, struktura pojęciowa fizyki jest tak ukształtowana, że działy te rozpoczynają każdy kurs fizyki i musimy sobie z tym jakoś radzić.

Jednym z trudniejszych zagadnień jest problem układów inercjalnych i nieinercjalnych oraz analiza sił opisujących zachowanie ciał w tych układach. Uczeń zdaje sobie sprawę z tego, że opis ruchu zależy od układu odniesienia. Bez trudu zaakceptuje fakt, że zarówno energia potencjalna ciała, jak i jego energia kinetyczna zależy od wyboru układu odniesienia, ale zaczyna się problem, gdy analizujemy działanie sił w zależności od wyboru układu odniesienia.

Rozważmy znany z życia codziennego przykład:

Otosiedzimy w samochodzie, który przyspiesza. (Rys. 1.) (Świadomie rozsunąłem wektory sił, by było widać ich punkty zaczepienia.)



Rys. 1.

Odczuwamy wówczas działanie siły przyciskającej nas do oparcia fotela.

Siłę tę nazwano siłą bezwładności. Na rysunku jest to wektor³ F_b . Pozostałe siły to siła nacisku naszego ciała na oparcie fotela (F_N) oraz – zgodna z III zasadą dynamiki – siła reakcji (F_R). Oczywiście, zgodnie z II zasadą dynamiki: $F_R = m a_{\text{układu}}$. Natomiast: $F_N = F_b = -m a_{\text{układu}}$.

Zauważmy, że w tym układzie wszystkie działające na nas siły się równoważą⁴ i jesteśmy w spoczynku względem samochodu, czyli układu, w którym się znajdujemy.

A jak opisz to ktoś stojący obok ruszającego samochodu? (Rys. 2.)

Powie on, że na pasażera działa siła nacisku F_N ze strony oparcia fotela, która – zgodnie z II zasadą dynamiki Newtona – nadaje pasażerowi przyspieszenie równe przyspieszeniu samochodu, a siła F_R jest reakcją na tę siłę.⁵ W tym opisie nie występuje siła bezwładności.

Jak widać, opisy obu obserwatorów znacznie się różnią, ale tylko w opisie pasażera pojawia się działanie, które nazwaliśmy siłą bezwładności.

Układ, w którym w opisie tego typu zjawisk istnieje potrzeba zastosowania siły bezwładności, nazwano układem nieinercjalnym. Natomiast układ obserwatora zewnętrznego, w którym opis tego zjawiska nie wymaga stosowania sił bezwładności nazywamy układem inercjalnym.

Weźmy inny przykład. Oto dziecko siedzi na krzeselku karuzeli łańcuchowej obracającej się z pewną prędkością kątową. Narysujmy wektory sił działających na to dziecko z punktu widzenia dziecka oraz obserwatora stojącego obok karuzeli. (Rys. 3.)

Lewa strona rysunku obrazuje opis dokonany w układzie obracającej się karuzeli. Występuje tu siła bezwładności (F_b) równa sile dośrodkowej z rysunku prawego, ale o przeciwnym zwrocie.⁶ Zatem: $F_b = -m a_d$. Przy czym a_d jest tu przyspieszeniem dośrodkowym krzeselka, na którym siedzi dziecko. Pozostałe siły to siła ciężkości F_C , F_W – wypadkowa siły bezwładności i siły ciężkości oraz reakcja węzłów F_R .

¹ Problem jest tym poważniejszy, że wymagana jest wiedza, którą uczeń jeszcze nie zdobył na lekcjach matematyki, jak np. funkcje i ich obrazy graficzne a także elementy trygonometrii.

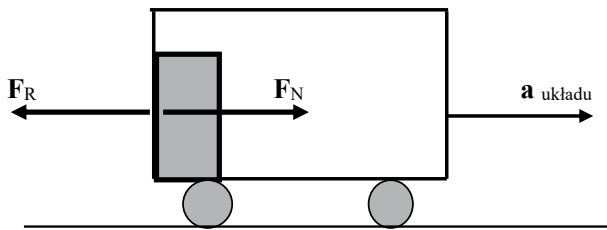
² Realizowano wówczas program zbliżony do programu przyrody, który był realizowany w szkole podstawowej przed ostatnią reformą.

³ Pogrubienie oznacza wielkość wektorową.

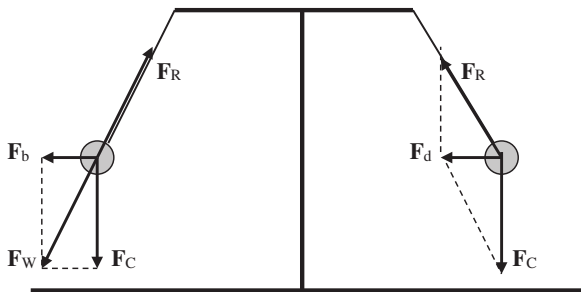
⁴ Pominąłem tu działającą na nas siłę ciężkości, reakcję na nią oraz siły tarcia, ale do naszych rozważań nie była konieczna ich analiza.

⁵ Zauważmy, że to która siła jest akcją a która reakcją zależy od wyboru układu odniesienia.

⁶ Siła ta bywa nazywana siłą odśrodkową bezwładności.



Rys. 2.



Rys. 3.

Zauważmy, że i w tym przypadku siły opisywane w układzie ruchomym (nieinercyjnym) wzajemnie się równoważą, a dziecko odczuwa jedynie zwiększony nacisk na krzesło karuzeli równy wypadkowej siły bezwładności i siły ciężkości.

Po prawej stronie rysunku przedstawiłem układ sił opisany przez obserwatora znajdującego się w układzie zewnętrznym – inercyjnym. W tym przypadku siły się nie równoważą, a siłą wypadkową jest siła dośrodkowa (F_d) zapewniająca ruch dziecka po okręgu.

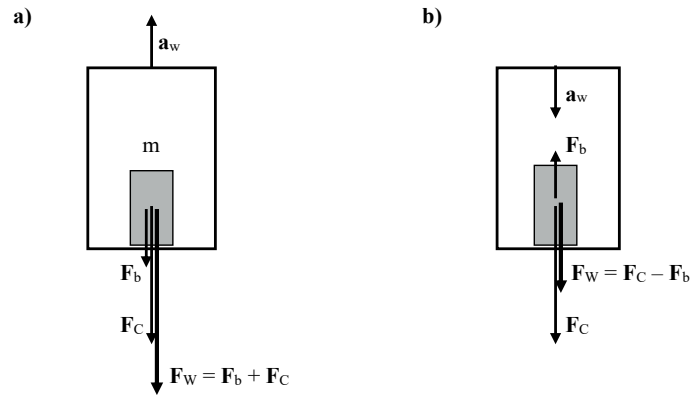
Zauważmy, że w przypadku użycia w opisie siły bezwładności uzyskaliśmy równowagę sił. Zauważył to d'Alembert,⁷ a zasada ta została potem nazwana jego imieniem. Zwykle formułowana jest następująco: „Suma działających na ciało sił wraz z siłą bezwładności równa jest zero”. Oczywiście dotyczy to opisu dokonanego z pozycji obserwatora znajdującego się w układzie nieinercyjnym.

I jeszcze jeden przykład:

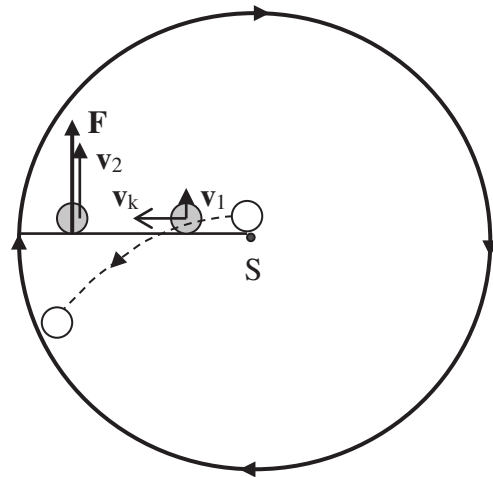
Oto jesteśmy w windzie, która rusza w górę. Odczuwamy wówczas chwilowy wzrost siły ciężkości.⁸ (Rys. 4a) Natomiast, gdy winda zatrzymuje się na wyższym piętrze (Rys. 4b), to odczuwamy chwilowe jej zmniejszenie.⁹ Jak to zjawisko opisać osoba znajdująca się w windzie, a jak będąca poza nią – w układzie inercyjnym? (Wektory sił rozsunąłem, by każdy z nich był widoczny, ale w rzeczywistości działają wzdłuż tej samej prostej.)

Oznaczenia: F_b – siła bezwładności, F_c – siła ciężkości, F_w – siła wypadkowa. W obu przypadkach: $F_b = -m a_{\text{windy}}$.

Opuściłem wektory reakcji podłoża, by nie komplikować rysunku. A siły te są równe co do wartości siłom wypadkowym, lecz mają przeciwne zwroty.



Rys. 4.



Rys. 5.

A jak to opisać osoba z układu inercyjnego? Otóż stwierdzi ona, że w przypadku a) zwiększa się nacisk podłoża na nasze stopy, by nadać nam przyspieszenie windy (a_w). Zatem: $F_N = m g + m a$. Natomiast w przypadku b) nacisk podłoża zmniejsza się o siłę nadającą naszemu ciału przyspieszenie (a) równe przyspieszeniu windy (a_w). Zatem: $F_N = m g - m a$.

Przyznam, że wolę opis dokonany z układu windy. Wolą go też uczniowie, bo lepiej odzwierciedla to, co wówczas czują. Z praktyki nauczycielskiej wiem, że dla uczniów na ogół bardziej przystępny jest opis tego typu zjawisk, gdy stosuje się siłę bezwładności, a więc w układzie nieinercyjnym.

A teraz co by było, gdyby nasza winda spadała z przyspieszeniem ziemskim? (Rys. 4b) Wówczas siła bezwładności zrównałaby się z siłą ciężkości i zniknąłby nasz nacisk na podłogę windy. Znaleźlibyśmy się więc w **stanie nieważkości**.¹⁰ Czy zatem moglibyśmy stwierdzić, że nasz układ jest nieinercyjny? Nie czujemy przecież

⁷ Jean de Rond d'Alembert (1717–1783) – francuski matematyk, członek Królewskiej Akademii Nauk w Paryżu. Warto zapoznać się z jego ciekawym życiorysem. Zob. A. K. Wróblewski, Historia fizyki, Wydawnictwo Naukowe PWN, Warszawa 2006, s. 196.

⁸ To samo odczuwamy, gdy winda zatrzymuje się na niższym piętrze.

⁹ To samo odczuwamy, gdy winda rusza w dół.

¹⁰ Byłoby podobnie, gdybyśmy znaleźli się w statku kosmicznym poruszającym się swobodnie (bez działania silników) w dowolnym polu grawitacyjnym. W obu przypadkach należałoby bowiem zbadać, czy ciało poruszające się ruchem jednostajnym i swobodnym względem tego układu, porusza się w tym układzie prostoliniowo. W obu tych przypadkach byłoby to bardzo trudne ze względu na ograniczone rozmiary kabin. Dodam, że jeżeli nasz układ poruszałby się swobodnie w jednorodnym polu grawitacyjnym, to zachowywałby się jak układ inercyjny.

żadnej siły i jesteśmy w spoczynku względem naszego układu odniesienia. Otóż to czy jesteśmy w układzie nieinercyjnym może bez trudu stwierdzić osoba obserwująca nas z pozycji układu inercyjnego. Bowiemy w każdym z wyżej opisanym przypadku stwierdzi, że poruszamy się z przyspieszeniem różnym od zera.

