

Fizyczne i fizjologiczne aspekty percepcji krótkich sygnałów akustycznych

Karolina Stelmach

Promotorzy pracy:

prof. dr hab. Piotr Zieliński

Instytut Fizyki Jądrowej PAN

oraz

dr hab. Tadeusz Kamisiński, prof AGH

Akademia Górniczo-Hutnicza AGH



Kraków

2023

Podziękowania:

Niniejsza rozprawa doktorska została wykonana w ramach Programu Operacyjnego Wiedza Edukacja Rozwój, nr projektu POWR.03.02.00-00-I004/16, współfinansowanego ze środków Unii Europejskiej.

Chciałabym złożyć serdeczne podziękowania wszystkim, którzy przyczynili się do powstania niniejszej pracy doktorskiej.

W szczególności:

Prof. dr hab. Piotrowi Zielińskiemu za wskazanie tematu pracy, wsparcie, cenne rady, pomoc i poświęcony czas, a także za zaproszenie do prowadzenia zajęć z przedmiotu „Fizyka i Fizjologia Dźwięku”, co umożliwiło mi zapoznanie się z szerokim wachlarzem literatury i technikami stosowanymi w syntezie dźwięku.

Dr. hab. Tadeuszowi Kamisińskiemu, prof. AGH za udostępnienie Laboratorium Akustyki Technicznej i jego wyposażenia do przeprowadzenia pomiarów niezbędnych do osiągnięcia celów badawczych mojej pracy a także udzielanie cennych konsultacji nt. otrzymanych wyników.

Dr. Jarosławowi Rubacha za pomoc w wykonaniu pomiarów w LAT.

Dr. Marcinowi Majce za pomoc w przygotowaniu i przeprowadzeniu testów behawioralnych.

Dr. dr. Marcinowi Majce i Pawłowi Sobieszczykowi za przekazanie informacji nt. wrażenia wysokości dźwięku skrajnie krótkich impulsów.

Pracownikom Katedry Fizyki Materii Skondensowanej Uniwersytetu w Białymstoku i Zakładu Fizyki Biomedycznej Uniwersytetu Warszawskiego.

Rodzinie i przyjaciółom za ciągłe wsparcie i motywację, zwłaszcza mężowi Adamowi za okazane wsparcie podczas przygotowywania rozprawy doktorskiej.

Wreszcie, chciałabym podziękować wszystkim, którzy we mnie wierzyli i wspierali, zarówno podczas mojej pracy badawczej, jak i w codziennym życiu.

Streszczenie

Rozprawa podejmuje temat ilościowego opisu wrażeń, jakie w słuchaczach wywołują skrajnie krótkie sygnały akustyczne. Podstawowa i znajdująca uzasadnienie w przeprowadzonych eksperymentach hipoteza polega na przypisaniu wartości oczekiwanej częstotliwości w widmach takich sygnałów roli, jaką częstotliwość sygnałów periodycznych odgrywa w wywoływaniu wrażenia wysokości dźwięku. Wykazano, że w przypadku sygnałów czysto gaussowskich ta wartość oczekiwana jest odwrotnie proporcjonalna do czasu trwania sygnału tj. szerokości funkcji Gaussa, a znana z mechaniki kwantowej relacja nieokreśloności ulega pewnej modyfikacji. Wykazano także, że przy założeniu stałego progu rozróżnialności tej efektywnej wysokości dźwięku, najmniejsza zauważalna różnica czasu trwania sygnału spełnia prawo Webera-Fechnera. Przeprowadzone na próbie 36 muzyków i 39 niemuzyków badania behawioralne potwierdziły w pewnym stopniu ten wynik z tą różnicą, że w granicy bardzo krótkich sygnałów najmniejsza zauważalna różnica nie zmierza do zera, lecz do pewnej wartości dodatniej, zależnej od doświadczenia badanych w pracy z dźwiękiem. Jest to więc przykład uogólnionego prawa Webera-Fechnera. Wykonane za pomocą symulatora ludzkich narządów słuchu (*Head and Torso Simulator, HATS*), zwanego sztuczną głową, pomiary przefiltrowanych przez taki model widm impulsów gaussowskich potwierdziły odwrotną proporcjonalność wartości oczekiwanej częstotliwości do czasu trwania impulsu mimo, że same widma mają skomplikowane kształty i wykazują szereg rezonansów pochodzących prawdopodobnie od mechanicznych części symulatora i/lub od elektronicznych generatorów i odbiorników dźwięku. Wydedukowana z wyników pomiarów na symulatorze zależność najmniejszej zauważalnej różnicy czasów trwania impulsów spełnia również uogólnione prawo Webera-Fechnera, lecz wartość granicy tej różnicy dla czasu trwania sygnałów zmierzającego do zera jest znacznie mniejsza, niż w przypadku uczestników badań behawioralnych. Jest to największa rozbieżność pomiędzy wynikami badań behawioralnych i pomiarami mechaniczno-elektronicznymi. Pochodzenie tej rozbieżności wydaje się ciekawym tematem dalszych badań. Mechanizmy percepcji skrajnie krótkich sygnałów dźwiękowych w różnych warunkach akustycznych może mieć znaczenie w rozwoju środków wyrazu muzyki instrumentalnej i elektroakustycznej.

Summary

The dissertation deals with a quantitative description of the sensations that extremely short acoustic signals evoke in the listeners. The basic hypothesis, which is justified by the conducted experiments, consists in assigning the expectation value of frequency in the spectra of such signals to the same role that the frequency of periodic signals plays in producing the impression of pitch. It has been shown that in the case of purely Gaussian signals, this expectation value is inversely proportional to the duration of the signal, i.e. to the width of the Gaussian function, and the uncertainty relation known from quantum mechanics undergoes some modification. It was also shown that assuming a constant threshold of discrimination of this effective pitch, the smallest noticeable difference in signal duration satisfies the Weber-Fechner law. Behavioural tests conducted on a sample of 36 musicians and 39 non-musicians confirmed this result to some extent, with the difference that in the limit of very short signals, the just noticeable difference does not tend to zero, but to a certain positive value, depending on the experience of the respondents in working with sound. So this is an example of the generalized Weber-Fechner law. Measurements of spectra of Gaussian pulses with the Head and Torso Simulator (HATS), also known as an artificial head, confirmed the inverse proportionality of the expectation value of frequency to the duration of the pulse, even though the spectra themselves have complicated shapes and show a number of resonances possibly from the mechanical parts of the simulator and/or from electronic sound generators and receivers. The dependence of the smallest noticeable difference in pulse durations, deduced from the results of the measurements with the simulator, also satisfies the generalized Weber-Fechner law, but the value of the limit of this difference for the duration of signals tending to zero is much smaller than in the case of participants in behavioural studies. This is the largest discrepancy between the results of behavioural studies and mechano-electronic measurements. The origin of this discrepancy seems to be an interesting topic for further research. Mechanisms of perception of extremely short sound signals in various acoustic conditions may be important in the development of the means of expression of instrumental and electroacoustic music.

Omówienie zawartości poszczególnych rozdziałów

Rozdziały 1 - 5 mają charakter wyboru wiadomości dotyczących dzisiejszej wiedzy i poglądów na temat fizycznych, fizjologicznych i psychofizycznych mechanizmów postrzegania zjawisk dźwiękowych przez człowieka. Wyboru tego dokonano tak, aby ułatwić czytelnikowi wyspecjalizowanemu w jednej z tych dziedzin zrozumienie przedstawionych dalej koncepcji i interpretacji wyników doświadczalnych.

W rozdziale 2 podano podstawy anatomiczne budowy ucha oraz znane w obecnym stanie wiedzy fizjologiczne funkcje poszczególnych jego części.

Rozdział 3 to krótki przegląd metod stosowanych w badaniach psychofizycznych, ze szczególnym uwzględnieniem testów behawioralnych. Najbardziej istotnym pojęciem przydatnym w części eksperymentalnej pracy jest tu „prawo psychofizyczne” określające zależność wrażenia od bodźca. Oczywiście przypomniano tu szeroko znane formuły matematyczne takich praw, tj. prawo Webera-Fechnera i prawo Stevensa. W wielu przypadkach prawo psychofizyczne wywodzi się od tego, w jaki sposób najmniejsza zauważalna różnica intensywności bodźca (*ang. just noticeable difference lub difference limen*) zależy od bezwzględnej wartości tej intensywności. We wzorach (5) - (7) zaargumentowano, że warunkiem obowiązywania prawa Webera-Fechnera jest niezależność wartości progowej zmiany wrażenia, oznaczanej w tej pracy jako $d\omega_{lim}$, dp_{lim} itp., od intensywności bodźca. To założenie jest stosowane w większości przeprowadzanych w tej pracy analiz.

W rozdziale 4 omówiono krótko oficjalne definicje cech dźwięku, m.in. zalecane przez Amerykański Narodowy Instytut Wzorców (ANSI). Najistotniejsze są tu wysokość dźwięku i barwa dźwięku, gdyż potrzebne są one do analiz zawartych w rozdziałach 7 - 10. Podkreślono znaczenie ilościowego opisu wysokości dźwięku i przytoczono wzór stosowany w urządzeniach i programach MIDI (wzór (10)). Jest to fizyczna miara wysokości dźwięku używana w dalszych częściach pracy.

Rozdział 5 przedstawia historyczne i obecnie funkcjonujące modele percepcji dźwięku oparte o obserwacje anatomiczne i fizjologiczne udoskonalane na od XIX w do naszych czasów. Ich najważniejszym wynikiem jest powiązanie umiejscowienia maksymalnej aktywności komórek rzęsatych ślimaka, będącej skutkiem jakiegoś silnie nieliniowego mechanizmu, z tonotopwą organizacją kory słuchowej mózgu.

Uzasadnienie badań podjętych w tej pracy znajduje się w rozdziale 6. Zebrano tam wyniki badań psychofizycznych dotyczących postrzegania bardzo krótkich sygnałów dźwiękowych. Zauważono, że w dziedzinie tonów prostych, najkrótsze badane impulsy zawierały od 1,6 [1] do 4 [2] okresów drgań. W tym kontekście obserwacje poczynione dla sygnałów jeszcze krótszych w pracy [3] wydają się pionierskie. Ich dokładniejsze ilościowe zbadanie było motywacją do wykonania testów behawioralnych oraz pomiarów przy użyciu modelu mechaniczno-elektronicznego stanowiących doświadczalną część niniejszej rozprawy.

W związku z sensacyjnym tytułem pracy [4] sugerującym, że ludzki słuch jest zdolny przewyciężyć ograniczenia wynikające z relacji nieokreśloności Heisenberga i Gabora, zbadano w rozdziale 7 iloczyn niepewności w dziedzinie czasu i częstotliwości dla impulsów czysto gaussowskich przy czasie trwania impulsu zmierzającym do zera. Ponieważ znana z mechaniki kwantowej zasada nieokreśloności dotyczy odchylenia standardowego od wartości oczekiwanej częstotliwości (lub – równoważnie - energii), prześledzono zachowanie tej wartości przy skracaniu czasu trwania sygnału σ i stwierdzono, że wykazuje ona proporcjonalność do $\frac{1}{\sigma}$ oraz, co więcej, że spełnia oryginalne prawo Webera-Fechnera. Postawiono więc hipotezę, że wartość oczekiwana częstotliwości bardzo krótkich sygnałów jest fizycznym korelatem częstotliwości odpowiedzialnej za wrażenie wysokości dźwięku (lepiej: efektywnej wysokości dźwięku [3]) relacjonowane przez większość słuchaczy.

Rozdział 8 zawiera opis i wyniki behawioralnego testu na zależność najmniejszej zauważalnej różnicy czasu trwania impulsu gaussowskiego od tegoż czasu trwania. Dopasowania analitycznych funkcji do otrzymanych czasów trwania wskazują na pewne uogólnione prawo Webera-Fechnera, w którym najmniejsza zauważalna różnica nie dąży do zera przy granicznym zmniejszaniu bodźca.

Mechaniczno-elektroniczny model zwany sztuczną głową został wykorzystany w rozdziale 9 do pozyskania widm takich samych impulsów gaussowskich poddanych przekształceniom (filtrowaniu, być może nieliniowemu) związanym z transmisją przez przewody o kształcie i własnościach lepkosprężystych odpowiadających ludzkiemu uchu. Mimo, że widma te odbiegały kształtem od pierwotnych/wejściowych widm gaussowskich, zależność wartości oczekiwanej częstotliwości od czasu trwania impulsu okazała się zadziwiająco podobna do funkcji $\frac{1}{\sigma}$ wykazanej dla impulsów wejściowych. W tym samym rozdziale zaproponowano sposób wydedu-

kowania analogu prawa Webera-Fechnera dla eksperymentalnych widm uzyskanych ze sztucznej głowy i wyprowadzono różne warianty tego prawa dla wybranych sposobów dopasowania krzywych do danych doświadczalnych.

Rozdział 10 stanowi próbę porównania i wychwycenia różnic w ilościowych rezultatach otrzymanych poprzez badania behawioralne i mechaniczno-elektroniczne. Istotna rozbieżność polega na stwierdzonej w badaniach behawioralnych stosunkowo dużej granicy najmniejszej zauważalnej różnicy czasów trwania impulsów gdy czas ten dąży do zera w porównaniu z odpowiednią granicą obliczoną dla sztucznej głowy.

Podsumowanie oraz wykaz najważniejszych wyników osiągniętych w przedstawianej rozprawie znajdują się w rozdziale 11.

Aneks zawiera wybrane widma uzyskane za pomocą sztucznej głowy dla różnych częstości próbkowania sygnałów dostarczanych przez różne typy słuchawek; dane te mogą być wskazówką w dalszych badaniach poruszanych tu zagadnień.

Spis treści

Streszczenie	4
Summary	5
Omówienie zawartości poszczególnych rozdziałów	6
1 Wstęp	14
2 Narządy słuchu człowieka	16
2.1 Budowa ucha i jego funkcje	16
2.1.1 Ucho zewnętrzne	16
2.1.2 Ucho środkowe	17
2.1.3 Ucho wewnętrzne	18
3 Metody oceny postrzegalności bodźców i ich różnic	22
3.1 Próg absolutny i próg różnicowy	22
3.2 Prawa psychofizyczne	24
3.2.1 Prawo Webera-Fechnera	25
3.2.2 Prawo Stevensa	27
4 Psychofizyczne podstawy percepcji cech dźwięku	28
4.1 Percepcja wysokości dźwięku	29
4.2 Wysokość dźwięku a okresowość sygnału	30
4.2.1 Barwa dźwięku.	34
4.2.2 Odczuwanie głośności.	35
4.2.3 Percepcja zjawisk czasowych	37
5 Mechanizmy i teorie słyszenia	39
5.1 Teoria miejsca	39
5.2 Teoria czasu	41
6 Dotychczasowe wyniki badań krótkich sygnałów dźwiękowych	43
7 Korelaty fizyczne wrażeń wywoływanych przez krótkie impulsy	48

7.1	Wysokość dźwięku bardzo krótkich tonów prostych z obwiednią gaussowską. Próba określenia fizycznych parametrów decydujących o tym wrażeniu w granicy małych czasów trwania.	48
7.2	Fizyczne korelaty wysokości i barwy dźwięku dla sygnałów czysto gaussowskich.	59
7.2.1	Wrażenie wysokości dźwięku.	59
7.2.2	Wrażenie barwy dźwięku.	62
8	Badania behawioralne	65
8.1	Wprowadzenie	65
8.2	Adaptacyjny test na wykrywanie najmniejszej zauważalnej różnicy czasu trwania sygnału czysto gaussowskiego.	66
8.3	Wyniki	72
9	Badania w komorze bezechowej	82
9.1	Wprowadzenie	82
9.1.1	Komorą bezechowa	82
9.1.2	Sztuczna głowa	82
9.2	Metody	84
9.3	Wyniki	88
9.3.1	Mikrofon rejestrujący sygnał z głośnika	90
9.3.2	Analiza reprezentacji wrażenia wysokości dźwięku.	92
9.3.3	Efektywna wysokość dźwięku krótkich impulsów a rozróżnialność czasu ich trwania	100
10	Porównanie wyników badań behawioralnych oraz badań w LAT	103
10.1	Fizyczne korelaty wrażenia wysokości dźwięku na podstawie porównania wyników testów behawioralnych i parametrów widm uzyskanych za pomocą sztucznej głowy.	103
10.2	Próby uzgodnienia prawa psychofizycznego dla czasu trwania impulsów uzyskanego z testów behawioralnych z analizami widm zarejestrowanych przez HATS.	108
11	Podsumowanie	114

12 Aneks	121
12.1 Pomiar HATS z użyciem głośnika.	121
12.2 Pomiar HATS z użyciem 4 różnych typów słuchawek.	125
12.3 Widma mocy otrzymane podczas pomiaru HATS - częstotliwość próbkowania 44100 Hz.	129
12.4 Widma mocy otrzymane podczas pomiaru HATS - częstotliwość próbkowania 192000 Hz.	139
13 Bibliografia	146
Spis rysunków	157
Spis tablic	165
Osiągnięcia naukowe	166

Wykaz ważniejszych skrótów i oznaczeń

DL Difference Limen

JND Just Noticeable Difference

EP Effective Pitch

AGH Akademia Górniczo-Hutnicza

EEG Elektroencefalografia

HATS Head, Torso and Ear Simulators

W – F Prawo Webera-Fechnera

Wykaz załączników

Załącznik 1 Oświadczenie uczestnika projektu o wyrażeniu zgody na przetwarzanie danych osobowych.

Załącznik 2 Ankieta osobowa.

Załącznik 3 Instrukcja wykonania testu.

1 Wstęp

Dźwięk - obok obrazu - jest podstawowym nośnikiem informacji docierających do istot żywych ze środowiska. Właściwa percepcja dźwięku i następująca po niej reakcja na bodźce dźwiękowe są niezbędne do przeżycia (pierwsze prawo biologiczne) oraz do przedłużania gatunku (drugie prawo biologiczne). Organizmy wyższe wykształciły zależne od swych uwarunkowań anatomicznych i fizjologicznych systemy porozumiewania się za pomocą sygnałów dźwiękowych. Przekaz bywa więc ostrzeżeniem przed niebezpieczeństwem, hasłem do wspólnego działania, a w końcu, na najwyższym poziomie, inspiracją przeżyć estetycznych i transcendentalnych. Aby te funkcje mogły być spełnione potrzebny jest jakiś kod pozwalający rozróżniać dźwięki. Fizyka dostarcza tu najbardziej systematycznego sposobu rozróżniania poprzez zdefiniowanie mierzalnych i/lub postrzegalnych parametrów liczbowych. Tak zazwyczaj zaczyna się edukacja w zakresie przedmiotu „zasady muzyki”. W swym klasycznym podręczniku „Zasady muzyki” [5] F. Wesołowski wyróżnia cztery cechy dźwięku: „a) wysokość, b) głośność, c) czas trwania, d) barwę”. Niektóre z nich łatwo odnieść do podstawowych wielkości fizycznych i nadać im odpowiednie jednostki, np. czas w sekundach. Inne jednak wymagają wiedzy o tym w jaki sposób parametry fizyczne wywołują interpersonalnie komunikatywne wrażenia. Np. głośność koreluje w pewien sposób z szybkością dostarczania energii narządowi słuchu, ale nie jest dosłownie tą szybkością. Stąd potrzeba definicji szeregu wielkości: natężenie dźwięku [$\text{J s}^{-1} \text{m}^{-2}$], poziom natężenia [dB], głośność [son], poziom głośności [fon]. Wielkości te wskazują drogę od miar czysto obiektywnych do parametrów uwzględniających gatunkowe i osobnicze mechanizmy percepcji. Podobnie jest z wysokością dźwięku. Ogólnie jest ona powiązana z okresem sygnału. Jednakże zdefiniowanie okresu nie jest zawsze jasne z powodu samej struktury sygnału, a dodatku skończony czas trwania sygnału uniemożliwia określenie okresu, co znalazło ilościowy wyraz w zasadzie nieokreśloności częstotliwość-czas Heisenberga i Gabora [6]. Powstaje więc pytanie jakim dźwiękom i w jakim zakresie ich parametrów można przypisać cechę wysokości dźwięku oraz jakie są tej cechy najistotniejsze korelaty fizyczne.

Głównym celem obecnej rozprawy jest badanie sygnałów skrajnie krótkich. Rozpoczynamy od najprostszych z punktu widzenia fizyki impulsów, w których

ciśnienie akustyczne dane jest funkcją rozkładu normalnego, zwaną też funkcją Gaussa. W niniejszej pracy przedstawię badania dźwięków krótkich, które spotykane są w przyrodzie, jednak nie zbadano wcześniej w jaki sposób następuje ich percepcja. Podczas badań nad tym tematem zauważono, że im bardziej skracamy czas trwania dźwięku tym bardziej wydaje się on wyższy. Zaproponowano więc pojęcie Efektywnej Wysokości dźwięku, która powstaje mimo, że nie da się wyróżnić okresu sygnału, lecz rośnie w opinii wielu słuchaczy, gdy skracany jest czas trwania impulsu w zakresie milisekundowym. Przeprowadzono badania behawioralne oraz badanie sygnałów przetworzonych przez mechaniczno-elektroniczny układ zwany „sztuczną głową”, która jest w istocie modelem wykorzystującym materiały o właściwościach sprężystych i lepko-sprężystych zbliżonych do tkanek ludzkich, szczególnie w okolicach ucha zewnętrznego, środkowego i wewnętrznego. Część elektroniczna sztucznej głowy pozwala zarejestrować do dalszej analizy sygnały akustyczne dochodzące z zewnątrz i przetworzone przez część mechaniczną. W niniejszej rozprawie doktorskiej przedstawię wyniki oraz wnioski z poczynionych nad fizyką i fizjologią dźwięków badań i analiz.

2 Narządy słuchu człowieka

Wśród pięciu zmysłów każdego człowieka wymieniamy: słuch, wzrok, węch, dotyk oraz smak. Są to zmysły kojarzące się nam z doświadczaniem życia, jednak głównym powodem dla którego człowiek został w nie wyposażony jest przetrwanie. Słuch pozwala nam usłyszeć zbliżające się zagrożenie, jest aktywny również w nocy, pracuje 24h na dobę. Niniejsza praca dotyczy zjawiska percepcji dźwięku, dlatego najpierw kluczowym jest poznanie jak działa system słyszenia człowieka.

2.1 Budowa ucha i jego funkcje

Według Urszuli Jorasz [7], system słyszenia człowieka dzieli się na trzy części: ucho (odbierające falę dźwiękową), drogę słuchową (przenosi i częściowo przetwarza informacje) oraz pola słuchowe kory mózgowej (miejsce gdzie powstaje wrażenie słuchowe). Struktura narządu ludzkiego słuchu przedstawiona jest w tabeli 1.

Opis układu słyszenia zaczniemy od ucha zewnętrznego.

Tablica 1: Struktura narządu słuchu [8].

Ucho zewnętrzne	Ucho środkowe	Ucho wewnętrzne
Małżowina uszna, przewód słuchowy zewnętrzny: (część chrzęstna i kostna)	Jama bębenkowa, trąbka słuchowa, komórki sutkowe, kosteczki słuchowe, mięsień napinacz bębena, mięsień strzemiączkowy, ujście trąbki Eustachiusza	Błędnik kostny, Kanał spiralny ślimaka, przedsionek, kanały półkoliste

2.1.1 Ucho zewnętrzne

Odbiorcami sygnałów z otoczenia są nasze uszy a dokładnie małżowina, która jest najbardziej zewnętrzną częścią ucha i występuje tylko u ssaków [9]. Docierająca do małżowiny fala dźwiękowa wpada do zewnętrznego przewodu słuchowego a dalej trafia do błony bębenkowej powodując jej wibrację. Przewód słuchowy

zewnątrzny ma od 2.5 do 3 cm długości oraz jest lekko wygięty ale o dość regularnej średnicy równej około 7 mm [10]. Kanał słuchowy zbudowany jest z tkanki chrzęstnej i z kości. Według danych empirycznych przewód słuchowy najbardziej wzmacnia częstotliwość równą 2500 Hz natomiast wlot kanału słuchowego również wzmacnia częstotliwość ale równą około 5 kHz. Połączenie kanału słuchowego oraz przewodu słuchowego wzmacnia częstotliwości między 1.5 kHz a 7.5 kHz [10]. Kanał słuchowy zapewnia również dodatkową ochronę dla elementów ucha środkowego i wewnętrznego a ze względu na bliskie mózgu umiejscowienie ucha, połączenia nerwowe między uchem a mózgiem są krótkie, co powoduje szybkie dostarczanie informacji [10]. Fala dźwiękowa przechodzi przez kanał słuchowy powodując wibrację błony bębenkowej. Całkowita powierzchnia błony bębenkowej to około 85 mm², grubość wynosi 0.1 mm a masa to około 14 mg [11]. Zakres ruchu błony bębenkowej przy najcichszych dźwiękach wynosi około 10⁻¹¹ m więc jest mniejszy od rozmiaru atomu wodoru [12]. Błona bębenkowa oddziela ucho zewnętrzne od środkowego.

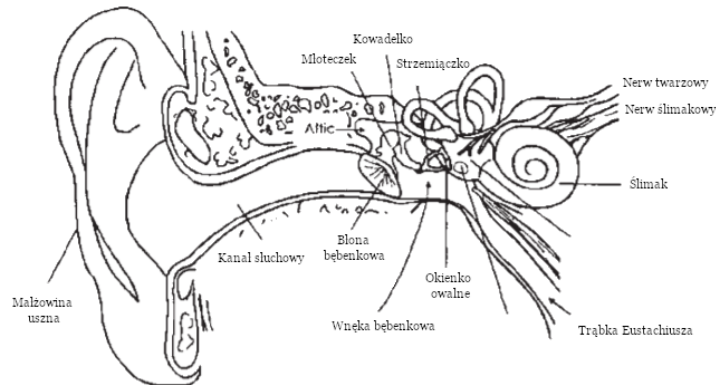
2.1.2 Ucho środkowe

Wnęka bębenkowa, zawiera najmniejsze w całym organizmie kosteczki słuchowe: młoteczek, kowadełko i strzemiączko. Przekazują one ruchy błony bębenkowej do okienka owalnego ślimaka, będącego już częścią ucha wewnętrznego. Ważną rolę w zabezpieczeniu ucha wewnętrznego przed uszkodzeniem wskutek nadmiernych ruchów odgrywają: mięsień napinacz błony bębenkowej (m. tensor tympani) oraz mięsień strzemiączkowy [13].

Pierwszy z nich zwiększa napięcie błony bębenkowej, a zatem usztywnia tę błonę, gdy amplituda jej ruchu staje się zbyt duża. Mięsień strzemiączkowy odchyła zaś podstawę strzemiączka od okienka owalnego, co również zmniejsza zakres ruchu u wejścia do ślimaka. Mięsień ten działa na zasadzie odruchu, zwanego odruchem strzemiączkowym. Odruch ten występuje przy poziomie natężenia dźwięku ok. 70-90 dB, lecz nie natychmiast. Zazwyczaj reakcja na zbyt głośny dźwięk pojawia się po ok. 100 ms dla dźwięków słabych i po 10 ms dla dźwięków o większym natężeniu [14]. Nie zapewnia więc ochrony przed silnym hałasem doświadczanym nagle. z tego powodu zaleca się artylerzystom i innym osobom spodziewającym się

nagłego silnego hałasu otwarcie ust dla wyrównania ciśnień zanim wystąpi odruch bębenkowy. Ponieważ sygnały dźwiękowe badane w tej pracy są krótsze niż 10 ms wydaje się, że obserwowana reakcja następuje przed uruchomieniem odruchu strzemiączkowego.

Na rysunku 1 przedstawiono przekrój ucha ludzkiego wraz z opisem poszczególnych elementów.



Rysunek 1: Widok ucha ludzkiego w przekroju [15].

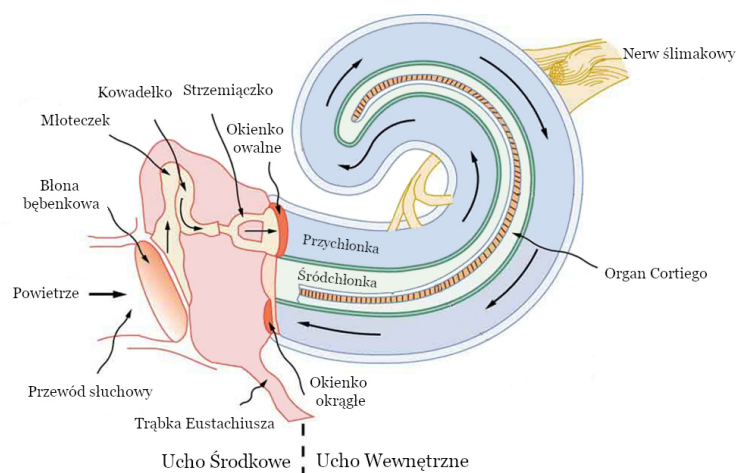
2.1.3 Ucho wewnętrzne

Ucho wewnętrzne jest wypełnione cieczą a więc przekazanie energii z ośrodka gazowego do ciekłego działa w następujący sposób. Drgania mechaniczne błony bębenkowej przekazywane są dalej przez układ kosteczek: młoteczka, kowadełka i strzemiączka, które są najmniejszymi kostkami w ludzkim organizmie. Za utrzymanie kosteczek odpowiadają więzadła i mięśnie, a dodatkowo młoteczek przyrośnięty jest do błony bębenkowej a strzemiączko do tkanki otaczającej ucho wewnętrzne okienkiem owalnym [15]. Mięśnie te, pozostają w ciągłym napięciu ale w sytuacji, kiedy ciśnienie akustyczne wzrasta powyżej 85 dB powstaje tzw. *od-ruch strzemiączkowy*. Mięśnie zaczynają kurczyć się w celu uniknięcia nadmiernego ruchu kosteczek a ciśnienie przekazywane dalej do ucha wewnętrznego jest redukowane [16].

Ucho wewnętrzne zbudowane jest z kanałów półkolistych, przedsionka oraz ślimaka. Dzięki kanałom półkolistym (które są integralną częścią ucha środkowego) utrzymywana jest równowaga, nie biorą one udziału w procesie słyszenia. Ślimak

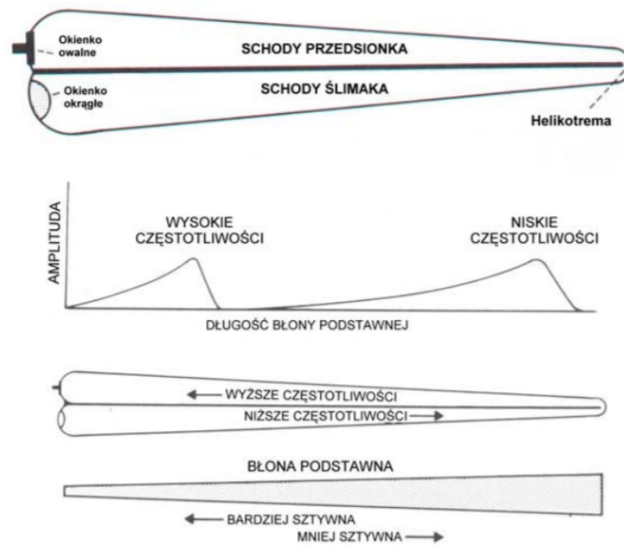
połączony jest z komorą przedsionka. Ślimak to spiralnie zwinięty kanał, który lekko zwęża się ku wierzchołkowi spirali. Zwinięcie ślimaka powoduje, że zajmuje on mniej miejsca, funkcja jaką dokładnie pełni zwinięcie nie jest jednak znana [11]. W ośrodku ślimaka znajdują się dwie błony: błona podstawna i błona Reisnera zwana też błoną przedsionkową, dzielące ślimaka na trzy komory: schody przedsionka, ślimaka i bębenek. Schody środkowe ślimaka zawierają narząd Cortiego, który odpowiada za powstawanie odpowiednich impulsów nerwowych.

Na rysunku 2 przedstawiono elementy ucha wewnętrznego i środkowego.

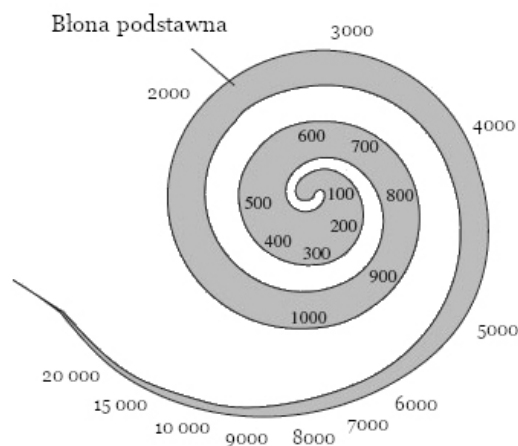


Rysunek 2: Ucho środkowe i wewnętrzne. Rysunek uproszczony [17].

Narząd (też: organ) Cortiego jest receptorem słuchowym o długości około 3.5 cm. Leży on na błonie podstawnej przewodu ślimakowego. Na rys. 3 przedstawiona jest błona podstawna, która zbudowana jest z poprzecznych włókien. U podstawy ślimaka jej szerokość wynosi ok. 0.04 mm, stopniowo zwiększa się aż do osiągnięcia wartości ok 0.52 mm u wierzchołka ślimaka [15]. Model błony podstawnej przedstawiony w postaci płaskiej widać na rys. 4. Błona podstawna wraz z organem Cortiego drga na skutek złożonych oddziaływań hydrodynamicznych. Najwyższe słyszane częstotliwości posiadają fale o większej długości niż długość ślimaka ze względu na to, że długość fali jest większa w ośrodku cieplem niż w powietrzu, co oznacza, że ciśnienie (w szerokim zakresie częstotliwości) jest podobne wzdłuż całej długości błony podstawnej.



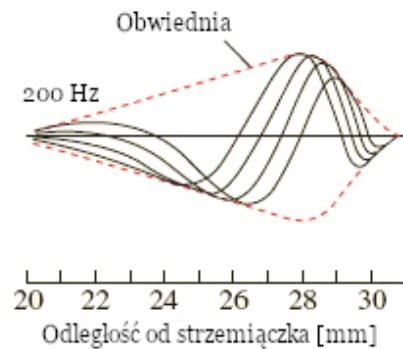
Rysunek 3: Schematyczne przedstawienie błony podstawnej [18].



Rysunek 4: Schematyczna mapa reprezentacji częstotliwości na błonie podstawnej pokazująca, że część błony podstawnej, która reaguje na dźwięk, zależy od danej częstotliwości dźwięku [19].

Błona podstawna działa jak swego rodzaju filtr. Jeśli pobudzimy błonę do drgania tonem prostym o częstotliwości 200 Hz, drgania błony narastają wzdłuż jej długości, następnie osiągają maksimum i maleją aż do zaniku tak jak zaprezentowano na rys. 5. Dźwięki o wysokiej częstotliwości powodują ruch małego obszaru

błony podstawnej w pobliżu strzemiączka, podczas gdy niskie częstotliwości powodują ruch prawie całej membrany. Jednak przemieszczenie piku membrany znajduje się w pobliżu wierzchołka. To pokazuje, że fala biegnąca zawsze przemieszcza się od podstawy do wierzchołka, a to jak daleko biegnie, zależy od częstotliwości stymulacji: niższe częstotliwości docierają dalej [15].



Rysunek 5: Obwiednia fali biegnącej utworzona przez ton 200 Hz. Kształt obwiedni jest opisany przez zbiór chwilowych lokalizacji (na rysunku pokazane cztery) śledzonych przez falę przemieszczającą się wzdłuż błony podstawnej [19].

3 Metody oceny postrzegalności bodźców i ich różnic

3.1 Próg absolutny i próg różnicowy

Różne systemy sensoryczne dostarczają informacji o fizycznych i chemicznych zmianach które mają miejsce w środowisku. Percepcja nie jest możliwa poniżej pewnego progu, wymaganego aby zaszła zmiana zauważona przez organizm. Ta minimalna wielkość bodźca, która ma zostać zaobserwowana czy odczuwalna nazywana jest *progiem absolutnym/bezwzględny*. Poniżej progu bezwzględnego zmiana nie może być zauważona. Tradycyjna psychofizyka oferuje kilka metod pozwalających na obliczenie tego progu [20].

Psychofizyka pojawiła się w momencie kiedy *Gustav Theodor Fechner* zaczął poszukiwać związku pomiędzy zewnętrznym i wewnętrznym światem. Żył on w latach 1801-1887. Zawdzięczamy mu utworzenie wielu metod używanych w psychofizyce po dziś dzień.

Jedną z metod jaką zaproponował Fechner jest właśnie sposób w jaki wykrywamy próg absolutny oraz metoda stałych bodźców (tłumaczenie z: Method of Constant Stimuli) [20]. Metoda ta polega na zgrubnym rozgraniczeniu obszarów, w których bodziec nie jest i jest odbierany, a następnie na wybraniu w tak oszacowanym obszarze przejściowym kilku stałych wartości wielkości bodźca najczęściej równoodległych na skali intensywności bodźca. Następnie generuje się bodźce które mają dokładnie te wybrane wartości i przedstawia się je uczestnikom badania. Osoby te odpowiadają za każdym razem czy wyczuwają generowany bodziec, czy też nie. Gdy uczestników badania jest wiele, liczba pozytywnych odpowiedzi w stosunku do pełnej liczby uczestników ma, jako funkcja intensywności bodźca, zazwyczaj kształt sigmoidalny, tzn. jest bliska zeru dla bodźców słabych, a po obszarze przejściowym wykazującym zwykle punkt przegięcia zbliża się do jedności w sposób nasycający. Za bezwzględny próg wrażliwości na bodziec (zauważalności bodźca) przyjmuje się albo punkt przegięcia tak uzyskanej krzywej, albo oszacowaną wartość intensywności bodźca, dla której krzywa zależności ułamka pozytywnych odpowiedzi przekracza 0,5. Krzywa ta nosi nazwę krzywej psychometrycznej [21].

Wymaga to od uczestnika badania dokonania setek analiz, jednak im więcej analiz tym bardziej szczegółową otrzyma się odpowiedź oraz występuje dużo większa szansa na znalezienie odpowiedniego progu absolutnego. Przy każdym pomiarze, uczestnik badania decyduje czy odczuwa bodziec. Zbierając informacje z każdej próby otrzymujemy dyskretny rozkład częstości występowania. Każda liczba przedstawia, ile razy dany impuls był zauważony. Użycie tej metody nie jest związane tylko z wrażliwością uczestnika badań na bodziec. Może także zależeć od tego w jaki sposób uczestnik będzie podejmował decyzję. Może on czekać dopóki nie będzie pewny swojej odpowiedzi ale również od razu odpowiadać ze 100%-tową pewnością. W metodzie SDT (*ang. Standard Detection Theory*) użyte są dwa parametry: pierwszy opisujący poziom wrażliwości oraz kolejny opisujący sposób w jaki podejmowana jest decyzja [22].

Badanie odczuć jest obarczone dużym błędem, ponieważ w tym przypadku percepcja bodźca zależna jest od wielu czynników. Najprościej opisując ważne są dwa elementy: reakcja badanego na bodziec czyli jego odczucia oraz to w jaki sposób przekaże informacje na temat swoich uczuć czyli tzw. raport o odczuciu.

Wspomniano już o *progu bezwzględnym*, jednak w niniejszej pracy badania behawioralne dotyczyły *progu względnego/różnicowego* (*ang. Difference Threshold*), czyli reprezentującego najmniejszą zmianę bodźca potrzebną aby zauważyć różnicę między dwiema różnymi jego wartościami. Jest to tzw. najmniejsza zauważalna różnica textit(*ang. JND - Just Noticeable Difference, lub DL – Difference Limen*). Wyznaczenie progu różnicowego dla zadanej wartości początkowej (wartości odniesienia, lub standardowej) bodźca można przeprowadzić w podobny sposób jak w przypadku progu bezwzględnego. Wybiera się w tym celu kilka wartości mniejszych i kilka wartości większych od wartości początkowej (standardowej) bodźca, a uczestnikom badania podaje się pary bodźców, z których jeden ma wartość początkową, a drugi którąś z wybranych wartości większych lub mniejszych. Względna liczba uczestników zauważających różnicę układa się zazwyczaj znowu w sigmoidalną krzywą psychometryczną: zbyt małe różnice pozostają niezauważone, a różnice dostatecznie duże zauważa większość uczestników. Różnica intensywności bodźca zauważana przez połowę (50%) uczestników jest estymatą najmniejszej zauważalnej różnicy w badanej populacji. Kiedy uczestnicy rozróżniają różnicę w górę i w dół, wartość odpowiadająca przechodzeniu przez zero, czyli tzw.

punkt subiektywnej identyczności (*ang. Point of Subjective Equality, PSE*) może się różnić od wartości standardowej bodźca. Różnicę między PSE i początkową (standardową) wartością bodźca nazywa błędem stałym (*Constant Error*).

Jak widać z powyższego opisu wyznaczenie zależności najmniejszej zauważalnej różnicy od wartości standardowej bodźca wymaga długotrwałych testów dla każdej wartości standardowej. Wykonanie takich badań dla rozpatrywanych w tej rozprawie krótkich sygnałów dźwiękowych byłoby bardzo interesujące. Trzeba jednak pamiętać, że uwaga uczestnika testów może się stępiać po długim okresie rozwiązywania tego samego nużącego zadania. Poza tym takie długotrwałe testowanie zwiększałoby koszty przedsięwzięcia. Dlatego w niniejszej pracy przyjęto prostszą procedurę, w której każdy uczestnik badania sam dokonuje zmian sygnałów porównywanych z sygnałem odniesienia (standardowym) i sam decyduje kiedy różnica sygnałów staje się dla niego zauważalna. Przypomina to trochę pracę stroiciela fortepianów, który sam decyduje, kiedy każda z trzech strun dla tego samego dźwięku jest tak samo nastrojona. Taka procedura należy do klasy metod adaptacyjnych [20], s. 13. Zestawianie ze sobą niewiele różniących się bodźców pozwala bardziej skutecznie rozróżnić ich intensywność, niż w przypadkowym przedstawianiu par, jak tego wymaga tworzenie krzywej psychometrycznej. Przekonującym tego przykładem jest rozróżnianie barw i stopnia szarości [23], zob. także przykłady na stronie [24].

Stosunek najmniejszej zauważalnej różnicy i standardowej wartości intensywności bodźca nosi nazwę ułamka Webera (*ang. Weber Fraction, [25]*).

3.2 Prawa psychofizyczne

W wielu obszarach badań interesującą wielkością jest zależność intensywności wrażenia od intensywności bodźca. Taka funkcja nosi nazwę prawa psychofizycznego (*ang. Psychophysical Law*) [20], s. 14. Oczywistą trudność sprawia tu ilościowe określenie wrażenia, gdyż to ostatnie bywa raczej trudno mierzalnym subiektywnym odczuciem. Pierwsze sformułowane prawo psychofizyczne zostało zaproponowane na podstawie bardziej uchwytneho pomiaru jakim jest zależność najmniejszej zauważalnej różnicy od intensywności bodźca.

3.2.1 Prawo Webera-Fechnera

Prawo Ernsta Heinricha Webera wg niektórych źródeł nazwane jako prawo Bouguera-Webera [26], wyraża relację między fizyczną miarą bodźca a reakcją zmysłów. Minimalna różnica jaka potrzebna jest do rozróżnienia dwóch bodźców zależy od wielkości wrażenia. Weber zauważył tę prawidłowość porównując wagę ciężaru unoszonego w ręce. Należy pamiętać, że odczuwanie bodźca to zawsze opinia subiektywna. Pierwsze sformułowanie prawa Webera (1834) brzmiało:

„Jeśli porównywane są wielkości bodźców, na naszą percepcję oddziałuje nie arytmetyczna różnica pomiędzy nimi, lecz stosunek porównywanych wielkości” [27].

Używając wprowadzonych powyżej terminów można to wyrazić w ten sposób, że zmiana wrażenia jest funkcją ułamka Webera.

Sformułowanie prawa Webera zostało zapisane 30 lat później w postaci matematycznej przez jego ucznia Gustava Fechnera, jak to przedstawia wzór (1):

$$dB_{DL} \propto B, \quad (1)$$

gdzie dB_{DL} oznacza najmniejszą zauważalną zmianę bodźca, a B ilościową intensywność bodźca. Możemy to odczytać tak, że najmniejsza zauważalna zmiana bodźca jest proporcjonalna do wielkości samego bodźca. Inaczej

$$dB_{DL} = \frac{B}{k}, \quad (2)$$

gdzie k jest stałym współczynnikiem. Oznacza to, że ułamek Webera jest stały i niezależny od intensywności bodźca.

$$\frac{dB_{DL}}{B} = k^{-1}. \quad (3)$$

Warto zauważyć, że w tym sformułowaniu nie występuje subiektywna wielkość wrażenia [28].

Prawo Webera-Fechnera zapisuje się często w postaci różniczkowej, co od-

powiada zacytowanemu zdaniu Webera [27]:

$$dw \propto k \frac{dB}{B}, \quad (4)$$

gdzie w oznacza subiektywne wrażenie, B fizyczną wielkość bodźca a k jest współczynnikiem proporcjonalności wyznaczanym doświadczalnie. We wzorze (1) zmiana wrażenia jest proporcjonalna do ilorazu Webera. Tu po raz pierwszy występuje wielkość wrażenia w . Jeżeli założymy, że najmniejsza zauważalna zmiana bodźca dB_{DL} jest wielkością małą oraz oznaczymy odpowiadającą jej zmianę wrażenia jako Δw_{lim} , to otrzymamy iloraz różnicowy

$$\frac{\Delta w_{lim}}{dB_{DL}} = \frac{k}{B} \quad (5)$$

równy w przybliżeniu pochodnej

$$\frac{dw}{dB} = \frac{k}{B}. \quad (6)$$

Zauważmy, że jeżeli równanie (5) ma odpowiadać pierwotnemu prawu Webera (3), to wielkość oznaczająca zmianę wrażenia wywoływaną przez najmniejszą zauważalną zmianę bodźca musi być stała, tj. niezależna od wielkości bodźca. Wielkość ta ma wpływ na wartość stałego ilorazu Webera

$$\frac{dB_{DL}}{B} = \frac{\Delta w_{lim}}{k} \quad (7)$$

Wydaje się, że jest to pewne założenie dotyczące mechanizmu wywoływania wrażeń przez bodźce, gdyż nie istnieje reguła zabraniająca istnienia zależności progu wrażenia od intensywności bodźca B . Równanie (6) jest równaniem różniczkowym, którego rozwiązanie jest funkcją logarytmiczną.

$$w(B) = k \ln(B) + C = k \ln\left(\frac{B}{B_0}\right), \quad (8)$$

gdzie $C = -k \ln(B_0)$ jest stałą całkowania.

Równanie (8) jest przykładem prawa psychofizycznego, często utożsamianego z prawem Webera-Fechnera. Orzeka ono, że wrażenie jest logarytmiczną funk-

cją bodźca. Prawo to ma uzasadnienie w pierwszym prawie biologicznym: bodźce słabe, bliskie progu bezwzględnego są odbierane z dużą czułością, wyrażoną przez pochodną $\frac{dw}{dB}$. Np. wydarzenia rozgrywające się blisko są dobrze różnicowane, a dalsze z mniejszą dokładnością, co pozwala reagować osobnikowi na zbliżające się zagrożenie. Podobnie dźwięki o głośności bliskie tła są odbierane z dużą dokładnością, a dźwięki bardzo głośne mniej dokładnie. Analogiczne prawo dotyczące postrzegania natężenia światła jest podstawą używanej w astronomii wielkości gwiazdowej *magnitudo* [29]. Różnica wielkości gwiazd $m_2 - m_1$ w skali magnitudo jest równa $-2,5 \log_{10} \frac{I_2}{I_1}$, gdzie I_2 i I_1 są jasnościami tych gwiazd w luksach. Zdziwiającą własnością ludzkiego słuchu jest to, że wrażenie nazywane wysokością dźwięku jest z bardzo dużą dokładnością i w szerokim zakresie logarytmiczną funkcją okresu sygnału.

3.2.2 Prawo Stevensa

Krzywa logarytmiczna nie jest jedyną funkcją o pochodnej malejącej ze wzrostem argumentu, co jest potrzebne do przetrwania osobnika. W r. 1957 Stanley Smith Stevens zaproponował uogólnienie zależności (8) do postaci zwanej prawem Stevensa [30]

$$w(B) = kB^N. \quad (9)$$

Ponieważ nigdy nie dysponujemy pełnym zakresem bodźców B w wielu przypadkach do tych samych danych możemy dopasować zarówno krzywą logarytmiczną wg równania (8) jak i potęgową (9) z wykładnikiem $0 < N < 1$. Prawo Stevensa obejmuje pewne reakcje psychofizyczne nie dające się opisać prawem Webera-Fechnera, np. postrzeganie długości odcinka, gdzie $N \approx 1$. W tym przypadku pochodna wrażenia względem bodźca jest stała, co oznacza, że najmniejsza zauważalna zmiana bodźca jest niezależna od B . Wrażenie wywołwane wstrząsem elektrycznym wykazuje wykładnik $N > 1$, tzn. najmniejsza zauważalna zmiana bodźca rośnie z B szybciej niż liniowo [zob. przykłady na stronie Appalachian State University/Boone, North Carolina [31]]. Próbę ujednoczenia praw Stevensa i Webera-Fechnera podjął Fan-Gang Zeng, w szczególności w zastosowaniu do związku głośności z natężeniem dźwięku [32].

4 Psychofizyczne podstawy percepcji cech dźwięku

W podstawowej nauce o muzyce wyróżnia się zazwyczaj cztery cechy dźwięku

„ a) wysokość, b) głośność, c) czas trwania, d) barwę” [5].

Trzeba od razu zaznaczyć, że nazewnictwo to dotyczy sygnałów akustycznych, w których wysokość dźwięku można w jakiś sposób określić. W innych przypadkach nauczyciele muzyki mówią o „szmerach”, choć można by tu wyróżnić także stuki, trzaski, syki i inne odgłosy, z których nie daje się utworzyć melodii. Najprostszy eksperyment polega na odtworzeniu jakiejś znanej melodii na samogłosce, np. a, lub u, oraz na spółgłosce bezdźwięcznej np. s, lub t. W drugim przypadku nie jest to możliwe. Istnieją też spółgłoski dźwięczne, np. z lub d; można z nich utworzyć namiastkę melodii, lecz efekt wydaje się jakoś zaszumiony. Niektórzy autorzy uważają każdy sygnał akustyczny za dźwięk [33], inni zaś dźwiękami nazywają tylko te sygnały, którym można przypisać wysokość dźwięku. Np. na rządowym portalu edukacyjnym Zintegrowana Platforma Edukacyjna [34] czytamy:

„Dźwiękami w języku potocznym nazywamy to, co może być zarejestrowane przez ludzkie ucho. Ścisłej należałoby nazwać to wrażeniami słuchowymi, a dźwięki są jednymi z nich. Ale co to jest? Otóż źródłem dźwięku są ciała drgające, które przekazują swoje drgania ośrodkowi sprężystemu (najczęściej jest to powietrze), inicjując w tym ośrodku falę dźwiękową, a ta rejestrowana jest przez zmysł słuchu.”

Sugeruje to, że ciała wykonujące ruch inny niż drgający, np. woda spływająca w wodospadzie, nie wytwarza dźwięków, a może nawet nie przekazuje swego ruchu ośrodkowi sprężystemu jakim jest powietrze.

Wymienione cechy dźwięku nazywane są czasem „atrybutami subiektywnymi” [20]. Nazwa ta jest o tyle adekwatna, że do określenia tych cech potrzeba osobnika, lub urządzenia rejestrującego sygnał akustyczny w sposób analogiczny do narządu słuchu człowieka lub zwierzęcia. Jednak przymiotnik „subiektywny” może też sugerować zmienność osobniczą podczas, gdy większość tych cech można ująć w sposób ilościowy za pomocą pewnych praw psychofizycznych, a wielkości

postrzeganych wrażeń bywają bardzo zbliżone dla różnych osobników jednego gatunku lub populacji, co jest obiektywnym zjawiskiem psychofizjologicznym, lub co najmniej interpersonalnym. Co więcej, odchylenia od występującej w populacji normy postrzegania różnych bodźców mogą służyć jako wskazówki diagnostyczne przy wykrywaniu chorób, zaburzeń i dysfunkcji [35].

Ponieważ w obecnej pracy przypisanie cech dźwięku sygnałom nie zawsze kwalifikowanym jako dźwięki ma istotne znaczenie, należy tu przypomnieć znane fakty i poglądy. W analizie skrajnie krótkich impulsów akustycznych najbardziej użytecznymi cechami dźwięku będą wysokość dźwięku i barwa dźwięku.

4.1 Percepcja wysokości dźwięku

Percepcję wysokości dźwięku tłumaczą teorie: miejsca i czasu. Teorie te różnią się od siebie, jednak badacze twierdzą, że obie są prawdziwe i częściowo, słuchając, wykorzystujemy mechanizmy znajdujące się w obu. Jednak przed omówieniem czym charakteryzuje się teoria miejsca i czasu należy zaznajomić się z pojęciem wysokości dźwięku. Wysokość dźwięku występuje, w sposób najmniej dyskusyjny, kiedy przebieg ciśnienia akustycznego ma charakter okresowy. Częstotliwością, tradycyjnie wyrażaną w Herzach - Hz jest liczba okresów w jednostce czasu. 1 Hz odpowiada jednemu cyklowi na sekundę. Za wzorcową częstotliwość przyjmuje się wysokość dźwięku „a” odpowiadającego 440 Hz [36].

Zagadnienie wysokości dźwięku sygnałów akustycznych o dobrze określonej okresowości zajmowało ludzi znacznie wcześniej niż odkrycie związku tej cechy dźwięku z drganiami okresowymi. Od czasów Pitagorasa wiadano, że różnice wysokości dźwięku, zwane w nauce o muzyce interwałami, zależą od stosunków wielkości, które później okazały się odwrotnie proporcjonalne do częstotliwości. Tymi wielkościami były długości strun. Wrażenie różnicy w odpowiedzi na iloraz wartości bodźca odpowiada dokładnie zależności wrażenia od bodźca w postaci logarytmicznej. Najbardziej wymownym tego przykładem jest numeracja klawiszy klawiatur muzycznych (*ang. keyboard*) w systemie MIDI (*ang. Music Instrument Digital Interface*) stosowana do dziś w programach komputerowych i w produkcji instrumentów. Numer m klawisza związany jest z odpowiadającą mu częstotliwo-

ścią f_m w Hz w następujący sposób

$$m = 12 \cdot \log_2\left(\frac{f_m[\text{Hz}]}{440}\right) + 69. \quad (10)$$

Oznacza to, że klawisz wyzwalający dźwięk a^1 (a razkreślne) o częstotliwości 440 Hz ma numer 69. Dźwięk o częstotliwości dwa razy większej, tj. 880 Hz ma numer o 12 większy. Oczywiście interwał między nimi to oktawa. Interwał odpowiadający sąsiednim klawiszom odpowiada stosunkowi częstotliwości $2^{1/12}$. Nosi on nazwę półtonu temperowanego, lub sekundy małej w stroju równomiernie temperowanym. Mimo, że strój równomiernie temperowany nie ma pełnego uzasadnienia fizycznego i fizjologicznego [37], jest on powszechnie stosowany w najprostszych instrumentach klawiszowych dostępnych na rynku. Pozwala on odtwarzać melodie i bardziej złożone utwory z zadowalającą dokładnością.

Jednostką wysokości dźwięku jest cent będący jedną setną półtonu temperowanego. Mając dany stosunek częstotliwości f_2/f_1 interwał w centach wyraża się wzorem

$$1200 \log_2\left(\frac{f_2}{f_1}\right) \quad (11)$$

[38]. Rzecz jasna logarytmiczna zależność wyrażona wzorem (10) odpowiada dokładnie prawu Webera-Fechnera. Jest ona najbardziej jasna i przekonująca w przypadku sygnałów sinusoidalnych.

4.2 Wysokość dźwięku a okresowość sygnału

Każdy przebieg okresowy można w jednoznaczny sposób rozłożyć na szereg Fouriera, tj. sumę funkcji sinusoidalnych o częstotliwościach będących naturalnymi wielokrotnościami najniższej częstotliwości, zwanej podstawową (*ang. fundamental*). Odwrotność częstotliwości podstawowej jest równa okresowi sygnału. Amplitudy poszczególnych sinusoid, zwanych alikwotami, stanowią widmo amplitudowe, a moduły kwadratów amplitud widmo mocy takiego dźwięku. Ponieważ sygnał akustyczny w postaci pojedynczej sinusoidy nazywa się w nauce o muzyce tonem prostym, okresowy przebieg ciśnienia akustycznego zawierający pewną liczbę tonów prostych nosi tam nazwę wielotonu harmonicznego. Gdy z jakichś powodów częstotliwości alikwotów nie są całkowitymi wielokrotnościami jednej częstotliwo-

ści podstawowej, mówimy o wielotonach nieharmonicznych; gdy przyczyną tego jest nieliniowość równań ruchu mówi się o anharmonizmie. Przykładami wielotonów nieharmonicznych są dźwięki fortepianu [39], w których wysokość dźwięku jest stosunkowo dobrze odczuwalna, oraz odgłosy obiektów dwu- i trójwymiarowych, np. bębnow i dzwonów, gdzie określenie wysokości dźwięku bywa trudne [40] zob. także [41]. Istnieją też sygnały akustyczne o widmach ciągłych. Do nich należą wymieniane w nauce o muzyce szmery, a także szumy wiatru, wodospadów, hałasy przemysłowe itp. Ciągłe widmo mają też sygnały o krótkim czasie trwania. Np. w idealnym sygnale w postaci funkcji δ Diraca zawierają się wszystkie rzeczywiste wartości częstotliwości z tą samą amplitudą. Od szumu odróżnia takie krótkie sygnały ścisły i niezmienny związek między fazami składowych sinusoid. Np. składowe impulsu w postaci funkcji δ Diraca mają fazy jednakowe.

Amerykański Narodowy Instytut Wzorców (ANSI) w swej definicji wysokości dźwięku unika odniesień do fizycznej postaci sygnału i definiuje ją tak:

„ta cecha wrażenia słuchowego, według której dźwięki można uporządkować w skali rozciągającej się od niskich do wysokich” [42].

Podobną definicję znajdujemy na stronie Amerykańskiego Towarzystwa Akustycznego [43]

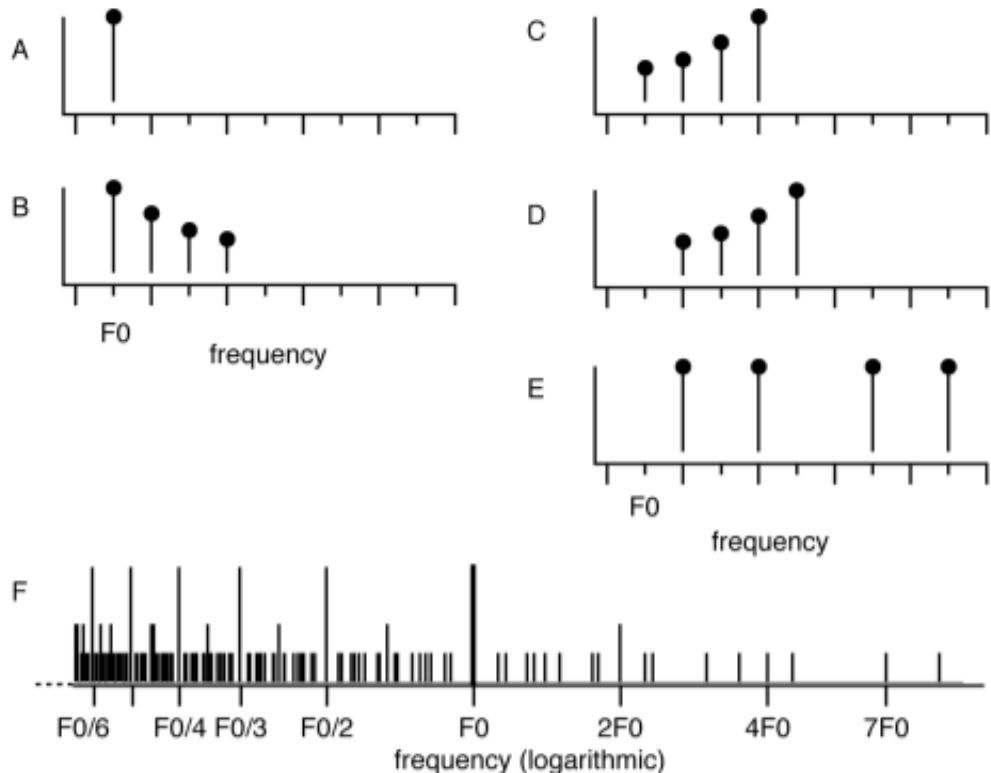
„11.01, wysokość dźwięku, Ta cecha wrażenia słuchowego, według której dźwięki są porządkowane w skali stosowanej dla melodii w muzyce.”

Rzecz jasna jedną z takich skal jest już przywołana numeracja klawiszy w systemie MIDI (10). Tam jednak mamy jednoznaczny związek wrażenia wysokości dźwięku z okresem wytwarzanego sygnału.

Najprostszą funkcją okresową jest sinusoida, czyli ton prosty. Wysokość dźwięku jest dla takich przebiegów ciśnienia akustycznego powszechnie odczuwana. Prawo psychofizyczne łączące wysokość dźwięku z częstotliwością było przedmiotem badań wielu naukowców. W znanej monografii [44] znajdujemy na rys. 6.1 zależność najmniejszej dostrzegalnej różnicy częstotliwości od samej częstotliwości. Jest to funkcja bliska liniowej przedstawiona w skali podwójnie logarytmicznej; współczynnik nachylenia wykresu podwójnie logarytmicznego jest bliski jedności, co odpowiada prawu Webera-Fechnera. Nieco dziwi stwierdzenie autora:

„The theoretical significance of this is not clear.” (s. 206).

Rysunek 6.5 z tej samej monografii przedstawia ułamek Webera (Relative Difference Limen of Frequency) dla częstotliwości tonu prostego. Dla dostatecznie długich sygnałów jest on słabo zależny od częstotliwości i mieści się w zakresie od 0,125% do 0,25% co odpowiada stałemu interwałowi od 2,16 do 4,32 centa (tj. setnych półtonu temperowanego, [38]). Jest to zgodne z prawem psychofizycznym wyrażonym wzorem (10) i w dodatku potwierdza użyteczność jednostki cent oraz wyznacza wielkość Δw_{lim} oznaczająca zmianę wysokości dźwięku wywołwaną przez najmniejszą zauważalną zmianę częstotliwości (zob. wzór (7)). Dla innych przebiegów okresowych wysokość dźwięku jest w zasadzie określona przez częstotliwość podstawową. Dobrze opisuje to podejście spektralne w publikacji Alaina de Cheveigné [45]. Podejście spektralne oparte jest na analizie Fouriera. Alain De Cheveigne [45] opisuje zjawisko na przykładach przedstawionych na rysunku 6.



Rysunek 6: Przykłady widm wywołujących wrażenie tej samej wysokości dźwięku. A: Widmo tonu prostego; B: Widmo wielotonu harmonicznego wywołującego tę samą wysokość dźwięku co widmo A; C: najwyższa harmoniczna ma najwyższą amplitudę jednak wysokość dźwięku nie zmienia się; D: Najniższa (podstawowa) harmoniczna ma amplitudę zerową lecz periodyczność sygnału jest taka sama - nie zmienia się wysokość dźwięku; E: Żaden z interwałów pomiędzy składowymi harmonicznymi nie odpowiada jego wysokości dźwięku, która jednak (dla niektórych słuchaczy) jest taka sama jak pozostałych tonów; F: Dla każdej składowej harmonicznnej dodano sub-harmoniczne, dzieląc częstotliwość składowej przez kolejne liczby naturalne, następnie dla każdej sub-harmonicznnej zwiększono odpowiadający przedział histogramu [45].

Alternatywne prawo psychofizyczne zostało zaproponowane przez S. Stevensa i J. Volkmana [46]. To prawo psychofizyczne nazywa się skalą wysokości dźwięku w melach nieco różniącą od skal używanych w muzyce. Dopasowanie krzywej do wyników eksperymentów S. Stevensa i J. Volkmana budzi do dziś dyskusje [1].

4.2.1 Barwa dźwięku.

Przytoczmy znów definicję Amerykańskiego Towarzystwa Akustycznego [47]:

„11.09 barwa. Ta wielowymiarowa cecha wrażenia słuchowego, która pozwala słuchaczowi ocenić, że dwa nieidentyczne dźwięki, podobnie prezentowane i mające tę samą głośność, wysokość, położenie przestrzenne i czas trwania, są niepodobne. Barwa jest związana z jakością dźwięku, często określaną przymiotnikami jakościowymi (np. jasny lub matowy). Adnotacja: Na barwę dźwięku duży wpływ mają jego zmieniające się w czasie cechy charakterystyczne, szczególnie podczas początkowej części (ataku), a także jego trwająca charakterystyka widmowa i czasowa. Barwa jest istotnym elementem identyfikacji źródła dźwięku (np. konkretnego instrumentu muzycznego) oraz sposobu jego wydobycia.”

Gdy mamy do czynienia z okresowymi sygnałami akustycznymi, barwę dźwięku określa jego widmo. Od razu pojawia się pytanie „jakie widmo?” Odpowiedź daje tu akustyczne prawo Ohma, poparte modelem szeregu rezonatorów (harfy) Helmholtza, orzekające, że narząd słuchu dokonuje analizy widmowej, ale nie jest czułe na fazy składowych harmoniczných. W skrócie mówi się, że ucho jest „głuche na fazy” [48] lub też o niezależności barwy dźwięku od faz składowych harmoniczných. Oznacza to, że we wrażeniu słuchowym sygnałów okresowych liczy się tylko widmo mocy. Nawet najprostsze programy do syntezy dźwięku pozwalają stwierdzić jak duże zmiany kształtu przebiegu okresowego nie zmieniają brzmienia, pod warunkiem, że są wynikiem tylko przesuwania faz bez naruszania amplitud składowych zob. rys. 4 w [49]. Jednym ze sposobów zmiany faz składowych harmoniczných jest odwrócenie strzałki czasu. W demonstracjach dydaktycznych pokazano, że brzmienie instrumentów, w których dźwięk jest z dobrym przybliżeniem stacjonarny (instrumenty dęte, organy, instrumenty smyczkowe) nie zmienia się znacząco przy takiej operacji. W przeciwieństwie do tego instrumenty szarpane, jak harfa lub gitara, a także fortepian są w praktyce nierozpoznawalne po odwróceniu strzałki czasu, zob. np. [50]. Przyczyną tego zjawiska jest szybka zmiana przebiegu

czasowego przy szarpnięciu lub uderzeniu struny. Kwestia zależności lub niezależności barwy od faz alikwotów była przedmiotem dyskusji S. Ohma i T.J. Seebecka, która istotnie zainspirowała H. Helmholtza [51]. Wiele nowszych eksperymentów można znaleźć na stronie [52] oraz w artykule [53]. W badaniach poziomu utleniania krwi (BOLD) metodą funkcjonalnego rezonansu magnetycznego (fMRI) w celu zidentyfikowania obszarów kory mózgowej odpowiedzialnych za percepcję cech dźwięku barwę dźwięku zmieniano modyfikując średnią ważoną widma (*ang. spectral centroid*) [54]. Oczywiście pozostaje otwarte pytanie czy barwę dźwięku można zdefiniować w przypadku przebiegów nieokresowych i widm ciągłych. Ponieważ jednak to pojęcie pojawiło się wśród odpowiedzi w badaniach behawioralnych przedstawionych w tej rozprawie, zaistniała potrzeba zdefiniowania takiej wielkości. Za różnicę barwy dźwięku przyjęto całkę z wartości bezwzględnej widm porównywanych dźwięków (rozdz. 7.2).

4.2.2 Odczuwanie głośności.

Miarą intensywności bodźca dźwiękowego jest ilość energii doprowadzanej do jednostki powierzchni organizmu, w tym do narządu słuchu, w jednostce czasu. Wielkość ta zwana natężeniem dźwięku wyrażana jest w jednostkach $\frac{J}{m^2s} = \frac{W}{m^2}$. Już dawno zauważono jednak, że wrażliwość na natężenie dźwięku (podobnie jak na natężenie światła, natężenie dotyku, stężenie substancji zapachowej...) maleje ze wzrostem natężenia. Z tego powodu wprowadzono bardziej realistyczną wielkość jaką jest poziom natężenia (*ang. Sound Intensity Level SIL*, lub *Sound Level SL*). Poziom SIL natężenia dźwięku I definiuje się jako

$$SIL = 10 \log_{10} \frac{I}{I_0} dB, \quad (12)$$

gdzie I_0 jest pewną umowną wielkością bodźca, a dB oznacza jednostkę decybel [55]. Jako natężenie odniesienia stosuje się wartość

$$I_0 = 10^{-12} \frac{W}{m^2}, \quad (13)$$

co odpowiada przeciętnemu progowi słyszalności tonu prostego o częstotliwości $1000Hz$. $20\mu P$.

Jednostka bela została wprowadzona w 1928 roku przez Bell System, aby możliwy był bardziej realistyczny opis percepcji natężenia dźwięku [56]. Skalę beli uogólniono w celu umożliwienia wyrażania logarytmów stosunków dowolnych wielkości. Z tego powodu beli, a jeszcze częściej decybeli, używa się jako jednostki wzmocnienia w elektronice i wszędzie tam, gdzie jakiś parametr zmienia się o wiele rzędów wielkości, gdyż łatwiej wtedy obliczać te rzędy wielkości, czyli logarytmy, niż posługiwać się olbrzymimi wartościami parametrów. Logarytm, zazwyczaj dziesiętny, stosunku danej wielkości do pewnej wartości odniesienia nazywa się poziomem tej wielkości. W naszym przypadku mamy więc poziom natężenia (mocy) dźwięku (*ang. sound intensity level, SIL*), lub czasem, poziom dźwięku (*ang. sound level SL*) obok poziomu amplitudy ciśnienia akustycznego (*ang. sound pressure level, SPL*). Najmniejsza różnica poziomów dźwięku rejestrowana przez ludzkie ucho to ok. 1 *dB*. Przybliżona stałość tej różnicy świadczy o adekwatności miary decybelowej w opisie wrażenia głośności. W przypadku dźwięku próg słyszenia jest poziomem odniesienia. Wzorcowo przyjmuje się próg słyszalności tonu prostego o częstotliwości 1000 *Hz*. Próg ten wynosi dla zdrowych młodych osób 2×10^{-5} Pa, co odpowiada natężeniu odniesienia $I_0 = 10^{-12} \frac{W}{m^2}$. Dla przykładu szept posiada moc około 100 razy większą niż próg słyszenia więc w belach mamy dla szeptu wartość $\log(\frac{100}{1}) = 2$. Dla porównania natężenie dźwięku promu kosmicznego (najgłośniejszy jak do tej pory spotykany odgłos) wynosi 14 *beli*. Pomimo ogromnego hałasu nie jest to wysoka wartość, dlatego w praktyce skalę podzielono na jednostki dziesięciokrotnie mniejsze i dlatego najczęściej używa się decybeli. Decybel jest logarytmiczną jednostką miary odpowiadającą $\frac{1}{10}$ bela, oznaczany jest jako *dB*. Jednostki decybela używa się w sytuacjach, gdzie porównywane są wielkości zmieniające się liniowo w bardzo szerokim zakresie - szczególnie zmiany względne np. podczas rejestracji głośności dźwięku lub wrażeń węchowych, które rejestrowane są zgodnie z prawem Webera-Fechnera [57].

Na głośność dźwięków ma również wpływ czas ich trwania. Ogólnie, zawsze gdy poziom jakiegoś bodźca odpowiada wywoływanemu przez ten bodziec wrażeniu, prawo Webera-Fechnera dla zależności wrażenia od bodźca jest spełnione. Wrażenie wyraża się wówczas jako pewien logarytm względnej wartości bodźca. Dźwięki krótsze niż 100 *ms* są odbierane jako dźwięki cichsze, jednak można zauważyć duże rozbieżności w badaniach na ten temat [44]. Jeśli czas trwa-

nia dźwięku wynosi mniej niż 80 ms , możemy zauważyć prawidłowość, że stała głośność odpowiada stałej energii dźwięku [58]. Różnica pomiędzy prawem Webera a prawem Stevensa wzięła się z tego czy bazujemy na analizie odczuć różnic czy na bezwzględnych wielkościach odczuć [59]. Alternatywą dla prawa Webera-Fechnera jest sformułowane przez Stanleya Smitha Stevensa prawo orzekające, że wrażenie jest proporcjonalne do pewnej potęgi bodźca [30]. To prawo potęgowe jest nieco bardziej ogólne, ale nie ma już w nim reguły, że ta sama względna zmiana bodźca powoduje taką samą zmianę wrażenia niezależnie od bezwzględnej siły bodźca. Ostatnio wykazano, że obydwa te prawa: Webera-Fechnera i Stevensa można ująć w jednym wzorze [32]. W badaniach przedstawianych w tej rozprawie słuchacze oceniają impulsy w postaci funkcji Gaussa z jednym wyraźnym maksimum. Impulsy te są znormalizowane do wartości tego maksimum. Nie jest to jedyny możliwy wybór. Np. można by było znormalizować impulsy do pola powierzchni pod funkcją Gaussa. Wtedy każdy impuls dostarczałby tę samą porcję energii. Jednak doświadczenie autorki i jej współpracowników wskazuje, że w takim przypadku impulsy o czasie trwania ok $0,5\text{ ms} - 20\text{ ms}$ byłyby praktycznie niesłyszalne. Byłoby oczywiście interesujące zbadać czy/jak zmieniają się wyniki uzyskanych tu pomiarów przy zmianie polaryzacji.

4.2.3 Percepcja zjawisk czasowych

W muzyce czas trwania dźwięku wiąże się z rytmem i artykulacją. Czas narastania i wybrzmiewania mają wpływ na barwę dźwięku jak to opisano w rozdziale 4.2.1. Czas trwania tonu prostego trwającego więcej niż 100 ms percepowany jest tak jak mówi jego fizyczna miara. Dźwięki krótsze niż 100 ms wydają się człowiekowi dłuższe od fizycznej miary. W badaniu przeprowadzonym przez [58] wykazano, że subiektywny czas trwania tonu wynoszącego 1000 Hz , o poziomie 60 dB oraz czasie trwania 30 ms odczuwany jest jako ton nieco dłuższy niż 30 ms . Dla porównania ten sam ton ale o czasie trwania 10 ms wydaje się być dwukrotnie dłuższy a o czasie 3 ms - trzykrotnie dłuższy. Wyniki badań zestawione na rys. 6.5 str. 210 w monografii B. C. J. Moore'a [44] wskazują, że ułamek Webera dla częstotliwości jest prawie stały dla tonów prostych trwających ponad 200 ms . Jednak ze skracaniem czasu trwania do $6,25\text{ ms}$ krzywa zależności ułamka Webera

od częstotliwości staje się coraz bardziej wygięta ku dołowi z minimum ok 2 kHz . Wskazuje to na odstępstwa od prawa Webera-Fechnera, a więc i od typowej skali muzycznej dla dźwięków krótkich. Obserwacje poczynione w artykule [3] wskazują, że odczuwana wysokość dźwięku wykazuje pewien wzrost, gdy czas trwania tonu prostego maleje poniżej 1 ms . Autorzy [58] zwracają również uwagę na różne odczucia czasu trwania jeśli mamy do czynienia z ocenieniem czasu trwania tonu oraz przerwy.

5 Mechanizmy i teorie słyszenia

Jest kilka teorii wyjaśniających w jaki sposób dokonywana jest percepcja dźwięku. Podstawową teorią rezonansową jest teoria Harfy Webera powstała w 1841 roku [60]. Definiuje ona błonę podstawną ślimaka jako konstrukcję z włókien. U podstawy ślimaka są one krótsze niż u jego szczytu. Włókna te porównuje się do strun harfy, gdyż przypominają kolejne jej struny. W roku 1863 teoria ta została rozwinięta przez Helmholtza i znana jest pod nazwą **teorii rezonansowej Helmholtza**. Wg niego, opisywane przez Webera, krótkie struny, odpowiedzialne są za przenoszenie tonów wysokich a długie - niskich. Włókna drgając pobudzają zakończenia nerwowe. Kolejną teorią jest teoria **telefoniczna Rutherforda** porównana do drgań membrany w słuchawce telefonu. Wg niej, zmiany impulsów nerwowych zachodzą tak jak zmiany prądu elektrycznego na wyjściu z mikrofonu a analiza dokonuje się później, w mózgu [18].

Dzisiaj uważa się, że przetwarzanie sygnałów akustycznych, tj. dobiegających do błony bębenkowej, a później do okienka owalnego przebiegów ciśnienia akustycznego, odbywa się w ślimaku. Jego anatomię dobrze poznał Herman von Helmholtz, gdy w początkach kariery pracował jako chirurg [61]. Pozwoliło mu to, już w okresie rozkwitu jego działalności badawczej w Heidelbergu, stwierdzić, że główną rolę w tym procesie odgrywają ruchy błony podstawnej przekazywane do nerwów słuchowych za pośrednictwem komórek rzęsatych narządu Cortiego. Odkrycie to znalazło się w fundamentalnym dziele „Nauka o wrażeniach dźwiękowych jako fizjologiczna podstawa teorii muzyki” (*Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik*) po raz pierwszy wydanym w r. 1863 [62].

5.1 Teoria miejsca

W Teorii miejsca przyjmuje się, że każdej częstotliwości składającej się na sygnał odpowiada pewne określone położenie wzdłuż osi ślimaka, w którym to położeniu-miejscu błona podstawna i otaczający ją płyn – przychłonka wykonują ruchy o największej amplitudzie [63]. Za odczuwanie obecności danej częstotliwości w sygnale dźwiękowym odpowiedzialne są więc miejsca silnych pobudzeń w narzą-

dzie Cortiego. Miejsce to jest jedno dla konkretnej częstotliwości tzw. *częstotliwość charakterystyczna*. Wysokość dźwięku określana jest zatem przez miejsca w których błona podstawna wykonuje najintensywniejsze ruchy i w ten sposób przekazuje sygnały do neuronów słuchowych. Wygenerowane impulsy neuronowe trafiają do określonego miejsca kory słuchowej [64], która wykazuje tonotopową organizację neuronów słuchowych, czyli tworzy mapę, na której przestrzennie ustawione są częstotliwości przetwarzane w mózgu. Za twórcę teorii miejsca uznaje się Hermana Helmholtza, choć znacznie wcześniej zakładano jej istnienie [65]. Koncepcja Helmholtza zakłada, że włókna błony podstawnej zachowują się analogicznie do rezonatorów akustycznych, np. strun harfy, tzn. pod wpływem padającego sygnału oscylującego wykonuje drgania o tej samej częstotliwości z amplitudą zależną od amplitudy wymuszenia i stałej tłumienia. Stąd częsta nazwa tego modelu: „model harfy”. Ponieważ w teorii miejsca, miejsce największego pobudzenia przesuwają się wzdłuż błony podstawnej w zakresie od około 50 Hz do 16 000 Hz, zob. rys. 4, nie bierze ona pod uwagę całego zakresu częstotliwości słyszanych [66].

Wiele dyskusji wywołały odkrycia Georga von Békésy’ego wskazujące, że w czasie percepcji tonu prostego porusza się znaczna część błony podstawnej i wypełniającej ślimak perylimfy, co zaprzeczyło założeniom Helmholtza o lokalizacji pobudzenia dla każdej częstotliwości z osobną, zob. [44] rys. 1.10. Wyjaśnienie przyniosły późniejsze pomiary wykonane dla żywych ssaków w odróżnieniu od preparatów prosektoryjnych Békésy’ego. Okazało się mianowicie, że aktywność zakończeń nerwów słuchowych jest znacznie lepiej zlokalizowana, w okolicach szerokich maksimum ruchów mechanicznych błony podstawnej. Ilustrują to tzw. krzywe strojenia (*ang. tuning curves*) przedstawiające próg pobudzenia danego miejsca ślimaka w zależności od częstotliwości podawanego tonu [67, 68] zob. też [44] rys. 1.11. W ten sposób teoria miejsca, przypisująca każdej pojedynczej częstotliwości dobrze określoną lokalizację na długości ślimaka odzyskała swoje uzasadnienie. Pozostało jednakże pytanie o fizyczny mechanizm przekształcający szerokie maksimum ruchów mechanicznych w wąską krzywą strojenia. Jednym z wyjaśnień jest koncepcja silnej nieliniowości odpowiedzi układu nerwowego prowadzącej do tzw. bifurkacji Hopfa [69], chociaż dyskusja na ten temat dalej trwa [70].

Bifurkacja– zjawisko skokowej zmiany własności modelu matematycznego przy drobnej zmianie jego parametrów (np. warunków początkowych procesu albo

warunków brzegowych). Szczególnie często spotykane i istotne jest to pojęcie przy rozwiązywaniu równań różniczkowych oraz badaniu fraktali (i teorii chaosu).

Nie jest wiadome w jaki sposób, za pomocą tej teorii, wyjaśnić percepcję dźwięków złożonych, szumów o wyczuwalnej wysokości oraz dźwięków nieharmonicznych. W przypadku wielotonów, które nie zawierają częstotliwości podstawowej (pierwszej harmonicznej lub kilku pierwszych harmonicznych) wysokość dźwięku ciągle odpowiada częstotliwości podstawowej tzw. *wysokości rezydualnej*. Podważa to akustyczne Prawo Ohma w jego pierwotnej postaci, na którym opiera się teoria miejsca i wg którego słuch analizuje widmo amplitudowe dźwięku i nie jest czuły na fazę składowych widmowych, czyli wysokość dźwięku może być słyszalna jedynie wtedy, kiedy fala zawiera w tej częstotliwości pewną moc. Aktualnie prowadzone są badania wykorzystujące związek ludzkiego układu słyszenia z akustycznym prawem Ohma dotyczące np. rozpoznawania mowy [71].

5.2 Teoria czasu

Jest to główna alternatywa dla teorii miejsca. Jest to teoria powiązana z teorią Volleya, w której włókna neuronów układu słuchowego operujące zgodnie z zasadą *phase-locking* czyli odwzorowaniem fazy drgania w fazie potencjału receptorowego reagują na dźwięk odwzorowując częstotliwość zmian potencjału receptorowego komórek rzęsatych [72]. Za słusznością teorii miejsca przeprowadzono badania. Jedno z nich dotyczyło percepcji szumu białego, którego widmo mocy jest płaskie i nie niesie informacji o rozkładzie aktywności wzdłuż narządu Cortiego. Uczestnicy badania wyczuwali wysokość dźwięku, za sprawą przerw występujących między 100 a 250 Hz [44]. Nasuwa to pewną obserwację, że słuch ludzki może wykrywać swego rodzaju rytm, bazując na odstępach czasowych pomiędzy kolejnymi impulsami neuronowymi. Narząd słuchu wykrywa najczęściej występujące przerwy i przetwarza odwrotność czasu przerwy na odczucie wysokości dźwięku. Odpowiedzialny za to jest centralny układ nerwowy [10].

Najbardziej słuszne wydaje się połączenie obu teorii. Istnieją dowody na potwierdzenie obu [45].

Od czasów Békésy'ego zwraca się uwagę na rolę hydrodynamiki ślimaka i zawartych w nim płynów. Ta dziedzina badań nazywa się mechaniką ślimaka

(*ang. cochlear dynamics*) [73]. Ciągłe pojawiają się nowe koncepcje na ten temat [74].

6 Dotychczasowe wyniki badań krótkich sygnałów dźwiękowych

Jak widać z przeglądu cech dźwięku w rozdziale 4, wysokość dźwięku jest o tyle wyjątkowa, że wiąże się z okresowością sygnału, co powoduje, że do określenia wysokości dźwięku powinniśmy używać w zasadzie sygnałów nieskończenie długich. Żadna bowiem funkcja różna od zera na odcinku o skończonej długości nie może być okresowa, a zatem częstotliwość takich sygnałów zawsze jest obarczona pewną niedokładnością. Z tego powodu kwestia postrzegania wysokości dźwięków bardzo krótkich jest w pewnym sensie problemem granic stosowności ogólnie przyjętych pojęć. W podstawowej monografii B.C.J. Moore'a [44] rozważa się czasową rozdzielczość narządów słuchu (rozdział 5, str. 169), zwracając uwagę, że

„Główna trudność w pomiarze czasowej rozdzielczości układu słuchowego polega na tym, że zmiany w przebiegu czasowym dźwięku są na ogół powiązane ze zmianami widma różnych charakteryzujących go wielkości”.

Problem widma chwilowego jest znany w elektronice. Próba jego rozwiązania jest odpowiednie oknowanie (*ang. windowing*) odcinków czasu i zakresów częstotliwości.

Zależność najmniejszej zauważalnej różnicy częstotliwości dla tonów prostych od czasu trwania tonu w zakresie $6,5\text{ ms} - 200\text{ ms}$ jest przedstawiona na rys. 6.5 (str. 210 [44]). Najkrótsze zbadane tam przebiegi zawierają ponad 1,6 okresu sinusoidy. Autor monografii słusznie zauważa, że

„im krótszy jest impuls tonu, tym szerszy zakres częstotliwości, w którym energia jest rozłożona.”.

Ciekawym wynikiem w porównaniu z dawniejszymi teoriami jest obserwacja:

„Przy krótkich czasach trwania, badani radzili sobie lepiej niż przewidywano dla wszystkich częstotliwości do 5 kHz . Poczynając od tej częstotliwości najmniejsze zauważalne różnice częstotliwości wykazywały ostry wzrost wartości.”.

Wskazuje to, że w zakresie do 5 kHz układ słuchowy działa tak, że do rozpoznania wysokości dźwięku potrzebuje odpowiedniej liczby okresów sygnału, co właśnie wyraża spadek względnej najmniejszej zauważalnej różnicy częstotliwości na rys. 6.5 w monografii [44]. W każdym razie w przypadku krótkich sygnałów akustycznych określenie ich wysokości dźwięku jest wątpliwe. Dlatego w literaturze stosuje się pojęcie „siły wysokości dźwięku” (*ang. pitch strength*). W artykule [42] siłę tę zmniejszają zaburzenia mowy. W przeprowadzonych w r. 2009 badaniach A. Rakowskiego i P. Rogowskiego wykazano, że siła wysokości dźwięku ocenianej przez osoby obdarzone słuchem absolutnym istotnie maleje, gdy czas trwania impulsu spada do dokładnie 4 okresów oscylacji, tj. w badanych dźwiękach do $4\text{--}7\text{ ms}$ [2].

W roku 2013 ukazał się artykuł autorów: J. N. Oppenheim oraz M.O. Magnasco sugerujący, że nieliniowość przetwarzania sygnałów przez ludzkie narządy słuchu sprawia, że narządy te są w stanie przewyciężyć ograniczenia wynikające z zasady nieokreśloności czas-częstotliwość [4]. Artykuł ten wzbudził wątpliwości [75] nie przyjęty w *Physical Review Letters*; szersze omówienie w [3].

Autorzy komentarza [76] zwrócili uwagę, że wielkości, których iloczyn powinien być ograniczony od dołu, przypominające postrzeganą jednocześnie najmniejszą zauważalną różnicą czasu pojawienia się sygnału i jego częstotliwości nie są, jak tego wymaga zasada nieokreśloności, niepewnościami częstotliwości i czasu. Autorzy tego komentarza uznali odwrotność czasu trwania całego „zadania 5” (ok. 2 s) przedstawionego na rys. 7 za niepewność w dziedzinie częstotliwości, a czasowy krok próbkowania sygnału za niepewność w czasie. Obliczony tak iloczyn jest rzeczywiście bardzo mały, ale sposób jego obliczania nie odpowiada chyba istocie zasady nieokreśloności podobnie jak procedura zastosowana w oryginalnym artykule Oppenheima i Magnasco [4]. Nie wiadomo więc czy małość tego iloczynu naprawdę oznacza, że jednoczesne rejestrowanie przez słuch różnic częstotliwości i czasu pojawienia się sygnału jest oczywiste. W dodatku sposób próbkowania sygnału nie może być miarą sprawności narządów słuchu. Autorzy komentarza [77], zauważyli, że sygnał podlegający „pomiarowi” przez narząd słuchu jest w procedurze Oppenheima i Magnasco (rys. 7) „mierzony” dwa razy: raz przez porównanie z „leading note”, a po chwili przez porównanie z „flanking note”. Nie jest to zatem test odpowiadający zasadzie nieokreśloności. Oczywiście dla osiągnięcia wyników przedstawionych w pracy [4] jakąś rolę odgrywa zapewne aktywność, być może

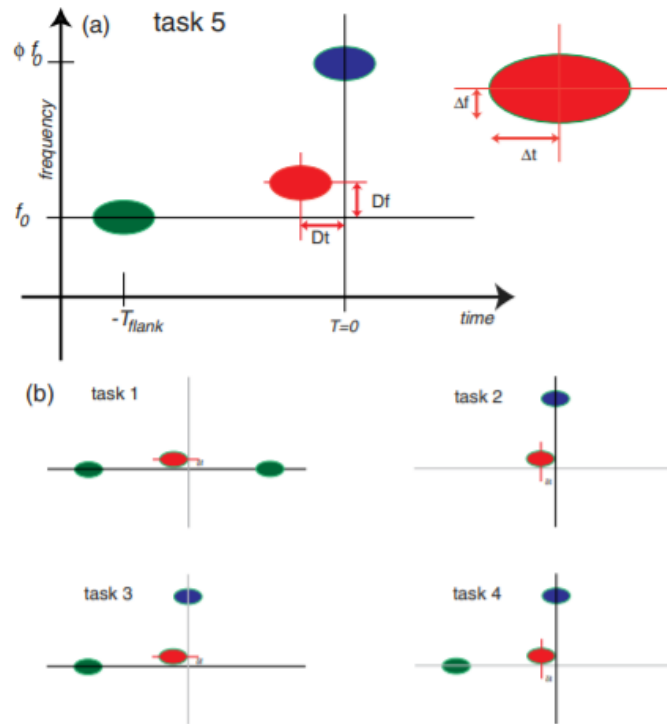


FIG. 1 (color online). Stimulus and task. (a) In our final task 5, subjects are asked to discriminate simultaneously whether the test note (red) is higher or lower in frequency than the leading note (green), and whether the test note appears before or after the flanking high note (blue). For each instance of the task, two numbers are generated (Dt and Df) and two Boolean responses (left-right, up-down) are recorded. (b) Tasks 1 through 4 lead to this final task: task 1 is frequency only (uses two flanking notes), task 2 timing only, task 3 is frequency only but with the flanking high note (blue) as a distractor, and task 4 is timing only, with the leading (green) note as a distractor

Rysunek 7: Badanie przeprowadzone przez autorów [4] wraz z opisem.

nieliniowa, ucha wewnętrznego i mózgu w czasie pomiędzy dwiema prezentacjami sygnału testowego. Trudno się jednak zgodzić, że czas trwania sygnału wynosi 2 s tak jak stwierdzają komentatorzy [76].

W swoim podstawowym sformułowaniu zasada nieokreśloności częstotliwość-czas orzeka, że iloczyn odchylenia standardowego Δt w dziedzinie czasu i odchylenia standardowego Δf w dziedzinie częstotliwości przy użyciu kwadratu modułu sygnału i odpowiednio widma mocy sygnału nie może być mniejszy niż $\frac{1}{4}\pi$

$$\Delta t \Delta f \geq \frac{1}{4}\pi \quad (14)$$

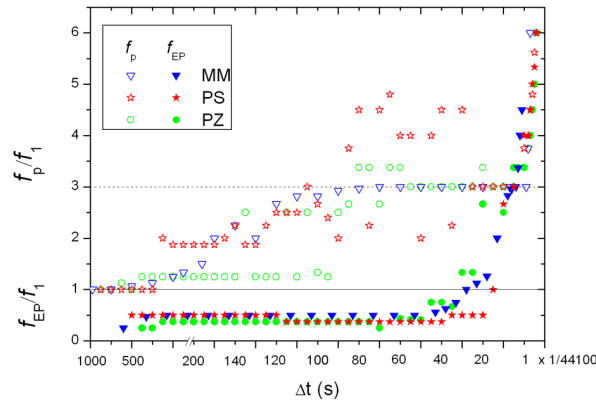
Jeżeli zamiast kwadratu modułu sygnału i widma mocy użyjemy samego sygnału i wartości bezwzględnej widma otrzymamy

$$\Delta t \Delta f \geq \frac{1}{2}\pi \quad (15)$$

W zasadzie nieokreśloności chodzi więc o to, że przy krótkim czasie trwania sygnału dokładność określenia częstotliwości spada. Wielkość Δt musi mieć związek z czasem trwania sygnału. Sygnały stosowane w [4] miały czas trwania 35 ms. W przypadku obwiedni gaussowskiej niepewność częstotliwości wg wzoru (14) jest wtedy niewiele większa niż 2 Hz, więc rozpoznawanie częstotliwości poprzez wysokość dźwięku powinna być całkiem dobra. Eksperyment Rakowskiego i Rogowskiego jest tu zapewne bardziej miarodajny [2]. Decydującą rolę w rozpoznawalności wysokości dźwięku zdaje się mieć liczba okresów zawartych w sygnale.