Rodzaje układów

Warto wreszcie powiedzieć nieco więcej o wspomnianych rodzajach układów. Często bowiem I zasadę dynamiki Newtona interpretuje się jako **postulat** istnienia układu inercyjnego. Spotkałem nawet taką jej postać: „Istnieje układ, w którym ciało spoczywa lub porusza się ruchem jednostajnym prostoliniowym, gdy nie działają na nie żadne siły lub też, gdy działające siły równoważą się.” Dodaje się wówczas stwierdzenie, że taki układ nazywamy układem inercyjnym oraz że żadna z w/w sił nie jest siłą bezwładności, która uważana jest za siłę **pozorną**, bo nie można określić jej źródła, jak to jest np. w przypadku oddziaływania grawitacyjnego czy elektromagnetycznego.

Dodam, że z powyższej definicji wynika, iż każdy układ, który spoczywa względem innego układu inercyjnego, lub porusza się względem niego ruchem jednostajnym i prostoliniowym jest też układem inercyjnym i takich układów może być (teoretycznie) nieskończenie wiele.

Powyższą postać I zasady dynamiki przytoczyłem za podręcznikami uniwersyteckimi, ale spotkałem ją również w podręczniku szkolnym. Uważam, że jest to dydaktycznie wątpliwe, gdyż taka jej postać jest raczej definicją układu inercyjnego i wywołuje poważne kłopoty interpretacyjne, choćby ten czy rzeczywiście **istnieje** układ inercyjny. Ale o tym napiszę nieco później. Natomiast osobiście wolę już postać zaproponowaną przez I. Newtona: „Każde ciało trwa w stanie spoczynku lub ruchu jednostajnego i prostoliniowego, dopóki siły przyłożone nie zmuszą ciała do zmiany tego stanu.”¹¹ Jakże pięknie ujawnia się tu cecha ciał obdarzonych masą zwana **bezwładnością**. Nic więc dziwnego, że zasadę tę nazwano właśnie **zasadą bezwładności**.

A teraz powrócę do pytania, czy istnieje układ w pełni inercyjny?

O ile stół bilardowy stojący w osiedlowym klubie może być układem inercyjnym dla poruszających się po nim bil czy stadion piłkarski dla graczy i piłki, to już powierzchnia Ziemi i ciała trwale z nią związane nie tworzą układu inercyjnego, bo – oprócz biegunów – wszystkie punkty z nim związane poruszają się po okręgu. Ziemia jako całość również takim układem nie jest, bo porusza się po orbicie eliptycznej wokół Słońca. A i Słońce wraz z Układem Planetarnym bierze udział w ruchu wirowym naszej Galaktyki. Wprawdzie przyspieszenie dośrodkowe tego ruchu jest bar-

dzo małe, ale niezerowe. A przecież każdy układ poruszający się względem układu inercyjnego z przyspieszeniem różnym od zera nie jest układem inercyjnym.

Na szczęście zjawiska, które na co dzień obserwujemy przebiegają na tak niewielkich fragmentach powierzchni Ziemi, że układy odniesienia z nimi związane można uznać za układy inercyjne. Inaczej jest, gdy obserwujemy zjawiska zachodzące na znacznych obszarach Ziemi jak np. zachodnie odchylenie się pasatów. Tu zaczynają dochodzić do głosu zjawiska, które opisujemy z użyciem sił bezwładności Coriolisa.

Dla uproszczenia opiszmy ruch kuli toczącej się prędkością (v_k) wzdłuż promienia obracającej się poziomo tarczy ze stałą prędkością kątową (ω) (Rys. 5).

Na tak toczącą się kulę działa siła (F) ze strony podpórki, gdyż kula musi nieustannie zwiększać swą prędkość obwodową (na rys. od v_1 do v_2 itd.).

Kula zatem „czuje”, że jest dociskana do owej podpórki. Siła ta to właśnie siła bezwładności Coriolisa (F_C). Jej **wartość**¹² jest proporcjonalna do prędkości kuli (v_k) oraz prędkości kątowej tarczy (ω). Zależy też oczywiście od masy tej kuli.

Dokładniej: $F_C = -F = -2m\omega \times v_k$, gdzie iloczyn $\omega \times v_k$ jest iloczynem wektorowym wektorów ω i v_k .

Gdyby nie podpórka ograniczająca swobodny ruch kuli, to kula ta zbaczałaby w kierunku przeciwnym do ruchu tarczy (linia przerywana). I tak właśnie zachowują się pasaty. Siła Coriolisa wywołuje też wschodzenie odchylenie swobodnie spadających ciał, ale jest ono na ogół niezauważalne, a jego opis jest identyczny z podanym wyżej.

Komentarz

Napisałem wyżej, że siły bezwładności uważa się za siły pozorne. Czy słusznie? Aby odpowiedzieć na to pytanie należy najpierw ustalić, co w fizyce uważa się za rzeczywiste, czyli obiektywnie istniejące.

Fizyk, chcąc badać przyrodę, musi założyć, że przyroda istnieje obiektywnie i jest poznawalna. Przyjmuje się więc, że realne są obiekty fizyczne, ich stany oraz relacje między nimi. Na tej podstawie tworzy się system opisujący tę rzeczywistość. System ten to zbiór wewnętrznie niesprzecznych i doświadczalnie weryfikowalnych pojęć i praw, opisujących właściwości ciał, zjawisk i procesów zachodzących w rzeczywistości przyrodniczej.

Podstawą języka fizyki są *pojęcia*, a więc nazwy ogólne ciał, zjawisk i procesów fizycznych. Ważną grupą tych pojęć są *wielkości fizyczne*, które są mierzalnymi cechami ciał, zjawisk i procesów fizycznych oraz relacji zachodzących między nimi. Zauważmy jednak, że wielkościom fizycznym nie odpowiadają żadne realne byty fizyczne.

Profesor Kurt Gödel¹³ udowodnił, że w każdym systemie logicznym muszą istnieć pojęcia niedefiniowalne oraz

¹¹ Tłumaczenie z łacińskiego oryginału w Principiach.

¹² Kierunek i zwrot tego wektora określa reguła śruby prawoskrętnej.

¹³ Kurt Gödel (1906–1978), matematyk austriacki, od 1940 r. w USA, członek Narodowej Akademii Nauk w Waszyngtonie. Powyższe stwierdzenie jest obecnie fundamentem logiki matematycznej.

¹⁴ Newton, zapytany, co to jest grawitacja, odpowiedział: „Nie wiem.” Oprócz realnie istniejących pól fizycznych – np. grawitacyjnego czy elektromagnetycznego tworzy się też matematyczne pola wielkości fizycznych jak np. pole potencjału itp.

nieweryfikowalne stwierdzenia. W fizyce takim pojęciami są np.: czas, przestrzeń czy pole fizyczne,¹⁴ a niedefiniowalnym jest np. stwierdzenie, że występują wzajemne oddziaływania ciał.¹⁵

W naszym przypadku zakładamy, że realnie istniejącymi obiektami są ciała fizyczne oraz zachodzące między nimi oddziaływania, które poznajemy i opisujemy analizując ich skutki. Natomiast siła to pojęcie utworzone dla opisu owych oddziaływań i ich skutków. Jest to wielkość fizyczna będąca miarą wzajemnych (mechanicznych) oddziaływań ciał. Jednym ze skutków oddziaływań jest np. zmiana ruchu ciała. Zjawisko to opisuje wielkość fizyczna zwana przyspieszeniem. Zależność między takimi wielkościami fizycznymi jak np. masa ciała i siła oraz wywołane nią przyspieszenie opisuje II zasada dynamiki Newtona.

Przeanalizujmy pod tym kątem opisane wyżej przypadki.

W pierwszym z nich pasażer czuje działanie przyciskające go do fotela. I jakkolwiek jest ono podobne do działania grawitacji, to jednak nie można wskazać jego źródła. Zauważmy też, że pasażer jest w spoczynku względem samochodu. Zgodnie więc z I zasadą dynamiki wszystkie oddziaływania związane z pasażerem powinny się w tym układzie odniesienia wzajemnie równoważyć. Pomijając oddziaływanie grawitacyjne i reakcję podłoża, musi on założyć, że pojawiło się działanie równe naciskowi fotela, lecz przeciwnie do niego zwrócone. Będąc zatem w tym układzie, jest on zmuszony do użycia pojęcia dodatkowej siły, którą z racji tego, że wiąże się z cechą ciał zwaną bezwładnością,¹⁶ nazwano *siłą bezwładności*. Podobnie jest w przypadku dziecka

na karuzeli. W tym przypadku ową siłę nazywa się zwykle *siłą odśrodkową bezwładności*.

A co stwierdzi osoba znajdująca się w windzie, która zaczyna swobodnie spadać w ziemskim polu grawitacyjnym. Zauważy ona zanik oddziaływania grawitacyjnego oraz brak wszelkich oddziaływań z windą. Zanik oddziaływania grawitacyjnego może owa osoba wyjaśnić jedynie pojawieniem się działania przeciwnego do grawitacyjnego. Podobnie jest też np. w statku kosmicznym swobodnie poruszającym się w niejednorodnym polu grawitacyjnym.

Czy zatem siła bezwładności jest pojęciem odnoszącym się do realnego działania? Należy raczej odpowiedzieć, że – nie. Czy zatem należałoby z niej zrezygnować? Można, ale wówczas opis zjawisk musi być dokonywany wyłącznie z pozycji układu inercjalnego. Jest to jednak na ogół trudniejsze i bardziej odległe od naszych codziennych obserwacji i odczuć.

W swoich wywodach posługiwałem się pojęciami „działanie” i „oddziaływanie”. Wygodniejsze jest jednak użycie wielkości fizycznej będącej miarą owego działania, czyli siły. Można wówczas opisać wszystkie cechy tego działania: że jest ono skierowane, że ma pewną wartość liczbową, że jest przyłożone do całego ciała lub jednego z jego punktu, ewentualnie do jednej z jego powierzchni itp., itd. Można wówczas również owe działania porównywać, sumować czy też analizować ich związki z innymi cechami ciał i zjawisk fizycznych, czyli z innymi wielkościami fizycznymi.

Waldemar Reńda

Olkusz, luty 2020

¹⁵ W przyrodzie występują oddziaływania grawitacyjne i elektromagnetyczne, które dokonują się poprzez pola oraz oddziaływanie bezpośrednie poprzez siły sprężystości, które są odmianą oddziaływań elektromagnetycznych zachodzących pomiędzy atomami lub cząsteczkami ciał.

¹⁶ Masę definiujemy zwykle jako miarę bezwładności ciała. I. Newton definiował ją jako ilość (ale nie licznosc) materii w ciele.

Kartka z kalendarza

200 lat elektromagnetyzmu

dokończenie ze str. 33

stolika Ampera. Lepsza właściwie nazwą byłoby waha-dło. Jest to urządzenie, które w wielu wersjach możemy spotkać dziś w pracowniach szkolnych. Jego zasadniczą częścią jest wiszący drut, może być zwinięty do postaci cewki. Należy go oczywiście podłączyć do źródła napięcia/prądu. Jeśli do takiej cewki zbliżymy przewodnik, lub drugi drut, przez który będzie płynąć prąd, to nasz drut ulegnie odchyleniu, przy czym kierunek tego odchylenia jest zależny od wzajemnej orientacji płynących prądów. To zjawisko przez wiele lat zresztą służyło do definiowania podstawowej jednostki układu SI – ampera.

Ampère prowadząc badania metodą prób i błędów podał wzór określający siłę oddziaływania dwóch przewod-

ników. W tym samym roku Jan Batiste Biot i Feliks Savart podali kolejny wzór, wiążący siłę oddziałującą na przewodnik z prądem i pole magnetyczne. W tym też roku powstał stworzony – jakże by inaczej – przez Ampera termin elektromagnetyzm.