Zainspirowani publikacją [4] autorzy krakowscy M. Majka i in.[3] wykonali wstępne proste pomiary na samych sobie jako najłatwiej dostępnej grupie testowej. Słuchali mianowicie tonów prostych ze zmienną co do szerokości obwiednią gaussowską. Badali odbierane przy tym wrażenie wysokości dźwięku w porównaniu z tonami o tej samej częstotliwości lecz trwającymi 1 s. W tych ostatnich sygnałach wysokość dźwięku jest łatwo rozpoznawalna, są to więc sygnały odniesienia (referencyjne). Obserwacja jest bardzo ciekawa, ponieważ okazuje się, że podczas skracania sygnału z obwiednią gaussowską, słyszane częstotliwości pokrywają się z wynikami otrzymanymi dla impulsu sinusoidalnego dla sygnałów dostatecznie długich. Gdy sygnał staje się coraz krótszy postrzegana wysokość dźwięku zdaje

się wzrastać. Wrażenie wysokości dźwięku pozostaje nawet dla skrajnie krótkich sygnałów, w których funkcja cosinus jest praktycznie stała. Wprowadzono więc jak się zdaje pożyteczne pojęcie efektywnej wysokości dźwięku, które zostanie wykorzystane w dalszych rozdziałach tej rozprawy. Wyniki opisanego eksperymentu przedstawione są na rysunku 8.



Rysunek 8: Efektywna Wysokość Dźwięku. Figury geometryczne bez wypełnienia dotyczą dźwięków cosinusoidalnych, figury geometryczne z wypełnieniem - dźwięków czysto gaussowskich. MM, PS oraz PZ - to uczestnicy badania [78].

Metoda określania wysokości dźwięku oparta była na rozpoznawaniu interwału pomiędzy dźwiękiem odniesienia, a badanym krótkim sygnałem. Jedną z badanych osób zgłaszała posiadanie słuchu absolutnego, więc nie potrzebowała dźwięku odniesienia, jednak nie wpłynęło to znacząco na wynik testu. Interesującą obserwacją, jest zgłaszanie przez testowane osoby wysokości dźwięku odpowiadających wyższym harmonicznym badanych krótkich tonów widocznych na rys. 8 w zakresie czasów trwania $(100 - 40)/44100s = 2.27ms - 0.9ms$. Wskazywałoby to na mechanizm rozpoznawania wysokości bardzo krótkich tonów na podstawie nieliniowego zjawiska jakim jest generowanie wyższych harmonicznych (anharmizm). Ponieważ liczba okresów drgań wyższych harmonicznych w czasie trwania sygnału jest znacznie większa niż tonu podstawowego, ten mechanizm mógłby być kandydatem do wyjaśnienia przewyższania przez słuch ograniczeń narzucanych przez zasadę nieokreśloności.

7 Korelaty fizyczne wrażeń wywoływanych przez krótkie impulsy

7.1 Wysokość dźwięku bardzo krótkich tonów prostych z obwiednią gaussowską. Próba określenia fizycznych parametrów decydujących o tym wrażeniu w granicy małych czasów trwania.

W swoim przeglądowym artykule Alain de Cheveigné [45] przedstawia różne możliwe modele wrażenia wysokości dźwięku. Są to w istocie sposoby zdefiniowania tego wrażenia na podstawie parametrów fizycznych sygnału. Dla tonu prostego wszystkie modele (fizyczne, psychofizyczne) sprowadzają się do określenia wysokości dźwięku jako funkcji jedyne go parametru jakim jest częstość, lub równoważnie odwrotność okresu. Oczywiście czas trwania takiego sygnału musi zawierać dostateczną liczbę okresów, aby okresowość mogła być określona. Dla wrażenia tego rodzaju autorzy używają pojęcia wysokości dźwięku związanej z okresem (*periodicity pitch* $F0$). W przypadku sygnałów bardziej złożonych niektórzy autorzy wprowadzają pojęcie częstości miejsca (*locus frequency*) odpowiadającej maksimum widma ([45], p. 4 section 2.4), tj. składowej fourierowskiej, dla której amplituda przyjmuje maksimum. Mówi się wtedy o widmowej wysokości dźwięku (*spectral pitch*) lub wysokość dźwięku składowej - w domyśle fourierowskiej (*partial pitch*) i oznacza symbolem f_{LOCUS} nawiązując do wyrosłej z odkryć Helmholtza [79] s. 119, 129 i Bekesy'ego [80] s. 406, [81] teorii miejsca, przypisującej wrażenie wysokości dźwięku miejscu najsilniejszego pobudzenia błony podstawnej ślimaka [82, 83]. Kiedy na sygnał sinusoidalny jest nałożona obwiednia o zmiennym, malejącym czasie trwania, mamy do czynienia z przejściem od sygnału okresowego do sygnału, w którym okresu nie daje się określić, a jednocześnie powstaje widmo o charakterze ciągłego pasma. Ilościowe sformułowanie tego zjawiska jest dobrze znane w fizyce kwantowej i teorii sygnału jako zasada nieokreśloności, odpowiednio Heisenberga [84] i Gabora [6]. Parametrem częstościowym podlegającym pomiarowi (i być może percepcji zmysłowej) jest wartość oczekiwana częstości w widmie sygnału. Oczywiście w przypadku długiego (w zasadzie nieskończonego) sygnału ta

wartość oczekiwana jest identyczna z częstotliwością sinusoidy.

Zasada nieokreśloności orzeka, że dla zadanego sygnału iloczyn odchyżeń standardowych w dziedzinie czasu i częstotliwości jest zawsze większy niż pewna wartość graniczna, przy czym nierówność przechodzi w równość dla obwiedni gausowskiej. W szczególności dla sygnału

$$x(t; f, \sigma) = \exp\left(-\frac{t^2}{2 \cdot \sigma^2}\right) \cdot \cos(2 \cdot \pi \cdot f \cdot t) \quad (16)$$

Widmo tego sygnału, czyli jego transformata Fouriera ma postać

$$X(\omega; f, \sigma) = \sqrt{\pi} \sqrt{2} \sigma \cdot \cosh(2\pi f \omega \sigma^2) e^{-\frac{\sigma^2(4\pi^2 f^2 + \omega^2)}{2}}, \quad (17)$$

gdzie ω jest pulsacją, tj. zmienną w transformacie Fouriera a f częstotliwością badanego sygnału.

Wartość oczekiwana pulsacji otrzymana dla $\omega > 0$ wynosi

$$\langle \omega_{\cos} \rangle(f, \sigma) = \frac{(2\pi^2 f \cdot \operatorname{erf}(\pi f \sigma \sqrt{2}) \sigma \cdot e^{2\pi^2 f^2 \sigma^2} + \sqrt{2} \sqrt{\pi}) e^{-2\pi^2 f^2 \sigma^2}}{\pi \sigma}. \quad (18)$$

Warto zauważyć, że z powodu parzystości funkcji cosinus wartość oczekiwana pulsacji obliczona w zakresie $-\infty < \omega < \infty$ byłaby zawsze równa zero.

Wariancję, tj. kwadrat niepewności pulsacji także można otrzymać w postaci analitycznej

$$\begin{aligned} \operatorname{Var}(f, \sigma) &= \langle (\omega - \langle \omega_{\cos} \rangle)^2 \rangle = \\ &= \frac{1}{\pi \cdot \sigma^2} (-4 \cdot \operatorname{erf}(\pi \cos f \cdot \sigma \sqrt{2})^2 \pi^3 f^2 \sigma^2 - 4 e^{-2\pi^2 f^2 \sigma^2} \cdot \\ &\cdot \operatorname{erf}(\pi \cos f \cdot \sigma \cdot \sqrt{2}) \pi^{\frac{3}{2}} \sqrt{2} \cdot f \cdot \sigma + 4 \cdot \pi^3 f^2 \sigma^2 - 2 \cdot e^{-4\pi^2 f^2 \sigma^2} + \pi) \end{aligned} \quad (19)$$

Pierwiastek z tej wielkości jest niepewnością (nieokreślonością) pulsacji

$$\Delta\omega(f, \sigma) = \sqrt{\operatorname{Var}(f, \sigma)}. \quad (20)$$

W granicy szerokich, tj. bardzo długich impulsów można obliczyć granicę iloczynu

$$\sigma \Delta \omega(f, \sigma) = \Delta t(f, \sigma) \Delta \omega(f, \sigma) \xrightarrow{\sigma \rightarrow \infty} 1 \quad (21)$$

albo używając częstości, tj. liczby cykli na jednostkę czasu

$$\sigma \Delta f(f, \sigma) = \Delta t(f, \sigma) \Delta f(f, \sigma) \xrightarrow{\sigma \rightarrow \infty} \frac{1}{2\pi}. \quad (22)$$

Powyższe wzory zostały otrzymane przy użyciu widma (17), które okazuje się funkcją nieujemną, a zatem nadaje się na rozkład prawdopodobieństwa.

W mechanice kwantowej oraz w wielu zastosowaniach analizy sygnałów i telekomunikacji przydatne jest widmo mocy, tj. kwadrat wartości bezwzględnej widma ze wzoru (17).

Odpowiednie wyrażenia przyjmują wtedy postać:

widmo mocy:

$$X_{pow}(\omega; f, \sigma) = 2\sigma^2 \pi \cosh(2\pi \cdot f \cdot \omega \cdot \sigma^2)^2 \cdot e^{-4\sigma^2(\pi^2 f^2 + \frac{\omega^2}{4})}, \quad (23)$$

wartość oczekiwana pulsacji:

$$\omega_{psr}(f, \sigma) = \frac{2 \cdot \operatorname{erf}(2\pi f \sigma) e^{4\pi^2 f^2 \sigma^2 \pi^{\frac{3}{2}}} f \sigma + 2}{\sqrt{\pi} \sigma (e^{4\pi^2 f^2 \sigma^2} + 1)}, \quad (24)$$

wariancja pulsacji względem wartości oczekiwanej:

$$A' = ((\alpha + 8\pi^{\frac{7}{2}} \sigma^2 f^2 + \pi^{\frac{3}{2}}) e^{4\pi^2 f^2 \sigma^2} + (-8\sqrt{\pi} + \pi^{\frac{3}{2}}))$$

$$Var_{pow}(f, \sigma) = \frac{A' \cdot e^{-4\pi^2 f^2 \sigma^2} + 8\pi^{\frac{7}{2}} \sigma^2 f^2 - 16\pi^2 \operatorname{erf}(2\pi f \sigma) + 2\pi^{\frac{3}{2}}}{(2 \cdot \sigma (e^{4\pi^2 f^2 \sigma^2} + 1) \pi^{\frac{3}{2}} \sigma (e^{-4\pi^2 f^2 \sigma^2} + 1))}. \quad (25)$$

Gdybyśmy zamiast oryginalnego sygnału użyli kwadratu jego modułu, tak jak to się stosuje w mechanice kwantowej, zamiast szerokości σ mielibyśmy mniejszą szerokość σ_{pow}

$$\sigma_{pow} = \frac{\sigma}{\sqrt{2}}. \quad (26)$$

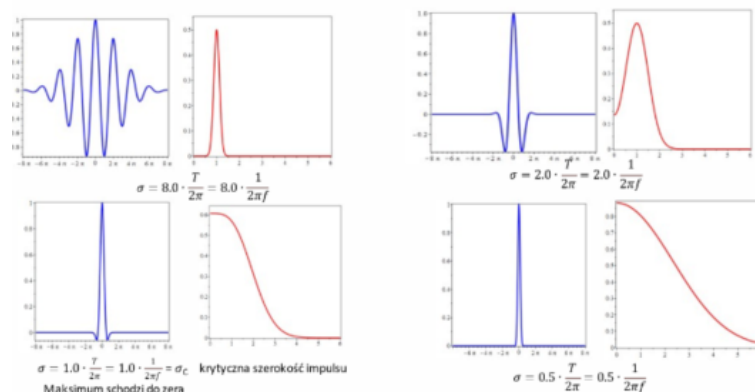
W granicy długoczasowych impulsów granica iloczynu niepewności wynosi

$$\sigma_{pow} \Delta_{pow} \omega(f, \sigma) = \Delta_{pow} t(f, \sigma) \Delta_{pow} \omega(f, \sigma) \xrightarrow{\sigma \rightarrow \infty} \frac{1}{2} \quad (27)$$

albo używając częstości, tj. liczby cykli na jednostkę czasu

$$\sigma_{pow} \Delta_{pow} f(f, \sigma) = \Delta_{pow} t(f, \sigma) \Delta_{pow} f(f, \sigma) \xrightarrow{\sigma \rightarrow \infty} \frac{1}{4\pi}. \quad (28)$$

Oczywiście wzory (21), (22), (27) i (28) odpowiadają znanym z podręczników relacjom nieokreśloności [85]. Rys. 9 ilustruje sygnały i widma (nie widma mocy) dla wybranych wartości σ .



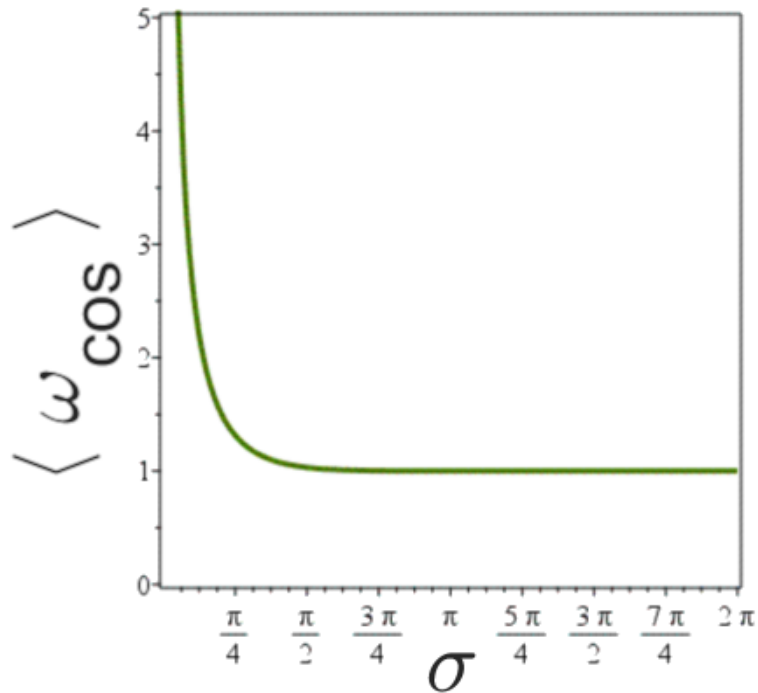
Rysunek 9: Cosinusoida z obwiednią gaussowską o malejącym parametrze szerokości σ . W dziedzinie czasu i częstości. Szerokość krytyczną σ_c zdefiniowano jako wartość przy której maksimum widma dochodzi do zera.

Należy zwrócić uwagę, że gdy parametr szerokości sygnału spada do wartości

$$\sigma_c(f) = \frac{T}{2\pi} = \frac{1}{2\pi f} \quad (29)$$

widmo nie ma już maksimum dla $f > 0$. Maksimum widma jest wtedy umiejscowione w $f = 0$. Nie ma więc w tym przypadku jasnego kandydata na częstość odpowiadającą wrażeniu wysokości dźwięku. W dalszym ciągu będziemy badać zachowanie wartości oczekiwanej, jako wielkości najbardziej znanej z mechaniki kwantowej i występującej w relacjach nieokreśloności. Rys. 10 ilustruje

zachowanie wartości oczekiwanej częstotliwości, gdy czas trwania sygnału maleje poniżej swej wartości krytycznej.



Rysunek 10: Wartość oczekiwana pulsacji w widmie sygnału cosinusoidalnego z obwiednią gaussowską o czasie trwania (szerokości) σ .

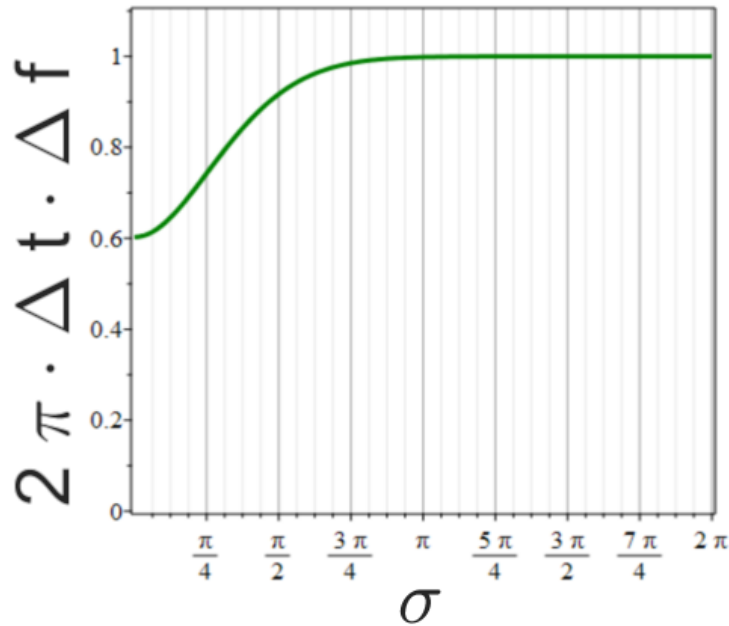
Warto zauważyć, że w granicy bardzo małych czasów trwania impulsu $\sigma \rightarrow 0$ wartość oczekiwana pulsacji zachowuje się jak $\frac{1}{\sigma}$:

$$\langle \omega_{\text{cos}} \rangle(f, \sigma) = \frac{\sqrt{2}}{\sqrt{\pi}} \sigma^{-1} + 2\pi^{\frac{3}{2}} f^2 \sqrt{2} \sigma + O(\sigma^3). \quad (30)$$

Iloczyn odchyłeń standardowych (niepewności) w dziedzinie czasu i częstotliwości

$$\Delta\omega \cdot \Delta t = 2\pi \cdot \Delta t \cdot \Delta f = 2\pi \cdot [s \cdot \sigma_c(f)] \cdot \Delta f \quad (31)$$

jest przedstawiony na Rys. 11



Rysunek 11: Iloczyn niepewności $\Delta\omega \cdot \Delta t = 2\pi \cdot \Delta f \cdot \Delta t$ jako funkcja czasu trwania impulsu σ wg wzoru (19 - 21).

Iloczyn ten zgodnie z zasadą nieokreśloności powinien przyjmować wartość 1, gdyż przy obwiedni gaussowskiej nierówność przewidywana przez relację nieokreśloności przechodzi w równość. Z Rys. 11 widać, że ta wartość jest osiągnięta, z dobrą dokładnością dla długości impulsu większego niż $\pi\sigma_c(f)$. W granicy bardzo krótkich impulsów $\sigma \rightarrow 0$, iloczyn ten dąży do wartości

$$\Delta\omega \cdot \Delta t = \frac{\sqrt{\pi - 2}}{\sqrt{\pi}}, \quad (32)$$

w szczególności:

$$\Delta\omega \cdot \Delta t = 0.6028... \text{ gdy } \sigma \rightarrow 0,$$

$$\Delta\omega \cdot \Delta t > 0.9986... \text{ gdy } \sigma > \pi,$$

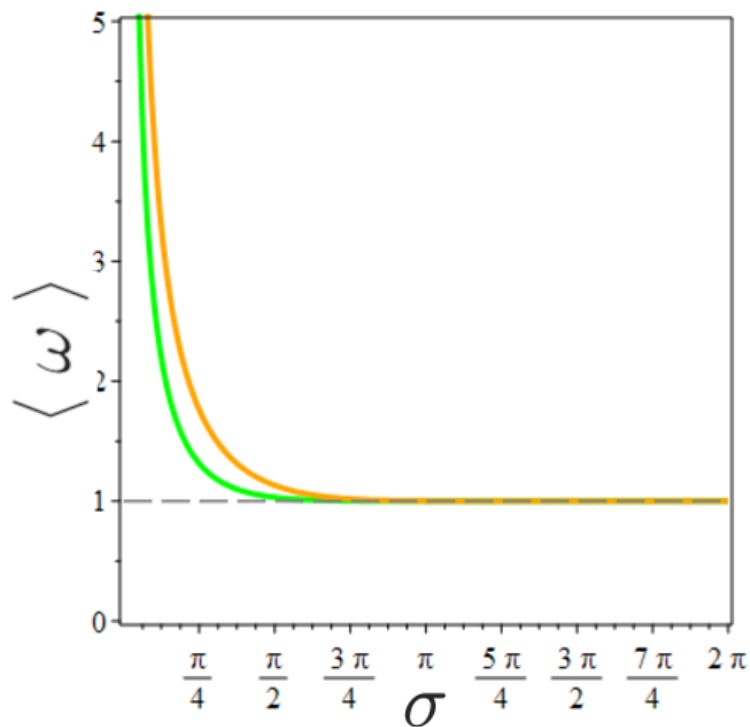
$$\Delta\omega \cdot \Delta t = 1 \text{ gdy } \sigma \rightarrow \infty$$

Wynika stąd, że zasada nieokreśloności w znanej z podręczników postaci jest spełniona tylko w tym zakresie impulsów z obwiednią gaussowską, gdzie wartość oczekiwana częstości jest bliska maksimum widma. Dla sygnałów krótszych

widzimy wartości mniejsze niż znana dolna granica.

W swoim pionierskim artykule nt. relacji nieokreśloności [84] Heisenberg zauważył, że relacje nieokreśloności nie zależą od fazy sygnału. Dla zastosowań w dziedzinie percepcji dźwięku informacja ta jest bardzo istotna ze względu na obserwację Ohma i Helmholtza, że narządy słuchu są nieczułe („głuche”) na fazę [86, 62, 87].

Aby sprawdzić zależność wartości oczekiwanych i iloczynów nieokreśloności od fazy na rys. 12 i 13 przedstawione jest porównanie wyników dla sygnału cosinusoidalnego z rys. 10 i 11 z analogicznymi wynikami dla sygnału sinusoidalnego, tj. przesuniętego w fazie o $\frac{\pi}{2}$. Dla porównania na rys. 12 widzimy wartość oczekiwaną pulsacji podzieloną przez pulsację pierwotnego sygnału $2\pi \cdot f$ poddawanego obwodni gaussowskiej, w zależności od szerokości σ tej obwodni dla sygnałów cosinusoidalnego i sinusoidalnego. Wartość oczekiwana w przypadku sygnału sinu-



Rysunek 12: Wartość oczekiwana dla sygnałów cosinusoidalnego (linia zielona) i sinusoidalnego (linia pomarańczowa) uzyskana przy użyciu widma mocy.

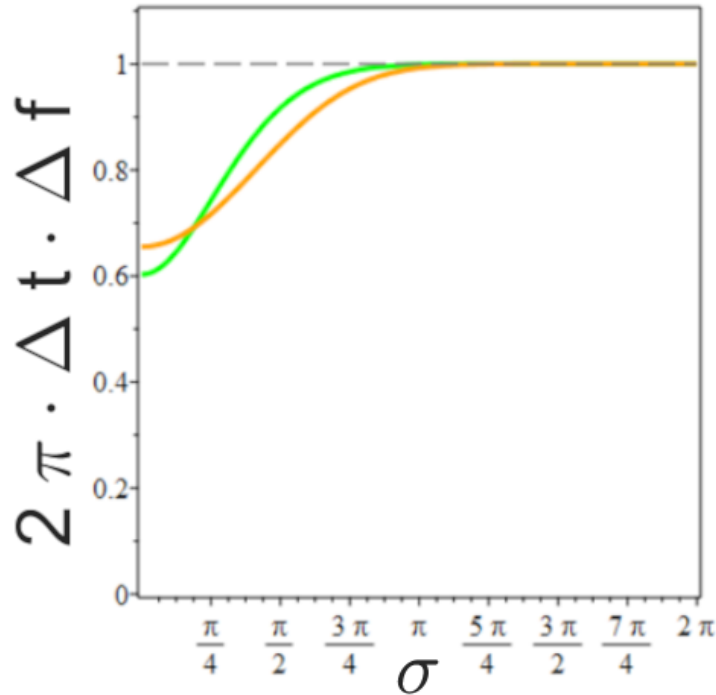
soidalnego wyraża się wzorem

$$\langle \omega_{sin} \rangle(f, \sigma) = \frac{2\pi f}{\operatorname{erf}(\pi f \sigma \sqrt{2})}, \quad (33)$$

a jej rozwinięcie dla najmniejszych czasów trwania także wykazuje zachowanie $\frac{1}{\sigma}$. Dla małych wartości σ mamy:

$$\langle \omega_{sin} \rangle(f, \sigma) = \sqrt{\frac{\pi}{2}} \sigma^{-1} + \frac{1}{3} \pi^{\frac{5}{2}} f^2 \sqrt{2} \sigma + O(\sigma^3). \quad (34)$$

Rys. 13 przedstawia porównanie iloczynów nieokreśloności dla sygnałów cos i sin. Widać zauważalną różnicę dla wartości σ mniejszych niż σ_c oraz wartość graniczną

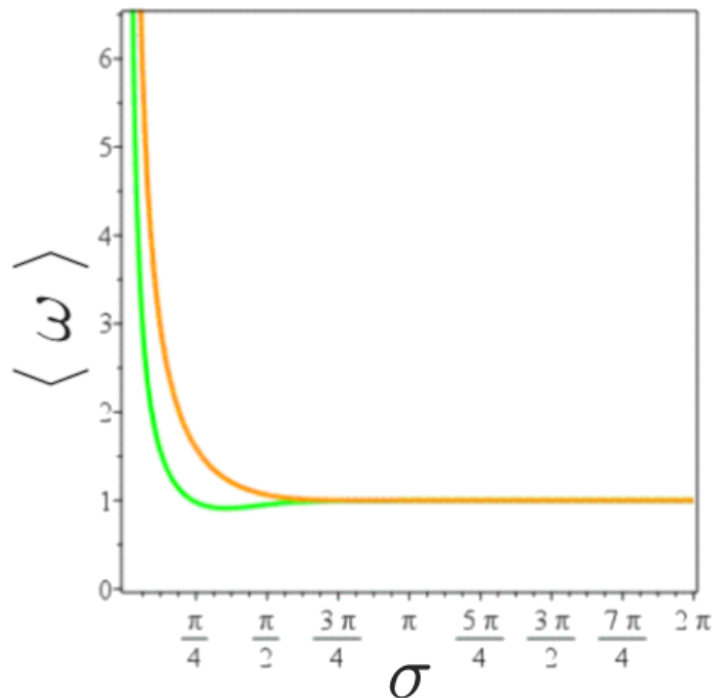


Rysunek 13: iloczyn niepewności czasu i pulsacji w funkcji σ dla sygnałów cosinoidalnych (linia zielona) i sinusoidalnych (linia pomarańczowa).

dla $\sigma \rightarrow 0$ dla sygnału sinusoidalnego.

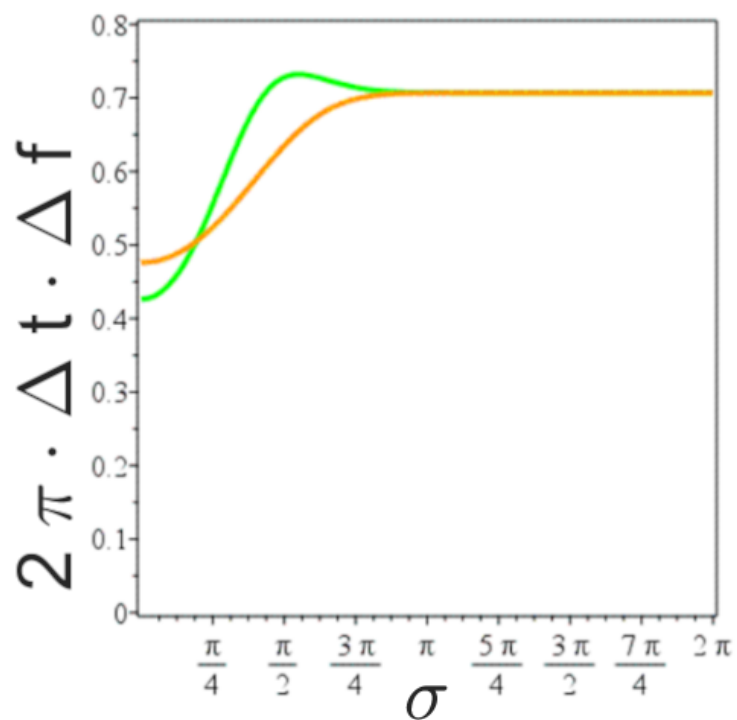
$$\frac{\sqrt{-2\pi + 8}}{2} \approx 0,65513637... \quad (35)$$

Oznacza to, że niezależność od fazy nie występuje, gdy czas trwania impulsu jest znacznie mniejszy niż $\pi \cdot \sigma_c(f)$. Powyższe wyniki zostały uzyskane przy wykorzystaniu rzeczywistego widma funkcji cosinus oraz wartości bezwzględnej czysto urojonego widma funkcji sinus. W różnych zastosowaniach zarówno w teorii sygnałów, jak i w mechanice kwantowej używa się jednak widma mocy, tj. kwadratu modułu. Dlatego na rys. 14 i 15 zaprezentowane są wielkości analogiczne do przedstawionych na rys. 12 i 13 lecz uzyskane z użyciem widm mocy.



Rysunek 14: Wartość oczekiwana pulsacji z widm sygnału cosinusoidalnego (linia zielona) i sinusoidalnego (linia pomarańczowa) z obwiednią gaussowską o szerokości σ otrzymana z użyciem widma mocy

Poniżej zachowanie graniczne wartości oczekiwanej dla $\sigma \rightarrow 0$. Okazuje się, że dla bardzo krótkich impulsów, czyli czysto gaussowskich powtarza się zachowanie $\frac{1}{\sigma}$, choć z różnymi współczynnikami.



Rysunek 15: Iloczyny nieokreśloności czasu i pulsacji gaussowskich sygnałów cosinusoidalnych i sinusoidalnych otrzymane z użyciem widm mocy.

$$\langle \omega_{cos} \rangle(f, \sigma) = \frac{2f \cdot \operatorname{erf}(2f\pi\sigma) \cdot e^{4\pi^2 f^2 \sigma^2 \pi^{\frac{2}{3}}} \cdot \sigma + 2}{\sqrt{\pi}\sigma(e^{4\pi^2 f^2 \sigma^2} + 1)} = \frac{1}{\sqrt{\pi}}\sigma^{-1} + 2\pi^{\frac{3}{2}}f^2/\sigma + O(\sigma^3) \quad (36)$$

$$\langle \omega_{sin} \rangle(f, \sigma)(f, \sigma) = \frac{2f\pi \cdot \operatorname{erf}(2f\pi\sigma)}{(e^{4\pi^2 f^2 \sigma^2} - 1)e^{-4\pi^2 f^2 \sigma^2}} \approx \frac{2}{\sqrt{\pi}}\sigma^{-1} + O(\sigma) \quad (37)$$

Zwraca uwagę niemonotoniczny przebieg wartości oczekiwanej i minimum dla impulsu cosinusoidalnego. Na rys. 15 widzimy porównanie iloczynów nieokreśloności uzyskanych przy użyciu widm mocy dla sygnałów cosinusoidalnego i sinusoidalnego. Odpowiednie wartości graniczne przy długości sygnałów zmierzających do zera $\sigma \rightarrow 0$ wynoszą odpowiednio:

$$\frac{\sqrt{2}\sqrt{1 - \frac{2}{\pi}}}{2} \approx 0,4262512332 \quad (38)$$

dla sygnału cosinusoidalnego i

$$\frac{\sqrt{3\pi^4 - 8^3}\sqrt{2}}{2\pi^2} \approx 0,4761937161 \quad (39)$$

dla sygnału sinusoidalnego.

Wnioski

Dla skrajnie krótkich gaussowskich sygnałów akustycznych wartość oczekiwana częstości zachowuje się odwrotnie proporcjonalnie do czasu trwania sygnału $\frac{1}{\sigma}$ zarówno dla sygnałów cosinusoidalnych (w granicy $\sigma \rightarrow 0$ czysto gaussowskich), jak sinusoidalnych. Iloczyny odchyłeń standardowych częstości i czasu dla czasów trwania sygnału σ mniejszych od $\pi \cdot \sigma_c(f)$ przyjmują wartości nieco mniejsze niż stałe znane z wiedzy podręcznikowej. W tym samym zakresie czasów trwania sygnału $\sigma < \pi \cdot \sigma_c(f)$ zarówno wartość oczekiwana częstości, jak i iloczyn odchyłeń standardowych wykazują zależność od fazy sygnału.

7.2 Fizyczne korelaty wysokości i barwy dźwięku dla sygnałów czysto gaussowskich.

W tym podrozdziale zajmiemy się postrzeganiem ogólnie znanych cech dźwięku, tu wysokości i barwy, na podstawie obiektywnych wielkości charakteryzujących sygnały czysto gaussowskie, tzn. w granicy szerokości obwiedni dążącej do zera $\sigma \rightarrow 0$ gdy sygnałem jest funkcja cosinus.

7.2.1 Wrażenie wysokości dźwięku.

W naszych badaniach używaliśmy sygnałów gaussowskich znormalizowanych do amplitudy w maksimum (eq. 40)

$$x(t; \sigma) = \exp\left(-\frac{t^2}{2\sigma^2}\right). \quad (40)$$

Zauważamy, że energia tj. pole powierzchni pod krzywą ilustrującą impuls zależy tu od parametru szerokości σ tj. od czasu trwania sygnału.

Zgodnie z akustycznym prawem Ohma oraz wnioskami z badań Helmholtza i jego następców zakładamy, że wrażenie dźwiękowe jest związane z widmem, tj. współczynnikami rozkładu sygnału na składowe sinusoidalne, czyli z transformatą Fouriera. W przypadku sygnału (40) transformata Fouriera jest szczególnie prosta

$$X(\omega, \sigma) = -e^{-\frac{\omega^2 \sigma^2}{\sqrt{2}\sqrt{\pi} \cdot \sigma}}. \quad (41)$$

Wysokość dźwięku przypisywana pojedynczej sinusoidzie, tj. tonowi prostemu jest proporcjonalna do częstości ω [20]. W przypadku wielotonu harmonicznego, wysokość dźwięku określa częstość podstawowa. Przyjmując, jak w poprzednim rozdziale, parametr częstościowy w postaci wartości oczekiwanej otrzymujemy dla niezniekształconego sygnału gaussowskiego

$$\omega_{sr} = \langle \omega \rangle = \sqrt{\frac{2}{\pi}} \cdot \frac{1}{\sigma}, \quad (42)$$

lub dla widma mocy

$$\omega_{srp} = \langle \omega \rangle_p = \frac{1}{\sqrt{\pi}} \cdot \frac{1}{\sigma}. \quad (43)$$

Widać, że mamy tu dokładną odwrotną proporcjonalność.

Założmy teraz, że postrzegana wysokość dźwięku p zależy od częstości średniej w sposób logarytmiczny, tzn.

$$p = A(\ln(\sqrt{\frac{2}{\pi}}) - \ln\sigma) \quad (44)$$

oraz, że do wykrycia różnicy tak zdefiniowanego wrażenia p potrzeba przekroczyć ten sam próg Δp_{lim} niezależnie od natężenia bodźca σ . Przyjmując, że próg Δp_{lim} jest mały w porównaniu z całkowitym wrażeniem p i odpowiada także małej zmianie bodźca możemy rozwinąć wyrażenie (44) w szereg Taylora

$$\Delta p_{lim} = -A \frac{d\sigma}{\sigma}. \quad (45)$$

Oczywiście odpowiadająca temu progowi zmiana bodźca jest jego najmniejszą postrzegalną różnicą (Discrimination Limen/Difference Limen/DL) $d\sigma_{DL}$. Założenie stałości progu Δp_{lim} prowadzi wprost do prawa Webera-Fechnera:

$$d\sigma_{DL} = -\frac{\Delta p_{lim}}{A} \sigma \quad (46)$$

wyrażającego proporcjonalność najmniejszej postrzegalnej różnicy bodźca do intensywności samego bodźca. Znak „-” pochodzi stąd, że postrzegana wysokość dźwięku maleje z szerokością, czyli z czasem trwania impulsu. Jest to w pełni analogiczne do zwykłego wrażenia wysokości dźwięku dla sygnałów okresowych, gdzie:

$$p(\omega) = 1200 \cdot \log_2 \omega + B = A \ln(\omega) + B = \frac{1200}{\ln 2} \cdot \ln(\omega) + B = \frac{-1200}{\ln 2} \ln(\sigma) + B', \quad (47)$$

gdzie A i B to stałe. Czynniki 1200 pochodzi z definicji jednostki cent, będącej jedną setną półtonu w stroju równomiernie temperowanym (zobacz rozdz. 4.1). Przy tych samych założeniach stałości progu wrażenia Δp_{lim} mamy tu:

$$d\omega_{DL} = \Delta p_{lim} \frac{\omega}{A} \quad (48)$$

tzn. najmniejsza zauważalna zmiana częstotliwości jest proporcjonalna do samej

częstotliwości.

Warto zauważyć, że każda potęgowa zależność wielkości w , której logarytm opisuje wrażenie W , od parametru bodźca b będzie prowadzić do prawa Webera-Fechnera. Rzeczywiście, jeżeli

$$w = b^\alpha \quad (49)$$

i

$$W = \ln w = \alpha \ln b, \quad (50)$$

to

$$db_{DL} = \alpha^{-1} \Delta W_{lim} b. \quad (51)$$

Różne wartości wykładnika α zostały zanotowane we wstępnych badaniach zależności efektywnej wysokości dźwięku od czasu trwania impulsu w [3]. **Wniosek** Gdy widmo impulsu gaussowskiego nie podlega żadnemu przekształceniu (filtracji), wówczas jeżeli za częstość odpowiadającą za wrażenie wysokości dźwięku uznamy średnią częstość widma (wartość oczekiwaną), średnia ta będzie odwrotnie proporcjonalna do czasu trwania impulsu ($\alpha = -1$), a prawo Webera-Fechnera będzie spełnione dokładnie w swej pierwotnej postaci. W ogólnym przypadku należy założyć, że częstość odpowiadająca za wysokość dźwięku jest pewną funkcją czasu trwania impulsu

$$\omega = \omega(\sigma). \quad (52)$$

Wówczas próg postrzegalności zmiany wysokości dźwięku dp_{lim} będzie

$$dp_{lim} = A \frac{d\omega}{\omega} = \frac{A}{\omega} \frac{d\omega}{d\sigma} d\sigma_{DL}. \quad (53)$$

Zatem

$$\Delta\sigma_{DL} = \frac{\omega}{A} \left(\frac{d\omega}{d\sigma} \right)^{-1} dp_{lim}. \quad (54)$$

W powyższym rozumowaniu przyjęto, że dla zauważenia różnicy potrzebna jest stała, niezależna od natężenia bodźca zmiana wielkości wrażenia, tj. próg Δp_{lim} . W przypadku wysokości dźwięku przypisanej widmu ciągłemu można użyć bardziej przekonującego kryterium. Mianowicie, widmo, z którego obliczamy średnią częstość pozwala też najczęściej wyznaczyć średnie kwadratowe odchylenie stan-

standardowe $\delta(\sigma)$

$$Var(\omega) = \frac{\int_0^\infty (\omega - \omega_{sr}(\sigma))^2 X(\omega; \sigma) d\omega}{\int_0^\infty X(\omega; \sigma) d\omega}, \quad (55)$$

$$\delta(\sigma) = \sqrt{Var(\omega)}. \quad (56)$$

Logiczną definicją najmniejszej zauważalnej różnicy $d\sigma_{DL}$ jest taka różnica parametru σ , która wartość średnią $\omega_{sr}(\sigma)$ zmieni o pewien ułamek K odchylenia standardowego $\delta(\sigma)$, tzn.

$$|\omega_{sr}(\sigma + d\sigma_{DL}) - \omega_{sr}(\sigma)| = K\delta(\sigma) \quad (57)$$

W przypadku niefiltrowanego sygnału gaussowskiego (40), (41) otrzymujemy:

$$\delta(\sigma) = \frac{\sqrt{\frac{\pi-2}{\pi}}}{\sigma}. \quad (58)$$

Wstawienie tego wyniku do równania (57) prowadzi do relacji:

$$d\sigma_{DL} = \sigma \frac{K\sqrt{\pi-2}}{\sqrt{2} - K\sqrt{\pi-2}}. \quad (59)$$

Oznacza to, że także przy takiej definicji prawo Webera-Fechnera dla rozróżniania krótkich sygnałów dźwiękowych jest dokładnie spełnione przy założeniu doskonałej percepcji widma w całym zakresie częstotliwości.

7.2.2 Wrażenie barwy dźwięku.

Jak wiadomo [88] barwa dźwięku nie daje się opisać jednym parametrem, lecz wynika z właściwości całego widma. Najlepszym przykładem są różnice brzmienia instrumentów wykonujących tę samą wysokość dźwięku, lub samogłoski śpiewane na tej samej wysokości. Aby nadać wrażeniu barwy wywoływanemu krótkimi sygnałami jakąś całościową charakterystykę uwzględniamy całe widmo. Nie mając możliwości ilościowego zdefiniowania barwy dźwięku możemy określić

jej zmianę pod wpływem zmiany parametru σ . Proponujemy wielkość:

$$\frac{\partial T}{\partial \sigma} = \int_0^{\infty} \left| \frac{\partial X(\omega; \sigma)}{\partial \sigma} \right| d\omega \equiv F_r(\sigma). \quad (60)$$

Obejmuje ona całe widmo. Wartość bezwzględna pozwala zsumować wszystkie zmiany widma i uniknąć sytuacji, gdy zmiana parametru σ prowadzi do wzrostu widma w pewnych zakresach częstości i jego zmniejszenia w innym zakresie tak, że widma nawet bardzo różniące się od siebie mogły by razem dać zerową zmianę barwy. Gdyby najpierw scałkować, a potem różniczkować, to dla widma mocy mielibyśmy pochodną całkowitej energii związanej z głośnością. Tu mamy zsumowane różnice w całym widmie. Z wykresów dla pierwotnych sygnałów gaussowskich wynika, że zerowanie się różnicy ma decydujące znaczenie. Zauważmy, że definiujemy tu szybkość zmian barwy dźwięku $F_r(\sigma)$ (Farbe rate), mimo, że nie definiujemy samej barwy dźwięku. Interesujące jest, że dla impulsów gaussowskich szybkość zmian barwy dźwięku $F_r(\sigma)$ można obliczyć analitycznie (zobacz wzór (63)).

Mając daną funkcję $F_r(\sigma)$ możemy stwierdzić, że sumaryczna zmiana barwy na progu wyczuwalności zmian

$$\Delta T_{lim}(\sigma) = F_r(\sigma) d\sigma_{DL} \quad (61)$$

w analogii do równań (45) i (54) dla wysokości dźwięku. Najmniejsza zauważalna zmiana czasu trwania sygnału $d\sigma_{DL}$ wyrazi się zatem jako

$$d\sigma_{DL} = \frac{1}{F_r(\sigma)} \Delta T_{lim}(\sigma), \quad (62)$$

gdzie znów zakładamy stały, niezależny od σ próg zmiany wrażenia $\Delta T_{lim}(\sigma)$. Prawo Webera będzie spełnione, gdy $F_r(\sigma) = \frac{1}{\sigma}$. W przypadku nieprzetworzonego sygnału gaussowskiego (40) o widmie (41) funkcja $F_r(\sigma)$ ma postać

$$F_r(\sigma) = \int_0^{\infty} \sqrt{2\pi} \left| 1 - \sigma^2 \omega^2 \exp\left(\frac{-\sigma^2 \omega^2}{2}\right) \right| d\omega = \frac{2\sqrt{\frac{2\pi}{e}}}{\sigma}. \quad (63)$$

Wniosek

Przyjmując definicję szybkości zmian barwy dźwięku z długością sygnału

wg wzoru (60) stwierdzamy, że prawo Webera-Fechnera dla najmniejszej zauważalnej różnicy barwy dźwięku jest spełnione w przypadku impulsów gaussowskich o nieprzetworzonym widmie. W kolejnych rozdziałach porównamy otrzymane tu analizy sygnałów pierwotnych z wynikami uzyskanymi za pomocą urządzeń sztucznej głowy. Z badaniami behawioralnymi dotyczącymi percepcji dźwięku wiąże się kilka problemów. Słuch ludzki jest bardzo skomplikowanym narządem i dokonanie wiarygodnego pomiaru jest praktycznie niemożliwe ze względu na to, że na percepcję dźwięku wpływa wiele czynników. Dla przykładu na poniższym rysunku (16) widzimy poglądowo jak rozchodzi się fala dźwiękowa kiedy sygnał rejestrowany jest przez mikrofon i przez człowieka. Autorzy [89] dokładnie opisują na czym polega problem dotyczący pomiarów różnego typu słuchawek.

Na zakończenie warto zauważyć, że gdyby we wzorze (60) użyto widma mocy, zamiast samego widma otrzymalibyśmy wartość stałą

$$F_r(\sigma) = [2 \cdot \operatorname{erf}(1) - 1] \pi^{\frac{3}{2}} + \frac{4\pi}{e}. \quad (64)$$

Zamiast prawa Webera-Fechnera otrzymalibyśmy niezależność najmniejszej zauważalnej różnicy od czasu trwania impulsu. Ta obserwacja może być przydatna w analizie wyników badań behawioralnych.

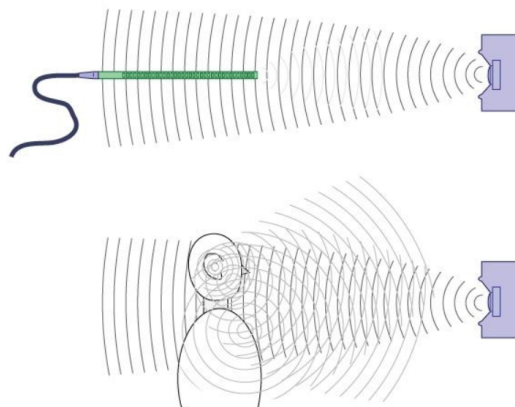
8 Badania behawioralne

8.1 Wprowadzenie

Aby przekonać się w jaki sposób ludzie odbierają wrażenia wywołane przez bodźce zewnętrzne przeprowadza się różnego rodzaju testy. Jest to najlepsza metoda zebrania informacji o reakcjach ludzi z uwzględnieniem ich własnych odczuć. Są to tzw. badania behawioralne oparte na relacji bodziec-reakcja. Znanych jest wiele form behawioryzmu, które wymieniono w tabeli 2 [90].

Behawioryzm	Behaviorists	Year	Field of practice
Radical	Skinner, Moore, Catania, Palmer	1945	Experimental Analysis of Behavior, Applied Behavior Analysis
Emergent	Killeen	1984	Learning theory, Cognitive Psychology, Neuropsychology
Theoretical	Staddon	2001, 2017	Learning theory, Cognitive Psychology, Neuropsychology
Contextualistic	Hayes	1992	Psychotherapy
Teleological	Rachlin	1994	Experimental Analysis of Behavior
Multiscale	Hineline, Shim	2001	Experimental Analysis of Behavior
Molar Multiscale	Baum	2002	Experimental Analysis of Behavior
Intentional	Foxall	2007	Behavioural economics/ Consumer psychology

Tablica 2: Lista behawioryzmów i ich charakterystyka [90].