Rok 1820 był więc rokiem, w którym narodziła się jedna z najważniejszych, jeśli nie najważniejsza gałąź fizyki stosowanej – elektromagnetyzm. Przeprowadzono wówczas nie tylko podstawowe eksperymenty, ale podano też główne zależności matematyczne. Był to rok niezwykłego przełomu, od którego mija dokładnie 200 lat.

LITERATURA

- [1] Andrzej Kajetan Wróblewski, Historia Fizyki, PWN, Warszawa 2006
- [2] George Gamow, Biografia Fizyki, Wiedza Powszechna, Warszawa 1967
- [3] Witold Jaszczuk, Elektromagnesy prądu stałego dla praktyków, Wydawnictwo BTC, Legionowo 2014.

Zastosowanie metody potencjałów węzłowych do obliczania obwodów elektrycznych

Czesław Surowiec

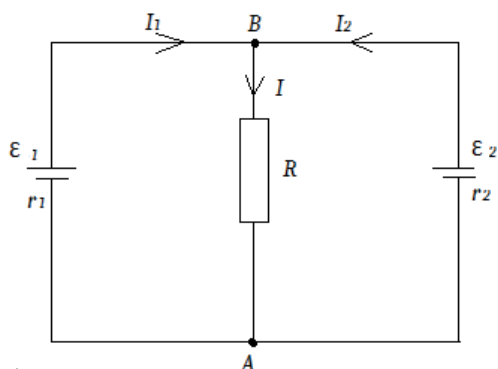
Do obliczeń obwodów elektrycznych zazwyczaj stosujemy równania Kirchhoffa. W przypadku złożonych obwodów z dużą ilością gałęzi, równania zawierają dużą ilość niewiadomych, co sprawia duże trudności w ich rozwiązaniu. Wygodnie jest w tym przypadku zastosować metodę potencjałów węzłowych, którą stosuje się w elektrotechnice. Metoda potencjałów węzłowych opiera się na:

Jako niewiadome przyjmujemy potencjały węzłów obwodu.

1. Do obliczenia natężenia prądu w gałęziach obwodu niezbędne jest obliczenie różnicy potencjałów między nimi.
2. Aby uprościć obliczenia różnicy potencjałów w gałęziach, przyjmujemy potencjał jednego z węzłów za zerowy.
3. Przyjęcie potencjału jednego z węzłów za zerowy zmniejsza ilość równań o 1.
4. Korzystając z prawa Ohma dla części obwodu, możemy wyrazić natężenia prądu w poszczególnych gałęziach obwodu przez potencjały węzłów, SEM źródeł prądu i opory.
5. Korzystając z I prawa Kirchhoffa uzyskujemy taką samą liczbę równań jak i niewiadomych.
6. Jeśli SEM jest zwrócona w stronę węzła o potencjale V , to przyjmujemy jej wartość za dodatnią, a jeśli w stronę potencjału 0 przyjmujemy za ujemną.
7. Jeśli obliczona wartość natężenia prądu jest ujemna, zwrot natężenia prądu jest przeciwny do obranego.
8. W przypadku większej liczby węzłów wygodnie jest obliczyć wartość V i obliczać natężenia prądów z prawa Ohma dla części obwodu.
9. Metodę tę możemy stosować tylko do obwodów liniowych (dla których zależność I od U jest liniowa).

Przykład 1:

Dla obwodu przedstawionego na rys. 1 oblicz:



Rys. 1

- a) potencjał punktu B
- b) natężenie prądu I.

Rozwiązanie:

Oznaczamy potencjały punktów A i B: $V_A = 0$, $V_B = V$. Różnica potencjałów $V_B - V_A = V_B - 0 = V$.

$$\text{Dla 1 gałęzi } V = \varepsilon_1 - I_1 \cdot r_1 \quad \text{stad } I_1 = \frac{\varepsilon_1 - V}{r_1}$$

$$\text{Dla 2 gałęzi } V = \varepsilon_2 - I_2 \cdot r_2 \quad \text{stad } I_2 = \frac{\varepsilon_2 - V}{r_2}$$

$$\text{Dla 3 gałęzi } V = I R \quad \text{stad } I = \frac{V}{R}$$

Korzystając z I prawa Kirchhoffa $I_1 + I_2 = I$ otrzymujemy

$$\frac{\varepsilon_1 - V}{r_1} + \frac{\varepsilon_2 - V}{r_2} = \frac{V}{R}$$

Rozwiązując to równanie względem V otrzymujemy

$$V = \frac{(\varepsilon_1 r_2 + \varepsilon_2 r_1) R}{r_1 \cdot r_2 + (r_1 + r_2) R}$$

Po podstawienie za V do I otrzymamy

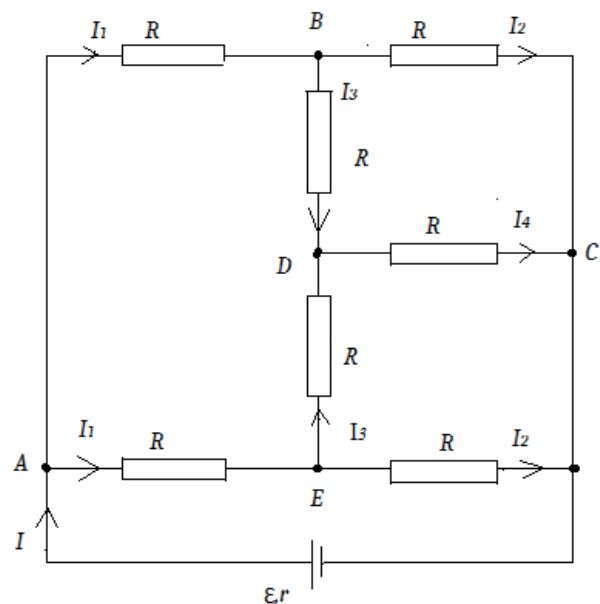
$$I = \frac{\varepsilon_1 r_2 + \varepsilon_2 r_1}{r_1 \cdot r_2 + (r_1 + r_2) R}$$

Przykład 2:

Dla obwodu (Rys. 2) oblicz natężenia prądów w obwodzie, jeśli każdy z oporów wynosi 1Ω a $\varepsilon = 14\text{ V}$. Przyjmij $r = 0$.

Rozwiązanie:

Przyjmujemy, że prądy przez opory płyną jak na rys. 2, oraz wartości potencjałów $V_C = 0$ oraz $V_A = \varepsilon$. Natężenia prądów w gałęziach obwodu AB i AF są jednakowe i wy-



Rys. 2

noszą I_1 , analogicznie mamy w gałęziach BC i EC oraz BD i ED, co możemy uzasadnić symetrią obwodu. Natężenia prądów możemy zapisać następująco:

$$I_1 = \frac{\varepsilon - V_B}{R}, I_2 = \frac{V_B}{R}, I_3 = \frac{V_B}{R}, I_4 = \frac{V_D}{R} \quad ①$$

Z I prawa Kirchoffa otrzymujemy

$$I_1 = I_2 + I_3, 2 I_3 = I_4 \quad ②$$

Podstawiając ① do ② otrzymujemy

$$\frac{(\varepsilon - V_B)}{R} = \frac{V_B}{R} + \frac{V_B - V_D}{R}$$

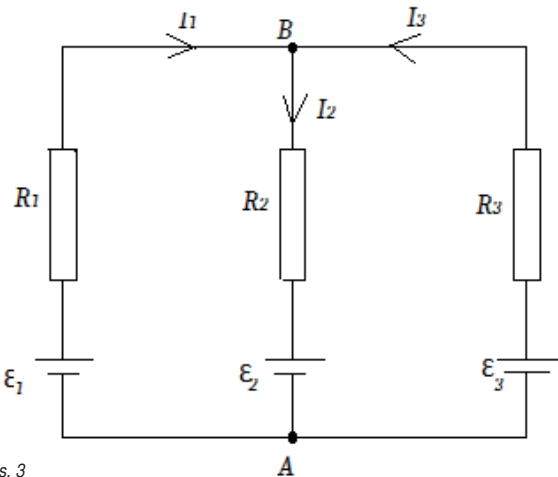
$$2 \left(\frac{V_B - V_D}{R} \right) = \frac{V_D}{R}$$

Rozwiązując ten układ równań otrzymujemy $V_B = 6V$ i $V_D = 4V$

Po podstawieniu wartości potencjałów do ① otrzymujemy $I_1 = 8 A, I_2 = 6 A, I_3 = 2A, I_4 = 4A$

Zadania:

- Dla obwodu przedstawionego na rys. 1 dla $\varepsilon_1 = \varepsilon_2 = 2V$ $r_1 = 0,1 \Omega, r_2 = 0,2 \Omega, R = 0.5 \Omega$ oblicz wartość:
 - Potencjału punktu B. (Odp.: $V = 1,8 V$)
 - Natężenia prądu I, I_1, I_2 . (Odp.: $I = 3,6 A, I_1 = 2,4 A$)
 - Dla jakiej wartości $\varepsilon_1 I = 0$. (Odp.: $I_2 = 1,2 A, \varepsilon_1 = -1 V$)
- Dla obwodu przedstawionego na rys. 3 oblicz natężenia prądów:



Rys. 3

I_1, I_2, I_3 dla $\varepsilon_1 = 30 V, \varepsilon_2 = 60 V, \varepsilon_3 = 180 V, R_1 = 3 \Omega, R_2 = 6 \Omega, R_3 = 12 \Omega$ dla wszystkich ogniw $r = 0$

Odp.: $I_1 = -\frac{1}{3} A, I_2 = \frac{7}{3} A, I_3 = 2A$

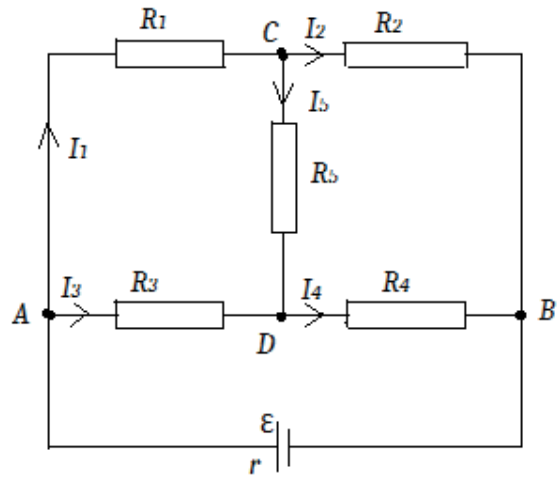
3. Dla obwodu przedstawionego na rys. 4 oblicz:

I_1, I_2, I_3, I_4, I_5 dla: $\varepsilon = 22V, r = 0, R_1 = 1 \Omega, R_2 = R_3 = R_4 = R_5 = 2\Omega$.

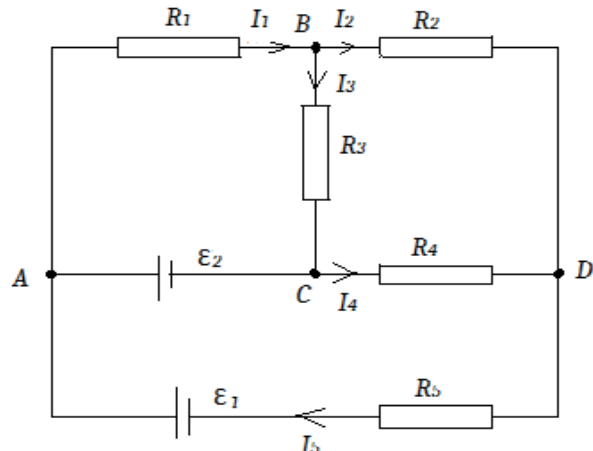
Odp. $I_1 = 8 A, I_2 = 7A, I_3 = 5A, I_4 = 6A, I_5 = 1A$.

4. Dla obwodu przedstawionego na rys. 5 oblicz:

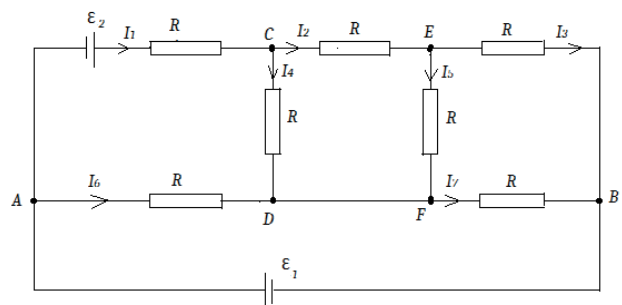
I_1, I_2, I_3, I_4, I_5 dla: $\varepsilon_1 = 65 V, \varepsilon_2 = 39 V, r_1 = r_2 = 0, R_1 = 20 \Omega, R_2 = R_3 = R_4 = 10 \Omega$.



Rys. 4



Rys. 5



Rys. 6

Odp.: $I_1 = 1,7 A, I_2 = 1,2 A, I_3 = 0,5 A, I_4 = 0,7 A, I_5 = 1,9 A$.

5. Dla obwodu przedstawionego na rys. 6, dla $\varepsilon_1 = 10 V, \varepsilon_2 = 30 V$ i $R = 1 \Omega$, pomijając opory wewnętrzne ogniw oblicz natężenia prądów w gałęziach obwodu.

Odp.: $I_1 = 20 A, I_2 = I_3 = I_4 = I_7 = 10 A, I_5 = I_6 = 0$.



Kosmiczny Teleskop Hubble'a (ang. HST – Hubble Space Telescope) to podstawowe narzędzie wykorzystywane w dzisiejszej astronomii obserwacyjnej. Taki stan rzeczy trwa prawie od samego początku działalności Teleskopu, czyli 24 kwietnia 1990 roku, kiedy to został on wyniesiony w przestrzeń kosmiczną przez prom Discovery w trakcie trwania misji STS-31.

30-ta rocznica wyniesienia w przestrzeń kosmiczną

Trzy dekady Teleskopu Hubble'a

Marcin Wesołowski

Jego nazwa, a w szczególności słowo „kosmiczny” wynika stąd, że porusza się on w przestrzeni kosmicznej po orbicie okołoziemskiej. Orbitę, po której porusza się Teleskop Hubble'a możemy zaliczyć do tzw. niskiej orbity okołoziemskiej (ang. LEO – Low Earth Orbit).

Pod pojęciem niskiej orbity okołoziemskiej należy rozumieć obszar przestrzeni kosmicznej przebiegającej od powierzchni Ziemi aż do Pasów Van Allena, czyli wysokości do 2000 km. Warto w tym miejscu dodać jako ciekawostkę, że nad niską orbitą okołoziemską znajduje się średnia orbita okołoziemiska (usytuowana jest ona na wysokości od 3000 km do 30000 km) oraz orbita geostacjonarna, która zapewnia utrzymanie stałej pozycji satelity nad dowolnym punktem równika Ziemi. Oznacza to, że orbita ta jest orbitą kołową położoną w płaszczyźnie równika. W przypadku Teleskopu Hubble'a wysokość od powierzchni Ziemi wynosi około 540 km.

Usytuowanie teleskopów w przestrzeni kosmicznej daje pewnego rodzaju przewagę w porównaniu do teleskopów znajdujących się na Ziemi. Warto w tym miejscu wspomnieć o dwóch kluczowych efektach oddziaływań, które wpływają na jakość otrzymywanych fotografii.

Po pierwsze, fotografie, które wykonujemy z przestrzeni kosmicznej przy pomocy teleskopów lub wysyłanych sond dają nam obraz bardziej szczegółowy w bezpośrednim porównaniu do zdjęć wykonywanych w naziemnych obserwatoriach. Po drugie, możemy zminimalizować wpływ sztucznego światła na jakość wykonywanych fotografii wybranych ciał niebieskich.

Porównując bezpośrednio wyniki obserwacji wykonywane Teleskopem Hubble'a do rezultatów, które uzyskujemy z Ziemi możemy stwierdzić, że Teleskop Hubble'a jest



Foto. 1. Teleskop Hubble'a (Foto. NASA, ESA).

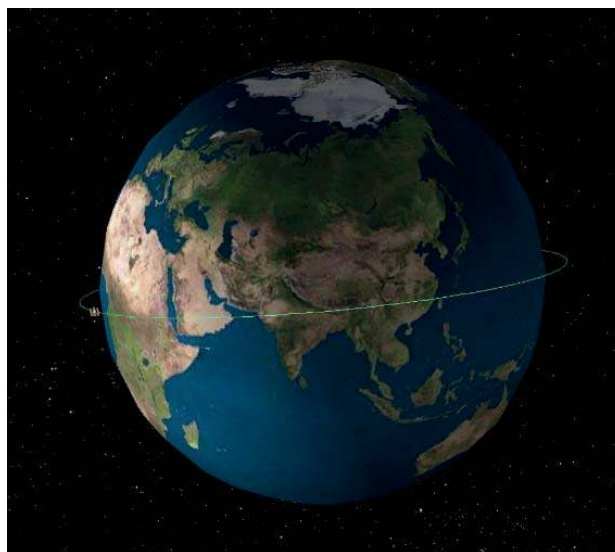


Foto. 2. Orbita Teleskopu Hubble'a – z lewej strony fotografii widoczny jest teleskop HST (Foto. NASA, ESA).

w stanie zarejestrować znacznie więcej szczegółów opisujących dany obiekt. W przypadku teleskopów naziemnych atmosfera ziemska i jej stan, a także aktualnie panujące warunki atmosferyczne stanowią nie lada wyzwanie dla obserwatorów nieba nocnego. Dlatego też obserwatoria astronomiczne te naziemne należy budować w miarę możliwości jak najwyżej.

Tę sugestię, która obowiązuje do dziś wysunął już wieki temu Isaac Newton, który zrozumiał, że ziemska atmosfera zaburza światło pochodzące z kosmosu. To zaburzenie polega na tym, że nasza atmosfera osłabia, częściowo pochłania i rozprasza światło dobiegające do Ziemi, które pochodzi od innych ciał niebieskich. Cały proces zjawisk zachodzących w atmosferze możemy opisać za pomocą pojęcia ekstynkcji atmosferycznej.

Za początek historii Teleskopu Hubble'a przyjmuje się rok 1946, jednakże pierwsze pomysły na skonstruowanie kosmicznego teleskopu były znane nieco wcześniej. Główne prace związane z konstrukcją Teleskopu Hubble'a i jego całym montażem przypadały na lata 70-te i 80-te XX wieku.

Dodajmy jeszcze, że Teleskop Hubble'a został ukończony w 1985 roku z zamiarem jego wyniesienia na orbitę jesienią 1986 roku. Niestety termin ten nie został dotrzymany. Na tą sytuację złożyło się kilka rzeczy:

- przede wszystkim katastrofa wahadłowca Challenger w dniu 28 stycznia 1986 roku w wyniku której zginęła cała siedmioosobowa załoga,
- problemy techniczne i budżetowe.

To wszystko sprawiło, że loty wahadłowców zostały zawieszono do czasu wyjaśnienia przyczyn katastrofy promu kosmicznego Challenger. Wobec tego w naturalny sposób misja wyniesienia Teleskopu Hubble'a w przestrzeń kosmiczną została opóźniona aż o cztery lata.

Dzięki wysiłkowi wielu naukowców w końcu udało się przywrócić program lotów kosmicznych. Podkreślmy jeszcze raz tą wyjątkową i niezwykle ważną datę w historii astronomii obserwacyjnej – 24 kwietnia 1990 roku o godzinie 14:33 czasu polskiego rozpoczęła się misja wyniesienia w przestrzeń kosmiczną Teleskopu Hubble'a.

W tabeli 1 przedstawiono podstawowe parametry fizyczne Teleskopu Hubble'a.

Pomimo niewątpliwie wielkiego sukcesu jakim było wyniesienie Teleskopu Hubble'a w przestrzeń kosmiczną, na samym początku doszło do rozczarowania rezultatami jakie otrzymywano. Okazało się bowiem, że zwierciadło główne miało nieprawidłowy profil a fotografie, które otrzymywano były po prostu nieostre i rozmyte.

Na podstawie pomiarów okazało się, że zwierciadło główne teleskopu było odchylone zaledwie o $2 \cdot 10^{-3}$ milimetra na obwodzie. Nawet tak małe odchylenie wprowadzało już znaczną aberrację sferyczną. W tym miejscu należy przypomnieć Naszym czytelnikom, że pod pojęciem aberracji sferycznej należy rozumieć pewnego rodzaju cechę danego układu optycznego (np. teleskopu) polegającą na tym, że w miarę oddalania się od osi układu optycznego promienie zaczynają odchyłać się od ogniska. Wówczas zamiast otrzymywać obraz punktowy otrzymujemy obraz rozciągly – plamkę.

Tabela 1. Charakterystyka fizyczna kosmicznego Teleskopu Hubble'a.

ORBITA	
Wysokość w perygeum (km)	537.00
Wysokość w apogeum (km)	540.90
Okres obiegu (min)	95.42
Kąt nachylenia (°)	28.47
Mimośród (-)	0.000283
Prędkość (km·h ⁻¹)	7.59
CZAS TRWANIA	
Początek misji (rok)	24.04.1990
Powrót do atmosfery (rok)	2030 – 2040
WYMIARY	
Kształt	cylindryczny
Średnica (m)	4.20
Długość (m)	13.20
Masa (kg)	10863
Teleskop główny	
Rodzaj	Teleskop Ritcheya – Chrétiena
Średnica (m)	2.40
Długość ogniskowej (m)	57.60
Obszar zbierania (m ²)	4.525

Tabela 2. Charakterystyka misji serwisowo – naprawczych teleskopu Hubble'a.

Nazwa misji	Rok	Cel podstawowe
Misja serwisowa 1 – STS 61	Grudzień 1993	– korekta układu optycznego, – instalacja kilku nowych instrumentów i nowego sprzętu
Misja serwisowa 2 – STS 82	Luty 1997	– montaż rejestratora półprzewodnikowego oraz naprawa izolacji termicznej
Misja serwisowa 3A – STS 103	Grudzień 1999	– wymiana żyroskopów oraz odnowa teleskopu
Misja serwisowa 3B – STS109	Marzec 2002	– instalacja nowych instrumentów naukowych np. Advanced Camera for Surveys – przegląd oraz konserwacja teleskopu
Misja serwisowa 4 – STS 125	Maj 2009	– zainstalowano zamienną jednostkę przetwarzania danych, naprawiono systemy ACS i STIS, zainstalowano ulepszone akumulatory wodorowo-niklowe i wymieniono inne elementy

W wyniku powstałego problemu, w celu jego eliminacji, rozważano nawet ponowne sprowadzenie teleskopu na Ziemię. Jednakże względy ekonomiczne, a przede wszystkim olbrzymi rozgłos w mediach sprawiły, że Teleskop Hubble'a jako pierwszy przyrząd astronomiczny został serwisowany i naprawiany bezpośrednio w przestrzeni kosmicznej. Dotychczas przeprowadzono pięć misji serwisowo-naprawczych Teleskopu Hubble'a, które przedstawiono w tabeli 2. Warto zaznaczyć, że konstrukcja Teleskopu Hubble'a, została tak zaprojektowana, aby można było dokonywać napraw, wymiany poszczególnych jego podzespołów, które uwzględniały także postęp technologiczny.