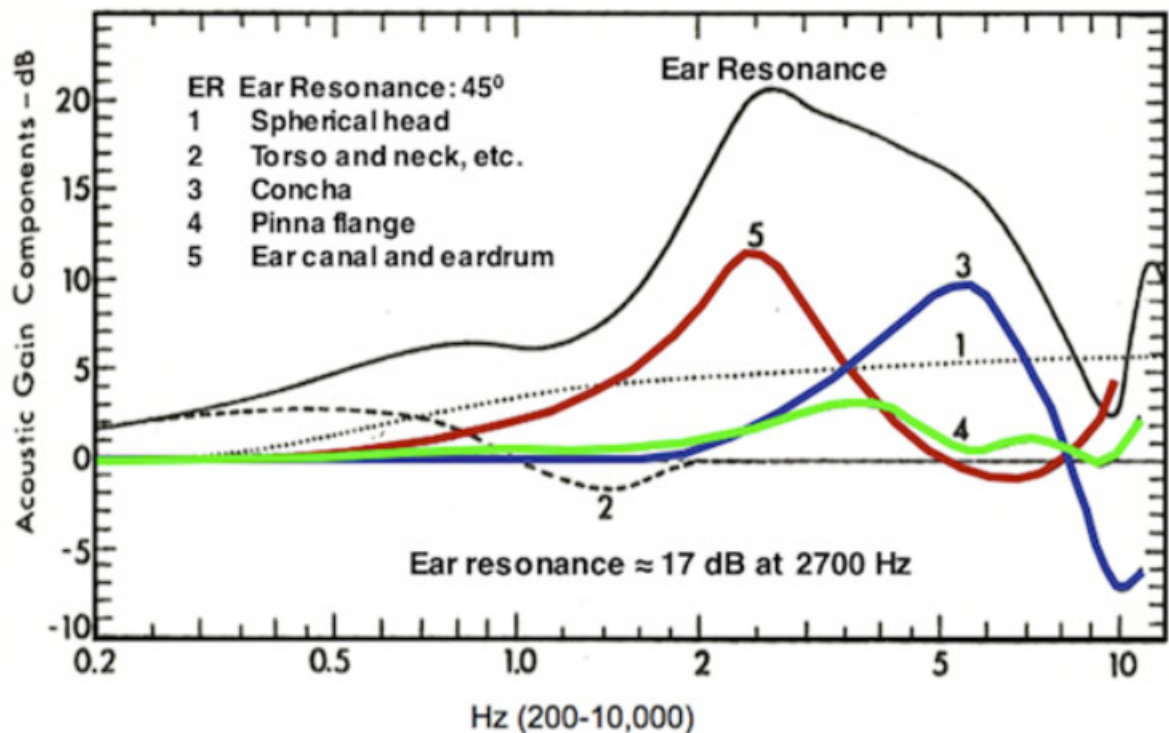
Rysunek 16: Poglądowy rysunek przedstawiający pomiar dźwięku rejestrowanego przez mikrofon oraz odbiór tego samego dźwięku przez człowieka. Geometria głowy, rozstaw uszu oraz kształt tułowia powodują wiele odbić i opóźnień [89].

8.2 Adaptacyjny test na wykrywanie najmniejszej zauważalnej różnicy czasu trwania sygnału czysto gaussowskiego.

Przeprowadzone tu badanie jest w pewnym sensie kontynuacją prac nad percepcją zjawisk czasowych znanych z literatury [91], lecz dotyczących bodźców trwających ponad 100 ms. W tym zakresie czasów trwania można raczej mówić o rytmie, natomiast zjawiska takie jak efektywna wysokość dźwięku lub efektywna barwa dźwięku nie występują. W badaniu opisanym w tej rozprawie zjawisko czasu trwania nie jest odbierane bezpośrednio, lecz poprzez zaproponowane w rozdz. 7 korelaty w postaci efektywnej wysokości i/lub barwy dźwięku. Zatem osoby testowane oceniały właśnie te cechy dźwięku, metodą nazwaną w artykule [91] porównawczą, w której

„Basically, a participant has to judge the relative duration of intervals presented successively and to indicate, by pressing the appropriate button, whether the second interval was shorter or longer than the first.”.

Nawet i w tej metodzie można by rozróżnić warianty w zależności od sposobu podawania par sygnałów podlegających rozpoznaniu. W obecnej rozprawie skorzystano z obserwacji



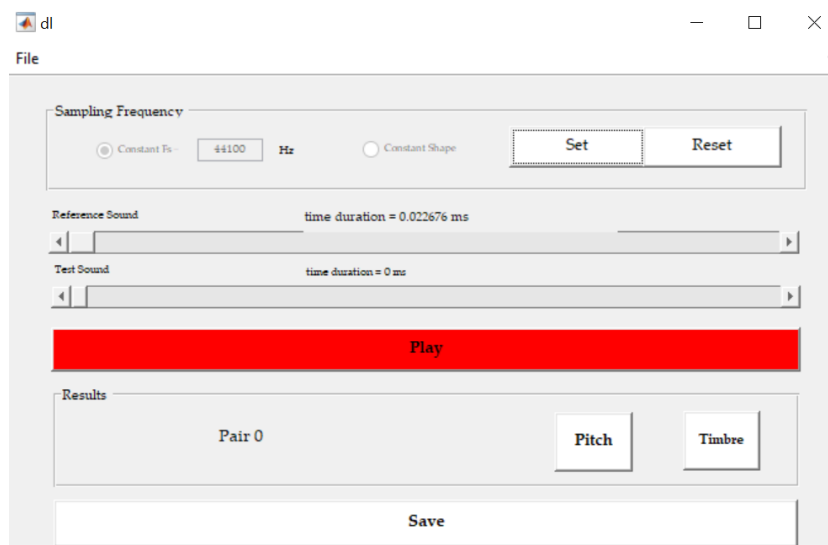
Rysunek 17: Schemat wpływu różnych czynników związanych z geometrią głowy, ucha i przewodu słuchowego na wzmocnienie lub osłabienie odbioru tonów prostych w zależności od ich częstotliwości [89].

„Duration discrimination is much better with the reminder method than with the roving one; in other words, discrimination is better when the standard interval, which is kept constant, is presented before a comparison interval, which is varied from trial to trial.” [91].

Zatem każda para porównywanych impulsów gaussowskich zaczynała się od impulsu wzorcowego (referencyjnego). Badana osoba sama dobierała drugi impuls z pary, aby ocenić kiedy wydaje się on identyczny, a kiedy inny niż impuls wzorcowy. Jest to więc przykład testu adaptacyjnego [20] str. 13. Oczywiście zebranie krzywej psychometrycznej dla każdego impulsu wzorcowego byłoby bardzo interesujące, lecz zarazem bardzo długotrwałe i męczące dla badanych osób. Zapewne warto wykonać takie badanie w przyszłości. Badania behawioralne odbyły się w Politechnice Krakowskiej. Badani uczestniczyli w teście przygotowanym w środowisku

MatLab. Test składał się z modułu o nazwie *Difference Limen*. Program testujący składa się z serii 25 dźwięków referencyjnych, różniących się czasem trwania. W trakcie każdej z serii testowej uczestnicy zmieniali czas trwania dźwięku testowego (maksymalnie 40 zmian). Interfejs użytkownika przedstawiony jest na rys. 18. Jak widać na rysunku 18, główną częścią programu są dwa suwaki oraz przycisk „Play”. Pierwszy suwak dotyczy dźwięku referencyjnego a drugi (poniżej) dźwięku testowego. Dźwięk referencyjny podczas każdego pomiaru ustawiany był na jedną wartość, a badany był zmieniany przez osobę testowaną. Operując suwakiem dotyczącym czasu trwania dźwięku testowego, każda osoba testowana mogła wyznaczyć moment, tj. numer dźwięku testowego, dla którego dźwięki testowy i referencyjny stawały się nierozróżnialne. W ten sposób otrzymano informację jaka różnica czasu trwania jest potrzebna, aby rozróżnić od siebie dźwięk testowy i dźwięk referencyjny.

Wartości czasów trwania impulsów referencyjnych σ wynosiły $n \cdot \frac{1000}{44100} ms = n \cdot 0,0226757...ms$ dla $n = 1, 2, \dots, 25$. Dla każdego impulsu referencyjnego słuchacz miał do wyboru 40 impulsów testowych, których czasy trwania wynosiły $(n + 0,15 \cdot m) \cdot \frac{1000}{44100} ms$ dla $m = 0, 1, 2, \dots, 39$. Analizując krok po kroku każdy



Rysunek 18: Interfejs programu testującego użytego podczas badań behawioralnych przeprowadzonych w Politechnice Krakowskiej.

dźwięk testowy, uczestnik badania decydował, w którym momencie tzn dla jakiej

wartości m różnica staje się wg niego wykrywalna. Kiedy podjął decyzję naciskał guzik „Pitch” jeśli słyszał różnice w wysokości dźwięku, lub przycisk „Timbre” jeśli słyszał różnice w barwie dźwięku. Ten sposób analizy dotyczył muzyków, którzy są wyczuleni na barwę dźwięku oraz zaznajomieni z nomenklaturą muzyczną. W przypadku niemuzyków podjęcie decyzji można było zakomunikować tylko używając przycisku „Pitch”. Jeśli ktoś omyłkowo nacisnął guzik „Timbre”, nie stanowiło to problemu, ponieważ później odczytując wynik testu informacja pojawiała się po prostu w innej kolumnie. Program testujący zwracał plik o rozszerzeniu .txt w którym w dwóch kolumnach zawarta była informacja nt.: w pierwszej numeru dźwięku referencyjnego, a w drugiej numeru pary przy której odnotowano różnicę. Przykładowe odpowiedzi zamieszczone są w tabeli 3. Dla najkrótszego sygnału $\sigma = \frac{1000}{44100} ms$, osoba badana zauważyła różnicę już w pierwszym sygnale testowym o $\sigma = \left(\frac{1000}{44100}\right) \cdot (1 + 0,15)ms$ zarówno w dziedzinie wysokości dźwięku, jak i barwy. Ta sama osoba zauważyła różnicę barwy dźwięku, dla sygnałów $\sigma = \frac{1000}{44100} \cdot (11)ms$ i $\frac{1000}{44100} \cdot (11 + 2 \cdot 0,15)ms$, ale różnicę wysokości dźwięku zauważyła dopiero gdy czas trwania dźwięku testowego wynosi $\sigma = \frac{1000}{44100} \cdot (11 + 10 \cdot 0,15)ms$. Oznacza to, że wybrana osoba jest bardzo czuła na barwę dźwięku, podczas, gdy czułość percepcji wysokości dźwięku jest znacznie mniejsza.

Większą z badanych grup (39 osób) była grupa niemuzyków, która nie potrafiła określić czy różnica zauważalna jest w barwie czy w wysokości dźwięku a przykładowe odpowiedzi dla 4 osób z grupy niemuzyków widać na wykresie 19. Uczestnik nr 3, o dość wyjątkowym sposobie postrzegania różnic pochodził z grupy muzyków.

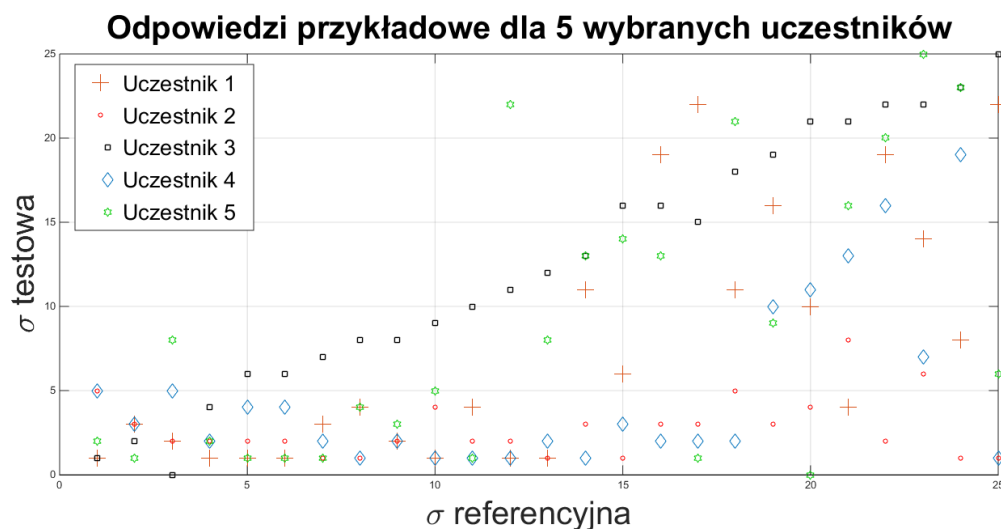
Tak przygotowane odpowiedzi z testów behawioralnych w łatwy sposób można skopiować do programu *Excel*, *MatLab*, *Maple* za pomocą których wykonywano dopasowania, analizy statystyczne itp.

Odpowiedzi uczestników były zaszyfrowane. Każdy uczestnik przystępując do badania musiał wypełnić *oświadczenie o wyrażeniu zgody na przetwarzanie danych osobowych na potrzeby projektu* (zał. 1), *ankietę osobową* w której podawał dane tj.: imię i nazwisko, numer telefonu, rok urodzenia, płeć oraz w przypadku muzyka instrument na którym praktykuje grę (zał. 2). Załącznik nr 3 przedstawia *instrukcję wykonania testu behawioralnego*. Każdy badany przed przystąpieniem do testu musiał dokładnie zapoznać się z instrukcją. Jeśli miał jakieś pytania czy wąt-

Tablica 3: Przykładowa odpowiedź uczestnika badań behawioralnych (nr 3 na wykresie 19). W kolumnie pierwszej przedstawiony jest numer dźwięku referencyjnego, w kolumnie drugiej numer pary dla której zarejestrowano różnicę w wysokości dźwięku, oraz w kolumnie trzeciej numer pary dla której zarejestrowana została różnica w barwie dźwięku.

Nr dźwięku referencyjnego n	Nr pary: pitch m	Nr pary: timbre m
1	1	1
2	2	3
3	3	1
4	4	2
5	6	2
6	6	2
7	7	2
8	8	2
9	8	3
10	9	7
11	10	2
12	11	3
13	12	3
14	13	1
15	16	3
16	16	5
17	15	5
18	18	4
19	19	6
20	21	6
21	21	5
22	22	3
23	22	4
24	23	6
25	25	5

pliwości mógł to zgłosić. Uczestnicy w każdej chwili mogli zrezygnować z badania. Każdy badany zajmował stanowisko komputerowe (rys. 21) na którym uruchomiony był test. Stanowiska wyposażone były w komputer oraz niezbędne do jego obsługi narzędzia tj.: klawiatura czy myszka. Dodatkowo każdy badany posiadał słuchawki nauszne z firmy Creative, model Hs-600 wyposażone w 50 mm głośniki



Rysunek 19: Odpowiedzi uzyskane podczas testu behawioralnego dla wybranych uczestników (uczestnicy 1, 2, 4, 5 - niemuzycy, 3 - muzyk, którego odpowiedź przedstawiona jest w tabeli 3).

dynamiczne oparte na magnesach neodymowych. Pasmo przenoszenia mieści się w zakresie od 20 Hz do 20 kHz , deklarowana skuteczność wynosi 102 dB , a opór 32 Ohm [92]. Słuchawki ważą 210 g . Ich rozmiar dostosowuje się do wielkości głowy użytkownika za pomocą regulowanych pałków jak widać na rys. 20. Muszle (część nauszna) okrywa całą małżowinę uszną. W razie potrzeby można regulować ich głośność za pomocą pilota w kształcie „łezki” umieszczonego na kablu.

Badanie przeprowadzono na dwóch grupach osób. Ze względu na doniesienia literaturowe nt. różnic pomiędzy percepcją dźwięku przez muzyków w porównaniu do osób nieposiadających muzycznych zdolności [93, 94, 95], postanowiono podzielić badaną grupę na dwie mniejsze wg kategorii:

- Muzycy (36 osób): osoby z wykształceniem muzycznym, praktykujące grę na instrumentach oraz posiadające zajęcia z kształcenia słuchu, podczas których ćwiczy się zagadnienia związane z percepcją dźwięku tj. rozróżnianie interwałów (odległości między dźwiękami), czy akordów. Muzycy deklarują również zdolność do rozróżnienia barwy dźwięku od jego wysokości.
- Laicy (niemuzycy, 39 osób): osoby bez wykształcenia muzycznego, bez wymienionych wyżej umiejętności.



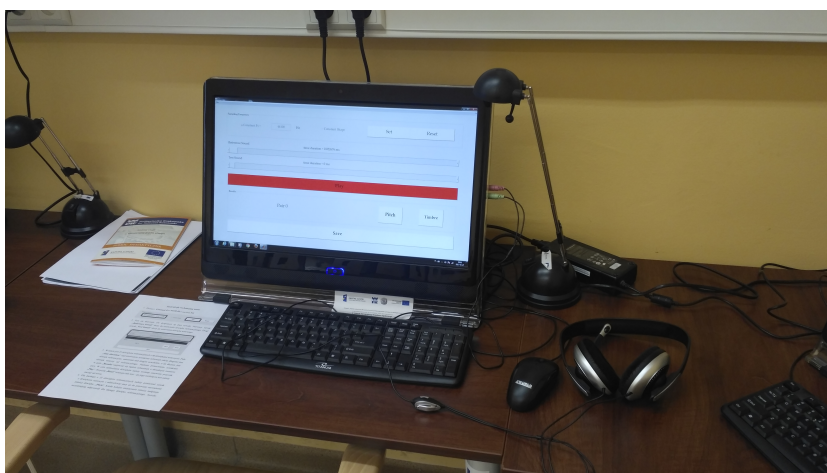
Rysunek 20: Słuchawki Creative, HS-600 wraz z mikrofonem [92].

Obie grupy były zróżnicowane pod względem wieku oraz płci. Uczestnicy testu nie otrzymywali wynagrodzenia za wzięcie udziału w badaniu.

Wszyscy uczestnicy badania wyrazili świadomą zgodę na udział w nim przed rozpoczęciem badania. Badanie zostało przeprowadzone zgodnie z Deklaracją Helsińską (Światowa Organizacja Medyczna, 1996) oraz protokołem zatwierdzonym przez Komitet Etyki Uniwersytetu Jagiellońskiego w Krakowie, Instytutu Psychologii, zajmującego się badaniami empirycznymi z udziałem ludzi (numer 77/2020).

8.3 Wyniki

Na rysunku 22 zebrano wyniki testu behawioralnego dla muzyków, którzy odpowiadali w kategorii wysokości dźwięku. Najmniejsza zauważalna różnica czasu trwania impulsu $\Delta\sigma_{DL}$ oceniana przez tę samą grupę na podstawie barwy dźwięku uwidoczniona jest na rys. 23. Z kolei rys. 24 przedstawia najmniejszą zauważalną różnicę $\Delta\sigma_{DL}$ zarejestrowaną przez grupę niemuzyków, którzy nie określali cechy dźwięku będącej podstawą ich decyzji. Na wszystkich tych rysunkach zaznaczono wartość średnią $\Delta\sigma_{DL}$ wraz z zakresem średniego kwadratowego odchylenia standardowego. Zatem rys. 22, 23 i 24 stanowią przykłady empirycznego prawa psychofizycznego. We wszystkich przypadkach prawa te przedstawione są funkcją

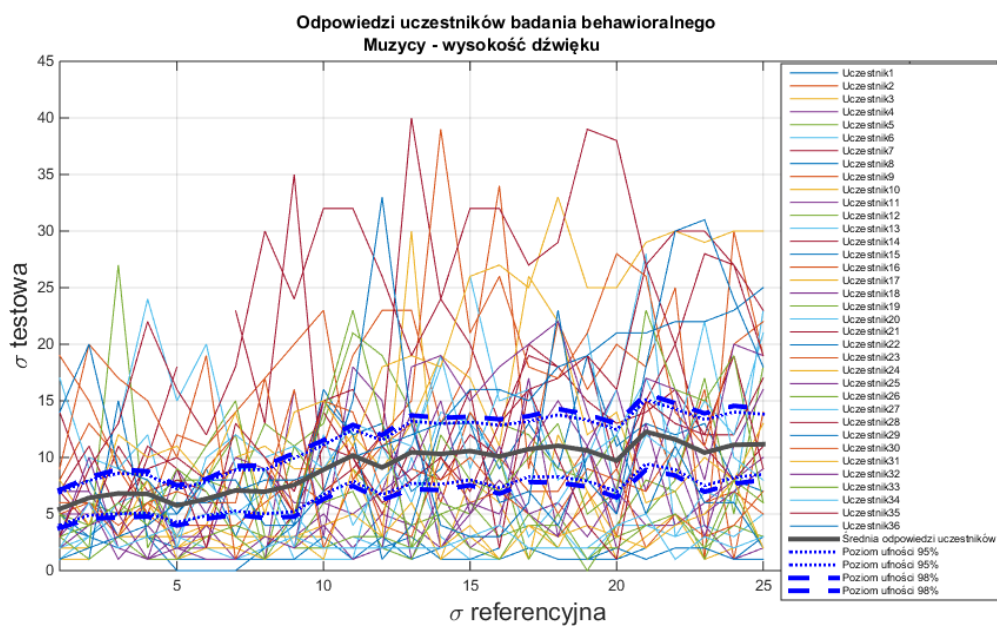


Rysunek 21: Stanowisko pomiarowe.

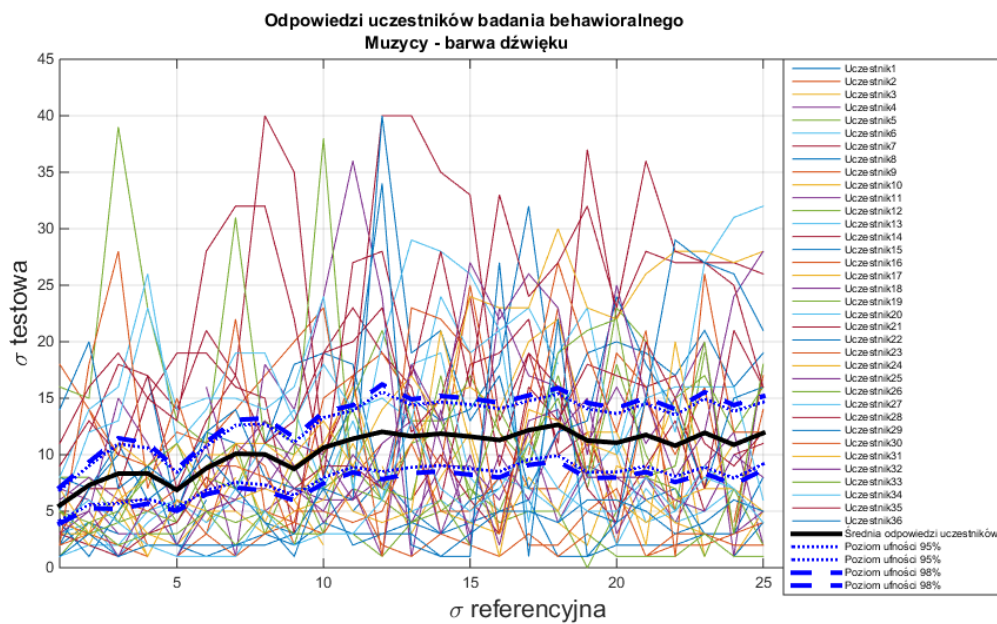
rosnącą, choć w odróżnieniu od prawa Webera-Fechnera krzywe nie przechodzą przez początek układu współrzędnych.

Porównanie przebiegu wartości średnich przedstawione na rys. 25 wskazuje, że muzycy zauważają mniejsze różnice czasów trwania impulsów niż niemuzycy. Wydaje się to zgodne z przewidywaniem i wskazuje to na rolę treningu w pracy z dźwiękami. Zwraca też uwagę wyraźnie większe nachylenie prawa psychofizycznego dla osób bez doświadczenia w słuchaniu i odtwarzaniu dźwięków. W grupie muzyków krzywa wysokości dźwięku leży pod krzywą barwy dźwięku co oznacza, że osoby te są bardziej wrażliwe na wysokość. Uczestnik badania, którego dane umieszczone są w tabelicy 3 jest tu wyjątkiem zauważającym różnicę w barwie dla bardzo małych różnic znacznie wcześniej niż różnicę w wysokości dźwięku, przy czym to ostatnie kryterium wskazuje na prawo psychofizyczne bardzo zbliżone do prawa Webera-Fechnera.

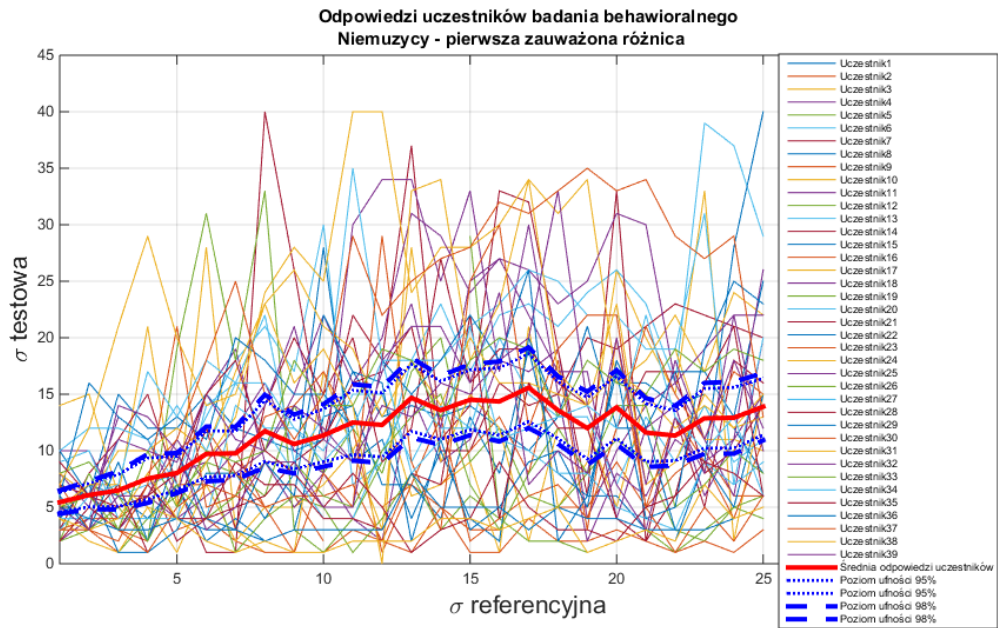
Otrzymane wyniki wskazują na prawo psychofizyczne przedstawione krzywą drugiego stopnia. Zadowalające przybliżenie liniowe, szczególnie widoczne w wynikach zarejestrowanych przez niemuzyków, otrzymuje się biorąc pod uwagę 17 pierwszych wartości. Ten efekt może mieć związek z obciążeniem niskich częstotliwości przez kartę dźwiękową komputera i/lub słuchawki. Być może podobne obciążenie zachodzi także w narządach słuchu.



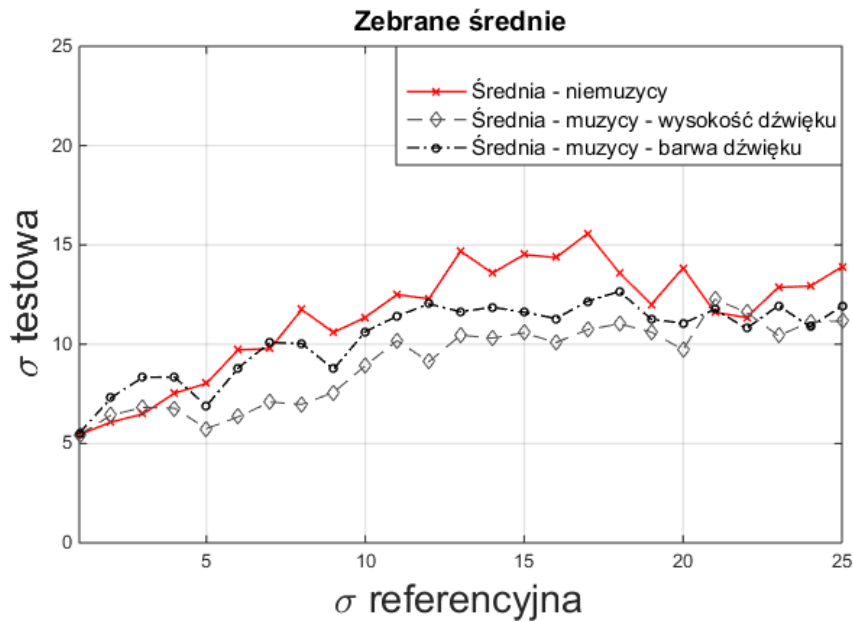
Rysunek 22: Odpowiedzi uczestników badania behawioralnego dla grupy muzyków, która dokonywała rozróżnienia pomiędzy dwoma dźwiękami na podstawie wysokości dźwięku.



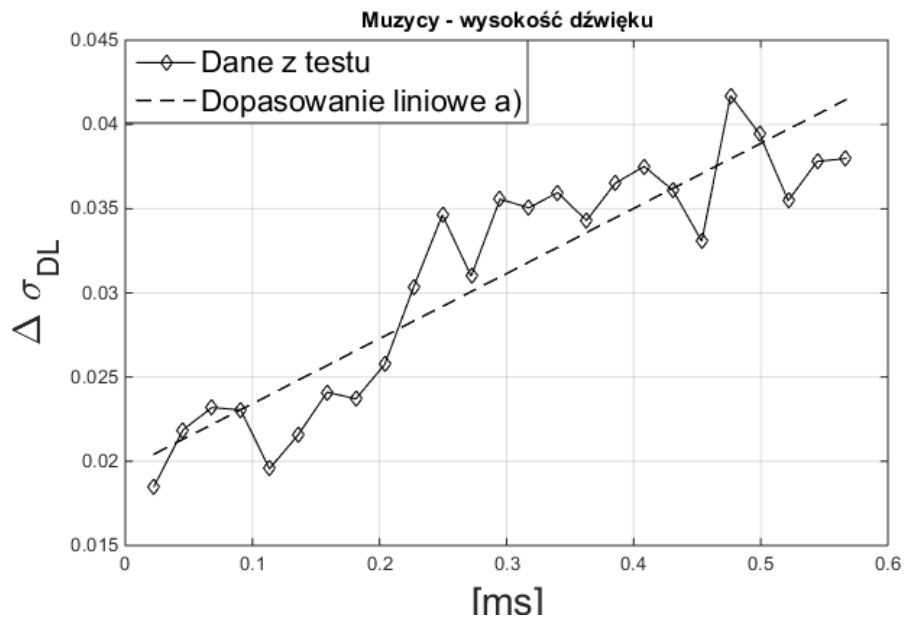
Rysunek 23: Odpowiedzi uczestników badania behawioralnego dla grupy muzyków, która dokonywała rozróżnienia pomiędzy dwoma dźwiękami na podstawie barwy dźwięku.



Rysunek 24: Odpowiedzi uczestników badania behawioralnego dla grupy niemuzyków - pierwsza zauważalna różnica.



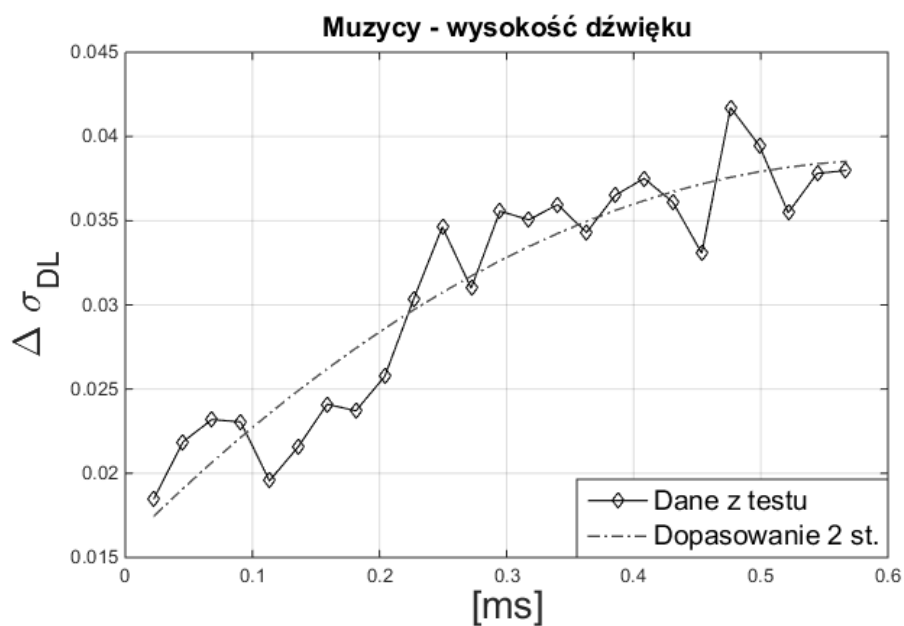
Rysunek 25: Wykres przedstawiający zebrane średnie z rysunków 22, 23 oraz 24.



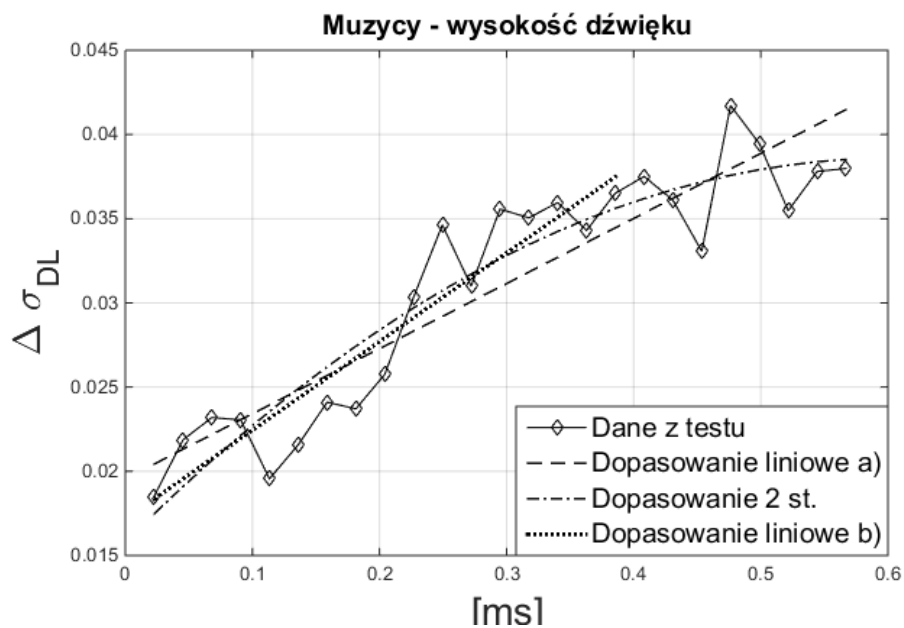
Rysunek 26: Wykres przedstawiający średnie z odpowiedzi muzyków dot. wysokości dźwięku oraz dopasowaną prostą $y = 0,039(0,003) \cdot x + 0,019(0,001)$ (dopasowanie liniowe a)).



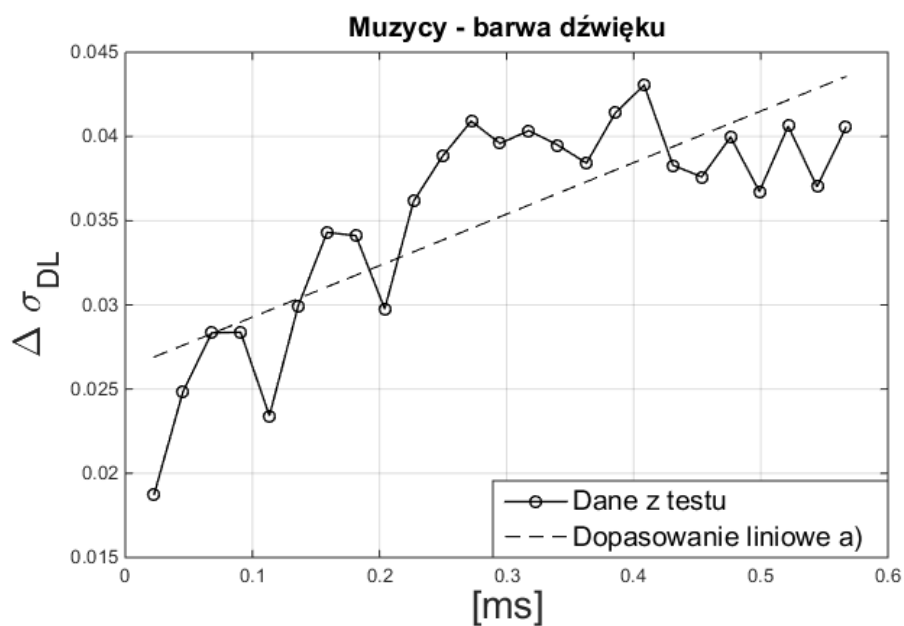
Rysunek 27: Wykres przedstawiający średnie z odpowiedzi muzyków dot. wysokości dźwięku oraz dopasowaną prostą dla pierwszych 17 punktów: $y = 0,052(0,004) \cdot x + 0,0171(0,0001)$ (dopasowanie liniowe b)).



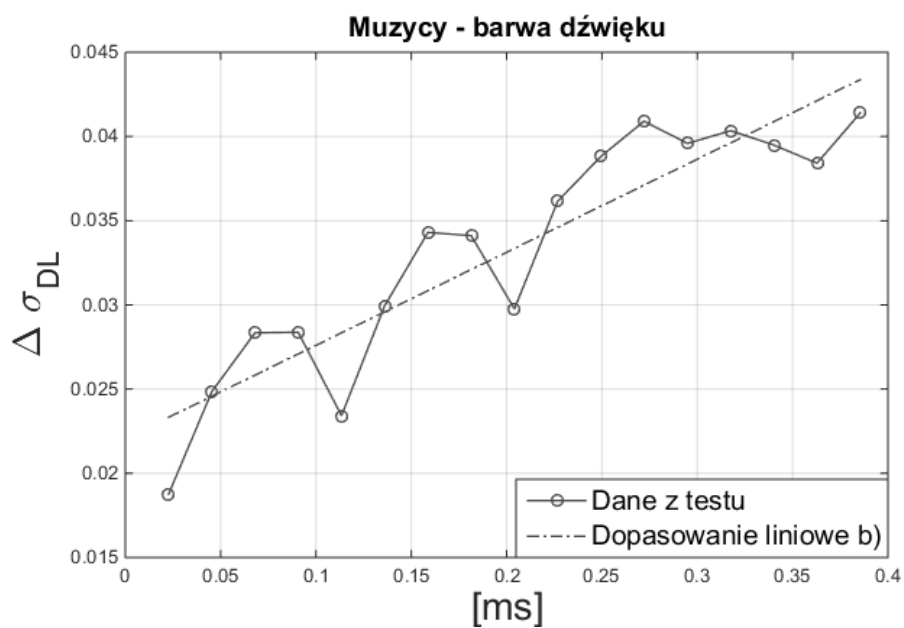
Rysunek 28: Wykres przedstawiający średnie z odpowiedzi muzyków dot. wysokości dźwięku oraz dopasowaną funkcją kwadratową $y = -0,062(0,01) \cdot x^2 - 0,075(0,09) \cdot x + 0,016(0,001)$ (dopasowanie 2 st.).



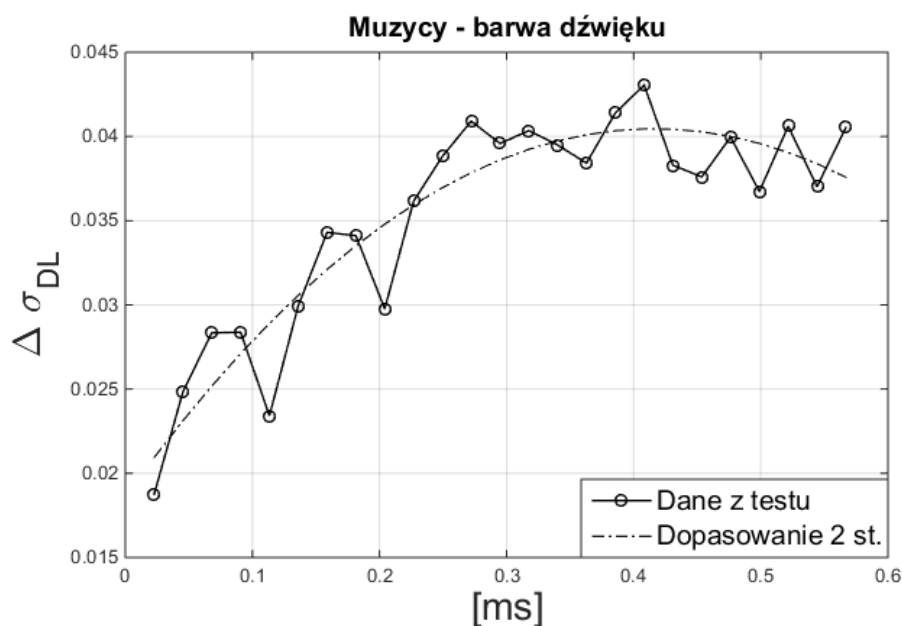
Rysunek 29: Zestawienie wszystkich dopasowań do średniej najmniejszej zauważalnej różnicy czasu trwania impulsu $\Delta \sigma_{DL}$ w funkcji czasu trwania σ stwierdzonej przez muzyków na podstawie wysokości dźwięku.



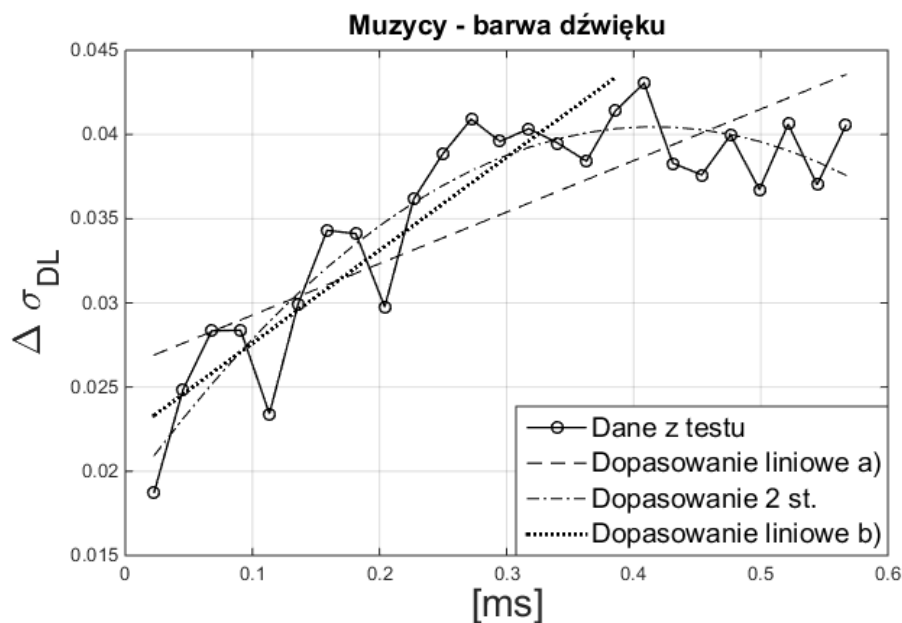
Rysunek 30: Wykres przedstawiający średnie z odpowiedzi muzyków dot. barwy dźwięku oraz dopasowaną funkcją liniową $y = 0,031(0,004) \cdot x + 0,026(0,002)$ (dopasowanie liniowe a)).



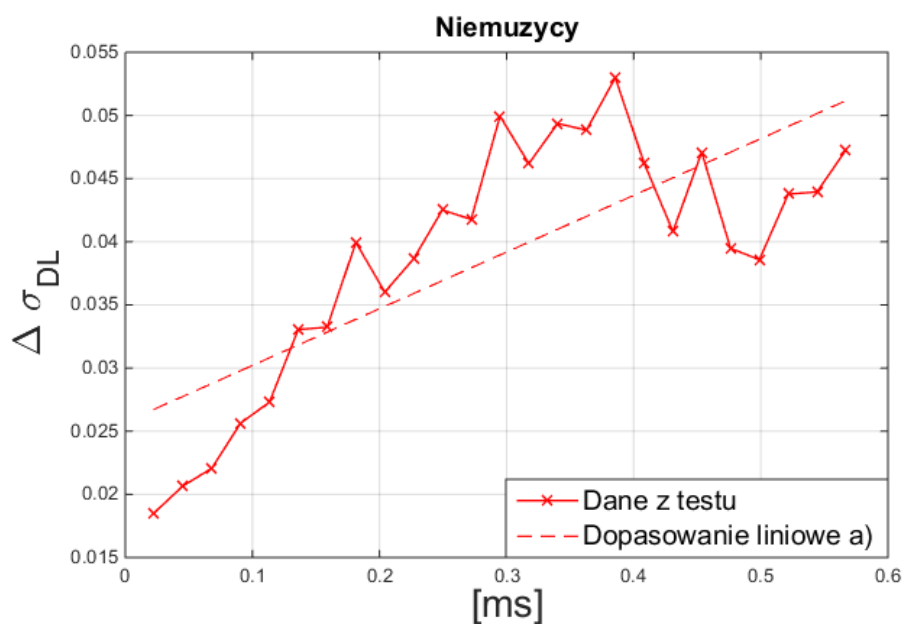
Rysunek 31: Wykres przedstawiający średnie z odpowiedzi muzyków dot. barwy dźwięku oraz dopasowaną funkcją liniową dla pierwszych 17 punktów. $y = 0,055(0,006) \cdot x + 0,022(0,001)$ (dopasowanie liniowe b)).



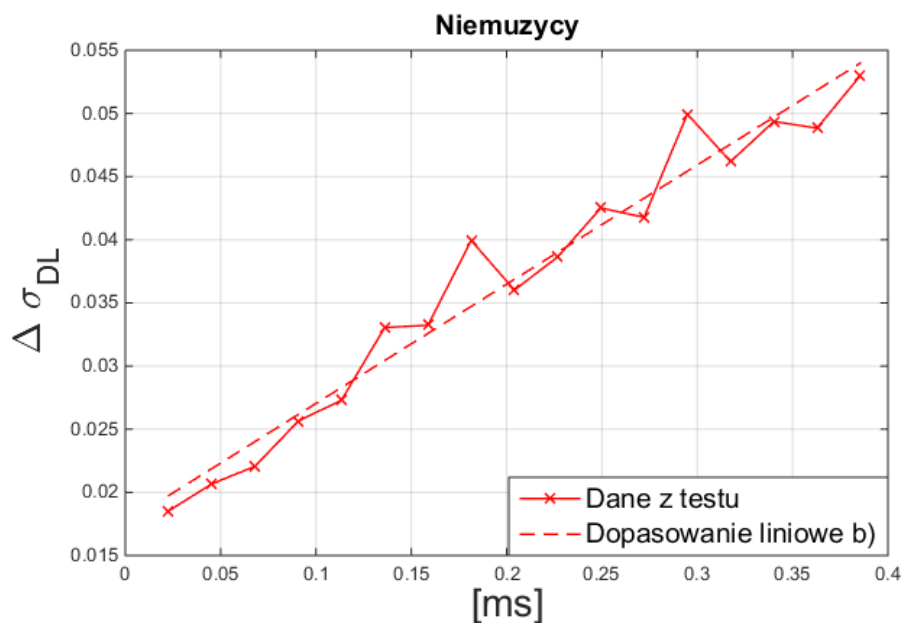
Rysunek 32: Wykres przedstawiający średnie z odpowiedzi muzyków dot. barwy dźwięku oraz dopasowaną funkcję kwadratową $y = -0,13(0,02) \cdot x^2 - 0,10(0,01) \cdot x + 0,019(0,001)$ (dopasowanie 2 st.).