Pomimo trzech dekad działalności Teleskop Hubble'a dzięki wspomnianym misjom serwisowy jest nadal jednym z najważniejszych teleskopów wykorzystywanych w astronomii. Poszczególne misje serwisowe niewątpliwie przyczyniły się do ogromnego sukcesu i zwiększenia



Foto. 3. Sznur perel powstały po uderzeniu komety Shoemaker-Levy 9 w powierzchnię Jowisza – fotografia wykonana przez Teleskop Hubble'a w dniu 17 maja 1994 roku (Foto. NASA).

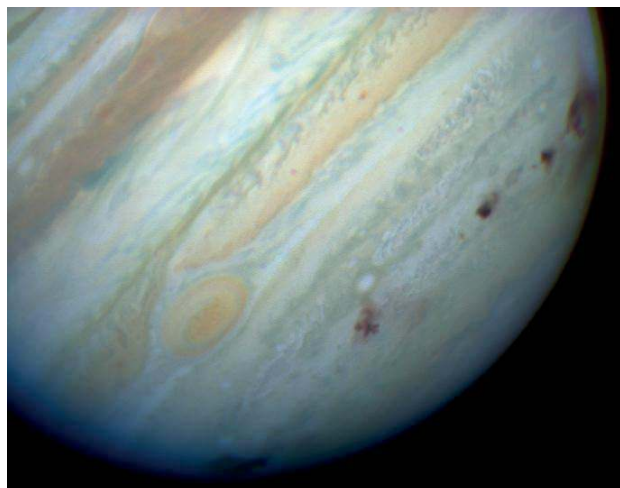


Foto. 4. Ślady na powierzchni Jowisza po kolizji z kometą Shoemaker-Levy 9 (Foto. NASA, ESA).

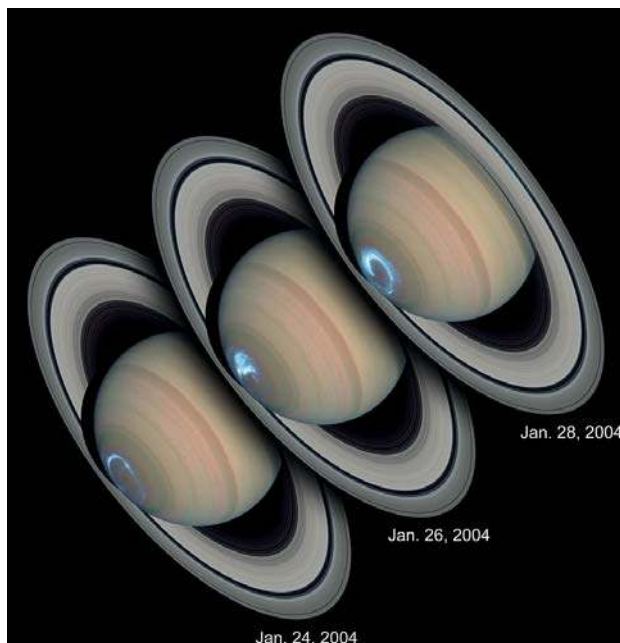


Foto. 5. Zorza polarna na Saturnie (Foto. NASA, ESA).

jego żywotności. Pomimo tych 30 lat działalności nadal w oparciu o Teleskop Hubble'a planowane są kolejne badania naukowe.

Na podstawie fotografii (3-10) przedstawiono najważniejsze osiągnięcia naukowe, które niewątpliwie zawdzięczamy Teleskopowi Hubble'a.

Wykorzystanie Teleskopu Hubble'a w badaniu Układu Słonecznego oraz Galaktyk przyczyniło się do zwiększenia naszej wiedzy o otaczającym nas Wszechświecie. Ponadto sfotografował on najgłębszy kosmos, ukazał dyski



Foto. 6. Mgławica Oriona (Messier 42, M42 lub NGC 1976) widoczna w świetle widzialnym (Foto. NASA, ESA).

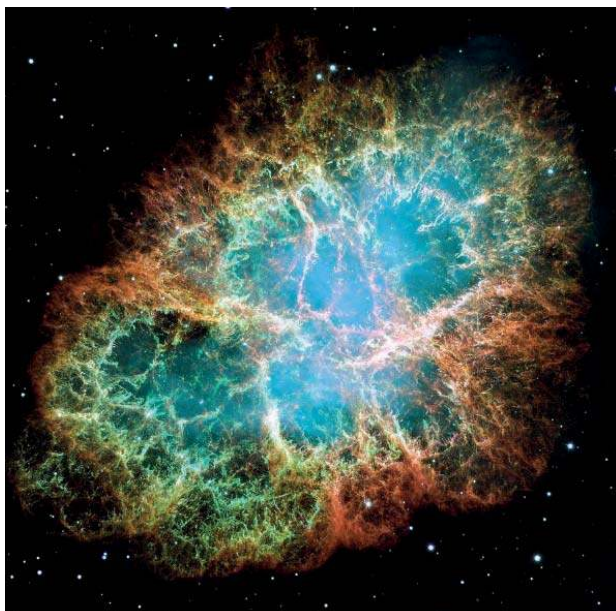


Foto. 7. Mgławica Kraba (Messier 1, M1 czy NGC 1952) – jedno z największych zdjęć jakie wykonał Teleskop Hubble'a (Foto. NASA, ESA).

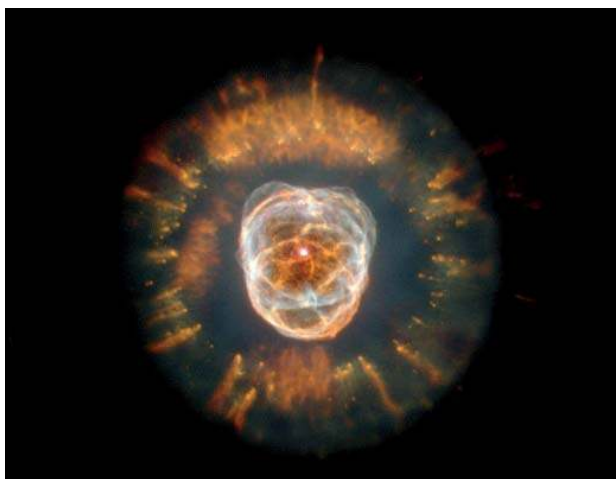


Foto. 8. Mgławica Eskimos (NGC 2392), która położona jest w gwiazdozbiore Bliźniąt (Foto. NASA, ESA).



Foto. 9. Mgławica planetarna Kocie Oko (NGC 6543) położona jest w gwiazdozbiore Smoka (Foto. NASA, ESA).

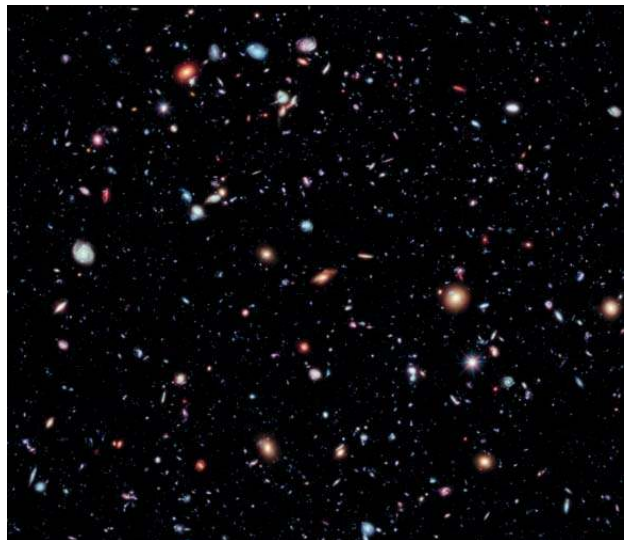


Foto. 10. To co zobaczył Teleskop Hubble'a – czyli jego ekstremalnie głębokie pole widzenia (Foto. NASA, ESA).

protoplanetarne, które mogą być odpowiedzialne za formowanie się planet pozasłonecznych. Zbadał on atmosferę wokół jednej z planet pozasłonecznych, a także jej skład chemiczny.

W przypadku Układu Słonecznego Teleskop Hubble'a badał odległe obiekty między innymi takie jak: Plutona i Eris – planety karłowate. W przypadku komet Teleskop Hubble'a przeprowadził bardzo dokładną obserwację uderzenia komety Shoemaker-Levy 9 o powierzchnię Jowisza (gazowego olbrzyma) w 1994 roku. Jednocześnie przy tej okazji wykonano bardzo dokładne fotografie Jowisza. W przypadku planet gazowych odkryto również zorze polarne nad północnym biegunem Saturna.

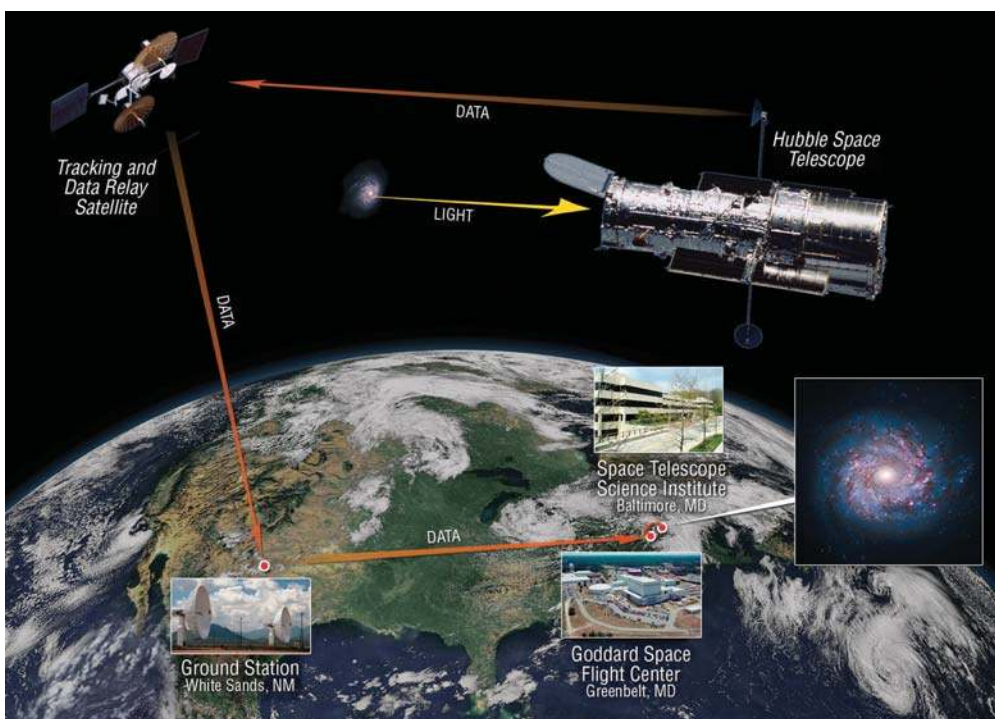
Na koniec naszych rozważań dodajmy jeszcze, że następcą Kosmicznego Teleskopu Hubble'a jest Kosmiczny Teleskop Jamesa Webba (ang. James Webb Space Tele-

scope, JWST), który ma zostać wyniesiony na orbitę najprawdopodobniej w dniu 31 marca 2021 roku. Jednakże teleskop JWST będzie wykonywał obserwacje astronomiczne w paśmie podczerwonym. Prawdziwym planowanym następcą Teleskopu Hubble'a ma być teleskop: Advanced Technology Large-Aperture Space Telescope (ATLAST).