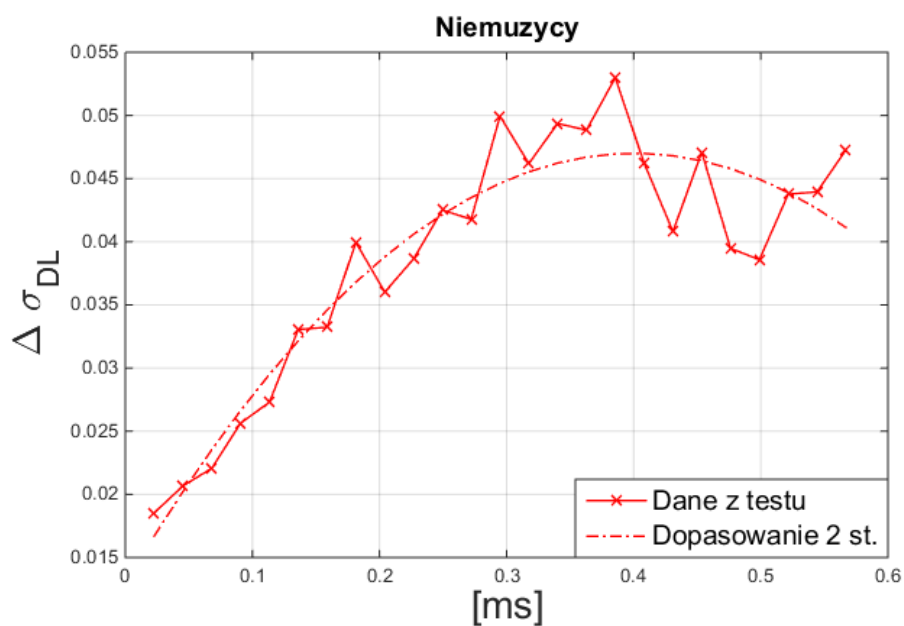
Rysunek 33: Zestawienie wszystkich dopasowań do średniej najmniejszej zauważalnej różnicy czasu trwania impulsu $\Delta \sigma_{DL}$ w funkcji czasu trwania σ stwierdzonej przez muzyków na podstawie barwy dźwięku.



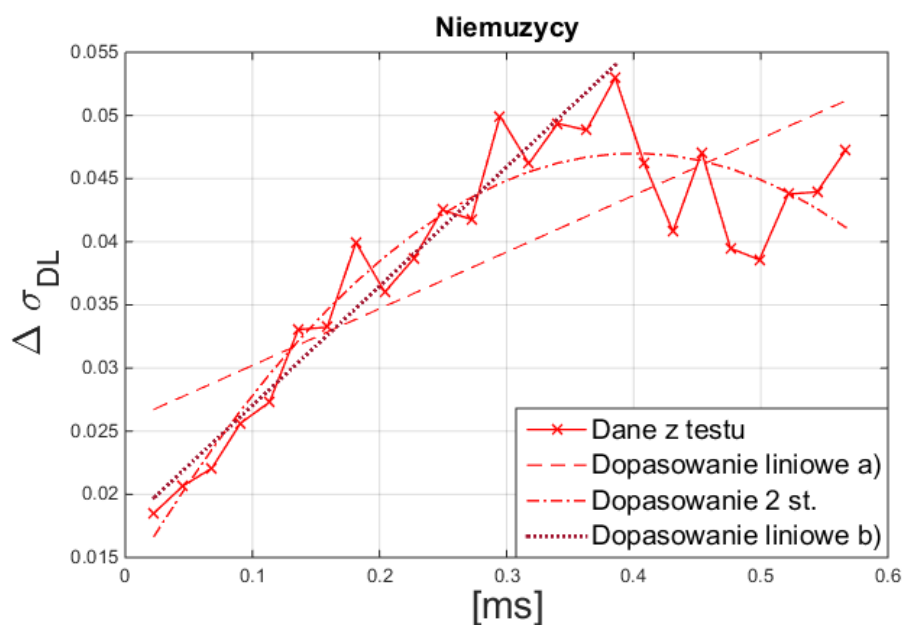
Rysunek 34: Wykres przedstawiający średnie z odpowiedzi niemuzyków oraz dopasowaną funkcją liniową $y = 0,045(0,007) \cdot x + 0,026(0,003)$ (dopasowanie liniowe a)).



Rysunek 35: Wykres przedstawiający średnie z odpowiedzi niemuzyków dla pierwszych 17 punktów oraz dopasowaną funkcję liniową $y = 0,094(0,005) \cdot x + 0,017(0,001)$ (dopasowanie liniowe b)).



Rysunek 36: Wykres przedstawiający średnie z odpowiedzi niemuzyków oraz dopasowaną funkcję kwadratową $y = -0,21(0,3) \cdot x^2 - 0,17(0,02) \cdot x + 0,013(0,002)$ (dopasowanie 2 st.).



Rysunek 37: Zestawienie wszystkich dopasowań do średniej najmniejszej zauważalnej różnicy czasu trwania impulsu $\Delta \sigma_{DL}$ w funkcji czasu trwania σ stwierdzonej przez niemuzyków.

9 Badania w komorze bezechowej

9.1 Wprowadzenie

9.1.1 Komora bezechowa

Komora bezechowa jest wykorzystywana do wielu dziedzin badań. Dzięki temu, że panuje tam idealna cisza można badać wszelkiego rodzaju dźwięki czy hałas, m.in: szum opon [96], hałas drogowy [97], zjawiska psychoakustyczne [98], testowanie różnego rodzaju urządzeń np. 5G [99]. Prowadzone są także badania nad modelowaniem komór bezechowych tak, aby spełniały określone wymagania np. komory stożkowe [100]. W naszym badaniu pod uwagę były brane bardzo krótkie impulsy akustyczne, odtwarzane w komorze bezechowej. W literaturze można spotkać również przykład badania nagrywania materiału dźwiękowego w komorze bezechowej, np. autorzy [101] porównywali zapis wielokanałowy źródła nagrany w komorze z nagraniem w studio.

Badania opisane w niniejszej rozprawie przeprowadzono w komorze bezechowej Akademii Górniczo-Hutniczej (AGH). Komora bezechowa AGH jest częścią Laboratorium Akustyki Technicznej. W komorze panują warunki pola swobodnego spowodowane umieszczeniem na ścianach elementów pochłaniających oraz rozpraszających. Komora bezechowa AGH zbudowana jest na bazie sześcianu, gdzie krawędź zewnętrzna wynosi około 10 m. Objętość użyteczna wynosi 342 m³. Masa komory to ok. 599 ton. Komora osadzona jest na wibroizolacji wspartej na 50 sprężynach, dzięki temu drgania z budynku i drogi są ograniczone. Do komory wchodzi się na rozciągniętą siatkę, która umieszczona jest pół metra nad klinami. Poziom ciśnienia akustycznego w komorze w ciągu dnia wynosi 1,5 dB a w ciągu nocy spada nawet poniżej 0 dB [102].

9.1.2 Sztuczna głowa

Podczas badań w komorze bezechowej użyto symulatora sztucznej głowy: *Head, Torso and Ear Simulators (HATS)*, typu 5128, produkowanego przez firmę *Brüel & Kjaer* (Rys. 38). Na rysunku 39 przedstawiono przekrój przez sztuczną głowę, gdzie możemy zobaczyć jak zaimplementowany jest układ słyszenia. Fala dźwiękowa wpada do małżowiny usznej ukształtowanej na podobieństwo prawdzi-

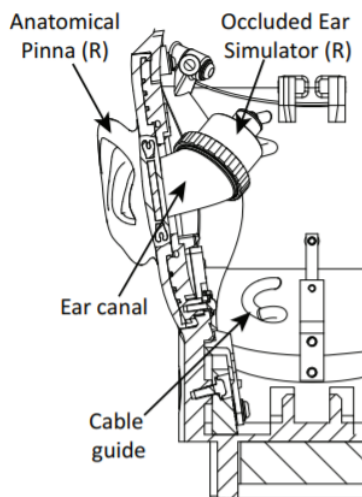
wej. Jest to realistyczna reprodukcja własności akustycznych spotykanych u dorosłego człowieka (dane uśrednione). Sztuczna głowa jest stworzona do elektroakustycznych badań *in situ* i pokrywa zakres częstotliwości słyszanych przez człowieka. Dzięki pomiarom za pomocą sztucznej głowy możliwa jest ocena zrozumiałości dwukierunkowej komunikacji zestawu słuchawkowego w obecności szumu tła i warunków, które mogą powodować sprzężenie zwrotne w systemie. Ułatwia również ocenę dokładności nagrywania mikrofonu w przypadku mowy ludzkiej ze zmianami poziomu czy kierunkowości. Co najważniejsze, pomiar odpowiedzi dźwiękowej słuchawek jest dokonywany w sposób w jaki sygnały dźwiękowe są odbierane przez człowieka. Pozwala to na identyfikowanie, korygowanie i niwelowanie błędów w konstrukcji słuchawek, które mogą prowadzić do artefaktów w odpowiedzi częstotliwościowej. Na sztucznej głowie można również przeprowadzać pomiary jakości dźwięku aparatów słuchowych z prawidłowym anatomicznym ułożeniem czy pomiary tłumienia i skutecznego uszczelnienia ochronników słuchu o prawidłowej geometrii przewodu słuchowego i anatomii głowy [104]. Przed przystąpieniem do wykonania pomiarów na sztucznej głowie sprzęt należy skalibrować. Tabela kalibracji dostarczana jest wraz z konkretnym modelem.

Wpływ połączenia tułowia, głowy, małżowiny usznej i przewodu słuchowego na sygnały dźwiękowe przenoszone drogą powietrzną można określić ilościowo za pomocą funkcji przenoszenia w swobodnym polu akustycznym. Jest to odpowiedź częstotliwościowa z pola swobodnego do błony bębenkowej, znana w telekomunikacji jako odpowiedź częstotliwościowa w polu swobodnym słuchacza lub



Rysunek 38: Sztuczna głowa: Head and Torso Simulator (HATS), typ 5128 [103].

funkcja przenoszenia związana z głową (HRTF) oraz w audiologii technicznej jako odpowiedź częstotliwościowa manekina [104].



Rysunek 39: Przekrój przez prawe ucho sztucznej głowy w symulatorach HATS. Zawiera małżowinę, kanał uszny oraz symulator ucha niedroźnego [104].

Badania przeprowadzono w trzech konfiguracjach:

- Mikrofon - głośnik, rys. 40;
- Sztuczna głowa - głośnik, rys. 41;
- Sztuczna głowa - słuchawki, rys. 42.

W konfiguracji pierwszej głośnik oddalony był od mikrofonu o 1 m, a oba przyrządy pomiarowe ustawione były na wysokości 1,25 m. Te same odległości zostały zachowane w przypadku pomiaru w konfiguracji drugiej tj. sztuczna głowa rejestrująca dźwięki z głośnika.

9.2 Metody

Podczas badań nad percepcją bardzo krótkich impulsów dźwiękowych używano różnego rodzaju sygnałów. Szczególnie, zajęto się analizą sygnałów coinusoidalnych oraz gaussowskich. Przedstawione są one na rys. 43.



Rysunek 40: Konfiguracja 1, mikrofon - głośnik [źródło własne].

Badane sygnały to gaussowskie impulsy akustyczne (wzór 65),

$$\exp\left(\frac{-t^2}{2(\Delta t)^2}\right), \quad (65)$$

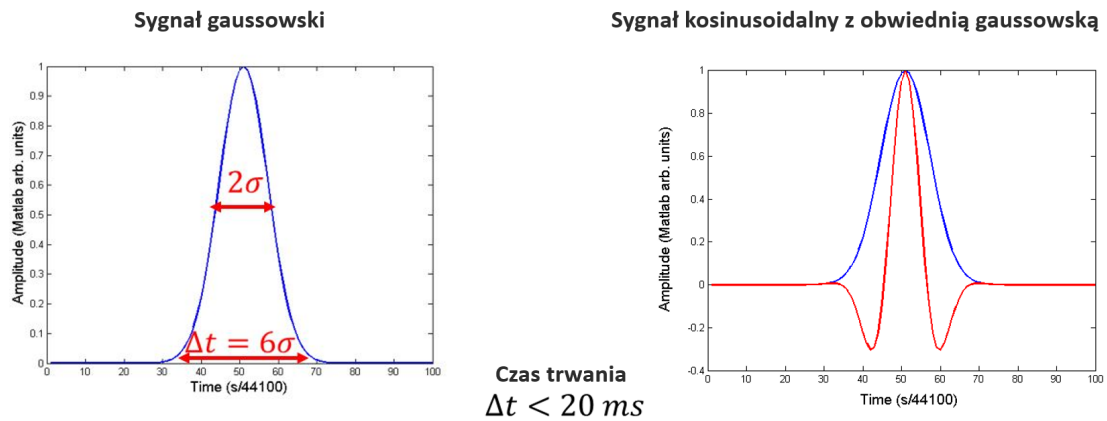
gdzie czas trwania dźwięku Δt jest z zakresu 0,013 ms - 27,21 ms. Zbadano również sygnały sinusoidalne o częstotliwości 440 Hz. Na rysunku 44 przedstawiono przebiegi generowanych impulsów w dziedzinie czasu.



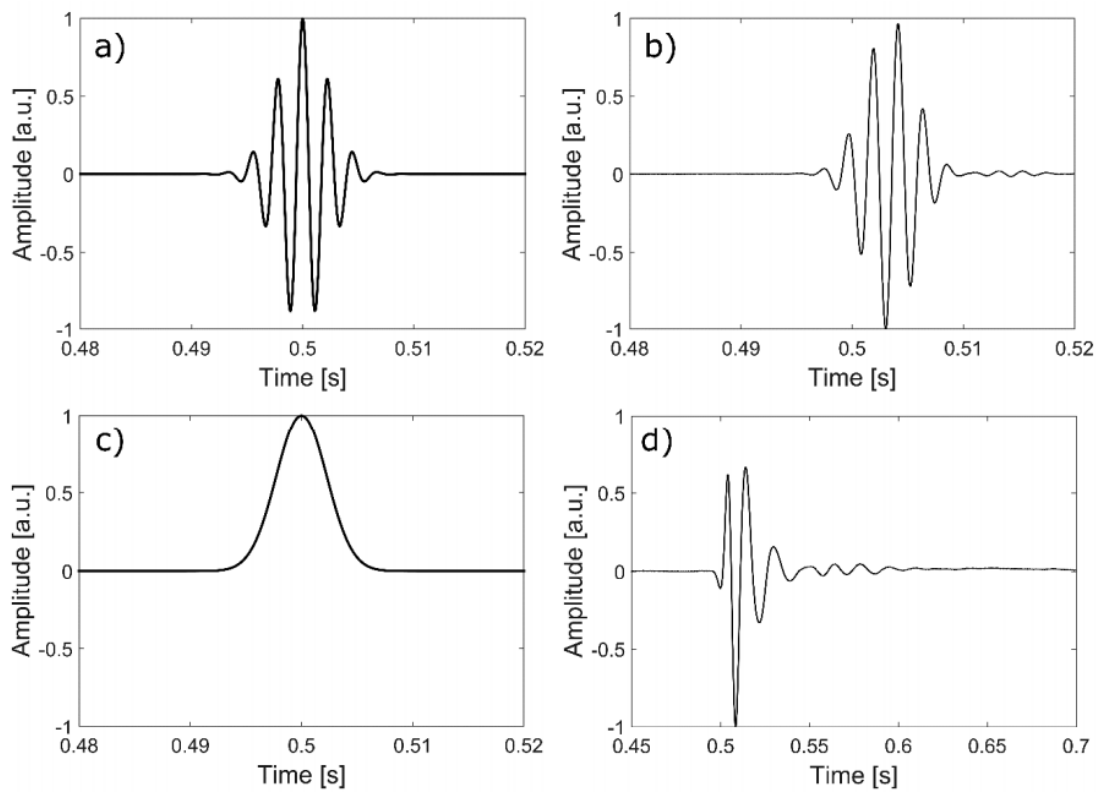
Rysunek 41: Konfiguracja 2, sztuczna głowa - głośnik [źródło własne].



Rysunek 42: Konfiguracja 3, sztuczna głowa - słuchawki [źródło własne].



Rysunek 43: Sygnal gaussowski (po lewej) oraz sygnal cosinusoidalny (po prawej)



Rysunek 44: Przebiegi impulsów generowanych w środowisku MatLab oraz odpowiedzi głośnika na zaproponowane wymuszenie: a) generowany sygnal cosinusoidalny z obwiednią gaussowską, b) odpowiedź głośnika na wymuszenie sygnałem a), c) generowany sygnal gaussowski, d) odpowiedź głośnika na wymuszenie sygnałem c). Czas trwania wszystkich impulsów wynosi 13,60 ms [105].

Pomiary przeprowadzone w komorze bezechowej pozwoliły na zebranie odpowiedzi zarejestrowanych przez sztuczną głowę. Dane dotyczące otrzymanych widm sygnałów zostały zebrane w dodatku (rozd. 12) zamieszczonym na końcu niniejszej rozprawy. Badaniom w komorze bezechowej poddano impulsy cosinusoidalne i gaussowskie. Poniżej prezentowany jest kod stworzony w środowisku Matlab za pomocą którego generowano dźwięki testowe - 200 impulsów testowych próbkowanych z częstotliwością 44100 Hz.

```

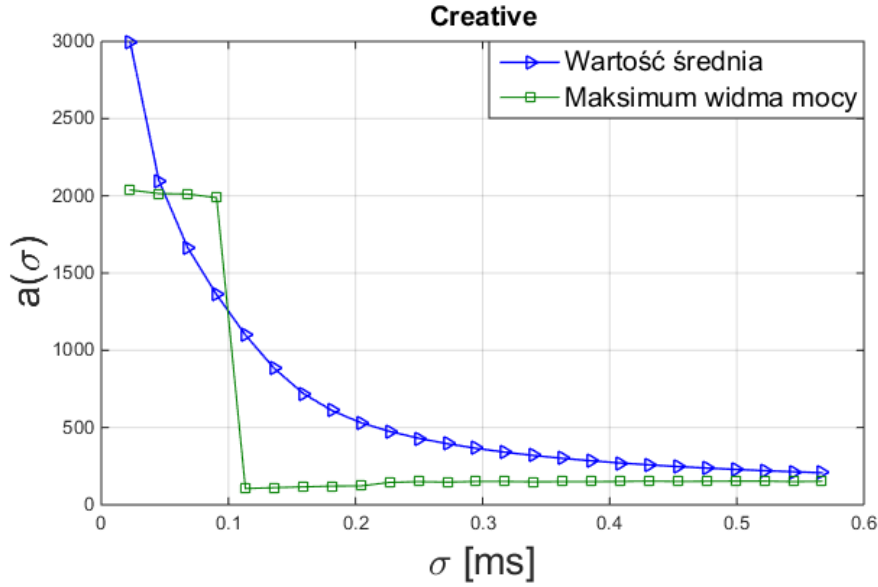
1  clc
2  close all
3  clear all
4  Fs = 44100;
5  t = -Fs/2:Fs/2;
6  sigma = 1:200;
7  f = 440;
8  y = cos(2*pi*f*t/Fs);
9  for j = 1:200
10     ref = exp(-(t.^2)/(2*(sigma(j)^2)));
11     audiowrite(['gauss_' num2str(j) '.wav'], ref, Fs);
12 end

```

9.3 Wyniki

Aby porównać wyniki testów behawioralnych z reakcjami sztucznej głowy, wygenerowano te same impulsy i dostarczono je przez ten sam typ słuchawek. Dla wszystkich impulsów zarejestrowano widma mocy $X(\omega; \sigma)$. Rozważono dwie cechy każdego widma: i) wartość oczekiwaną częstotliwości - jako reprezentację efektywnej wysokości dźwięku

$$a(\sigma) = \frac{\sum_{\omega} \omega \cdot X(\omega; \sigma)}{\sum_{\omega} X(\omega; \sigma)}. \quad (66)$$



Rysunek 45: Wartości częstotliwości uzyskane z widma mocy dla słuchawek Creative wyliczone jako maksimum widma mocy oraz wartość oczekiwana

oraz ii) barwę dźwięku $F_r(\sigma)$

$$F_r(\sigma) = \frac{\partial T}{\partial \sigma} = \int \left| \frac{\partial X(\omega; \sigma)}{\partial \sigma} \right| d\omega \quad (67)$$

biorąc pod uwagę stopień zmienności barwy przy zmianie σ . Następnie obliczono minimalną zmianę $\Delta\sigma_K$ czasu trwania σ potrzebną do zapewnienia odpowiedniego przesunięcia wartości oczekiwanej widma $a(\sigma)$ o odchylenie standardowe

$$s(\sigma) = \sqrt{\frac{\sum_{\omega} |\omega - a^2| \cdot X(\omega; \sigma)}{\sum_{\omega} X(\omega; \sigma)}}, \quad (68)$$

pomnożone przez współczynnik eksperymentalny.

Następnie obliczono minimalną zmianę $\Delta\sigma_K$ czasu trwania σ potrzebną do zapewnienia odpowiedniego przesunięcia wartości oczekiwanej widma $a(\sigma)$ o odchylenie standardowe

$$s(\sigma) = \sqrt{\frac{\sum_{\omega} |\omega - a|^2 \cdot X(\omega; \sigma)}{\sum_{\omega} X(\omega; \sigma)}}, \quad (69)$$

pomnożone przez współczynnik eksperymentalny.

Obliczeń dokonano za pomocą poniższego kodu:

```
1  clc
2  close all
3  clear all
4
5  for i=1:25
6  filename=['gauss/gauss' num2str(i) '.wav'];
7  [y,Fs]=wavread(filename);
8  T=(length(y)/Fs); Y=fft(y); w=(0:length(Y)-1)/T;
9  W=w(1:(length(w)/2));
10 Y=abs(Y(1:(length(Y)/2))).^2;
11
12 a(i)=(sum(W*Y)/sum(Y));
13 odch(i)=sqrt(sum((abs(W-a(i))).^2)*Y)/sum(Y);
14 aio(i)=a(i)-0.1*odch(i);
15 end
16 for i=1:25
17 [value(i) diff(i)]=min(abs(a-aio(i)));
18 diffx(i)=diff(i)-i;
19 end
```

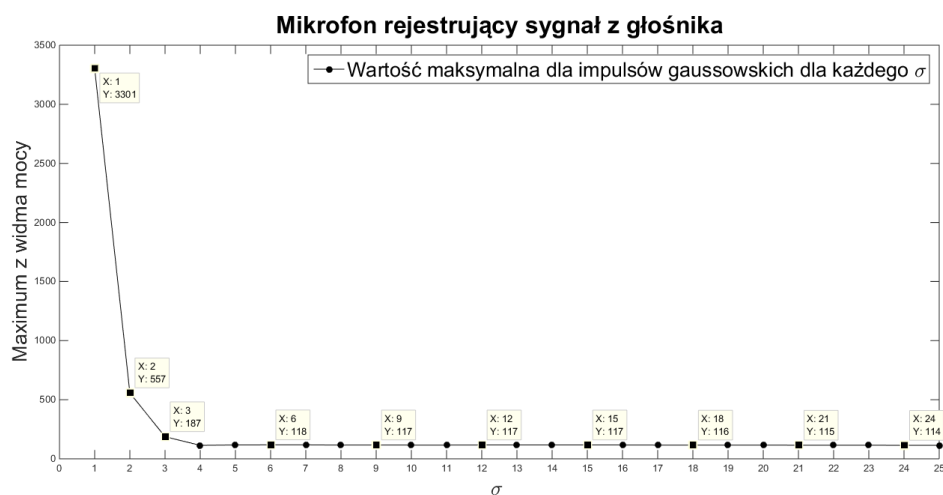
9.3.1 Mikrofon rejestrujący sygnał z głośnika

Na wykresie 47 widzimy wartości maksymalne widm otrzymanych przez nagranie impulsów gaussowskich przez mikrofon G.R.A.S 46AE [106] przedstawiony na zdjęciu 46.

Powyższe maksima zostały wyznaczone na podstawie otrzymanych widm mocy, które zobaczyć można na rys. 49. Wykazują one szereg maksimów a w tabeli 4 można zobaczyć jak kształtują się charakterystyczne piki. Widzimy, że dla dźwięków krótkich wartości pików w widmach są bardzo zbliżone, aż do momentu kiedy dźwięk jest wystarczająco długi i inne składowe nie wybijają się tak wyraźnie na pierwszy plan (od $n = 9$ czyli $\sigma = 0,204ms$ ku dłuższym dźwiękom).



Rysunek 46: Mikrofon G.R.A.S 46AE użyty podczas badań w komorze bezchewej.



Rysunek 47: Maksymalna wartość dla każdego σ z zakresu 1 – 25 zarejestrowana przez mikrofon.

Numer σ	I pik [Hz]	II pik [Hz]	III pik [Hz]
1	125	550	871
2	135	557	881
3	150	554	873
4	114	545	870
5	117	550	871
6	118	545	870
7	118	547	870
8	118	546	814
9	116	-	-
10	126	-	-

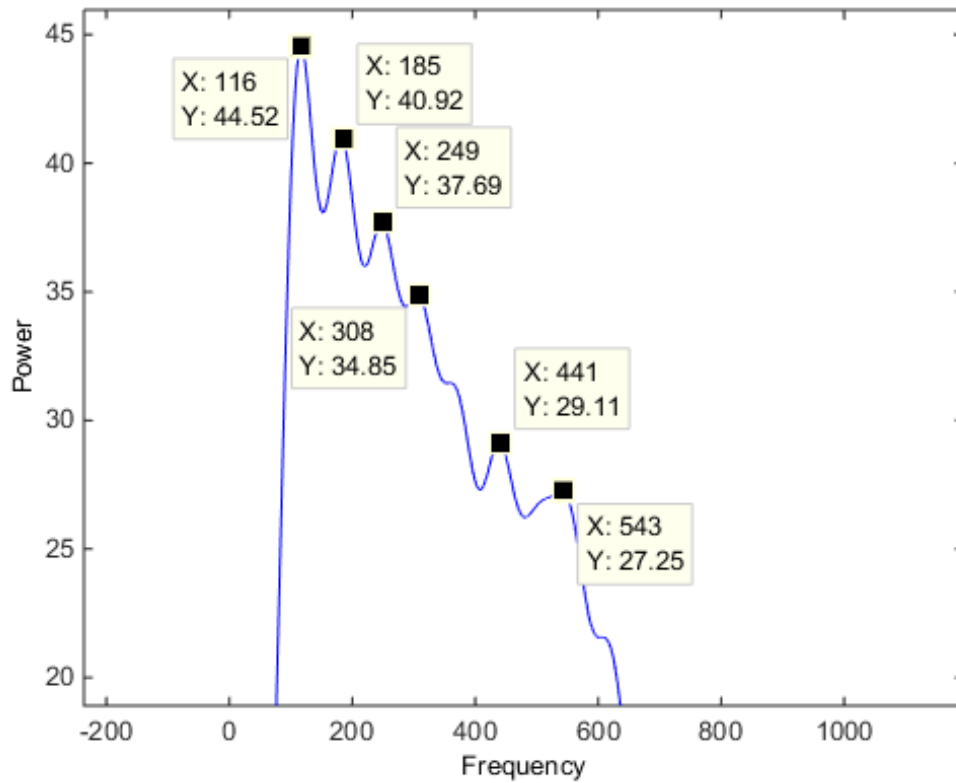
Tablica 4: Tabela przedstawiająca pierwsze 3 wyraźne maksima dla dźwięków o podanych σ zmierzonych za pomocą mikrofonu rejestrującego sygnał z głośnika.

Dla dźwięku nr 9 o $\sigma = 0,204ms$ nadal obserwujemy charakterystyczny „grzebień”, jednak nie wydaje się on już tak bardzo dominować jak można zobaczyć na przybliżonym widmie dla tego dźwięku na rys. 48. Widać, że wartości wymienione w tabeli nadal są wyraźnie widoczne.

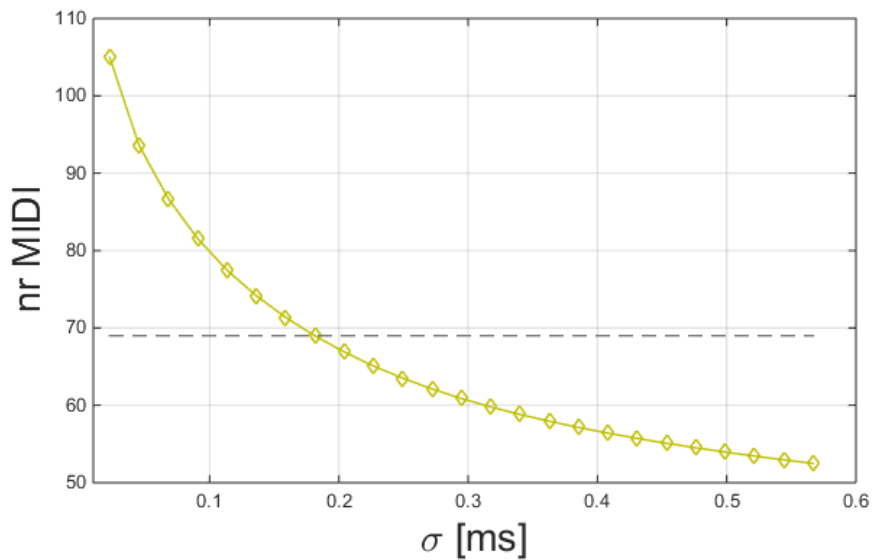
9.3.2 Analiza reprezentacji wrażenia wysokości dźwięku.

Zależność wartości oczekiwanej $a(\sigma)$ (zob wzór 66) z widm otrzymanych za pomocą sztucznej głowy od szerokości impulsu sigma są przedstawione na rys. 50a. Przebieg krzywej przypomina funkcję $\frac{1}{\sigma}$, ale aby się przekonać o stopniu tego podobieństwa na rys. 50b przedstawiono iloczyn $\omega_{sr} \cdot \sigma$. Widać, że iloczyn ten jest bliski stałej i zmienia się w zakresie niewiele większym niż 15% (79,59 – 96,09), zatem odstępstwo od funkcji $\frac{1}{\sigma}$ jest niewielkie.

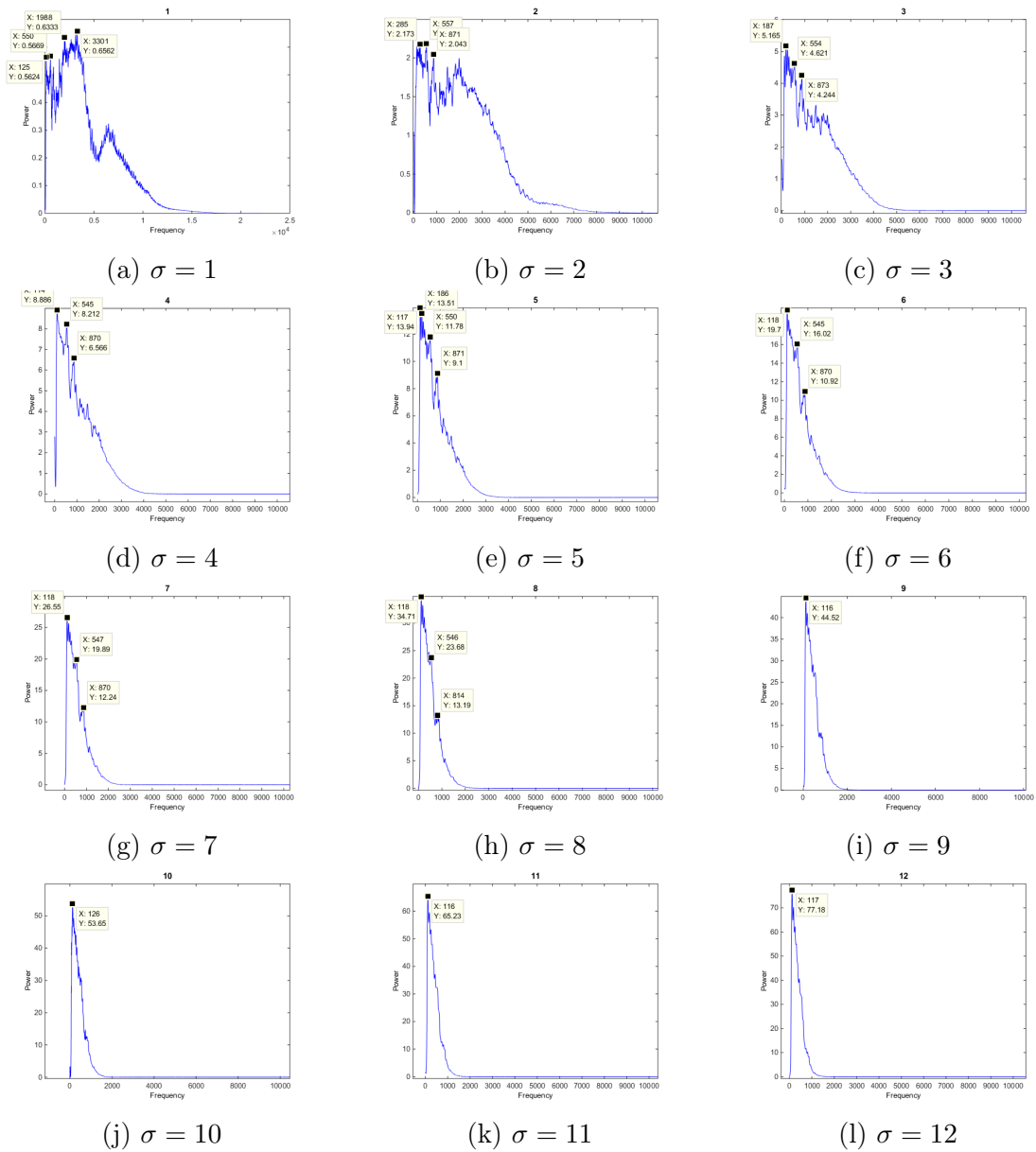
Wartości oczekiwane częstotliwości wyznaczone dla sztucznej głowy (rys. 50a) nietrudno przeliczyć na wysokość dźwięku wg wzoru (10). Przedstawiono to na rys 54. Widać, że te wysokości dźwięku leżą w zakresie bliskim używanego w muzyce instrumentalnej i wokalne. Jest to odpowiednik efektywnej wysokości dźwięku (zob. [78] i rysunek 8).



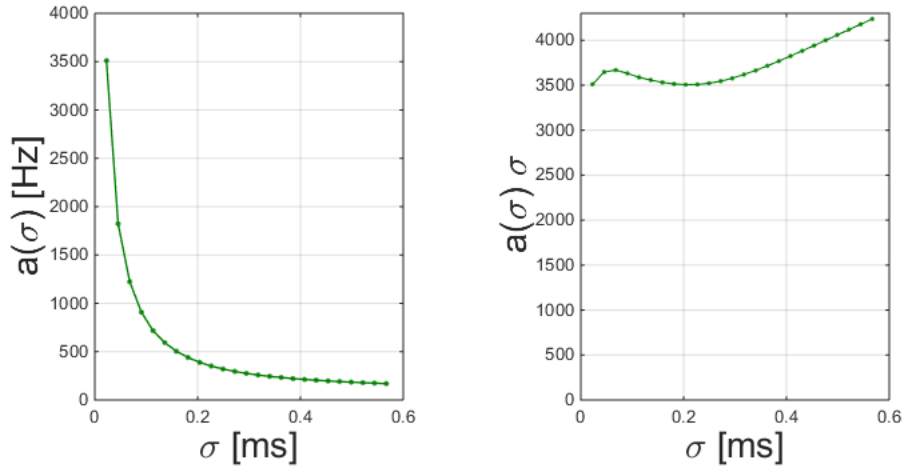
Rysunek 48: Przybliżenie (zoom) na widmo mocy dźwięku $n = 9$, ($\sigma = 0,204ms$).



Rysunek 51: Wysokość dźwięku w skali MIDI obliczona z wartości oczekiwanej częstotliwości widm otrzymanych dla sztucznej głowy jako funkcja czasu trwania impulsu σ . Linia przerywana odpowiada wartości 69, tj. klawiszowi a^1 i razkreślne (a^1).



Rysunek 49: Widma impulsów gaussowskich otrzymane z mikrofonu rejestrującego krótkie impulsy z głośnika dla σ z zakresu od 1 do 12.



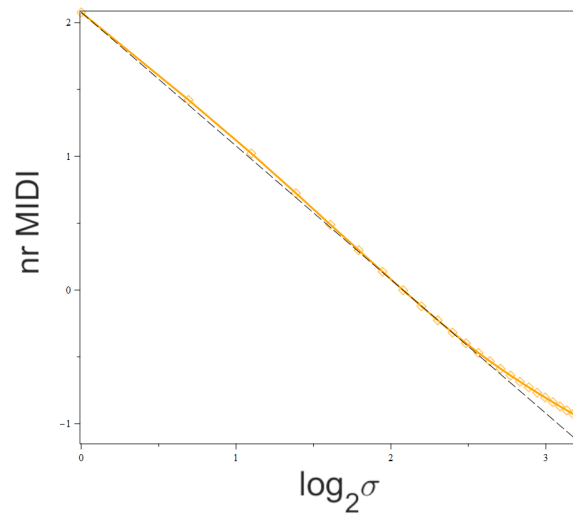
(a) Wartość oczekiwana $a(\sigma)$ (wzór (66)) częstości jako funkcja szerokości impulsu obliczona z widm zarejestrowanych przez sztuczną głowę. (b) Ta sama wartość oczekiwana pomnożona przez szerokość impulsu σ .

Rysunek 50: Wartość oczekiwana częstości.

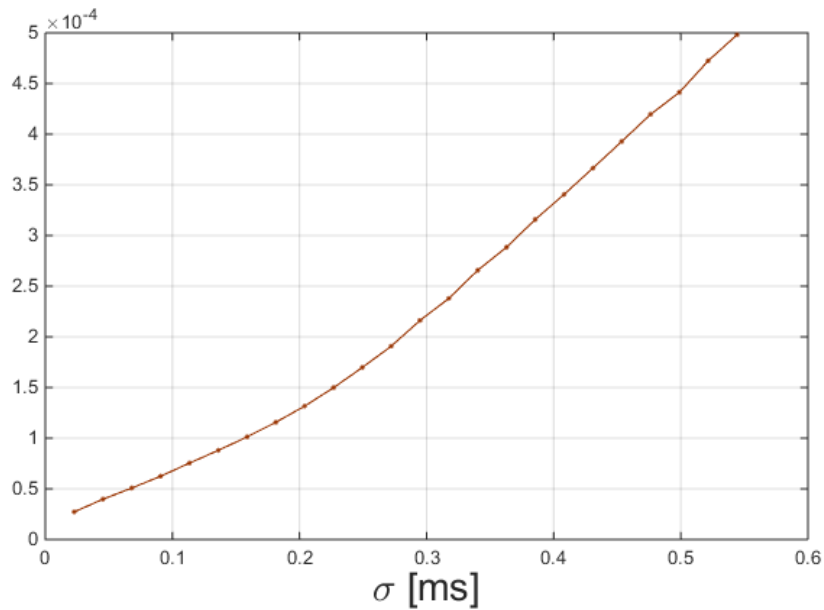
Jeszcze jedną wizualizacją potęgowego prawa wykrytego dla wartość oczekiwanej widm impulsów w funkcji czasu trwania jest wykres podwójnie logarytmiczny zależności z rysunku 50a. Przedstawiono ją na rysunku 52. Jak widać, współczynnik kierunkowy tej prostej jest bardzo bliski jedności, co wskazuje, że wykładnik $\frac{1}{A\gamma a_p} \approx 1$ we wzorze (88). Obliczając pochodną numeryczną z wykresu 50a i podstawiając do wzoru (54) otrzymujemy krzywą widoczną na rys. 54. Odpowiada ona najmniejszej zauważalnej różnicy wysokości dźwięku z dokładnością do czynnika dp_{lim} ponieważ sztuczna głowa nie może dostarczyć informacji nt. progu rozdzielczości wysokości dźwięku. Przebieg tak otrzymanego prawa psychofizycznego jest rosnący, co przypomina prawo Webera-Fechnera, lecz składa się z dwóch segmentów prostoliniowych (rys. 53). Na rys. 54 widzimy dopasowanie tych dwóch segmentów. Zwraca uwagę znacznie niższy próg $\Delta\sigma_{DL}(\sigma \rightarrow 0)$ w stosunku do nachylenia pierwszego odcinka prostoliniowego niż to wykazały badania behawioralne. Przeciwna też jest tendencja przy zwiększaniu σ : w przypadku sztucznej głowy nachylenie wzrasta, podczas, gdy dla badanych osobników maleje.

Na wykresie 53 widać linię dwu-segmentową lub prawo potęgowe. Wykres przedstawiony za pomocą skali podwójnie logarytmicznej z dopasowaną prostą

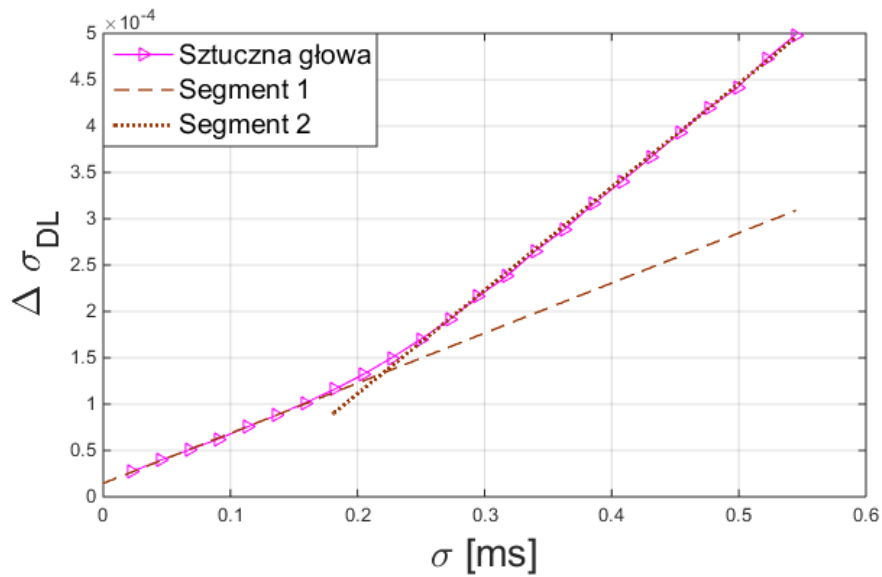
widzimy na rys. 52.



Rysunek 52: Podwójnie logarytmiczny wykres zależności przedstawionej na rysunku 50a. Dopasowana prosta ma wzór: $-0,99(0,01) \cdot x + 2,08(0,01)$. Nachylenie bardzo bliskie jedności wskazuje na proste prawo potęgowe z wykładnikiem bliskim -1 .

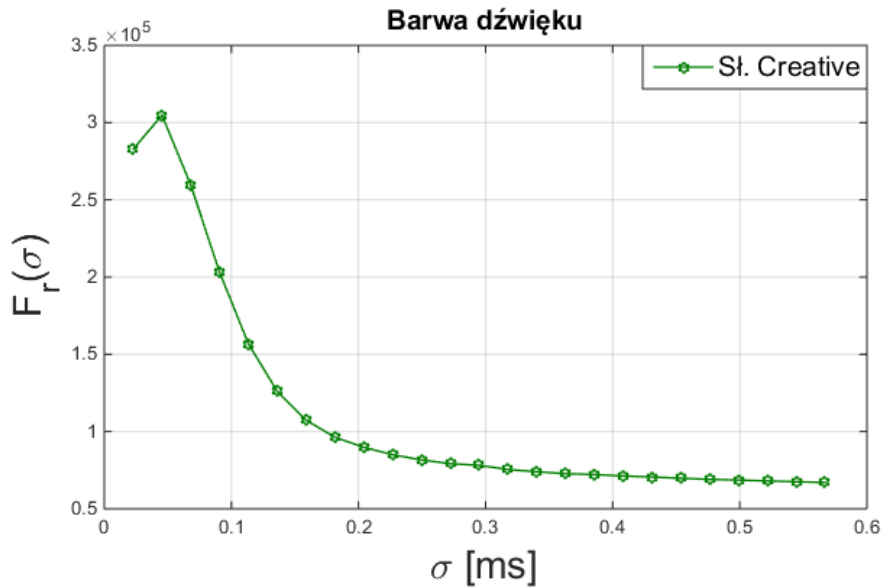


Rysunek 53: Wykres przedstawiający prawo Webera-Fechnera dla sztucznej głowy bez czynnika dp_{lim} .

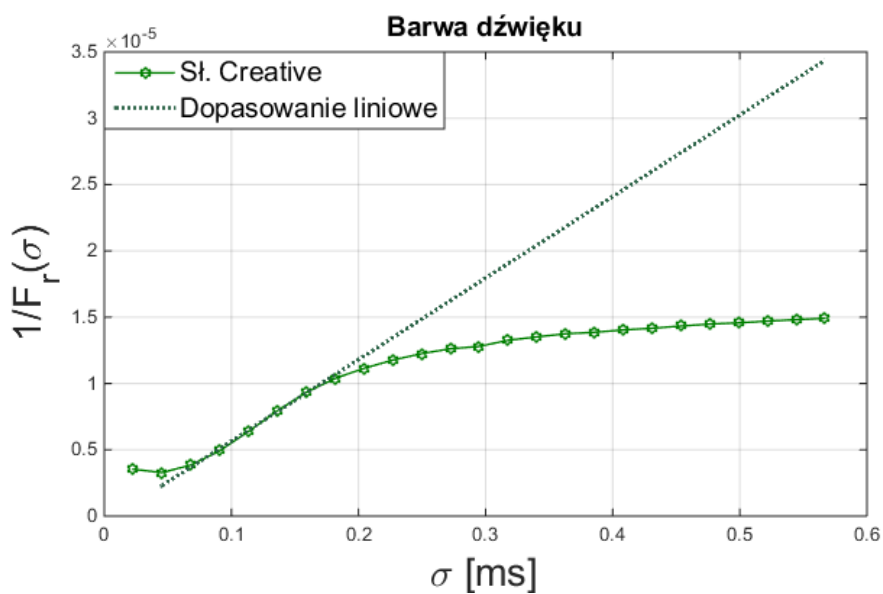


Rysunek 54: Prawo psychofizyczne z dokładnością do dp_{lim} otrzymane z danych z rys. 50a wg wzoru (54), wraz z dopasowaniami liniowymi dla dwóch segmentów. Segment dla małych σ : $\Delta\sigma_{DL} = 0,000532(6,7 \cdot 10^{-6})\sigma + 0,0000151(6,0 \cdot 10^{-7})(ms)$, segment dla $\sigma > 0,25ms$: $\Delta\sigma_{DL} = 0,00111(0,00001)\sigma - 0,000112(4 \cdot 10^{-6})(ms)$.