Pomimo tego nadal trzymajmy mocno kciuki za Kosmiczny Teleskop Hubble'a, aby mógł działać jak najdłużej a z pewnością jeszcze nie raz nas zaskoczy.

dr Marcin Wesołowski

Kolegium Nauk Przyrodniczych, Instytut Nauk Fizycznych,
Uniwersytet Rzeszowski,
Centrum Innowacji i Transferu Wiedzy Techniczno-Przyrodniczej Uniwersytetu Rzeszowskiego.



Fot. 11. Ta grafika ilustruje, w jaki sposób obserwacje z Teleskopu Hubble'a, skonwertowane na dane, są przesyłane z teleskopu na ziemię za pośrednictwem satelitarnego systemu śledzenia i przekazywania danych (TDRSS). Źródło: NASA



Promieniowanie kosmiczne (1938-2000)

Zaskakująco wielkie pęki

Tadeusz Wibig

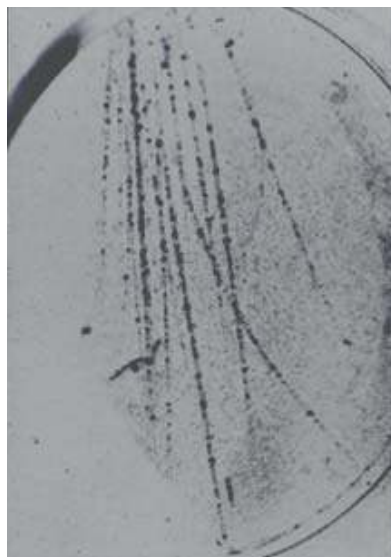
Teoretyczne rozwiązanie zagadki wielokrotnych śladów, które można nazwać „małymi pękami” rejestrowanymi przez Rossiego i widzianych na fotografiach z komory mgłowej przez Skobielcyna i Balcketta i Occhialiniego na początku lat trzydziestych pojawiło się w roku 1937 w pracy Bhabhy i Heitlera i znane jest jako teoria kaskadowa. Znalazła ona pełne i dość zaskakujące potwierdzenie już w roku 1938, kiedy to Pierre Auger i Roland Maze odkryli „wielkie pęki atmosferyczne” (Extensive Air Showers). Idea ich eksperymentu była dokładnie taka sama, jak w doświadczeniu Rossiego z licznikami Geigera-Mullera ustawionymi w wierzchołkach trójkąta. Modyfikacja polegała w zasadzie jedynie na skali.

Dwa liczniki G-M ustawione jeden nad drugim pracujące w koincydencji jako swoisty teleskop dawały

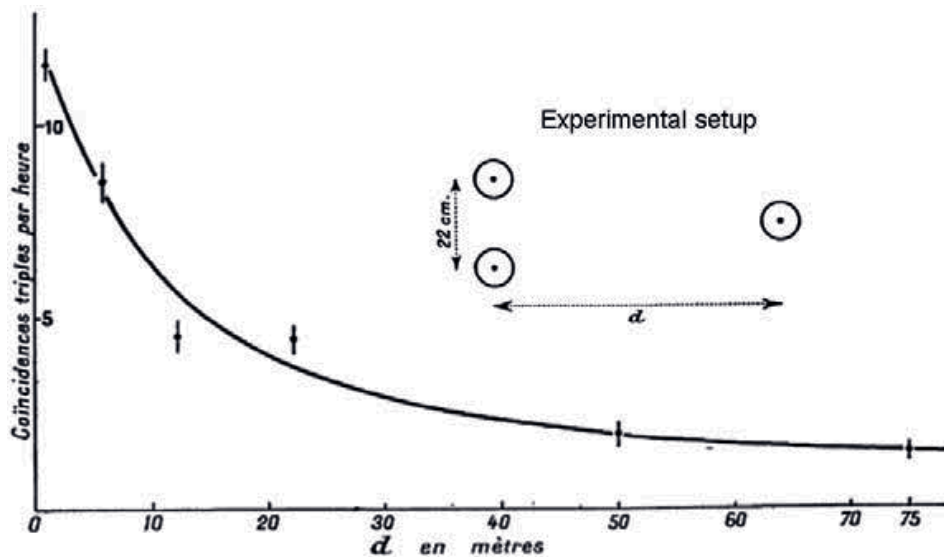
sygnał o przejściu w tym miejscu z grubsza pionowo co najmniej jednej naładowanej cząstki wysokiej energii. Trzeci licznik G-M odsuwano od pionowej pary liczników rejestrując jednoczesne przejście przezeń innej cząstki.

Osiągnięcie Augera i Maze możliwe było jedynie dzięki wielkim umiejętnościom technicznym tego drugiego, który zmodyfikował tak układ Rossiego, by przypadkowe zliczenia pojawiały się na tyle rzadko, by mierzyć wiarygodnie sygnały fizycznej koincydencji pojawiające się zaledwie raz na kilka godzin.

Gdy przesuwany licznik znajduje się blisko pary liczników teleskopu pojawienie się potrójnej koincydencji nie było niczym nadzwyczajnym. Zauważył to już Rossi i widzieli też Skobielcyn, Balckett i Occhialini w komorach Wilsona, Carmichael w komorze jonizyjnej, jednak żaden z nich nie spodziewał się, by kaskady cząstek miały rozmiary mierzone w metrach.



Rysunek 1. „Mały pęk” zarejestrowany przez Blacketta i Occhialiniego w 1933 roku.



Rysunek 2. Schemat i wyniki eksperymentu Augera i Maze'a.

Auger i Maze postanowili sprawdzić, jak wielkie bywają kaskady cząstek w powietrzu. Mając precyzyjny układ pomiarowy po rozsunięciu liczników na kilka metrów ciągle jeszcze zdarzały się przypadki fizycznej korelacji rejestracji w tak oddalonych punktach. Przy odległości kilkunastu metrów Maze, ciągle widząc zliczenia potrójnej koincydencji, zaczął poszukiwać technicznej usterki swojego układu pomiarowego i mimo wielu wysiłków nic takiego nie znalazł. Auger, który był już wtedy uznanym naukowcem uznał, że Maze błędu nie popełnił i szybko, na palcach oszacował, że gdyby zsumować energię wszystkich cząstek, jakie musiały w tym momencie dotrzeć na powierzchnię ziemi w przewidzianej przez Bhabhe i Heitlera komicznej kaskadzie, to musiałaby ona być co najmniej milion razy większa od energii wyzwolanej podczas rozpadów radioaktywnych jąder promieniotwórczych. Czegoś takiego nikt się nie spodziewał!

Natychmiast, w tym samym roku, wynik Augera i Maze'a został potwierdzony między innymi przez Kolhooorstera. Sam Auger zaś stwierdził istnienie wielkich pęków atmosferycznych o rozmiarach co najmniej setek metrów. Energie cząstek je inicjujących u szczytu atmosfery musiały być miliardy razy większe niż typowe energie rozpadów promieniotwórczych. Natychmiast pojawiło się pytanie, skąd biorą się takie cząstki, gdzie i jak powstają, i czym właściwie są.

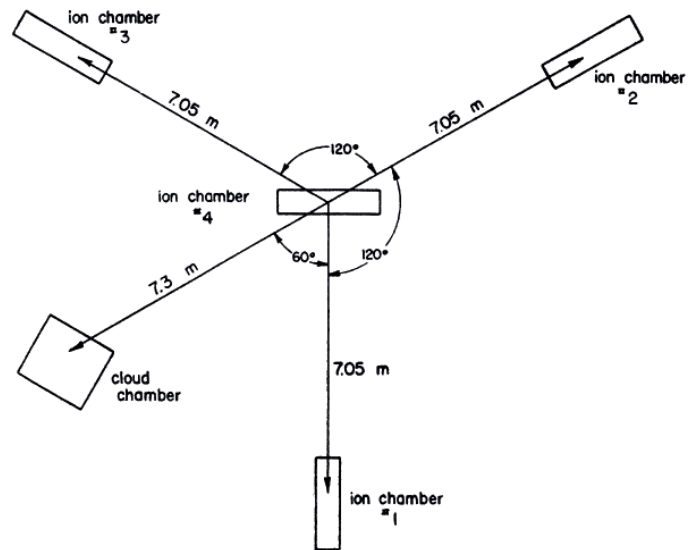
Po przerwie spowodowanej II wojną światową przystąpiono do systematycznych badań. Przy rozsądnym założeniu, że wtórne cząstki rozkładają się symetrycznie względem osi będącej przedłużeniem kierunku pierwotnie przylatującej cząstki, stosując niespecjalnie wyrafinowane metody matematyczne, jeśli zmierzmy, ile cząstek opadło w trzech różnych punktach na ziemi, da się określić, gdzie padła oś pędu, i w miarę dokładnie scałkować gęstość cząstek w całym pęku i wyznaczyć jego wielkość, a przez to energię pierwotnej cząstki.

Metodę tę zastosowano w eksperymencie zbudowanym przez Williamsa z grupy Rossiego zatrudnionego wtedy w MIT na Mt. Evans, Echo Lake w Kolorado w roku 1948. Trzy komory jonizacyjne umieszczono w wierzchołkach trójkąta równobocznego o boku 6 metrów (i jeszcze jedną w środku).

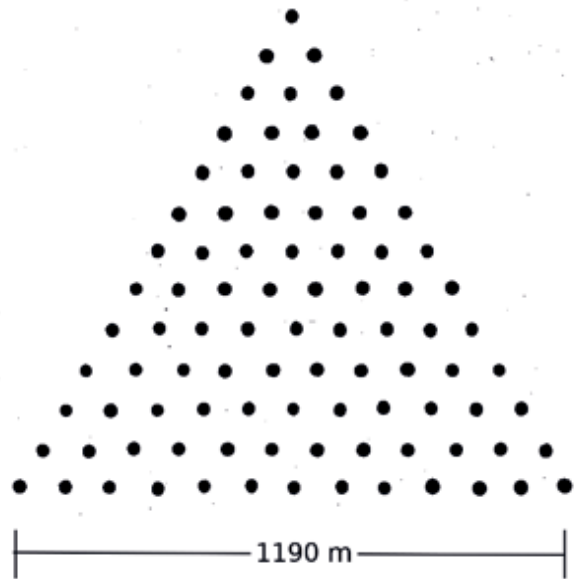
Na początku lat 50. w Harwell w Wielkiej Brytanii zbudowano wielką aparaturę wielkopękową, opartą na prawie 80 stacjach pomiarowych złożonych z sześciu liczników G-M każda rozmieszczonych w obrębie trójkąta równobocznego o boku o długości prawie 1200 metrów. Uzupełniono ją potem (1958) w wodne liczniki Czernkova mające uzupełniać binarną informację liczników G-M (trafiony/nietrafiony) informacją analogową o ilości światła wygenerowanego w dużym zbiorniku czystej wody przez poruszające się szybciej od światła w wodzie naładowane cząstki wielkiego pędu.

Odpowiedzią Amerykanów była aparatura wielkopękowa Agassiz, gdzie 12 detektorów scyntylacyjnych o powierzchni około metra kwadratowego każdy rozmieszczono na obszarze mniej więcej kilometra kwadratowego. Aparatura ta w roku 1957 zarejestrowała rekordowy pęk o energii 3×10^{18} eV.