Zarejestrowane przez sztuczną głowę widma impulsów pozwalają obliczyć numerycznie szybkość zmiany barwy dźwięku wg wzoru (67). Wynik tej procedury jest przedstawiony na rys. 55. Postępując w sposób wyjaśniony przy wyprowadzaniu wzoru (62) możemy otrzymać zależność najmniejszej zauważalnej różnicy czasu trwania impulsu od samego czasu trwania z dokładnością do progu rozróżnialności barwy $\Delta T_{lim(\sigma)}$. Nie mamy sposobu wyznaczenia tego czynnika dla sztucznej głowy. Na rysunku 56 przyjęto, że jest on stały i wynosi $\Delta T_{lim(\sigma)} = 1$. Jest to zatem analog wykresów 30 - 32 otrzymanych dla muzyków rozróżniających impulsy na podstawie barwy dźwięku. Zauważamy jakościowe podobieństwo tych wykresów. Linia prosta obrazuje dopasowanie do części liniowej zależności otrzymanej dla sztucznej głowy. Dopasowanie to ma postać $y = 61,5(0,02)\sigma - 0,49(0,3)$, co prawie idealnie odpowiada prawu Webera-Fechnera.



Rysunek 55: $F_r(\sigma)$ obliczone dla słuchawek Creative wg wzoru (67).



Rysunek 56: Zależność najmniejszej zauważalnej różnicy czasu trwania czasu impulsu $\Delta\sigma_{DL}(\sigma)$ od czasu trwania σ otrzymana na podstawie wzoru (67) i (62), z dokładnością do czynnika $\Delta T_{lim}(\sigma)$ przyjętego za stały, otrzymane z widm mocy zarejestrowanych przez sztuczną głowę przy użyciu tych samych słuchawek co w badaniach behawioralnych (Creative).

9.3.3 Efektywna wysokość dźwięku krótkich impulsów a rozróżnialność czasu ich trwania

Na zakończenie tego rozdziału podamy kilka wariantów związków jakie mogą łączyć matematyczną postać zależności najmniejszej zauważalnej różnicy czasu trwania impulsu od samego czasu trwania tj. funkcję $\Delta\sigma_{DL}(\sigma)$, z efektywną wysokością dźwięku $p(\sigma)$ postrzeganą dla takiego impulsu, zakładając, że rozróżnianie czasów trwania następuje właśnie na podstawie różnic efektywnej wysokości dźwięku.

Ogólnie mamy więc dwie funkcje tego samego argumentu σ : $a(\sigma)$ i $\Delta\sigma_{DL}(\sigma)$. Idealnym źródłem informacji byłyby dwa różne badania behawioralne dotyczące niezależnie obydwu tych wielkości dla tych samych osobników i/lub dla grup (kohort) osobników. Oczywiście wartości otrzymane w nawet takim idealnym eksperymencie mogą być zależne od metody uzyskiwania danych. Obecnie dysponujemy danymi nt. funkcji $\Delta\sigma_{DL}(\sigma)$ otrzymanej jak to opisano w rozdziale 8 oraz dość fragmentaryczne wyniki postrzegania efektywnej wysokości dźwięku (opisane w [3]). Nawet w samej definicji wysokości dźwięku dla sygnałów sinusoidalnych i/lub periodycznych istnieją kontrowersje. W tym kontekście należy porównać ściśle logarytmiczną skalę przyjmowaną przez muzyków i producentów instrumentów muzycznych (np. numeracja klawiszy w systemie MIDI (równanie (10)), choć i tu zdarzają się odstępstwa, np. wynikająca z nieharmoniczności strun fortepianu krzywa Railsbacka [39] ze skalą melową [107, 108, 109] oraz skalą barkową [110]. Rozważmy więc obydwie funkcje, aby prześledzić założenia czynione w poszczególnych krokach rozumowania:

$$\Delta\sigma_{DL}(\sigma) = f(\sigma), \quad (70)$$

$$p(\sigma) = g(\sigma). \quad (71)$$

Przyjmując, że zarówno najmniejsza zauważalna różnica bodźca $\Delta\sigma_{DL}$, jak i próg rozróżnialności wrażenia dp_{lim} są na tyle małe, że ich iloraz można przybliżyć pochodną otrzymujemy

$$\frac{dp}{d\sigma} = -\frac{dp_{lim}}{\Delta\sigma_{DL}} = g'(\sigma), \quad (72)$$

zatem

$$\Delta\sigma_{DL}(\sigma) = -dp_{lim}(\sigma)g'(\sigma)^{-1}. \quad (73)$$

Znak „-” w równaniach (72) i (73) pochodzi z obserwacji i modeli, które wskazują, że dp_{lim} i $\Delta\sigma_{DL}$ są przeciwnych znaków: wzrost czasu trwania sygnału powoduje zmniejszenie wysokości dźwięku i na odwrót.

Mamy zatem równanie różniczkowe

$$g'(\sigma) = -\frac{dp_{lim}(\sigma)}{f(\sigma)}. \quad (74)$$

Warunkiem jego rozwiązalności jest znajomość funkcji $f(\sigma)$ i $dp_{lim}(\sigma)$. W przypadku $dp_{lim}(\sigma) = const = dp_{lim}$ oraz $f(\sigma) \sim \sigma$ otrzymujemy prawo Webera-Fechnera.

W obecnej rozprawie przyjmujemy, że wrażenie efektywnej wysokości dźwięku powstaje poprzez pośrednią wielkość $\omega(\sigma) = 2\pi f(\sigma)$, tj. pewną częstość związaną z widmem odbieranego sygnału. W przypadku impulsów czysto gaussowskich wybór tej wielkości nie jest wcale oczywisty i z trudem daje się pogodzić z klasyfikacją widmowej wysokości dźwięku z pracy A. de Cheveigné [45].

Kiedy funkcja $\omega(\sigma)$ jest znana równanie różniczkowe (74) przechodzi w następujące równanie

$$\frac{dp}{d\omega} \frac{d\omega}{d\sigma} = -\frac{dp_{lim}(\sigma)}{f(\sigma)}, \quad (75)$$

Potrzebna jest wiedza zarówno o funkcji $\omega(\sigma)$ jak i w jaki sposób przekłada się ona na wysokość dźwięku $p(\omega)$. We wzorach (49-51) widzieliśmy czysto logarytmiczną zależność

$$p(\omega) = -A \ln(\omega^\gamma), \quad (76)$$

ale można też przyjąć zależność potęgową, którą implikuje prawo Stevensa, jako alternatywa do prawa Webera-Fechnera (zob. rozdz. 3.2.2 i [32, 111])

$$p(\omega) = -k\omega^n. \quad (77)$$

W przypadkach (76) i (77) otrzymujemy równania różniczkowe na niewiadomą funkcję $\omega(\sigma)$

Symbol	Funkcja $f(\sigma)$	Rozwiązanie równania (78)
a0	$f(\sigma) = a_p \sigma$	$\omega(\sigma) = C \cdot \sigma^{-\frac{dp_{lim}}{\gamma A a_p}}$
a1	$f(\sigma) = a_p r$	$\omega(\sigma) = C \cdot \exp\left(\frac{-\sigma \cdot dp_{lim}}{\gamma A \cdot a_p}\right)$
a2	$f(\sigma) = a_p(r + \sigma)$	$\omega(\sigma) = C \cdot (\sigma + r)^{\left(\frac{-1 \cdot dp_{lim}}{\gamma A \cdot a_p}\right)}$
a3	$f(\sigma) = a_p(r + \sigma)^{1+\beta}$	$\omega(\sigma) = C \cdot \exp\left(\frac{(r+\sigma)^{-\beta} \cdot dp_{lim}}{\gamma A a_p \beta}\right)$
a4	$f(\sigma) = a_p(r + \sigma^{1+\beta})$	$\omega(\sigma) = C \cdot \exp\left(-\frac{dp_{lim} \sigma \operatorname{LerchPhi}\left(-\frac{\sigma^{1+\beta}}{r}, 1, \frac{1}{1+\beta}\right)}{\gamma A a_p r^{1+\beta}}\right)$

Tablica 5: Przykłady rozwiązań równań dla funkcji $f(\sigma)$ sugerowanych przez wyniki obecnych badań behawioralnych

Symbol	Funkcja $f(\sigma)$	Rozwiązanie równania (79)
b0	$f(\sigma) = a_p \sigma$	$\omega(\sigma) = \left(-\frac{C \cdot k \cdot a_p + dp_{lim} \ln(\sigma)}{k \cdot a_p}\right)^{\frac{1}{n}}$
b1	$f(\sigma) = a_p r$	$\omega(\sigma) = r^{-\frac{1}{n}} \left(-\frac{C \cdot k a_p r - dp_{lim} \sigma}{k a_p}\right)^{\frac{1}{n}}$
b2	$f(\sigma) = a_p(r + \sigma)$	$\omega(\sigma) = \left(-\frac{C \cdot k a_p - dp_{lim} \ln(r + \sigma)}{k a_p}\right)^{\frac{1}{n}}$
b3	$f(\sigma) = a_p(r + \sigma)^{1+\beta}$	$\omega(\sigma) = \left(-\frac{C \beta \cdot k a_p + (r + \sigma)^{-\beta} dp_{lim}}{\beta k a_p}\right)^{\frac{1}{n}}$
b4	$f(\sigma) = a_p(r + \sigma^{1+\beta})$	$\omega(\sigma) = r^{-\frac{1}{n}} \left(\frac{1}{k a_p (1 + \beta)} H^{\frac{1}{n}}\right)$ $H = \left(-dp_{lim} \sigma \operatorname{LerchPhi}\left(-\frac{\sigma^{1+\beta}}{r}, 1, \frac{1}{1+\beta}\right) + C \cdot k a_p \cdot r(1 + \beta)\right)$

Tablica 6: Przykłady rozwiązań równań dla funkcji $f(\sigma)$ sugerowanych przez wyniki obecnych badań behawioralnych

$$\frac{1}{\omega} \frac{d\omega}{d\sigma} = -\frac{dp_{lim}(\sigma)}{A \gamma f(\sigma)} \quad (78)$$

$$\omega^{n-1} \frac{d\omega}{d\sigma} = -\frac{dp_{lim}(\sigma)}{n k f(\sigma)} \quad (79)$$

W tabeli 5, przedstawiono kilka przykładów rozwiązań tych równań dla funkcji $f(\sigma)$ sugerowanych przez wyniki obecnych badań behawioralnych. Zakładamy przy tym stałą wartość dp_{lim} . Badania behawioralne sugerują postać funkcji $f(\sigma) \Delta \sigma_{DL}$, która w granicy $\sigma \rightarrow 0$ dąży do pewnej stałej dodatniej wartości r .

10 Porównanie wyników badań behawioralnych oraz badań w LAT

10.1 Fizyczne korelaty wrażenia wysokości dźwięku na podstawie porównania wyników testów behawioralnych i parametrów widm uzyskanych za pomocą sztucznej głowy.

Jak to przedstawiono w rozdz. 7.2 najmniejsza zauważalna różnica czasu trwania impulsu jest w przybliżeniu liniową funkcją samego czasu trwania impulsu,

$$\Delta\sigma_{DL} = -dp_{lim}a_p(\sigma + r) = -\alpha(\sigma + r) \quad (80)$$

Zakładamy tu stały, niezależny od wysokości dźwięku próg rozróżnialności wysokości dp_{lim} , który występuje w oryginalnej wersji prawa Webera-Fechnera. W przypadku wysokości dźwięku jest to najmniejszy wykrywalny słuchem interwał dźwięków podawanych po kolei [112, 113] (discrimination, w odróżnieniu od selectivity, tj. rozróżniania dźwięków podawanych jednocześnie). Zakładając, że rozróżnienie impulsów następuje na podstawie wrażenia przypisywanej im wysokości dźwięku oraz, że podobnie jak dla dłuższych sygnałów, wysokość dźwięku związana jest z pewną częstotliwością (kołową) musimy uznać, że częstotać ta jest funkcją czasu trwania impulsu

$$\omega = \omega(\sigma) \quad (81)$$

Przyjmując, że uogólnione prawo Webera-Fechnera (80) obowiązuje dla infinitezmalnych różnic $\Delta\sigma_{DL}$ i dp_{lim} , możemy je zapisać w postaci różniczkowej

$$\frac{\partial p}{\partial \sigma} = -\frac{1}{\alpha(\sigma + r)} \quad (82)$$

Używając zależności (81) otrzymujemy równanie

$$\frac{\partial p}{\partial \omega} \frac{\partial \omega}{\partial \sigma} = -\frac{1}{\alpha(\sigma + r)} \quad (83)$$

Zależność $p(\omega)$ w przypadku dłuższych sygnałów ma charakter logarytmiczny (np. numery klawiszy w systemie MIDI). Stosując używaną często jednostkę wysokości dźwięku cent, zależność tę można zapisać

$$p(\omega) = A \ln(\omega), \quad (84)$$

gdzie

$$A = \frac{1200}{\ln(2)} \text{centów}, \quad (85)$$

tak, że podwojenie częstości podwyższa wysokość dźwięku o 1200 centów, tj. o dwa-
naście półtonów równomiernie temperowanych, czyli o oktawę. Trzeba pamiętać,
że czynnik A ma charakter umowny i przyjęty jest dla wygody stroicieli i innych
pracujących w dziedzinie praktycznej akustyki. Warto zauważyć, że logarytmiczna
zależność (84) będzie także spełniona gdy logarytm będzie wzięty z pewnej potęgi
częstości

$$p(\omega) = A \ln(\omega^\gamma) = \gamma A \ln(\omega). \quad (86)$$

Zmieni się tylko czynnik A na γA . Oczywiście tradycyjna zależność (84) odpo-
wiada $\gamma = 1$. Ta niezależność relacji między częstością a wysokością dźwięku od
wykładnika γ nie była chyba dotąd szczegółowo badana. Być może należy ją wiązać
z różną zdolnością rozróżniania wysokości dźwięku u różnych osobników.

Wstawienie zależności (86) do związku (83) prowadzi do równania różnicz-
kowego

$$\frac{1}{\omega} \frac{\partial \omega}{\partial \sigma} = - \frac{dp_{lim}}{A \gamma a_p (\sigma + r)}, \quad (87)$$

które ma rozwiązanie (zob. Tablica 5, a3)

$$\omega(\sigma) = (\sigma + r)^{-\frac{1}{A \gamma a_p}} \quad (88)$$

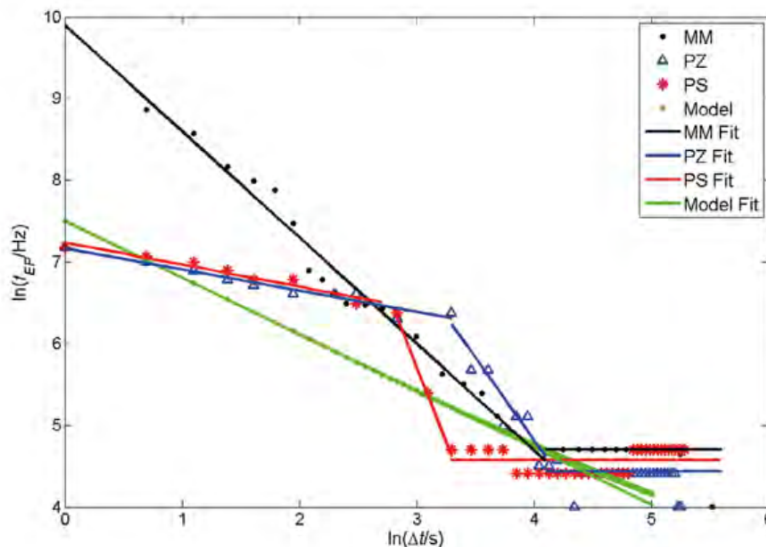
Mamy więc potęgową zależność częstości odpowiedzialnej za wysokość dźwięku
od czasu σ trwania impulsu. Jak widać wykładnik γ także się tu pojawia. Zatem
można rozważać dwa odrębne zagadnienia

- potęgową zależność dopowiedzianej za wysokość dźwięku częstości długich
sygnałów

- zależność takiej częstości wyodrębnionej z sygnałów bardzo krótkich, w których nie można wyróżnić obwiedni i sygnału okresowego od czasu trwania impulsu.

W obecnej rozprawie zajmujemy się tym drugim problemem, pamiętając, że właściwie jest on potencjalnie związany z pierwszym. Pamiętając, że współczynnik a_p pochodzi z testu behawioralnego oraz stała A ma charakter umowny, ciekawe byłoby porównanie ilorazu $\frac{dp_{lim}}{\gamma}$ dla różnych osobników. Oczywiście najlepszą metodą jest tu eksperyment. Wskazówki na ten temat znajdujemy w [3] s. 29, gdzie niewielka liczba testowanych oceniała wysokość dźwięku postrzeganą przy percepcji sygnałów czysto gaussowskich.

▼ FIG. 2: Pitch attributed to purely Gaussian pulses (envelopes only) as a function of duration Δt .



Rysunek 57: Logarytmiczno-logarytmiczny wykres wysokości dźwięku przypisanej przez autorów pracy [3] sygnałom gaussowskim o różnym czasie trwania. Czas trwania impulsu jest tu oznaczony Δt , zaś częstotliwość f_{EP} została obliczona na podstawie zgłaszanych przez testowanych interwałów muzycznych względem ustalonego dźwięku odniesienia.

Bardzo dobrze przybliżoną linią prostą przedstawioną na wykresie log-log zależność tej wysokości od czasu trwania impulsu świadczy o tym, że potęgowa zależność typu wzoru (88) znalazła tu potwierdzenie. Warto zwrócić uwagę, że współczynnik nachylenia prostej wydaje się zależeć od osobnika. Interesującą obserwacją

są bardzo zbliżone wykresy uzyskane przez dwóch uczestników (PS i PZ) mających doświadczenie w śpiewie chóralnym. Obydwaj posiadają słuch względny, dlatego wysokość dźwięku oceniali poprzez porównanie wrażenia odnoszonego z impulsu gaussowskiego z pewnym dźwiękiem odniesienia. Według relacji obydwu testowanych ocena odbieranego interwału odbywała się niezależnie i zupełnie inaczej: PZ „odgadywał” interwał dość szybko a PS potrzebował więcej czasu i powtórzeń par dźwięk odniesienia – impuls gaussowski [Paweł Sobieszczyk, Piotr Zieliński, informacja prywatna]. Uczestnik MM, z doświadczeniem w grze na skrzypcach, nie potrzebował dźwięku odniesienia, ponieważ dysponuje słuchem absolutnym [Marcin Majka, informacja prywatna]. Czytelnik tej rozprawy może sam przeprowadzić taki test na podstawie programu przygotowanego w oprogramowaniu Matlab - dołączonego do niniejszej rozprawy. Jedynym wymaganiem jest posiadanie zainstalowanego oprogramowania MATLAB na komputerze. Rzecz jasna statystyka tych obserwacji jest zbyt mała, aby stwierdzić, że doświadczenie chóralne i słuch względny prowadzą do identycznego wykładnika, a słuch absolutny do innego. Badania tych zależności byłyby bardzo interesujące. Oczywiście ten rodzaj wysokości dźwięku różni się od wysokości dźwięku sygnałów okresowych. Dlatego lepiej nazywać to wrażenie efektywną wysokością dźwięku (effective pitch - EP).

Jak wyjaśniono w rozdz. (7.1) poważnym kandydatem na częstość wyznaczającą efektywną wysokość dźwięku jest wartość oczekiwana częstości obliczana przy użyciu dodatniej strony widma. Z tego powodu obliczono zależność wartości oczekiwanej częstości $f = \frac{\omega}{2\pi}$ [Hz] od szerokości σ impulsu używając widm mocy, zarejestrowanych za pomocą „sztucznej głowy” czyli Head and Torso Simulator (HATS), type 5128. Wynik przedstawiony jest na rys. 50a; rys. 50b przekonuje jak dobrym przybliżeniem jest zależność typu $f \propto \frac{1}{\sigma}$. Należy zwrócić uwagę, że wartości częstości na rys. 50a pokrywają zakres ok. 170 Hz – 3500 Hz, tj. zakres dźwięków często używanych w muzyce jak przedstawiono w tabeli 7.

Częstotliwość	Skala MIDI	Oznaczenie oktaw
164,81 Hz	E3	e
3520,00 Hz	A7	a ⁴

Tablica 7: Używane w muzyce nazwy dźwięków odpowiadających przedziałowi częstości przedstawionych na rys. 50a.

Jak zaznaczono w rozdziale 9.3.3, w wyniku przeprowadzonych badań behawioralnych dysponujemy doświadczalnymi wynikami dla funkcji $f(\sigma)\Delta\sigma_{DL}(\sigma)$, czyli prawem psychofizycznym, podczas gdy pomiary za pomocą sztucznej głowy dostarczyły zależności wartości oczekiwanej częstotliwości $a(\sigma) = \langle\omega\rangle$. Argumenty przedstawione w rozdziale 7.1 oraz w opisie rysunku 57 [3] wskazują, że ta ostatnia wielkość jest w uzasadniony sposób kandydatem na fizyczny parametr częstotliwościowy decydujący o wrażeniu wysokości wg równania (86).

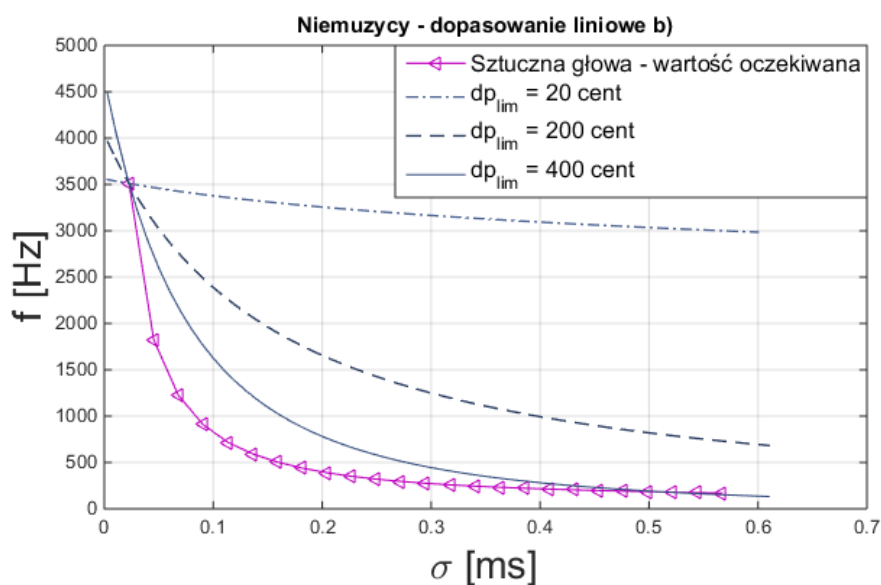
W rozdziale 9.3.3 został zaproponowany związek, jaki powinien zachodzić między $f(\sigma)\Delta\sigma_{DL}$ i $a(\sigma) = \langle\omega\rangle$ dla jednego „układu odbiorczego” przy założeniu, że próg rozróżniania wysokości dźwięku dp_{lim} jest dla tego układu takim samym interwałem muzycznym w całym zakresie rozpatrywanych częstotliwości. Pozwala to uchwycić podobieństwa i różnice pomiędzy percepcją czysto gaussowskich bardzo krótkich sygnałów przez ludzkie narządy słuchu i przez sztuczną głowę. Oczywiście próg rozróżniania wysokości dźwięku dp_{lim} dla sztucznej głowy jest nieznan, zatem pozostaje on parametrem dopasowania. Natomiast wykładnik γ z równania (86) jest bliski jedności wobec charakteru krzywej z rys. 50a. Istotna różnica pomiędzy sztuczną głową, a badanymi słuchaczami tkwi w wartości r z równania (82) i wynosi:

- $r_{SztuGł} = 0,0284$ dla sztucznej głowy
oraz
- $r_{niem} = 0,578$ dla niemuzyków
i
- $r_{muzWys} = 0,329$ dla muzyków rozróżniających impulsy na podstawie wysokości dźwięku.

Fakt, że najmniejsza zauważalna różnica czasu trwania impulsu jest niższa dla muzyków wydaje się łatwo wytłumaczalny treningiem słuchowym, ale tak duża rozbieżność pomiędzy wartościami parametru r dla urządzeń elektronicznych i istot ludzkich powinna być może przedmiotem do dalszych badań.

10.2 Próby uzgodnienia prawa psychofizycznego dla czasu trwania impulsów uzyskanego z testów behawioralnych z analizami widm zarejestrowanych przez HATS.

Krzywa z rys. (53) została wykreślona dla $dp_{lim} = 1$. Rysunek 58 przedstawia wartość oczekiwaną częstości uzyskaną dla sztucznej głowy (znaną z rys. 50a) oraz, dla porównania, analogiczną wartość oczekiwaną częstości obliczoną z równania 88 dla różnych wartości dp_{lim} . Wartości dp_{lim} 20, 200 i 400 centów są



Rysunek 58: Próba odtworzenia częstości odpowiadającej efektywnej wysokości dźwięku na podstawie linowego dopasowania b) do danych otrzymanych przez niemuzyków wg równania (88).

równe odpowiednio $\frac{1}{5}$ półtonu temperowanego, sekundzie wielkiej temperowanej i tercji wielkiej temperowanej. Krzywe na rysunku 58 są dobrane tak, aby przechodziły przez punkt $\frac{1000,0}{44100} = 0,0226\dots$ i $f = 3510,08\text{Hz}$, tzn. początkowy punkt danych ze sztucznej głowy (zob. rys. 50a). Aby uzyskać choćby jakościowe podobieństwo, np. zakres przebieganych częstotliwości musimy założyć dość duży próg rozróżniania interwałów dp_{lim} . Byłoby to niezgodne z obserwacjami dla dźwięków o określonej periodyczności, ale jest być może uzasadnione dla efektywnej wysokości dźwięku zapostulowanej dla naszych krótkich impulsów. Głównym powodem

trudności uzgodnienia wyników widocznych między badaniami behawioralnymi i ze sztucznej głowy jest rozbieżność w wartości parametru r .

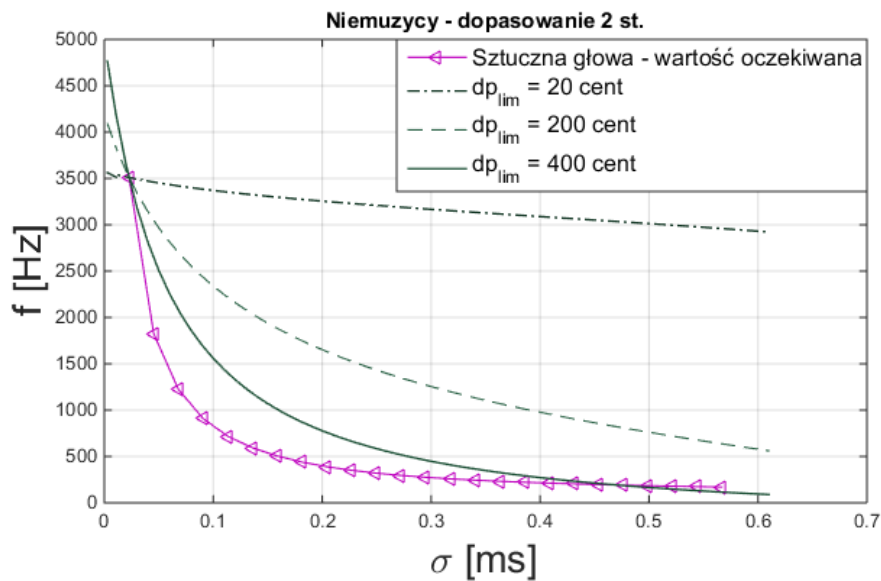
Dopasowanie drugiego stopnia dla najmniejszej zauważalnej różnicy czasu trwania impulsu dla niemuzyków, zob. (rys. 36), prowadzi do równania różniczkowego

$$\frac{1}{\omega} \frac{d\omega}{d\sigma} = - \frac{dp_{lim}}{A\gamma} \frac{1}{-a\sigma^2 + b\sigma + c}, \quad (89)$$

którego rozwiązaniem jest funkcja

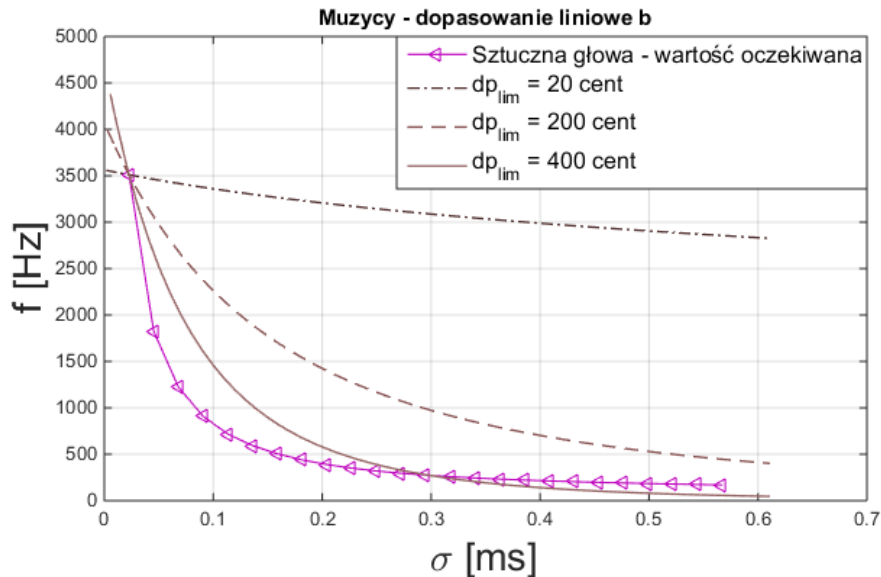
$$\omega(\sigma) = C \cdot \exp\left(\frac{-2dp_{lim} \arctang\left(\frac{2a\sigma - b}{\sqrt{b^2 + 4ac}}\right)}{A\gamma\sqrt{b^2 + 4ac}}\right), \quad (90)$$

gdzie C jest dowolną stałą całkowania. Na rysunku 59 przedstawiono porównanie częstotliwości obliczonych ze wzoru (90) dla trzech wartości progu rozróżnialności wysokości dźwięku dp_{lim} przy użyciu wyników testów behawioralnych niemuzyków z wartościami oczekiwanymi częstotliwości dla sztucznej głowy (zob. rys 50a). Dopasowanie drugiego stopnia nieznacznie zwiększa zakres częstotliwości obejmowanych dla poszczególnych wartości dp_{lim} . Analogiczne porównania wyników testów



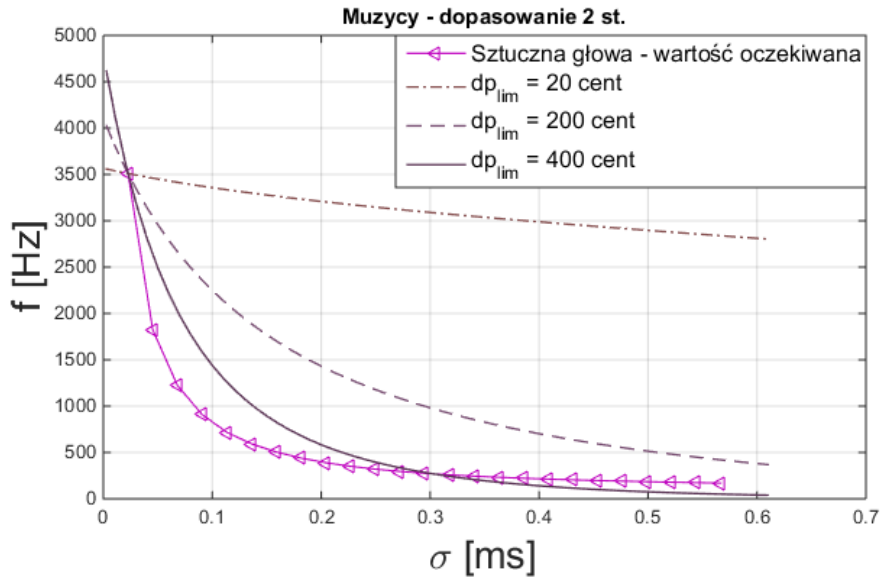
Rysunek 59: Próba odtworzenia częstotliwości odpowiadającej efektywnej wysokości dźwięku na podstawie dopasowania 2 st. do danych otrzymanych przez niemuzyków, wg równania (90).

behawioralnych muzyków rozróżniających impulsy na podstawie postrzeganej wysokości dźwięku są widoczne na rysunku 58 dla dopasowania liniowego oraz na rysunku 61 dla dopasowania drugiego stopnia. Ponieważ wartość parametru r jest w tym przypadku mniejsza, zakres częstotliwości pokrywany przy tych samych progach dp_{lim} okazuje się nieco szerszy.



Rysunek 60: Próba odtworzenia częstotliwości odpowiadającej efektywnej wysokości dźwięku na podstawie liniowego dopasowania b) do danych otrzymanych przez muzyków na podstawie różnic w wysokości dźwięku porównywanych impulsów wg wzoru (88).

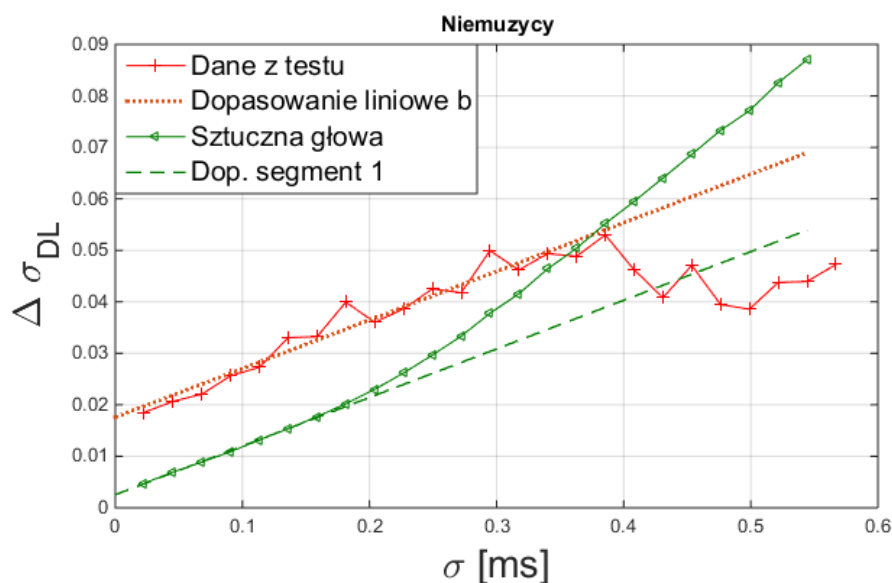
Jak widać, współczynnik kierunkowy tej prostej jest bardzo bliski jedności, co wskazuje, że wykładnik $\frac{1}{A\gamma a_p} \approx 1$ we wzorze (88). Obliczając pochodną numeryczną z wykresu 50a i podstawiając do wzoru (54) otrzymujemy krzywą widoczną na rys. 54. Odpowiada ona najmniejszej zauważalnej różnicy wysokości dźwięku z dokładnością do czynnika dp_{lim} ponieważ sztuczna głowa nie może dostarczyć informacji nt. progu rozdzielczości wysokości dźwięku. Przebieg tak otrzymanego prawa psychofizycznego jest rosnący, co przypomina prawo Webera-Fechnera, lecz składa się z dwóch segmentów prostoliniowych (rys. 53). Na rys. 54 widzimy dopasowanie tych dwóch segmentów. Zwraca uwagę znacznie niższy próg $\Delta\sigma_{DL}(\sigma \rightarrow 0)$ w stosunku do nachylenia pierwszego odcinka prostoliniowego niż to wykazały badania behawioralne. Przeciwna też jest tendencja przy zwiększaniu σ : w przypadku



Rysunek 61: Próba odtworzenia częstotliwości odpowiadającej efektywnej wysokości dźwięku na podstawie dopasowania drugiego stopnia do danych otrzymanych przez muzyków na podstawie różnic w wysokości dźwięku porównywanych impulsów wg wzoru (90)

sztucznej głowy nachylenie wzrasta, podczas, gdy dla badanych osobników maleje.

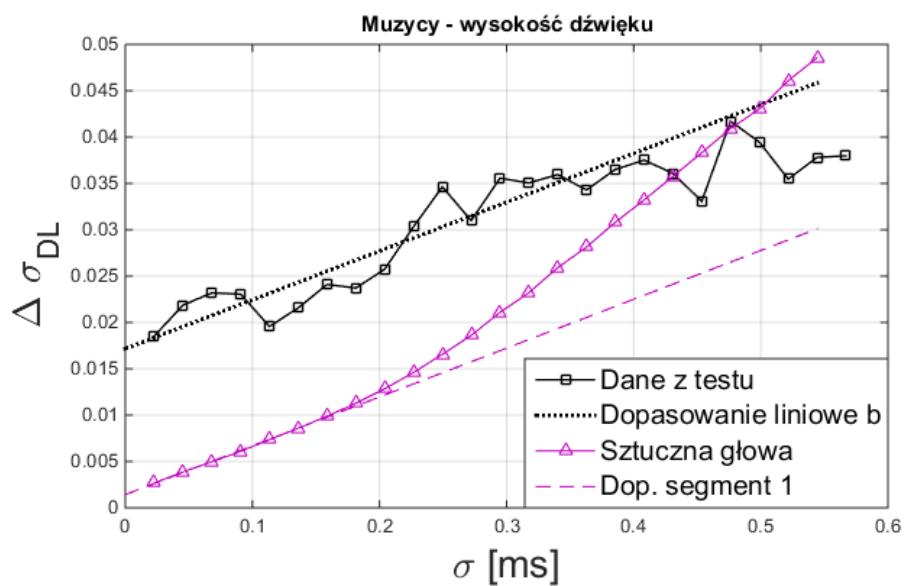
W tym podrozdziale wykorzystamy krzywą z rysunku 53 w sposób odwrotny, niż to było w poprzednim podrozdziale. Przyjmiemy mianowicie za punkt wyjścia prawo psychofizyczne (zob. wzór (70)) uzyskane z dopasowań do wyników testów behawioralnych, a hipotetyczne prawo psychofizyczne wydedukowane z widm zarejestrowanych przez sztuczną głowę (wg wzoru (54)) poddamy modyfikacjom możliwym dzięki swobodzie wyboru parametrów niemożliwych do określenia dla urządzenia mechaniczno-elektronicznego (zob. wzór (88)). Ponieważ wykładnik γ jest bliski jedności $\gamma \approx 1$, A jest ustalonym czynnikiem umownym, pozostają parametry ap , r i dp_{lim} , które dobieramy, tak, aby nachylenie liniowych dopasowań do wyników testów behawioralnych i zmodyfikowanego prawa psychofizycznego dla sztucznej głowy było takie samo. Używamy do tego dopasowania dla pierwszych 17 wartości σ dla badań behawioralnych (dopasowanie liniowe b) oraz pierwszy segment dopasowania z rys. 54. Wyniki takiej procedury w przypadku odpowiedzi niemuzyków przedstawione są na rysunku 58. Dla uzyskania jednakowych nachyleń dopasowań liniowych na rys. 6259 wartość dp_{lim} okazała się równa



Rysunek 62: Porównanie prawa psychofizycznego dla czasu trwania impulsu otrzymanego dla niemuzyków (zob. rys. 35) z wyprowadzonym na podstawie wzoru (54) analogiem tego prawa dla sztucznej głowy przy założeniu jednakowego nachylenia liniowych dopasowań do otrzymanych wyników.

ok. 174 centów, tj. niewiele mniej niż sekunda wielka temperowana. Jest to wynik bardzo rozsądny dla wrażenia wysokości dźwięku. Jednakże wartość graniczna najmniejszej zauważalnej różnicy $\Delta \sigma_{DL}(\sigma \rightarrow 0)$ dla sztucznej głowy pozostaje prawie 7 razy mniejsza niż analogiczna wartość otrzymana w badaniach behawioralnych.

Podobna procedura porównania została wykonana dla wyników testu behawioralnego muzyków zgłaszających rozróżnianie czasów trwania impulsów na podstawie wysokości dźwięku. Przedstawiono to na rysunku 63. Tym razem próg d_{lim} wyniósł ok. 97 centów, czyli mniej niż półton temperowany. Wydaje się to być wytłumaczalne większym doświadczeniem słuchowym muzyków. Jednocześnie jednak stosunek wartości granicznych $\Delta \sigma_{DL}(\sigma \rightarrow 0)$ wzrósł do ponad 12 na korzyść badań behawioralnych.



Rysunek 63: Porównanie prawa psychofizycznego dla czasu trwania impulsu otrzymanego dla muzyków stwierdzających różnice trwania impulsów na podstawie wysokości dźwięku (zob. rys. 26) z wyprowadzonym na podstawie wzoru (54) analogiem tego prawa dla sztucznej głowy przy założeniu jednakowego nachylenia liniowych dopasowań do otrzymanych wyników.

11 Podsumowanie

W rozdziałach 8 i 9 koncepcje dotyczące fizycznych mechanizmów percepcji bardzo krótkich impulsów gaussowskich sformułowane w rozdziale 7 zostały skonfrontowane z dwoma rodzajami doświadczeń: i) na żywych istotach ludzkich relacjonujących świadomie swoje odczucia oraz ii) na mechaniczno-elektronicznym modelu, którego kształty i własności lepkosprężyste mają parametry zbliżone do ciała ludzkiego, a w szczególności do ucha zewnętrznego i środkowego. W modelu tym jednak ucho wewnętrzne zastąpione jest mikrofonem. Brakuje w nim zatem całego systemu przetwarzania sygnałów i ich widm, który w żywym organizmie jest funkcją układu nerwowego. Pamiętając o tej różnicy dwóch podejść należy docenić jakościowe podobieństwo obydwu rodzajów wyników. Mająca swój początek w mechanice kwantowej hipoteza o znaczeniu wartości oczekiwanej częstotliwości w widmie mocy jako parametru częstotliwościowego związanego z wrażeniem wysokości dźwięku wydaje się znajdować tu swoje potwierdzenie. Trudniejsze w opisie ilościowym wrażenie barwy dźwięku także wykazuje podobieństwa w wynikach otrzymanych w obu rodzajach eksperymentów. Najbardziej widoczna różnica dotyczy znacznie wyższej granicy przy $\sigma \rightarrow 0$ stwierdzona dla analogów prawa Webera-Fechnera w przypadku wrażeń uczestników badań behawioralnych w porównaniu z wielkościami wydedukowanymi dla sztucznej głowy. Ta rozbieżność wydaje się ciekawym tematem badań w przyszłości. Podobnie interesujące byłoby zbadanie wrażenia efektywnej wysokości dźwięku dla większej liczby osób, niż to opisano w artykule [3]. W tej dziedzinie pewne znaczenie mogą mieć warunki akustyczne pomieszczenia, w którym odbywa się odsłuch, co też przekłada się na możliwości wykorzystania efektów skrajnie krótkich sygnałów w twórczości artystycznej. Osobnym zagadnieniem jest złożony charakter widm uzyskanych w wyniku przetwarzania przez sztuczną głowę. Widma te wykazują szereg rezonansów o nie do końca poznanym pochodzeniu. Mogą to być rezonanse zarówno sztucznej głowy, jak i pochodzące z głośników, słuchawek i mikrofonów. Pomieszczenie odsłuchu może zapewne także wytwarzać podobne rezonanse. Pole dalszych prac wydaje się więc dość szerokie. Warto zatem na koniec wymienić główne wnioski i spostrzeżenia wypływające z przedstawionych tu pomiarów i analiz. Poniżej przedstawiamy je w punktach.

Najważniejsze wyniki rozprawy:

- Wyprowadzenie analitycznych wzorów na iloczyn niepewności czasu i częstotliwości dla sygnałów cosinusoidalnych i sinusoidalnych z obwiednią w postaci funkcji Gaussa (rozkładu normalnego) dla szerokości funkcji Gaussa zmierzającej do zera (rozdział 7, rys. 11, 13 i 15).
- Wykazanie, że przy bardzo krótkich impulsach gaussowskich stwierdzona przez W. Heisenberga niezależność relacji nieokreśloności od fazy przestaje obowiązywać (rozdział 7, rys. 11, 13 i 15).
- Stwierdzenie, że wartość oczekiwana częstotliwości w widmach bardzo krótkich impulsów gaussowskich wykazuje odwrotną proporcjonalność do czasu trwania impulsu, tj. szerokości funkcji Gaussa (rozdział 7.1, wzory (30) i (34) oraz rysunki 10, 12 i 14).
- Wykazanie, że przyjęcie wartości oczekiwanej widma za częstotliwość odpowiedzialną za efektywną wysokość dźwięku prowadzi do prawa Webera-Fechnera (rozdział 7.2.1, wzory (46) i (59)).
- Wypracowanie metody ilościowego opisu barwy dźwięku i szybkości jej zmian dla widma ciągłego (rozdział 7.2.2 wzór (60)) i wykazanie obowiązywania prawa Webera-Fechnera w rozróżnianiu czasów trwania na tej podstawie pod warunkiem użycia widma amplitudowego (rozdział 7.2, wzór (63)).
- Przygotowanie testu behawioralnego na wyznaczenie najmniejszej zauważalnej różnicy czasu trwania impulsu jako funkcji tego czasu trwania, przeprowadzenie tego testu dla grupy 36 muzyków i 39 niemuzyków (rozdział 8.2) oraz wyznaczenie parametrów uogólnionego prawa Webera-Fechnera (rozdział 8.3, rysunki 22 – 37).
- Zaprojektowanie i wykonanie pomiarów widm skrajnie krótkich impulsów gaussowskich poddanych przetworzeniu przez mechaniczno-elektroniczny symulator narządu słuchu (sztuczną głowę, rozdziały 9.2 i 9.3).
- Stwierdzenie zależności bliskiej $\frac{1}{\sigma}$ dla wartości oczekiwanych częstotliwości widm uzyskanych w eksperymencie ze sztuczną głową (rozdział 9.3) , rysunek

(50) i wyznaczenie na tej podstawie analogu prawa Webera-Fechnera (9.3.2 rysunki 53 i 54).