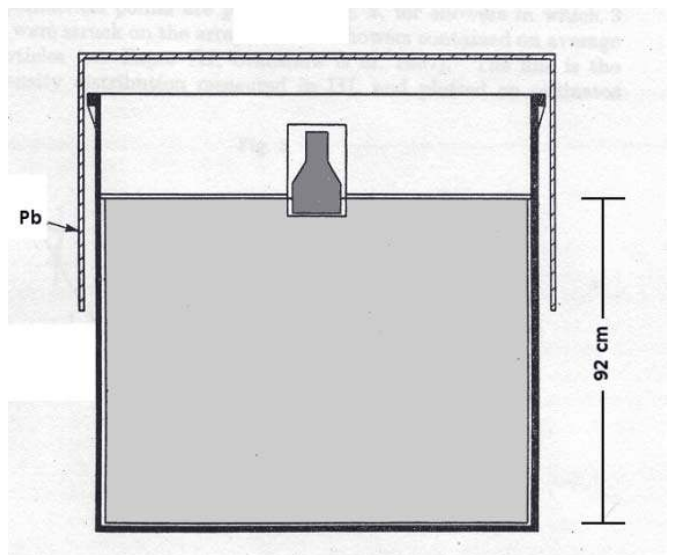
W końcu detektory ze stacji Agassiz przeniesiono w góry Boliwii do El Alto, by zbadać, jak wielkie pęki rozwijają się na dużych wysokościach w atmosferze,



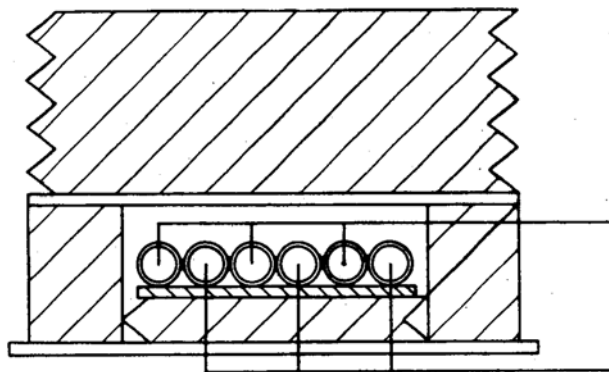
Rysunek 3. Aparatura pękowa w Echo Lake.



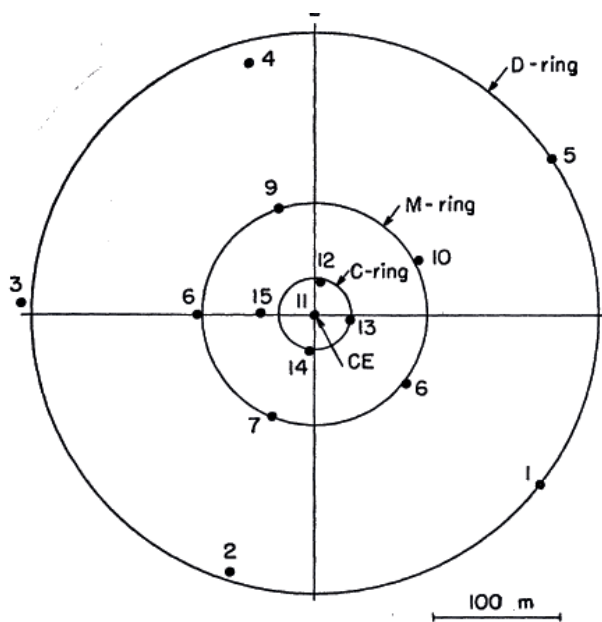
Rysunek 4. Schemat aparatury w Harwell.



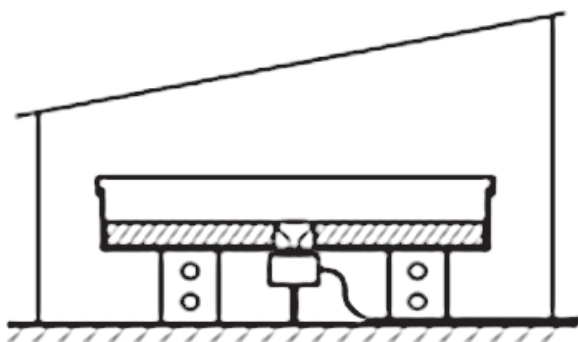
Rysunek 5. Jedna ze stacji detekcyjnych w Harwell.



Rysunek 6 detektor czerenkowski z Harwell.



Rysunek 7. Stacja Agassiz.

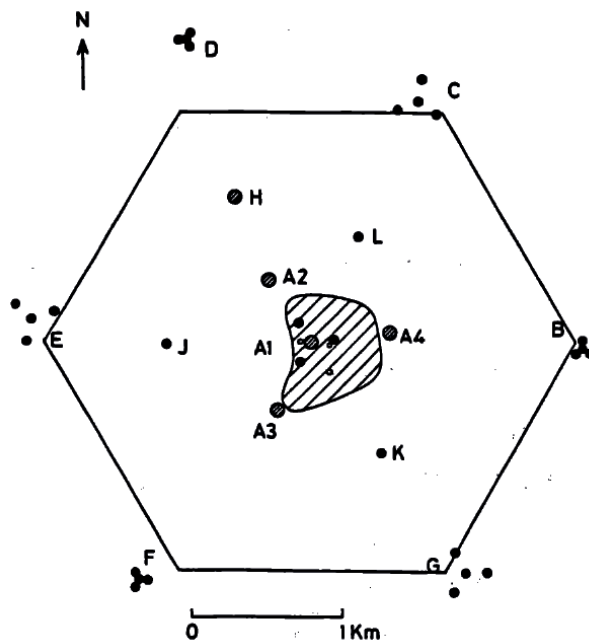


Rysunek 8. Detektor scyntylicyjny w stacji Agassiz.

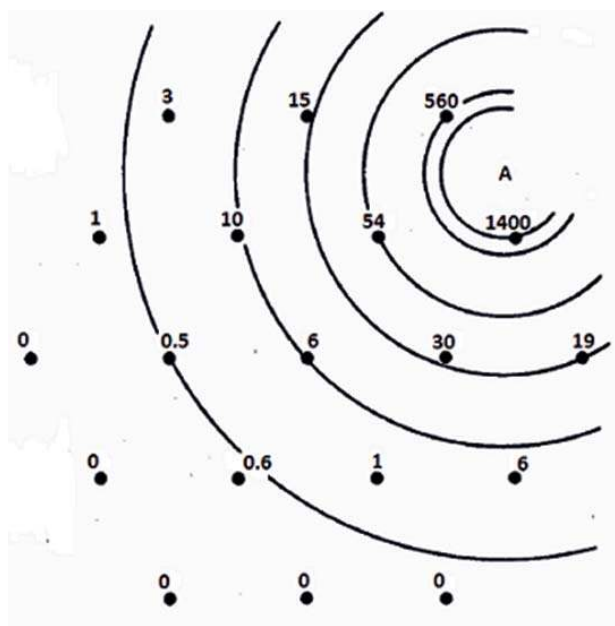
a sam eksperyment ze stacji Agassiz przeniesiono do Volcano Ranch niedaleko Albuquerque w Nowym Meksyku. Detektory Volcano Ranch umieszczono symetrycznie w sześciokącie o rozmiarze około 3 km. Eksperyment ten pracował ćwierć wieku od roku 1960.

Anglicy odpowiedzieli równie wielką aparaturą w Haverah Park, gdzie eksploatowano technikę detektorów czerenkowskich, jak te w Harwell. Aparatura ta pracowała od początku lat 60 przez prawie 30 lat.

Najbardziej spektakularnego odkrycia dokonano jednak w Ameryce. W lutym 1962 roku zarejestrowano pęk zawierający jakieś 50 miliardów cząstek, którego energię osza-



Rysunek 9. Aparatura Volcano Ranch.

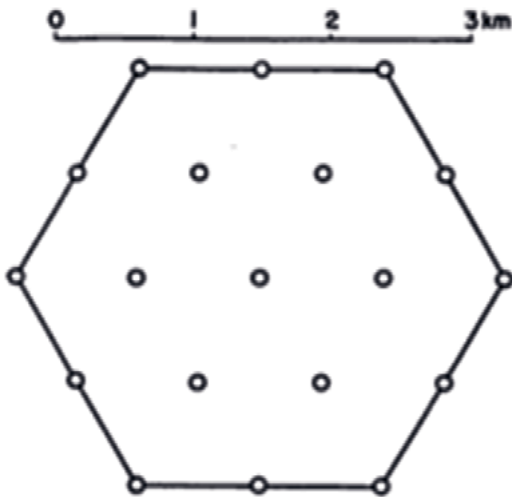


Rysunek 10. Aparatura wielkopękowa w Haverah Park.

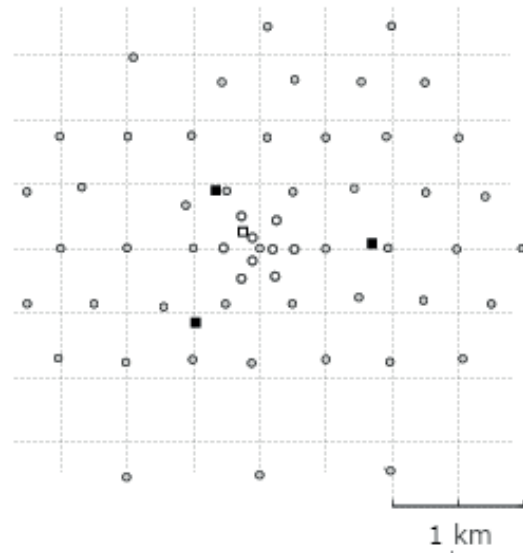
cowano początkowo na 10^{20} eV. Tę historyczną rejestrację pokazano na rysunku 10. Liczby przy detektorach pokazują zmierzoną gęstość cząstek pękowych w danym miejscu.

Wydaje się, że nie ma w tym nic nadzwyczajnego. Rejestrowano już wcześniej pęki zawierające miliardy cząstek, nigdy jednak tak wielkiego jak 2-4834.

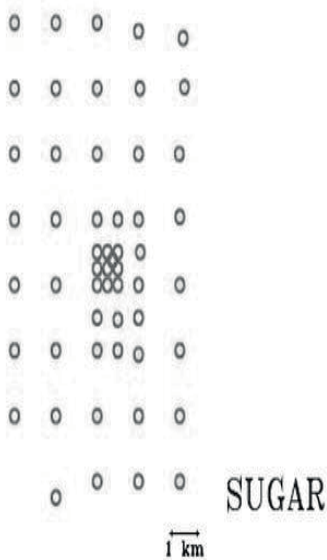
Odkrycie to jednak nabrało znaczenia po roku 1964. W tym roku dwóch radioastronomów Penzias i Wilson zauważyła, że ciągle odbierają na fali 7.35 cm emisję nie wiadomo skąd i jest jej tyle, jakby wysyłało ją ciało doskonale czarne o temperaturze 3.5 ± 1 K. Wyglądało tak, jakby cały wszechświat wypełniony był tym promieniowaniem. I okazało się, że rzeczywiście tak jest. Penzias i Wilson dostali za swoje odkrycie w roku 1978 nagrodę Nobla, ale nie to jest najważniejsze. Jeśli wszędzie pełno jest tych bardzo zimnych, ale mających jakąś energię fotonów to, jak zauważyli Rosjanie Zatsepin i Kuzmin i niezależnie od nich Amerykanin Greisen, protony promieniowania kosmicznego



Rysunek 11. Rejestracja w Volcano Ranch wielkiego pęku o numerze 2-4834.



Rysunek 13. Aparatura wielkopękowa w Jakucku.



Rysunek 12. Eksperyment SUGAR.

o bardzo wielkich energiach powinny zderzać się z tymi fotonami i jak pokazały proste rachunki, powinny bardzo szybko tracić swą bardzo wielką energię. Na tyle szybko, że właściwie cząstek o energiach 10^{20} eV nie powinno w ogóle być. Efekt ten nazwano od nazwisk odkrywców obcięciem GZK.