- Wyprowadzenie analitycznych wzorów na zależność częstotliwości odpowiedzialnej za efektywną wysokość dźwięku dla wybranych postaci zależności najmniejszej zauważalnej różnicy czasu trwania od tegoż czasu trwania (rozdział 9.3.3, tablice 5 i 6).
- Oszacowanie progu rozróżnialności efektywnej wysokości dźwięku przez grupy osób uczestniczących w testach behawioralnych, aby uzyskać zakres częstotliwości pokrywany przez pomiary na sztucznej głowie (rozdział 10.2) oraz podobnych progów dla sztucznej głowy, tak aby odtworzyć pewne cechy uogólnionego prawa Webera-Fechnera uzyskanego w badaniach behawioralnych (rozdział 10).

**OŚWIADCZENIE UCZESTNIKA PROJEKTU O WYRAŻENIU ZGODY
NA PRZETWARZANIE DANYCH OSOBOWYCH**

W związku z przystąpieniem do badań pt. „Test for discrimination of duration times of Gaussian pulses” wyrażam zgodę na przetwarzanie w. w. danych osobowych.

Oświadczam, iż przyjmuję do wiadomości, że:

1. administratorem tak zebranych danych osobowych jest Instytut Fizyki Jądrowej im. H. Niewodniczańskiego Polskiej Akademii Nauk, z siedzibą przy ul. Radzikowskiego 152, 31-342 Kraków;
2. moje dane osobowe będą przetwarzane wyłącznie w celu realizacji projektu badawczego pt.: „Test for discrimination of duration times of Gaussian pulses”;
3. moje dane osobowe mogą zostać udostępnione innym podmiotom wyłącznie w celu realizacji projektu pt.: „Test for discrimination of duration times of Gaussian pulses”;
4. podanie danych jest dobrowolne, odmowa ich podania jest równoznaczna z brakiem możliwości wzięcia udziału w badaniach w ramach Projektu;
5. mam prawo dostępu do treści swoich danych i ich poprawiania.

.....
MIEJSCOWOŚĆ I DATA

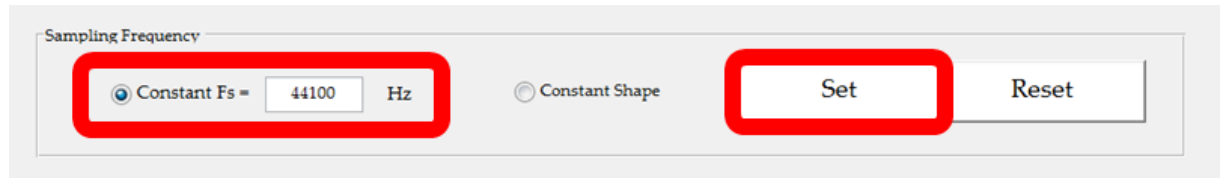
.....
CZYTELNY PODPIS UCZESTNIKA PROJEKTU

ANKIETA OSOBOWA

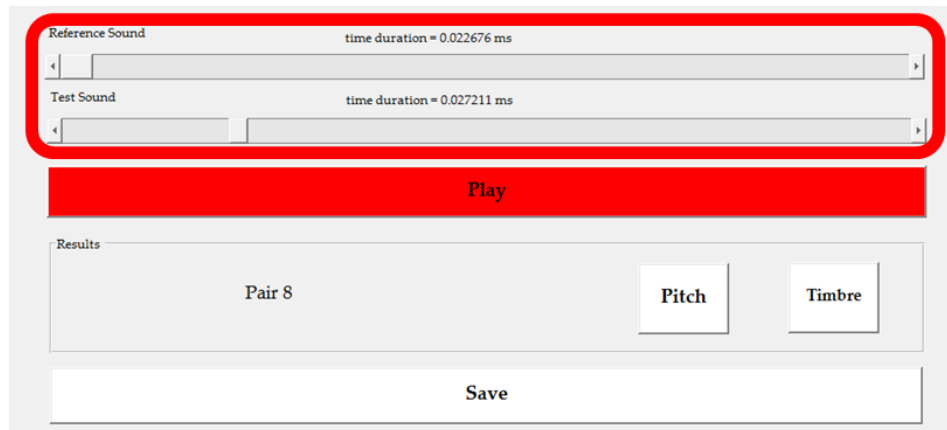
IDENTYFIKATOR	dl0001
IMIĘ I NAZWISKO	
TELEFON	
Rok urodzenia	
Płeć	
Instrument główny	

Instrukcja wykonania testu

1. Zaznacz **Constant Fs= 44100 Hz** i naciśnij **Set**

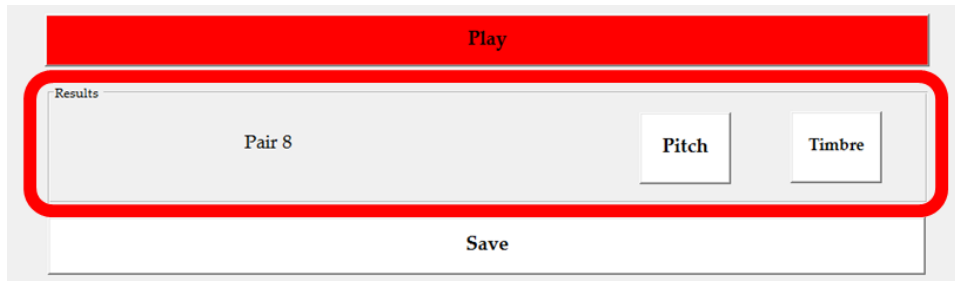


2. Test się aktywuje. Do dyspozycji są dwa suwaki. Pierwszy suwak „**Reference Sound**” służy do ustawienia dźwięku referencyjnego. Drugi suwak „**Test Sound**” służy do ustawienia dźwięku testowego.

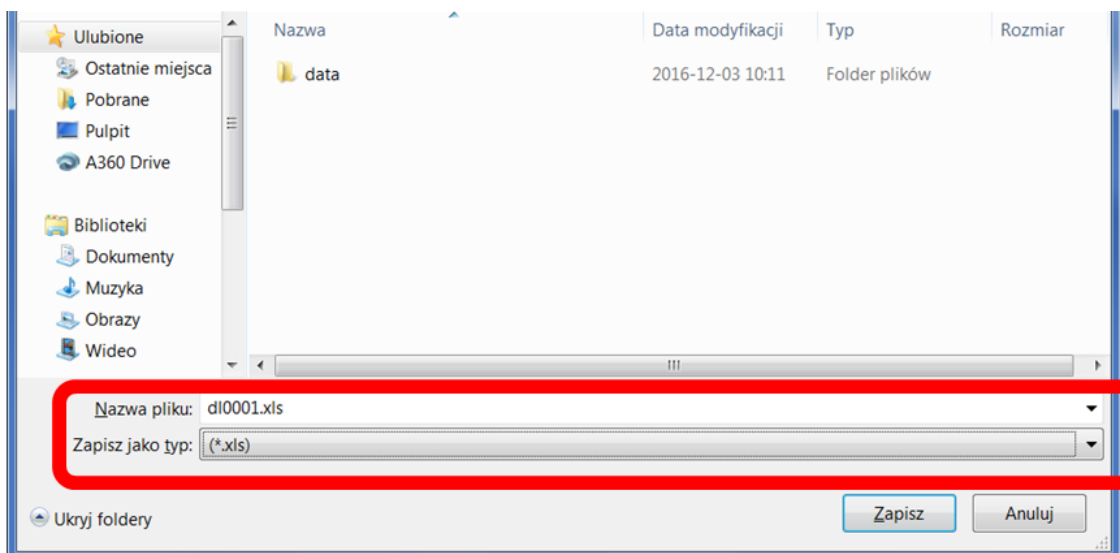


3. W teście jest 25 dźwięków referencyjnych z 40 dźwiękami testowymi. Pole „**time duration**” nad pierwszym suwakiem informuje jakiej długości jest dźwięk referencyjny, natomiast nad drugim suwakiem, o ile dłuższy jest dźwięk testowy od referencyjnego. Podczas przesuwania suwakami w polu „**Results**” pojawiać się będzie informacja o aktualnym numerze pary. W celu odsłuchania dźwięków należy wcisnąć czerwony przycisk „**Play**”. Przycisk „**Reset**” resetuje cały test – po jego wciśnięciu test należy zacząć od nowa.
4. Dla każdego z 25 dźwięków referencyjnych należy przesuwać suwak z dźwiękiem testowym i odsłuchiwać pary aż do pierwszej zauważonej różnicy w wysokości „**Pitch**” i/lub barwie „**Timbre**” dźwięku. Każde

kolejne zaznaczenie różnicy w wysokości lub barwy dźwięku nadpisuje wcześniejszą odpowiedź dla danego dźwięku referencyjnego. Suwaki umożliwiają powrót do wcześniejszych par dźwięków i ich ponowne odsłuchanie.



5. Po skończonym teście należy zapisać wyniki do pliku przyciskiem „Save” znajdującym się na dole ekranu. Domyślna nazwa to „dl.xls”. Prosimy wpisać jako nazwę swój IDENTYFIKATOR, np.: dl0001.xls dla identyfikatora dl0001.



12 Aneks

Podczas pierwszego eksperymentu jakim było zastosowanie słuchawek używanych przez uczestników badania na sztucznej głowie zauważono charakterystyczny „grzebień” (rys. 48), którego pochodzenie nie jest oczywiste. Pojawiły się dodatkowe pytania związane z jakością sprzętu użytego podczas badań behawioralnych a następnie zdecydowano o kontynuacji badań w komorze bezechowej i przetestowaniu różnych typów słuchawek z użyciem sztucznej głowy. W niniejszym aneksie chciałabym przedstawić wykresy na których zaprezentowano zbiorcze porównanie różnych typów słuchawek dla impulsów gaussowskich. Impulsy gaussowskie generowane do niniejszej analizy mają częstotliwość próbkowania równą 44100 Hz oraz 192000 Hz. Przedstawię również pomiary przeprowadzone bez użycia słuchawek, gdzie impulsy odtwarzano za pomocą głośnika.

12.1 Pomiar HATS z użyciem głośnika.

Do pomiaru użyto głośnika GENELEC 8020C. Całą specyfikację techniczną dotyczącą tego urządzenia można znaleźć w cytowanym odniesieniu do strony internetowej [114].

Na rysunku 64 przedstawiono maksymalną wartość otrzymaną dla każdego z impulsów gaussowskich z zakresu od $\sigma = 1 - 25$ dla konfiguracji w której sztuczna głowa „odsłuchiwała” impulsy generowane przez głośnik.

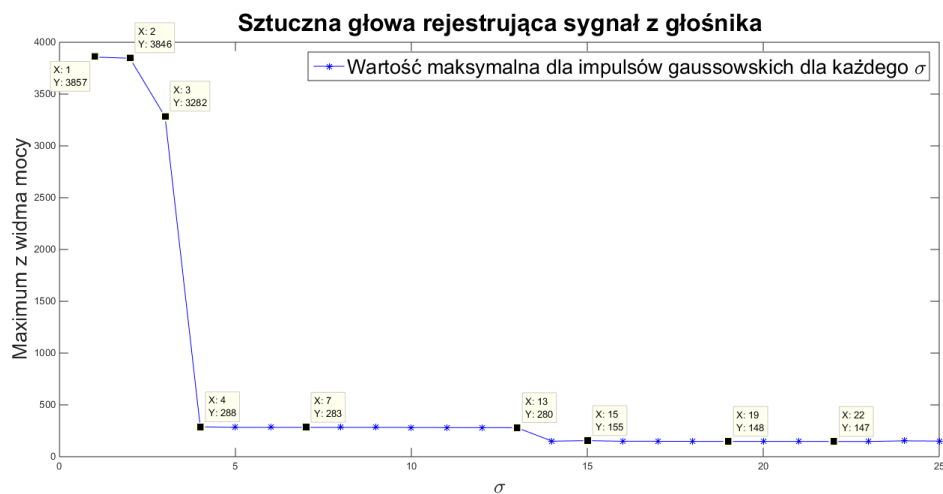
Odczytanie wartości maksymalnej nie sprawiło większych trudności jednak uwagę zwracają charakterystyczne maksima, które wyraźnie widać na widmach mocy załączonych do niniejszej pracy. Na rysunkach 65, 66 oraz 67 widzimy widma otrzymane dla impulsów gaussowskich. w celu analizy do każdego wykresu dodano etykiety pokazujące jakie wartości przyjmowane są dla konkretnego impulsu.

Na wykresach zauważono wyraźne piki, które przesuwają się w miarę wydłużania dźwięku. Dla przykładu w tabeli 8 pokazano, że piki o zbliżonych wartościach występują dla impulsów różnej długości. Piki te przestają być dominującymi dla dźwięków najkrótszych, jednak ciągle można zauważyć je na wykresie widma (np. rys. 65b czy 65c). Pierwsze trzy piki na potrzeby niniejszych rozważań nazywamy „grzebieniem” lub „zębami”.

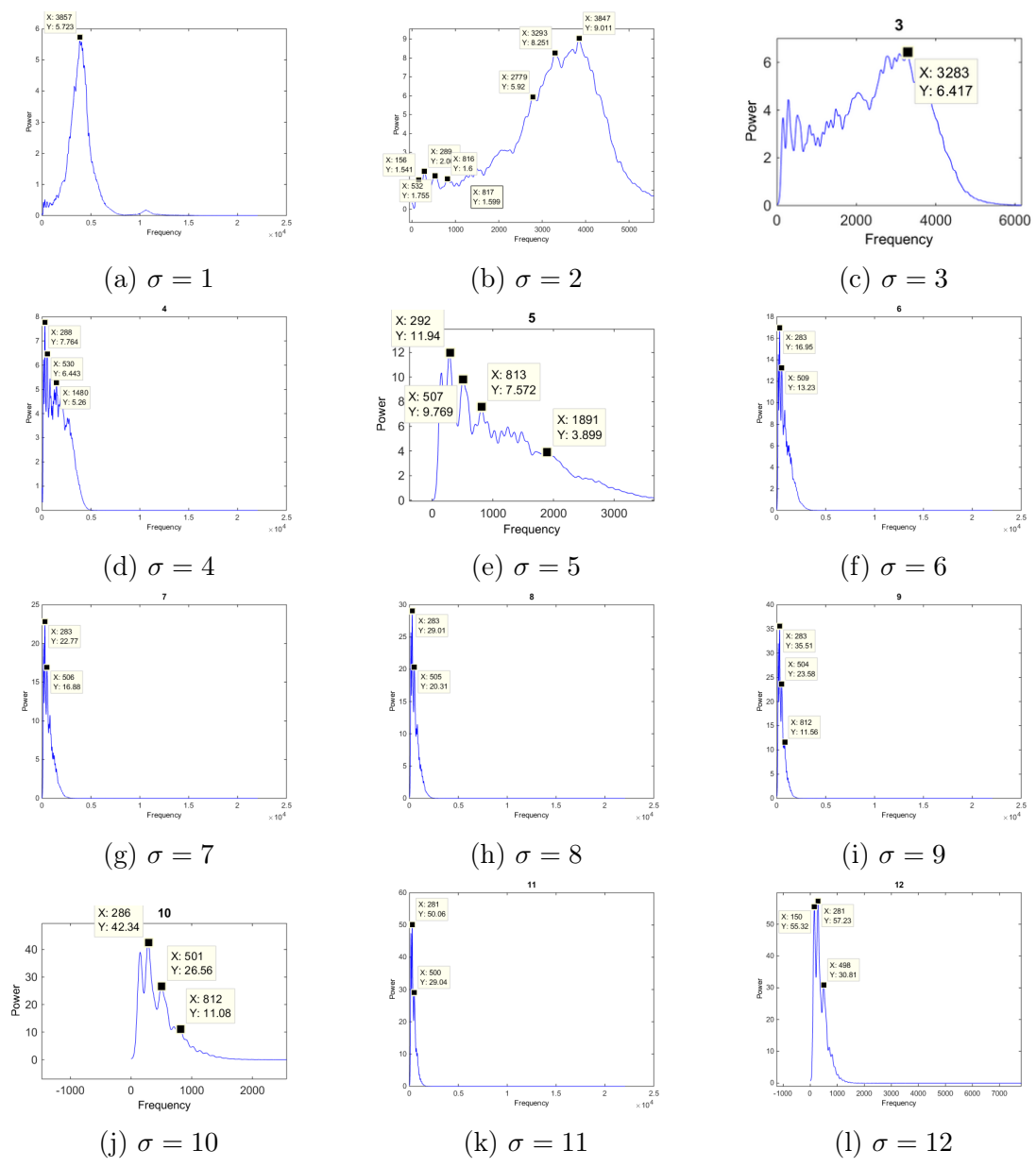
Numer σ	I pik	II pik	III pik	IV pik
2	156	289	532	816
5	150	292	507	813
10	149	286	501	812
15	155	273	463	-

Tablica 8: Tabela przedstawiająca pierwsze 4 wyraźne maksima dla dźwięków o $\sigma = 2, \sigma = 5, \sigma = 10$ i $\sigma = 15$ zmierzonych w konfiguracji: HATS - głośnik.

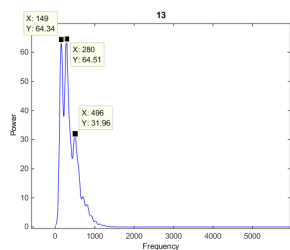
Próba analizy: Gdy wartość σ jest mała, rys: 65a, 65b 65c, maksima wydają się być fizycznymi gdyż leżą wyżej niż początkowy „grzebień”. Natomiast dla większych wartości σ nie widać już fizycznych maksimów tylko sam „grzebień”. Możliwe, że „grzebień” pojawia się ze względu na rezonans któregoś z urządzeń użytych do pomiaru.



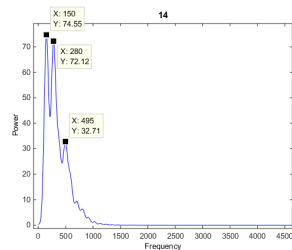
Rysunek 64: Maksymalna wartość dla każdego σ z zakresu 1 – 25 zarejestrowana przez sztuczną głowę



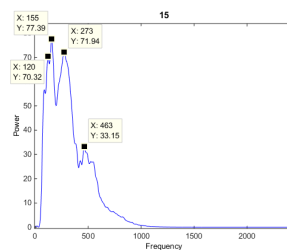
Rysunek 65: Widma impulsów gaussowskich otrzymane dla sztucznej głowy (HATS) słuchającej impulsów z głośnika dla σ z zakresu od 1 do 12. Próbkowanie impulsów: 44100Hz.



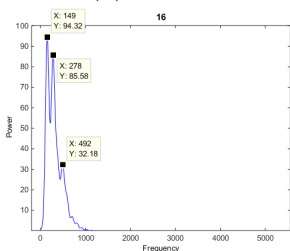
(a) $\sigma = 13$



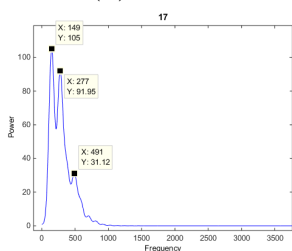
(b) $\sigma = 14$



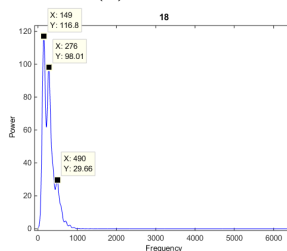
(c) $\sigma = 15$



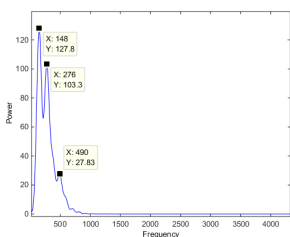
(d) $\sigma = 16$



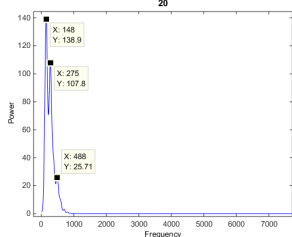
(e) $\sigma = 17$



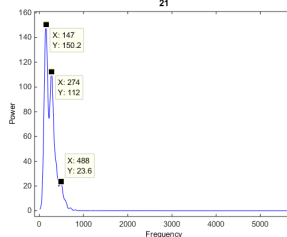
(f) $\sigma = 18$



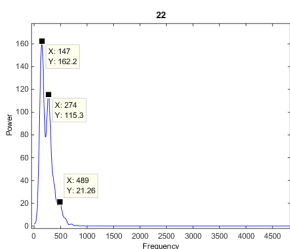
(g) $\sigma = 19$



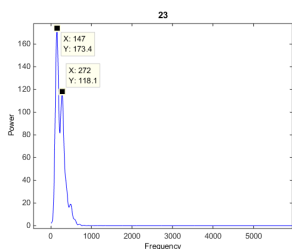
(h) $\sigma = 20$



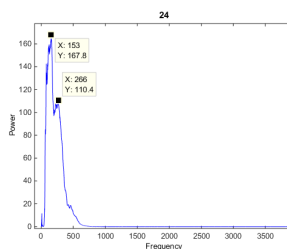
(i) $\sigma = 21$



(j) $\sigma = 22$

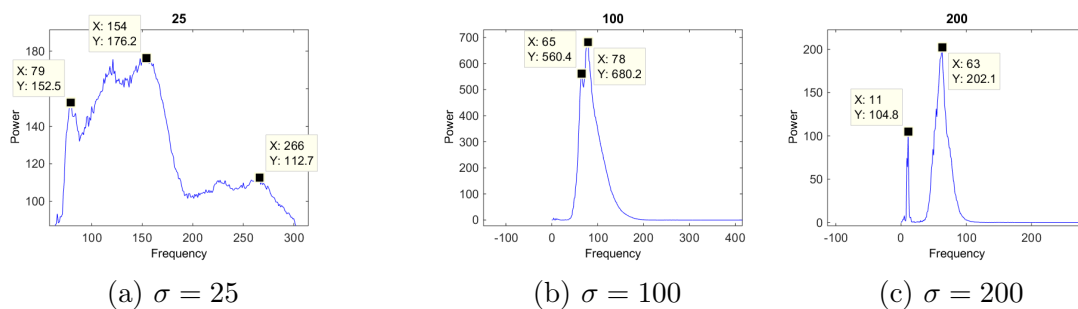


(k) $\sigma = 23$



(l) $\sigma = 24$

Rysunek 66: Widma impulsów gaussowskich otrzymane dla sztucznej głowy (HATS) słuchającej impulsów z głośnika dla σ z zakresu od 13 do 24. Próbkowanie impulsów: 44100Hz.



Rysunek 67: Widma impulsów gaussowskich otrzymane dla sztucznej głowy (HATS) słuchającej impulsów z głośnika dla wybranych σ . Próbkowanie impulsów: 44100Hz.

Na rys. 71 widzimy wartości maksymalne otrzymane z widma dla σ w zakresie od 1-50. Czyli dla pierwszych 50-ciu próbek dźwiękowych. Na kolejnym rysunku (72) widzimy ten sam zestaw słuchawek i próbkowań ale z wyliczoną wartością oczekiwaną z każdego widma i w nieco mniejszym zakresie bo dla pierwszych 25 dźwięków.

12.2 Pomiar HATS z użyciem 4 różnych typów słuchawek.

Badania przeprowadzono z użyciem słuchawek:

- Creative HS-600 - opisane szczegółowo w rozdziale 8.2, prezentowane na zdjęciu 20.
- Bayerdynamic DT770 - słuchawki zamknięte co oznacza, że mocno izolują użytkownika od otoczenia zewnętrznego a ich używanie będzie bardzo komfortowe w głośnych miejscach. W tym typie słuchawek głośnik pracuje w przestrzeni zamkniętej co oznacza, że potrzebna jest konstrukcja umożliwiająca wytłumienie wszelkich odbić (np. od obudowy). Słuchawki takie są pozbawione wentylacji dlatego dłuższy odsłuch może spowodować dyskomfort.

Dane techniczne DT-770 80 Ohm

- Przetworniki dynamiczne, zamknięte
- Konstrukcja wokółuszna
- Pasmo przenoszenia: 5 – 30 000 Hz



Rysunek 68: Słuchawki Bayerdynamic DT770, [115].

- Nominalne natężenie dźwięku: 105 dB
 - Impedancja: 80 Ohm
 - Waga: 300 g
 - Połączenie jack 3,5 mm stereo z przejściówką jack 6,3 mm
- Bayerdynamic DT880 - słuchawki półotwarte. W porównaniu do słuchawek DT770 Pro dźwięk w tym przypadku jest bardziej trójwymiarowy.

Dane techniczne DT-880 80 Ohm

- Konstrukcja wokółuszna
 - Pasmo przenoszenia: 5 – 35 000 Hz
 - Nominalne natężenie dźwięku: 96 dB
 - Impedancja: 250 Ohm
 - Waga: 10,1 Oz
 - Połączenie mini-jack (3,5 mm) i adapter 1/4
- Bayerdynamic DT990 - słuchawki otwarte. Za głośnikiem znajduje się ażurowa osłona, dzięki której możliwy jest przepływ powietrza poruszanego membraną. Takie rozwiązanie powoduje, że głośnik nie napotyka na rosnące ciśnienie oraz odbicia od obudowy spotykane w konstrukcjach z obudową zamkniętą. Dzięki temu membrana oddaje nawet najsubtelniejsze detale. Słuchawki otwarte bardzo dobrze sprawdzają się podczas długiego słuchania, dzięki otwartej obudowie łatwiejsza jest wentylacja ucha ale co za tym



Rysunek 69: Słuchawki Bayerdynamic DT880, źródło: [116].

idzie izolacja akustyczna praktycznie nie występuje, dlatego ważne jest aby w pomieszczenie w którym słuchamy była cisza. Pomiar przeprowadzony był w komorze bezchowej AGH gdzie nie mieliśmy z tym problemu. Widma mocy otrzymane dla krótkich dźwięków gaussowskich widzimy na rys. 80.

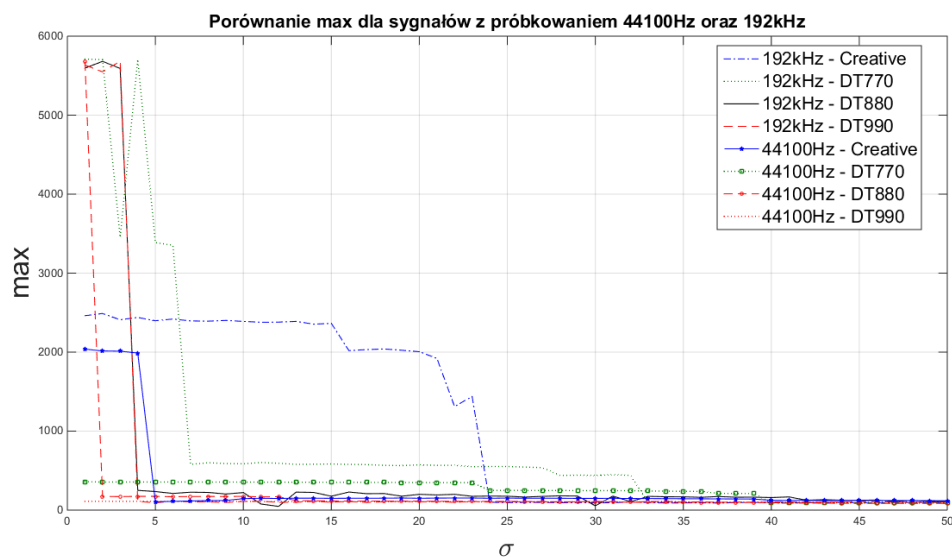


Rysunek 70: Słuchawki Bayerdynamic DT990, źródło [117].

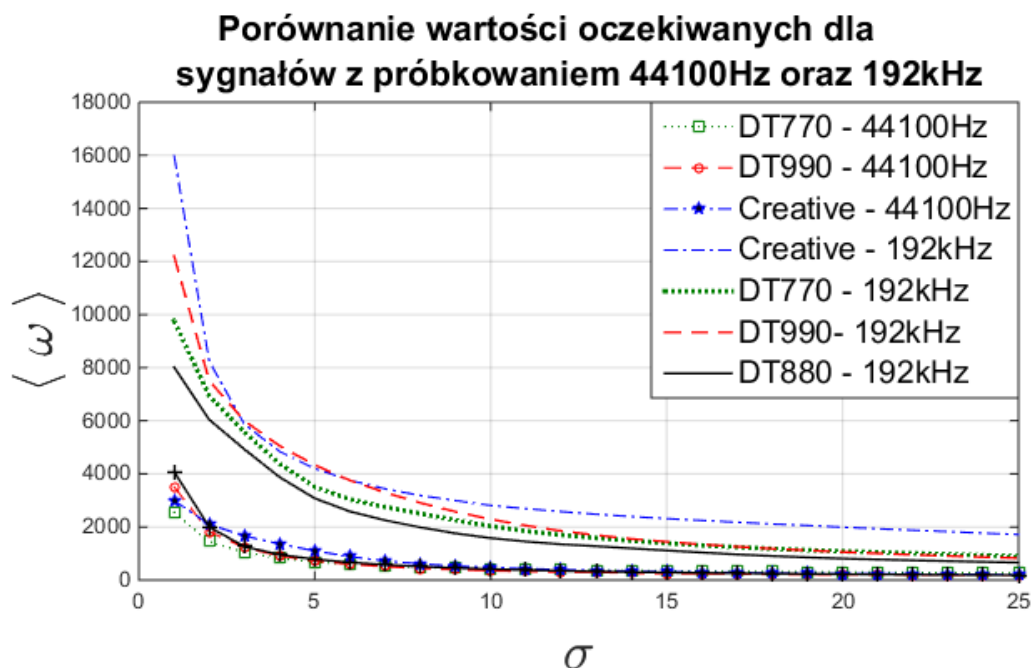
Dane techniczne DT-990 250 Ohm

- Typ: Otwarte, dynamiczne
- Pasmo przenoszenia: 5Hz - 35kHz
- Impedancja: 250 Ohm

- Nominalny SPL: 96 dB
- Nominalne zniekształcenia THD: <0,2
- Maksymalna moc wejściowa: 100mW
- Waga: 250g



Rysunek 71: Maksymalna wartość dla każdego σ z zakresu 1 – 50 zarejestrowana przez HATS słuchającą sygnałów z 4 różnych typów słuchawek.



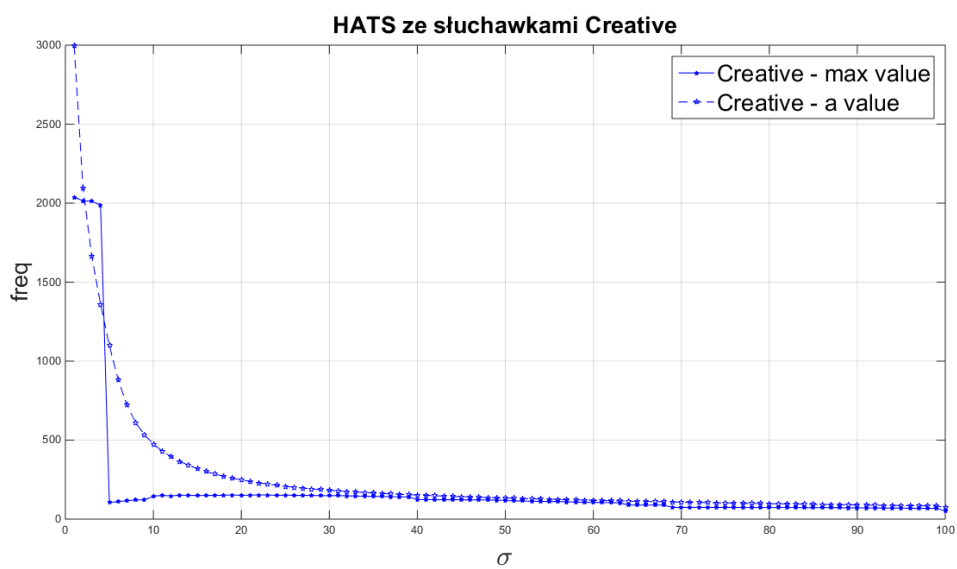
Rysunek 72: Wartość oczekiwana dla każdego σ z zakresu 1 – 25 zarejestrowana przez HATS słuchającą sygnałów z 4 różnych typów słuchawek.

12.3 Widma mocy otrzymane podczas pomiaru HATS - częstotliwość próbkowania 44100 Hz.

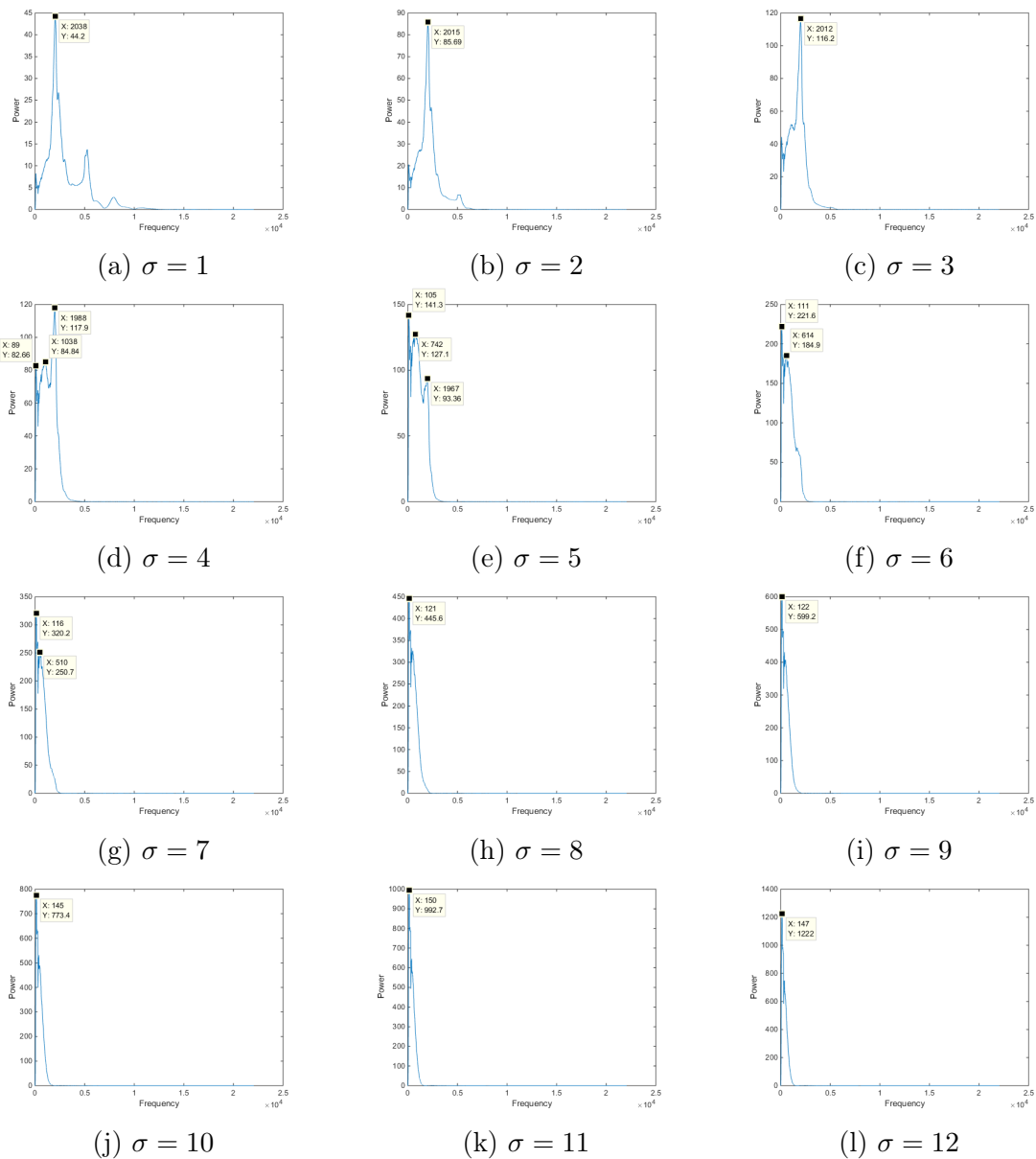
W niniejszym rozdziale przedstawię widma mocy otrzymane dla różnego typu słuchawek, gdzie odtwarzane były impulsy gaussowskie. Przedstawiam również wykresy zależności maksimum widma od σ oraz wartości średniej od σ . Częstotliwość próbkowania przygotowanych sygnałów wynosi 44100 Hz.

Na rys. 73 przedstawiono maksymalną wartość oraz wartość oczekiwaną otrzymaną dla widma mocy dla każdego σ dla słuchawek Creative.

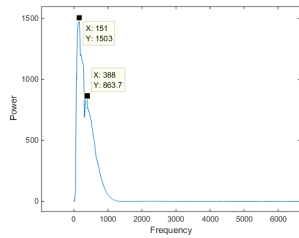
Na rys. 75 przedstawiono widma mocy otrzymane dla sztucznej głowy słuchającej krótkich impulsów akustycznych ze słuchawek Creative, model Hs-600. Przedstawione widma dotyczą dźwięków od $\sigma = 1$ do $\sigma = 12$.



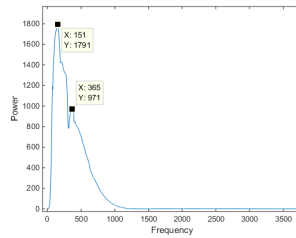
Rysunek 73: Maksymalna wartość dla każdego σ z zakresu 1 – 100 zarejestrowana przez sztuczną głowę słuchającą sygnałów ze słuchawek Creative.



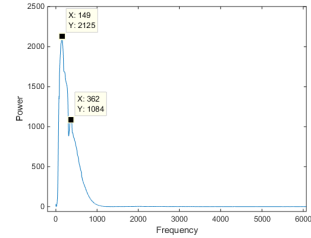
Rysunek 74: Widma impulsów gaussowskich otrzymane dla sztucznej głowy (HATS) słuchającej impulsów dla σ z zakresu od 1 do 12 dla słuchawek Creative, model Hs-600. Próbkowanie impulsów: 44100Hz.



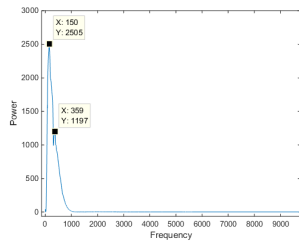
(a) $\sigma = 13$



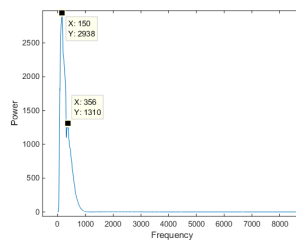
(b) $\sigma = 14$



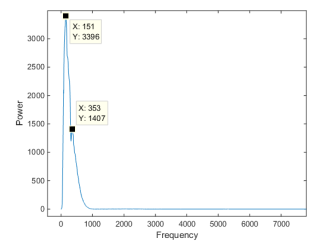
(c) $\sigma = 15$



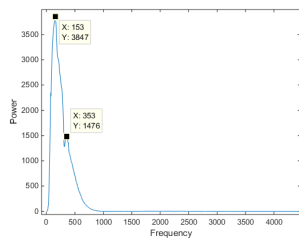
(d) $\sigma = 16$



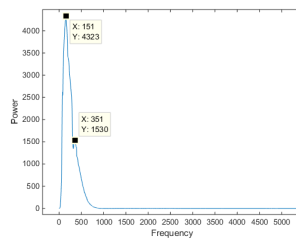
(e) $\sigma = 17$



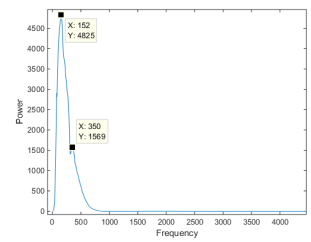
(f) $\sigma = 18$



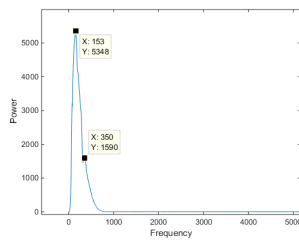
(g) $\sigma = 19$



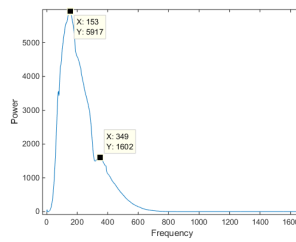
(h) $\sigma = 20$



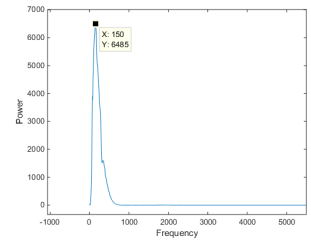
(i) $\sigma = 21$



(j) $\sigma = 22$

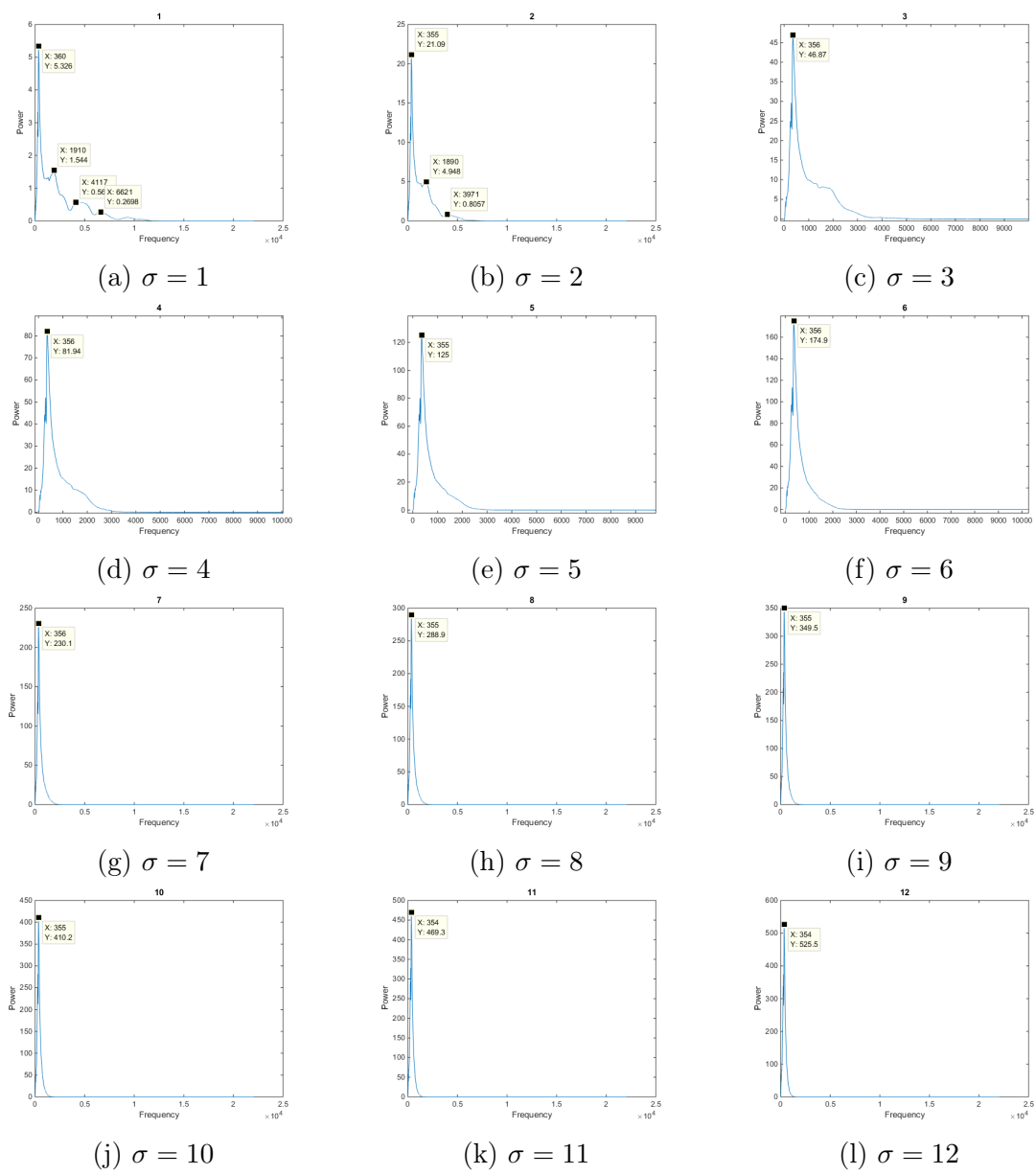


(k) $\sigma = 23$

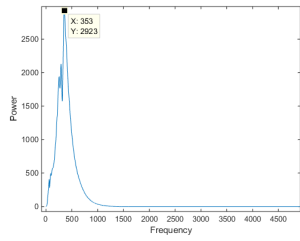


(l) $\sigma = 24$

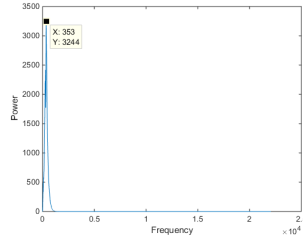
Rysunek 75: Widma impulsów gaussowskich otrzymane dla sztucznej głowy (HATS) słuchającej impulsów dla σ z zakresu od 13 do 24 dla słuchawek Creative, model Hs-600. Próbkowanie impulsów: 44100Hz.



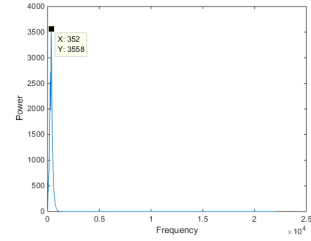
Rysunek 76: Widma impulsów gaussowskich otrzymane dla sztucznej głowy (HATS) słuchającej impulsów ze słuchawek DT770 σ z zakresu od 1 do 12. Częstotliwość próbkowania 44100 Hz.