Odkrycie z Volcano Ranch pozostawało w sprzeczności z tym, co wiemy o oddziaływaniach

cząstek elementarnych. Nikt nie mógł podważyć wyniku pomiaru. O ile pomyłka (błąd?) o kilkanaście, nawet kilkadziesiąt procent w wyznaczaniu energii pęku byłby do pomysłenia, to jednak tu sprzeczność wydawała się nie do pokonania

Czyżby w fizyce działo się coś dziwnego? Każdy taki sygnał powoduje, że rzesze fizyków rzucają się na taki temat i powstają różne, dziwne czasem teorie (a może jednak teoria względności nie działa do końca tak, jak powiedziała Einstein, albo coś równie dziwnego – czemu nie!), ale z drugiej strony wzmocniono wysiłki eksperymentalne, by sprawdzić, jak to jest w istocie, bo w fizyce w szczególności, jedna jaskółka (2-4834) wiosny od razu nie czyni. Wynik z Volcano Ranch należało potwierdzić, albo też wykazać, że jest to jednak jednorazowa pomyłka, fluktuacja, nieszczęśliwy zbieg okoliczności.

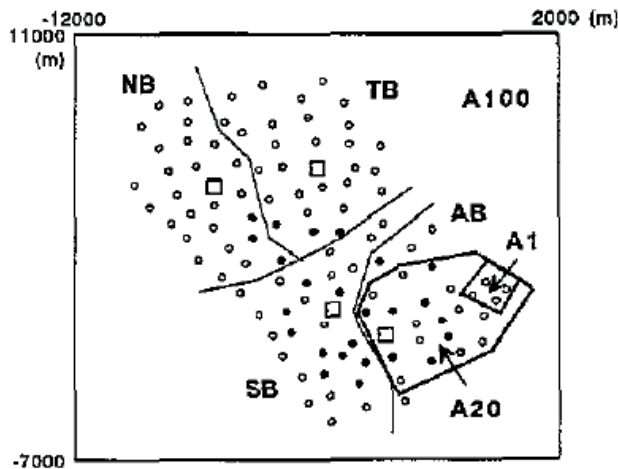
W końcu lat 60. do wyścigu amerykańsko-angielskiego w budowie największych aparatur pękowych i poszukiwaniu cząstek o najwyższej energii dołączyli Australijczycy, którzy skonstruowali eksperyment SUGAR, (Sydney University Giant Air-shower Recorder). 47 stacji pomiarowych, każda składająca się z dwóch wielkich, po 6 metrów kwadratowych każdy, zbiorników ciepłego scyntylatora pokrywała obszar $\sim 100 \text{ km}^2$.

O ile liczniki w Volcano Ranch i w Haverah Park ustawione były na powierzchni ziemi, aby nie przeszkadzało w dotarciu do nich wielkopękowym elektronom, to w Australii tamtejsze wielkie zbiorniki zakopano w ziemi na kilka metrów, na tyle głęboko, aby typowe elektrony (i fotony), produkowane w atmosferze i tworzące zdecydowaną większość każdego pęku, dotrzeć do nich nie mogły.

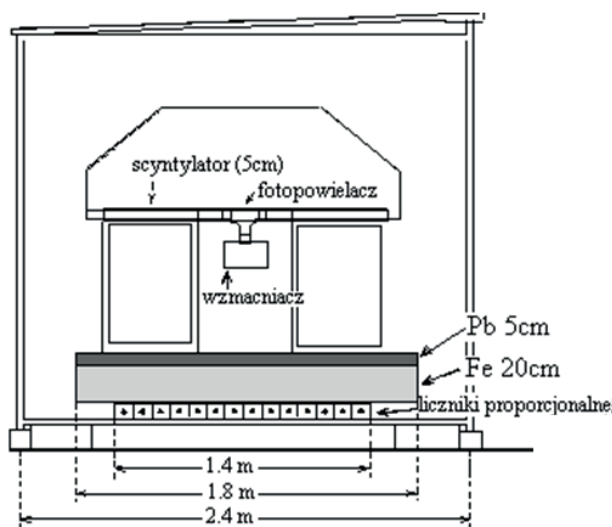
Co zatem obserwowały detektory w SUGAR? Obserwowały zupełnie inne cząstki elementarne. Te same, które w eksperymencie Bothego i Kolhörstera przebiegały się przez sztabkę złota i powodowały koincydencje w doświadczeniach Rossiego z wielocentymetrowymi warstwami ołowiu. Nazywamy je dziś mionami (i oznaczamy grecką literą μ). Stanowią one mniej więcej 10% wszystkich cząstek w pęku (zależy od typu cząstki, jaka pęk zainicjowała i od wysokości w atmosferze, na jakiej jesteśmy).

Inną ciekawostką aparatury australijskiej było to, że stacje nie były ze sobą w żaden sposób połączone. Każda z nich działała samodzielnie i ciekawe przypadki, gdy w obu scyntylatorach pojawiał się wystarczająco jasny błysk, rejestrowano przypadek na zwykłej magnetofonowej taśmie, Co jakiś czas ktoś taśmy te wymieniał i potem porównywano zapisy w dzisiejszej terminologii „off-line”. Aby ustalić, czy jakieś stacje zarejestrowały ten sam wielki pęk niezbędna była ich czasowa synchronizacja. Przy braku systemu GPS i systemu serwerów czasu NTP problem rozwiązano budując własny maszt radiowy i wysyłając z niego do wszystkich stacji własne sygnały czasu. Aparatura SUGAR pracowała trochę ponad 10 lat.

Od lat 70 w wyścigu ku najwyższym energiom biorą też udział Rosjanie. W Jakucku nad brzegiem Leny zbudowali oni aparaturę, która rozbudowywana mierzyła wielkie pęki na wiele różnych sposobów. Głównie były to pomiary elektronów licznikami scyntylacyjnymi, ale także mionów licznikami osłoniętymi przed składową elektronową i pomiary czerenkowskie światła powstającego w atmosferze, robione w bezksiężycowe noce z użyciem otwartych, skierowanych w niebo rur z czułymi detektorami światła umieszczonymi na końcu, ale też kamerami rejestrującymi światło fluorescencyjne powsta-



Rysunek 14. Japońska aparatura w Akeno.



Rysunek 15. Detektor naziemny aparatury Akeno /AGASA.

jące przy wzbudzeniu cząstek atmosfery (azotu N_2) przez przelatujący pęk.

Wszystkich przebili jednak Japończycy. W małej miejscowości Akeno w 1979 roku wybudowano aparaturę pękową na powierzchni 1 km^2 (A1). Składała się ona ze 156 detektorów scyntylacyjnych o powierzchni 1 m^2 . Dodatkowo zainstalowano tam 25 m^2 detektora mionowego, zbudowanego z długich i wąskich liczników proporcjonalnych ($5 \text{ m} \times 10 \text{ cm} \times 10 \text{ cm}$), przykryte dwumetrową warstwą betonu.

W kolejnym kroku w 1984 r. powstała aparatura o powierzchni 20 km^2 (A20). Składało się na nią 23 detektorów elektronów rozstawionych tym razem w odstępach 1 km.

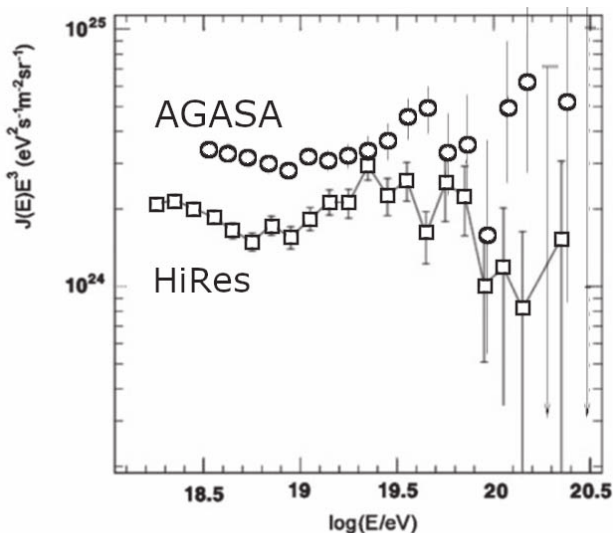
W 1990 roku uruchomiono Akeno Giant Air Shower Array (AGASA) obejmującą powierzchnię 100 km^2 i składającą się w sumie z 111 detektorów naziemnych o powierzchniach $2,2 \text{ m}^2$, rozstawionych co mniej więcej 1 km (A100) i 27 osłoniętych betonem detektorów mionowych.

Wyniki uzyskane przez aparaturę AGASA skonfundowały większość fizyków zajmujących się promieniowaniem kosmicznym. Wskazywały one na istnienie sporej liczby przypadków o energiach przewyższających energię dozwoloną przez obcięcie GZK.

Najwyższą energię zarejestrowano w maju 2001 roku i wynosiła ona $\sim 2,5 \times 10^{20} \text{ eV}$.



Rysunek 16. Teleskopy eksperymentu Fly's Eye.



Rysunek 17. Wyniki eksperymentów AGASA i HiRes (stereo).

Wszystko wskazywało na to, że coś z naszym rozumieniem fizyki wszechświata jest nie tak.

Po raz kolejny sprawą zajęli się Amerykanie. Od lat osiemdziesiątych ubiegłego wieku na pustyni w stanie Utah funkcjonował eksperyment Fly's Eye (później, po rozbudowie w latach 90. nazywany High Resolution Fly's Eye albo po prostu HiRes). Nazwa „Oko muchy” odpowiada technice pomiaru pęków. Nie polegała ona na rejestracji cząstek przez detektory umieszczone na powierzchni ziemi, a na patrzeniu w górę i poszukiwaniu błysków światła fluorescencyjnego powstającego przy przechodzeniu przez atmosferę jednocześnie bardzo wielu cząstek naładowanych. W górę patrzyły pojedyncze bardzo czułe czujniki światła, umieszczone w ogniskach teleskopów schowanych w okrągłych obudowach, ustawionych na dwóch wzniesieniach ponad pustynią w Dugway. W sumie takich teleskopów było 22 (w oku lewym – Five Mile Hill) i 42 (w oku prawym – Camel's Back Ridge).

Oczywiście eksperyment działać mógł tylko w bezchmurne i bezmgłowe noce, ale za to był stosunkowo tani w realizacji. W maju 1997 r. eksperyment rozbudowano zastępując zwykłe oka much oczami o wysokiej rozdzielczości i tak mierzono w Utah pęki do roku 2006.

Wyniki eksperymentów w Utah nie zgadzały się specjalnie w tymi uzyskanymi przez Japończyków. Jakieś obcięcie widma przed energią 10^{20} eV jest widoczne, choć te kilka przypadków, w tym pęk z 1991 roku o największej dotąd zarejestrowanej energii $3,2 \times 10^{20} \text{ eV}$ mogą niepokoić.

W XXI wieku postanowiono zbudować dwa wielkie eksperymenty: Pierre Auger Observatory w Argentynie i Telescope Array w Stanach. Miały one dać definitywne odpowiedzi na wszystkie pytania dotyczące wysokoenergetycznego końca widma energetycznego promieniowania kosmicznego.

Tylko u nas! Pod jednym adresem www.dspress.com.pl

Prenumerata **w wersji drukowanej**

– nie ponosicie dodatkowych opłat za przesyłkę!

E-Prenumerata

– pliki PDF
– wysyłamy bezpośrednio na Twój komputer!

Roczniki archiwalne **w wersji drukowanej**

– do wyczerpania nakładu...

Wydania darmowe

– zamów, Czytaj i wiedz więcej!

Wydania specjalne

– najciekawsze zbiory artykułów, zadań i scenariuszy lekcji!

Wszystkie wydania od 2009 roku
w wersji PDF – bez ograniczeń!

Wydania specjalne

w wersji elektronicznej – pliki PDF



Ceny już
od 10 zł!



Szczegóły i formularz zamówienia: www.aspress.com.pl/specjalne/

eptasa.pl 906/27 3058