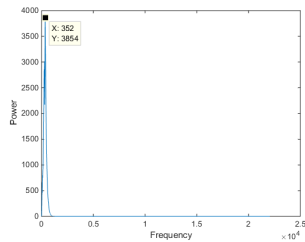
(a) $\sigma = 13$



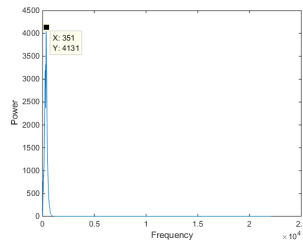
(b) $\sigma = 14$



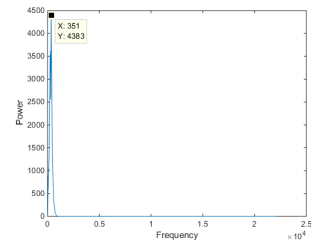
(c) $\sigma = 15$



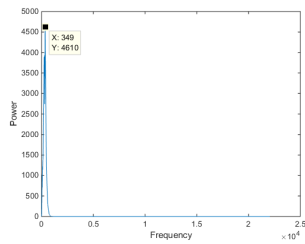
(d) $\sigma = 16$



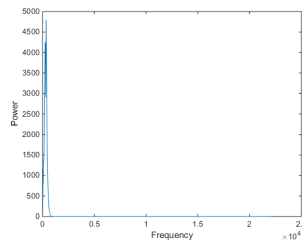
(e) $\sigma = 17$



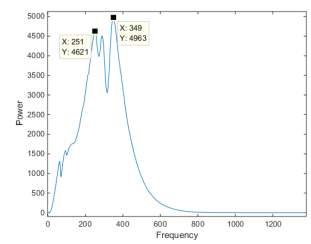
(f) $\sigma = 18$



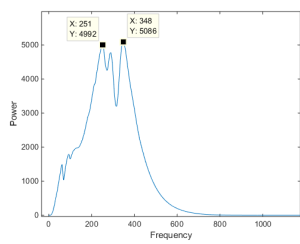
(g) $\sigma = 19$



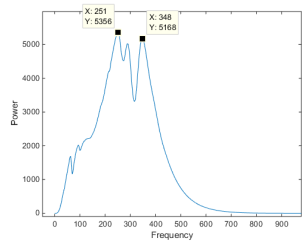
(h) $\sigma = 20$



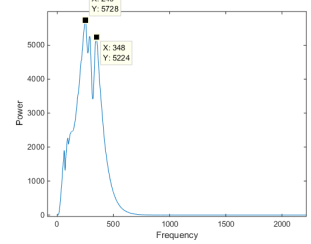
(i) $\sigma = 21$



(j) $\sigma = 22$

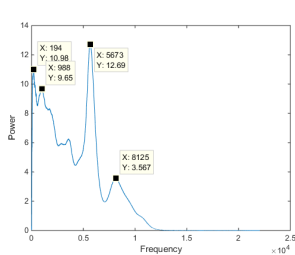


(k) $\sigma = 23$

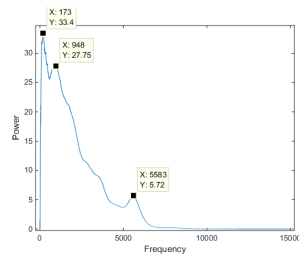


(l) $\sigma = 24$

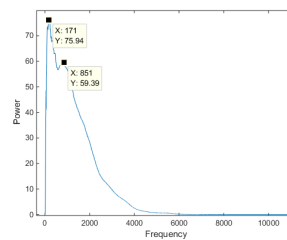
Rysunek 77: Widma impulsów gaussowskich otrzymane dla sztucznej głowy (HATS) słuchającej impulsów ze słuchawek DT770 σ z zakresu od 13 do 24. Częstotliwość próbkowania 44100 Hz.



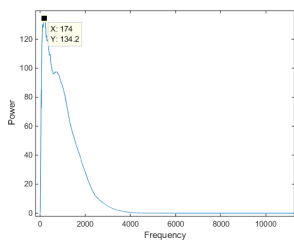
(a) $\sigma = 1$



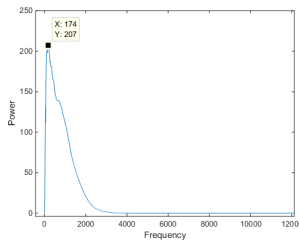
(b) $\sigma = 2$



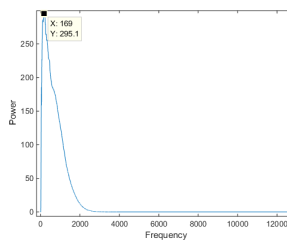
(c) $\sigma = 3$



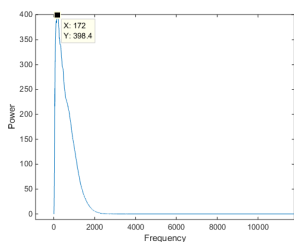
(d) $\sigma = 4$



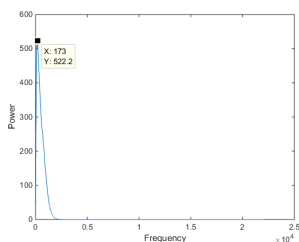
(e) $\sigma = 5$



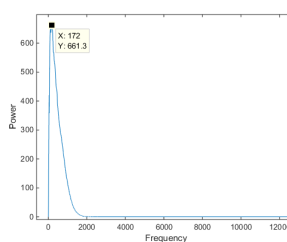
(f) $\sigma = 6$



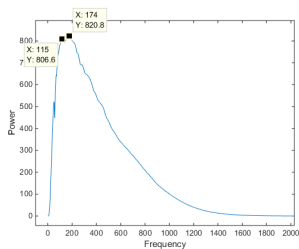
(g) $\sigma = 7$



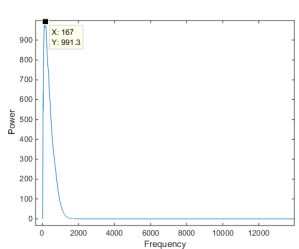
(h) $\sigma = 8$



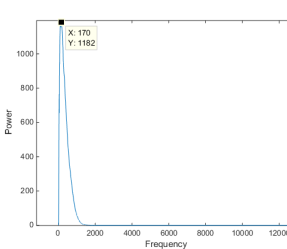
(i) $\sigma = 9$



(j) $\sigma = 10$

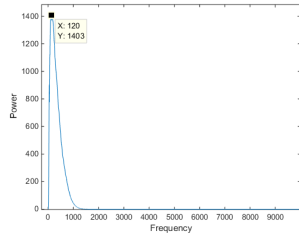


(k) $\sigma = 11$

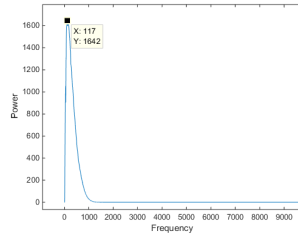


(l) $\sigma = 12$

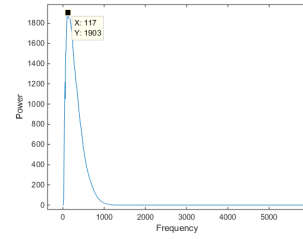
Rysunek 78: Widma impulsów gaussowskich otrzymane dla sztucznej głowy (HATS) słuchającej impulsów ze słuchawek DT880 σ z zakresu od 1 do 12. Częstotliwość próbkowania 44100 Hz.



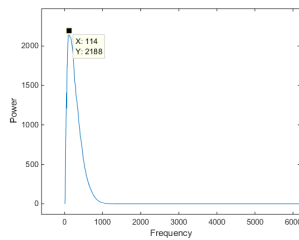
(a) $\sigma = 13$



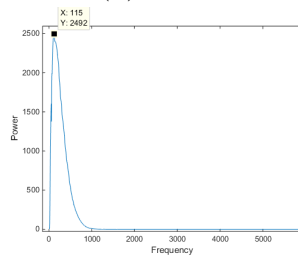
(b) $\sigma = 14$



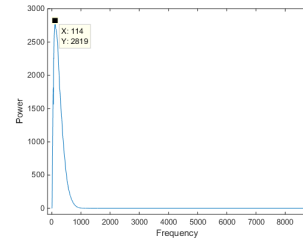
(c) $\sigma = 15$



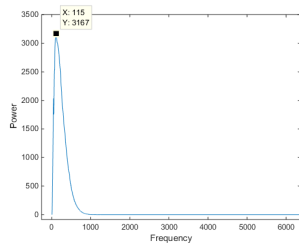
(d) $\sigma = 16$



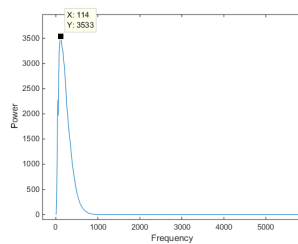
(e) $\sigma = 17$



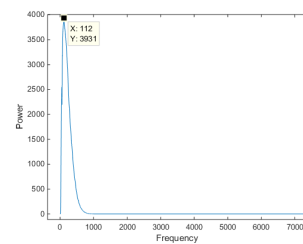
(f) $\sigma = 18$



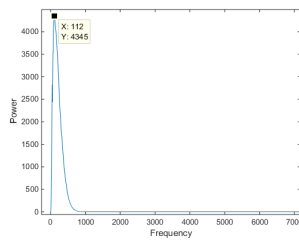
(g) $\sigma = 19$



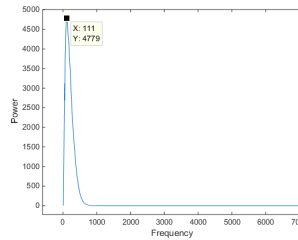
(h) $\sigma = 20$



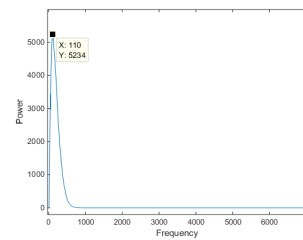
(i) $\sigma = 21$



(j) $\sigma = 22$

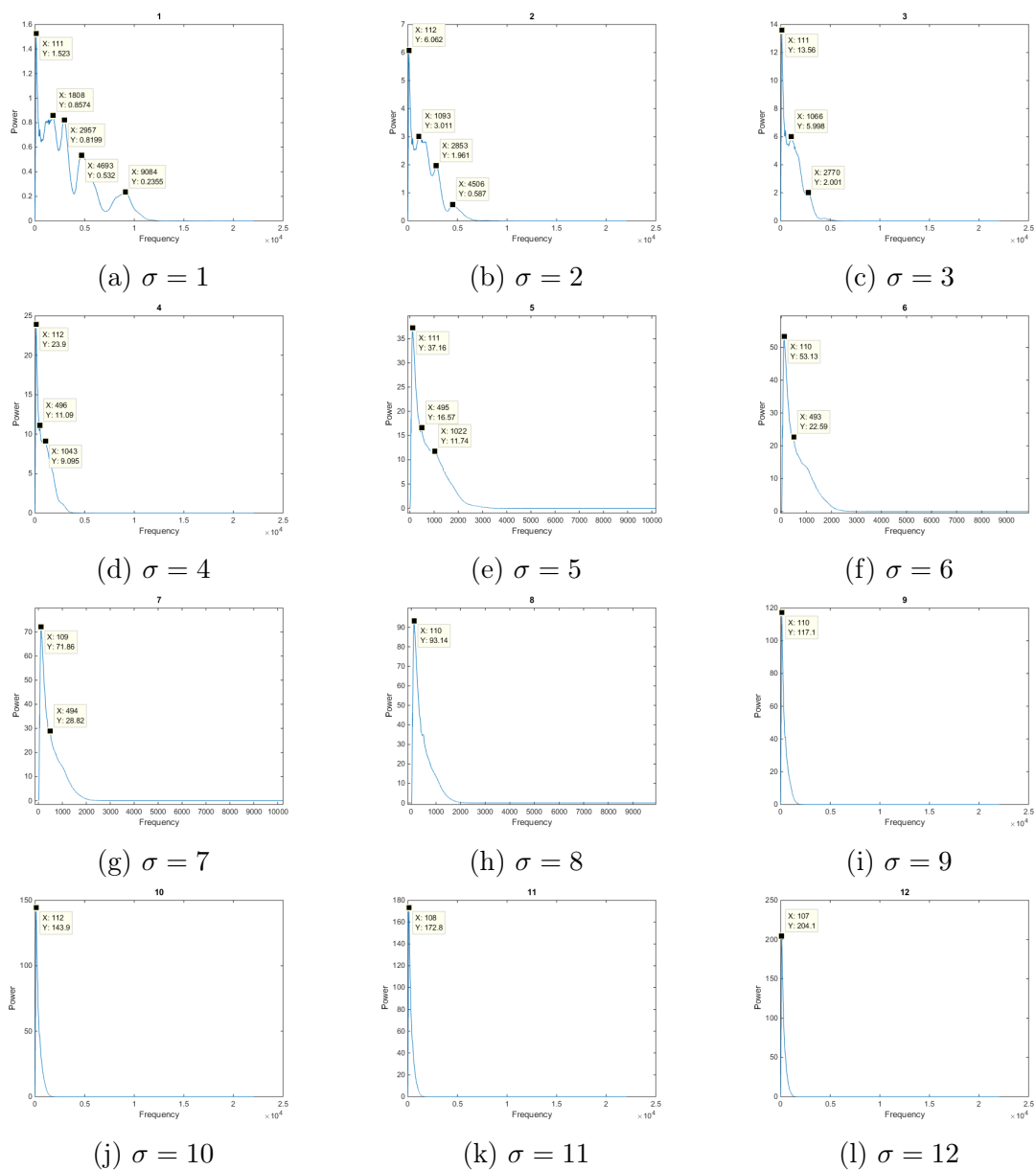


(k) $\sigma = 23$

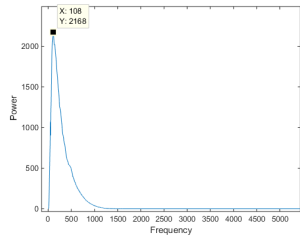


(l) $\sigma = 24$

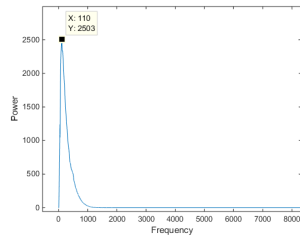
Rysunek 79: Widma impulsów gaussowskich otrzymane dla sztucznej głowy (HATS) słuchającej impulsów ze słuchawek DT880 σ z zakresu od 13 do 24. Częstotliwość próbkowania 44100 Hz.



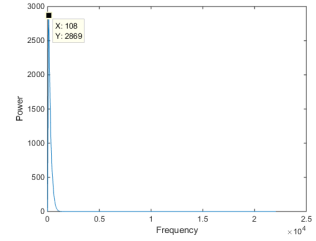
Rysunek 80: Widma impulsów gaussowskich otrzymane dla sztucznej głowy (HATS) słuchającej impulsów ze słuchawek DT990 σ z zakresu od 1 do 12. Próbkowanie impulsów: 44100Hz.



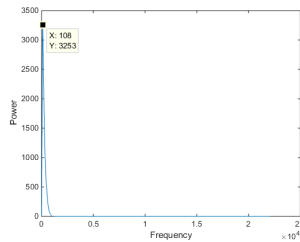
(a) $\sigma = 13$



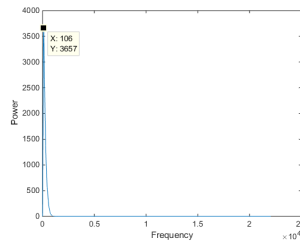
(b) $\sigma = 14$



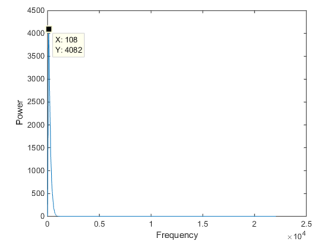
(c) $\sigma = 15$



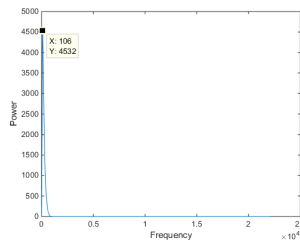
(d) $\sigma = 16$



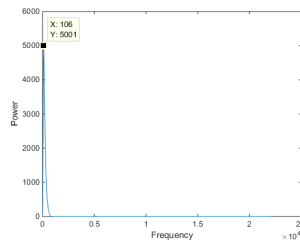
(e) $\sigma = 17$



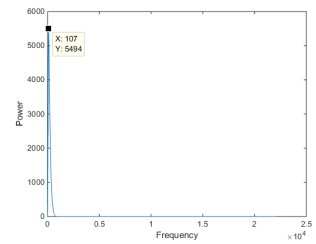
(f) $\sigma = 18$



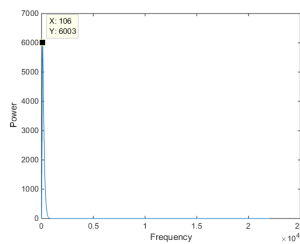
(g) $\sigma = 19$



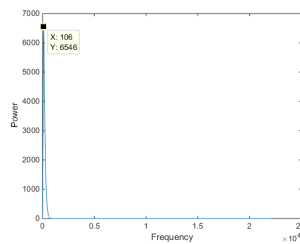
(h) $\sigma = 20$



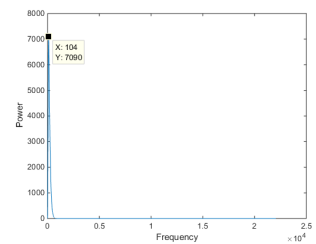
(i) $\sigma = 21$



(j) $\sigma = 22$



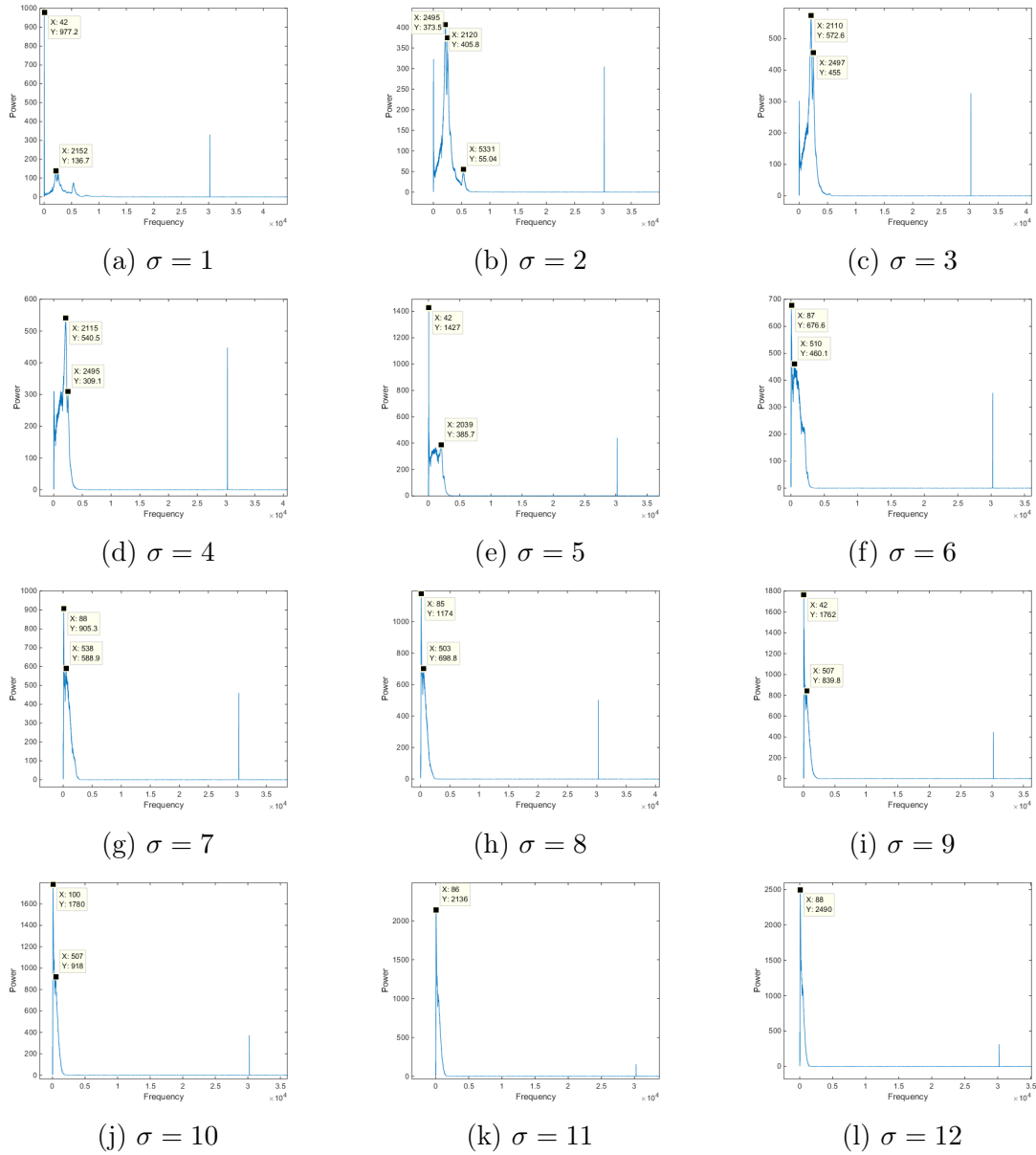
(k) $\sigma = 23$



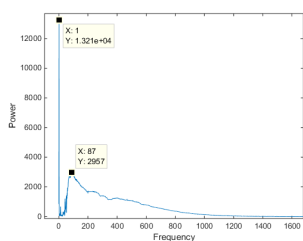
(l) $\sigma = 24$

Rysunek 81: Widma impulsów gaussowskich otrzymane dla sztucznej głowy (HATS) słuchającej impulsów ze słuchawek DT990 σ z zakresu od 13 do 24. Próbkowanie impulsów: 44100Hz.

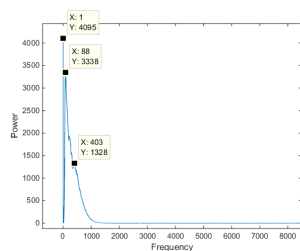
12.4 Widma mocy otrzymane podczas pomiaru HATS - częstotliwość próbkowania 192000 Hz



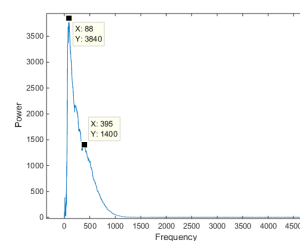
Rysunek 82: Widma impulsów gaussowskich otrzymane dla sztucznej głowy (HATS) słuchającej impulsów ze słuchawek Creative σ z zakresu od 1 do 12. Częstotliwość próbkowania dźwięków 192000 Hz



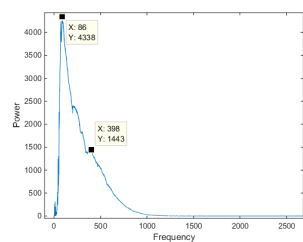
(a) $\sigma = 13$



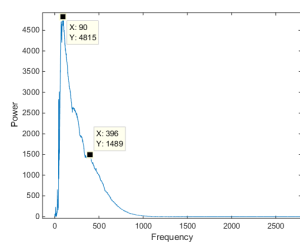
(b) $\sigma = 14$



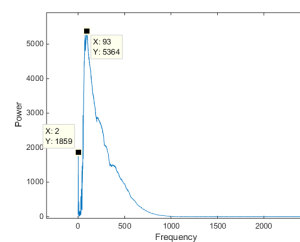
(c) $\sigma = 15$



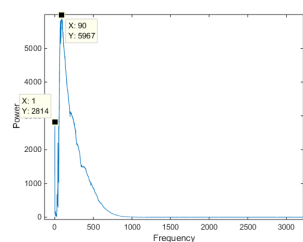
(d) $\sigma = 16$



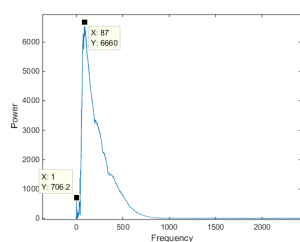
(e) $\sigma = 17$



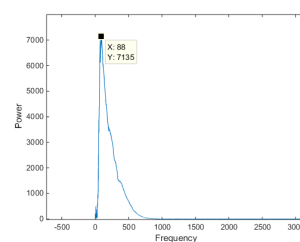
(f) $\sigma = 18$



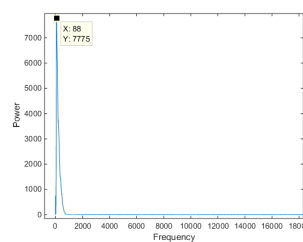
(g) $\sigma = 19$



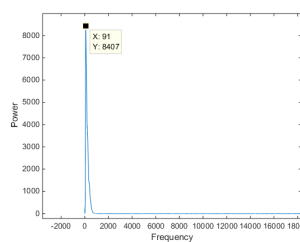
(h) $\sigma = 20$



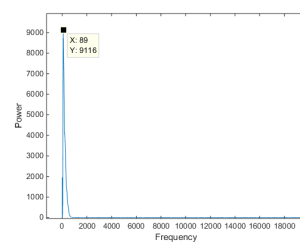
(i) $\sigma = 21$



(j) $\sigma = 22$

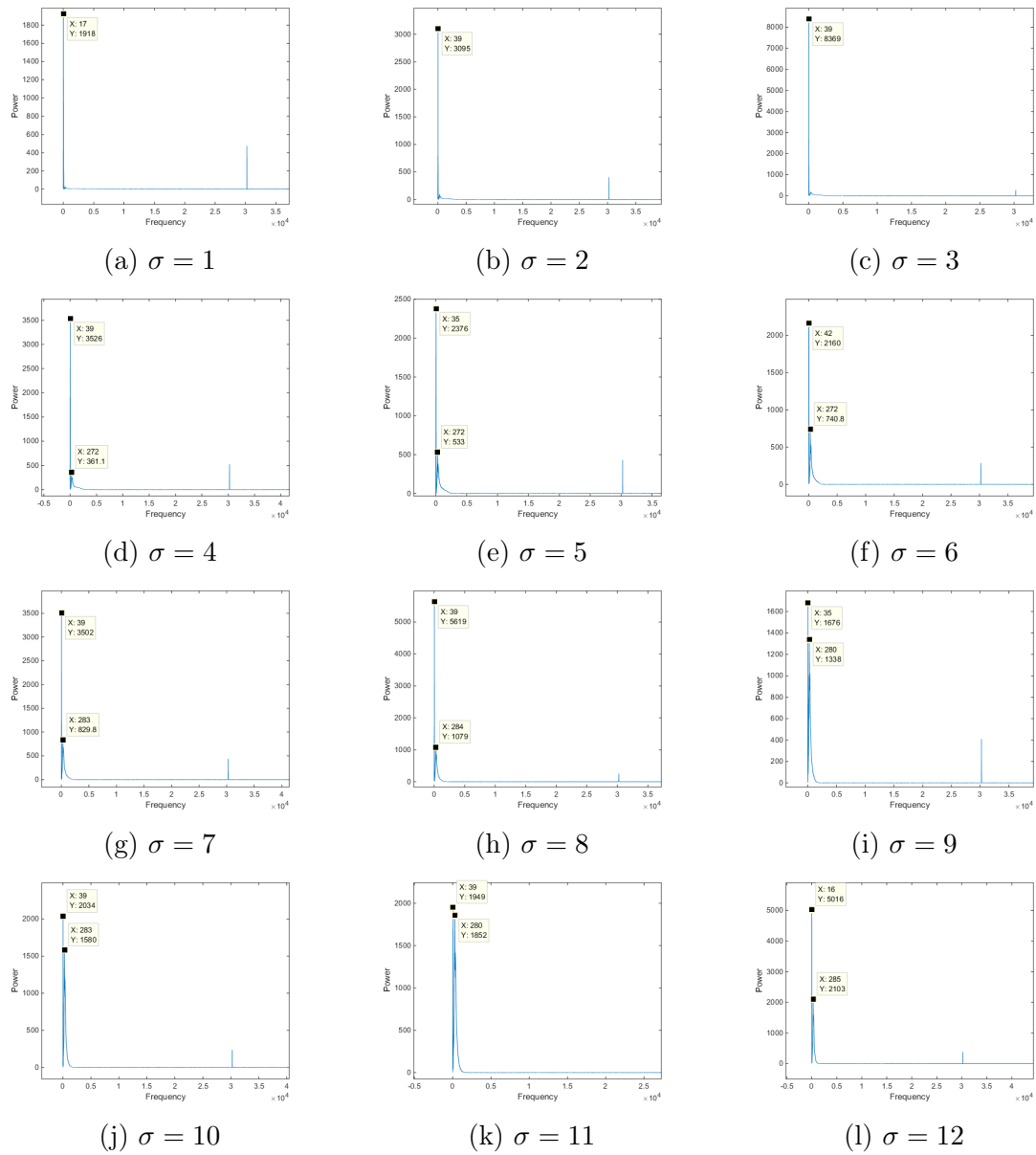


(k) $\sigma = 23$

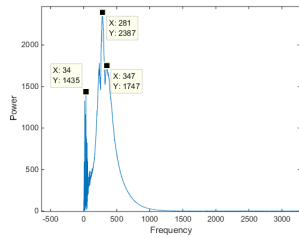


(l) $\sigma = 24$

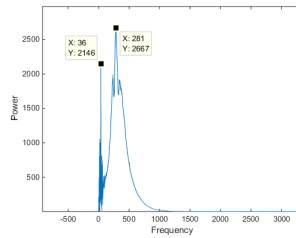
Rysunek 83: Widma impulsów gaussowskich otrzymane dla sztucznej głowy (HATS) słuchającej impulsów ze słuchawek Creative σ z zakresu od 13 do 24. Częstotliwość próbkowania dźwięków 192000 Hz



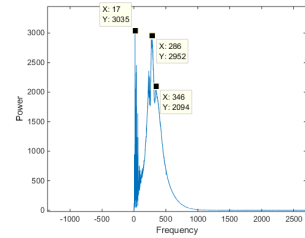
Rysunek 84: Widma impulsów gaussowskich otrzymane dla sztucznej głowy (HATS) słuchającej impulsów ze słuchawek DT770 σ z zakresu od 1 do 12. Częstotliwość próbkowania dźwięków 192000 Hz.



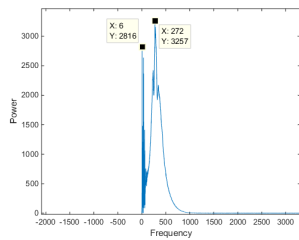
(a) $\sigma = 13$



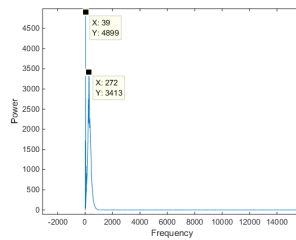
(b) $\sigma = 14$



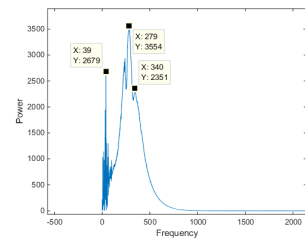
(c) $\sigma = 15$



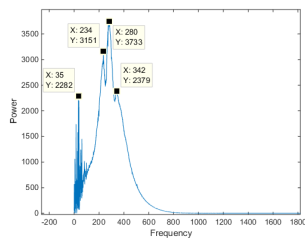
(d) $\sigma = 16$



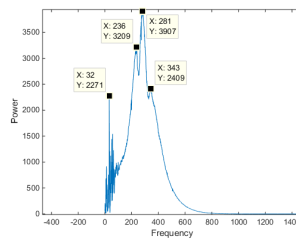
(e) $\sigma = 17$



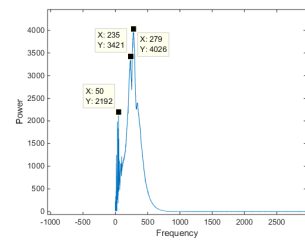
(f) $\sigma = 18$



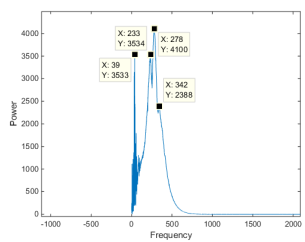
(g) $\sigma = 19$



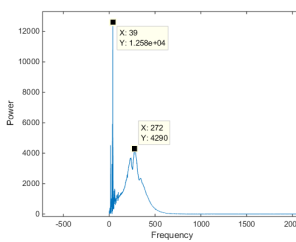
(h) $\sigma = 20$



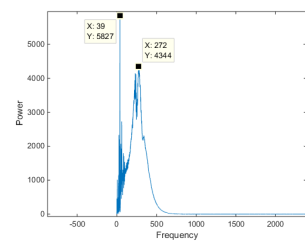
(i) $\sigma = 21$



(j) $\sigma = 22$

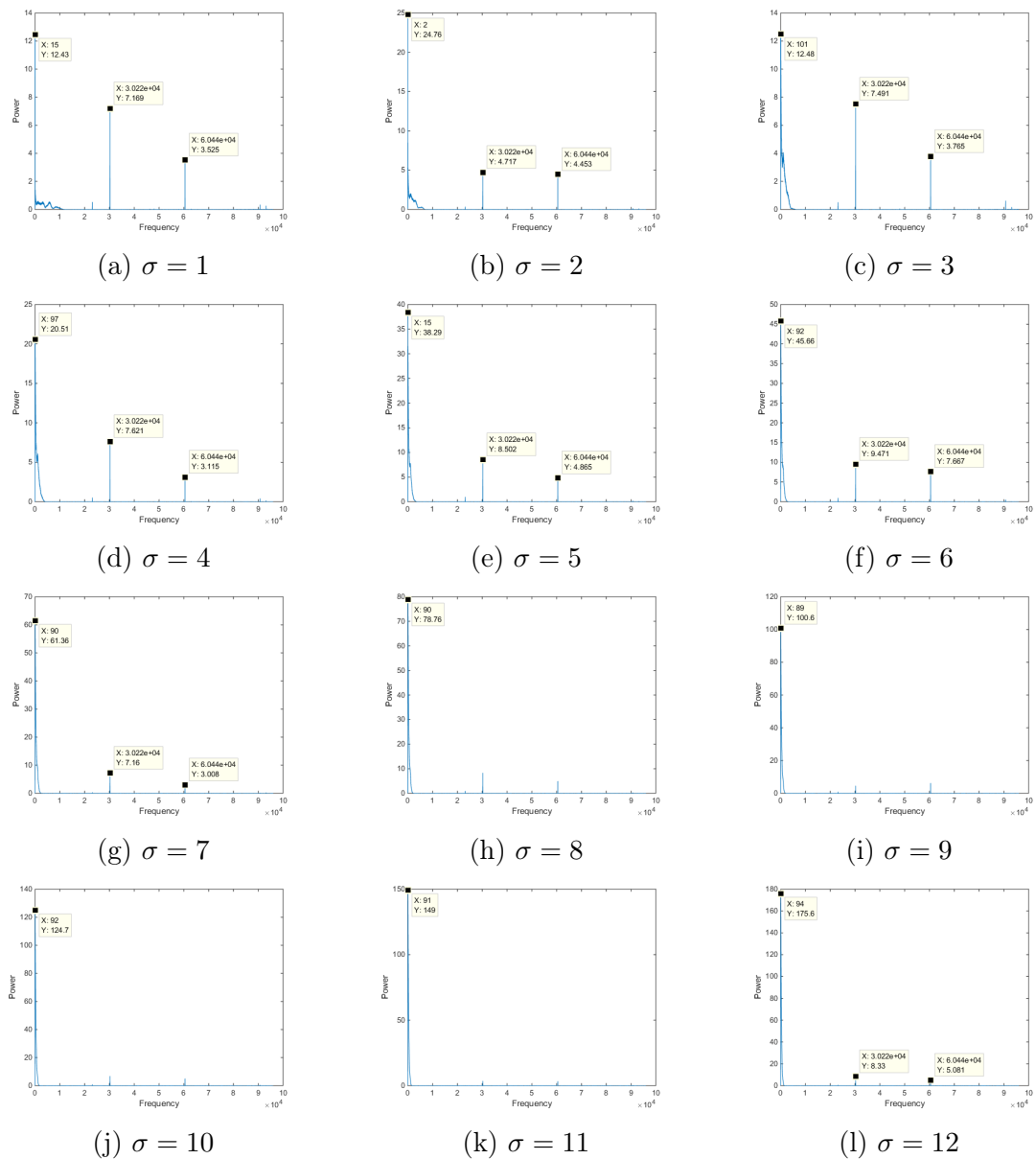


(k) $\sigma = 23$

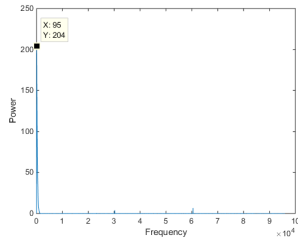


(l) $\sigma = 24$

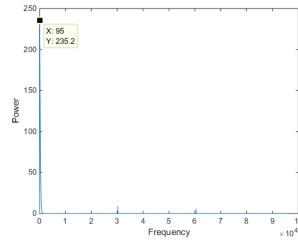
Rysunek 85: Widma impulsów gaussowskich otrzymane dla sztucznej głowy (HATS) słuchającej impulsów ze słuchawek DT770 σ z zakresu od 13 do 24. Częstotliwość próbkowania dźwięków 192000 Hz.



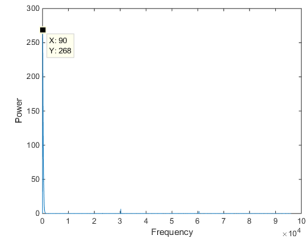
Rysunek 86: Widma impulsów gaussowskich otrzymane dla sztucznej głowy (HATS) słuchającej impulsów ze słuchawek DT990 σ z zakresu od 1 do 12. Częstotliwość próbkowania dźwięków 192000 Hz.



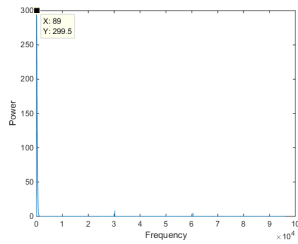
(a) $\sigma = 13$



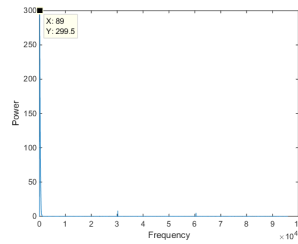
(b) $\sigma = 14$



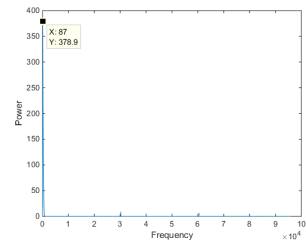
(c) $\sigma = 15$



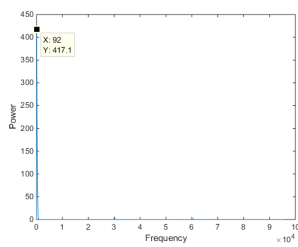
(d) $\sigma = 16$



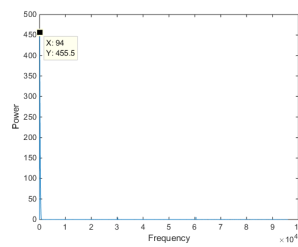
(e) $\sigma = 17$



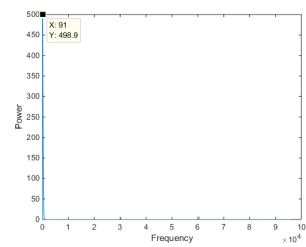
(f) $\sigma = 18$



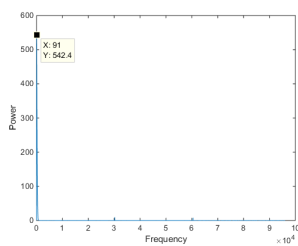
(g) $\sigma = 19$



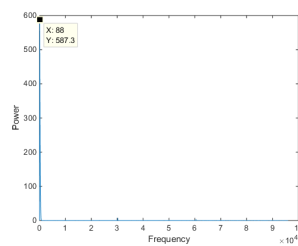
(h) $\sigma = 20$



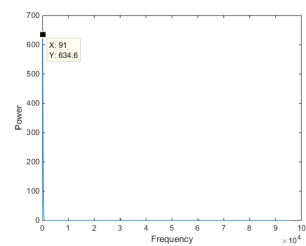
(i) $\sigma = 21$



(j) $\sigma = 22$



(k) $\sigma = 23$



(l) $\sigma = 24$

Rysunek 87: Widma impulsów gaussowskich otrzymane dla sztucznej głowy (HATS) słuchającej impulsów ze słuchawek DT990 σ z zakresu od 13 do 24. Częstotliwość próbkowania dźwięków 192000 Hz.

13 Bibliografia

- [1] S. Umesh, L. Cohen, and D. Nelson. Fitting the mel scale. *1999 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings. ICASSP99 (Cat. No.99CH36258)*, 1:217–220, 1999.
- [2] A. Rakowski and P. Rogowski. Investigating the pitch strength of short pure-tone pulses in middle frequency range through their chroma recognition by absolute-pitch listeners. *Interdisciplinary Studies in Musicology*, 8:271–278, 2018.
- [3] Marcin Majka, Paweł Sobieszczyk, Robert Gębarowski, and Piotr Zieliński. Hearing overcome uncertainty relation and measure duration of ultrashort pulses. *Europhysics News*, 46 1:27–31, 2015.
- [4] J. N. Oppenheim and M. O. Magnasco. Human time-Frequency Acuity Beats the Fourier Uncertainty Principle. *Physical Review Letters*, 110:044301, 2013.
- [5] Franciszek Wesołowski. *Zasady muzyki*. Polskie Wydawnictwo Muzyczne, 1986. ISBN: 83-224-0250-3.
- [6] Dennis Gabor. Theory of communication. *Journal of Institution of Electrical Engineers*, 93(3):429–457, 1946.
- [7] Urszula Jorasz. *Wykłady z psychoakustyki*. Wydawnictwo Naukowe UAM, 1998. ISBN: 83-232-0768-2.
- [8] Edward Hojan. *Protetyka Słuchu*. Wydawnictwo Naukowe UAM, 2017.
- [9] William E. Brownell. How the Ear Works - Nature's Solutions for Listening. *Volta Review*, 99(5):9–28, 1997.
- [10] Piotr Kleczkowski. *Percepcja Dźwięku*. Wydawnictwa AGH, Kraków, fabryka inżynierów edition, 2013.
- [11] William A. Yost. *Fundamentals of Hearing. An Introduction*. Emerald Group, 5th edition, 2008.

- [12] E.B. Goldstein and J. Brockmole. *Sensation and Perception*. Cengage Learning, 2016. ISBN: 978-1305580299.
- [13] M. Ugarteburu, R. H. Withnell, L. Cardoso, A. Carriero, and C. Richter. Mammalian middle ear mechanics: A review. *Front. Bioeng. Biotechnol. Sec. Biomechanics*, 10:983510, 2022.
- [14] T. Gierek and A. Śląska Kaspera. Mięsień strzemiączkowy – aktualne poglądy na temat anatomii i fizjologii. *Otolaryngologia Polska*, LXI(1):29, 2007.
- [15] Stanley A. Gelfand. *Hearing. An Introduction to Psychological and Physiological Acoustics*. Informa healthcare, Great Britain, 5th edition, 2010.
- [16] Edward Ozimek. *Dźwięk i jego percepcja. Aspekty fizyczne i psychoakustyczne*. Wydawnictwo Naukowe PWN, 2002.
- [17] Website: Jobilize. Ucho środkowe i wewnętrzne. https://www.jobilize.com/ocw/mirror/col11406/m42297/figure_18_06_07.jpg, May 2020.
- [18] Website: Podstawy fizjologii i patologii słuchu. <https://sound.eti.pg.gda.pl/student/pp/roz2.pdf>, March 2023.
- [19] Website: Open Learn. Błona podstawna, uproszczenie. <https://www.open.edu/openlearn/science-maths-technology/biology/hearing/content-section-3.3>, May 2020.
- [20] Simon Grondin. *Psychology of Perception*. Springer, 2016. ISBN: 978-3-319-31789-2.
- [21] F. A. Wichmann and F. Jäkel. *Methods in psychophysics. Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience*. 2018. ISBN: 9781119170167.
- [22] Neil A. Macmillan and C. Douglas Creelman. *Detection Theory: A user's guide*. Cambridge University Press, 1991. ISBN: 05-213-6892-8.
- [23] J. Albers. *Interaction of Color*. New Haven and London: Yale University Press, 1963.

- [24] Website: Perceived magnitude. <https://graphworkflow.com/decoding/perceived-magnitude/>, March 2023.
- [25] Website: Dictionary of American Psychological Association. <https://dictionary.apa.org/weber-fraction>, March 2023.
- [26] Claude Bonnet. *Manuel pratique de psychophysique*. Armand Colin, 1986. ISBN: 22-003-1208-3.
- [27] Website: Prawo webera-fechnera. <https://nyquista.pl/slownik/prawo-webera-fechnera/>, March 2023.
- [28] D. Algom. The weber–fechner law: A misnomer that persists but that should go away. *Psychological Review*, 128(4):757–765, 2021.
- [29] J. M. Kreiner. *Astronomia z astrofizyką*. Państwowe Wydawnictwo Naukowe, Warszawa, 1988. ISBN: 9788301076467.
- [30] S. S. Stevens. On the psychophysical law. *Psychological Review*, 64(3):153–181, 1957.
- [31] Website: Psychology 3203. http://www.appstate.edu/~steelekm/classes/psy3203/Topics/topics_Sumr07.html, March 2023.
- [32] Fan-Gang Zeng. A unified theory of psychophysical laws in auditory intensity perception. *Frontiers in Psychology*, 11:1459, 2020.
- [33] Dariusz Czaja. Szmer, szepty i krzyki. muzyka wenecka. *Acta Universitatis Vratislaviensis, Prace Kulturoznawcze XIII*, 12:57, 2012.
- [34] Website: Fale dźwiękowe jako przykład fal harmoniczych. Cechy dźwięków: wysokość i głośność. <https://zpe.gov.pl/a/fale-dzwiekowe-jako-przyklad-fal-harmoniczych-cechy-dzwiekow-wysokosc-i-glosnosc/D1C0rWy07>, March 2023.
- [35] Oliver Sacks. *Muzykofilia. Opowieści o muzyce i mózgu*. Wyd. Zysk i S-ka, 2009. ISBN: 978-83-7506-322-6.

- [36] Piotr Zieliński. *Fizyka i fizjologia dźwięku*. Politechnika XXI wieku - Program rozwojowy Politechniki Krakowskiej – najwyższej jakości dydaktyka dla przyszłych polskich inżynierów, 2011.
- [37] Marcel Zentner and Rebecca L. Shiner (Eds). *Handbook of Temperament*. Guilford Publications, 2015. ISBN: 978-1-462-52499-0.
- [38] Website: Cent (music). [https://en.wikipedia.org/wiki/Cent_\(music\)](https://en.wikipedia.org/wiki/Cent_(music)), March 2023.
- [39] N. Giordano. Explaining the railsback stretch in terms of the inharmonicity of piano tones and sensory dissonance. *The Journal of the Acoustical Society of America*, 138(4):2359–2366, 2015.
- [40] J. H. Eggen and A. J. M. Houtsma. The pitch perception of bell sounds. *Annual Progress Report*, 21, 1986.
- [41] Website: Virtual pitch and stretch tuning. <https://www.hibberts.co.uk/virtual-pitch-and-stretch-tuning/>, March 2023.
- [42] R. Shrivastava, M. Randall, and D. A. Eddins. Pitch strength of normal and dysphonic voices. *The Journal of the Acoustical Society of America*, 131(3):2261, 2012.
- [43] Website: pitch. <https://asastandards.org/terms/pitch/>, March 2023.
- [44] B. C. J. Moore. *An introduction to the psychology of hearing*. Academic Press, 6th edition, 2013. ISBN: 0-12-398331-2.
- [45] Alain d’Cheveigné. Pitch perception models. In *Pitch. Neural Coding and Perception*, chapter New York. Springer, 2005.
- [46] S. S. Stevens and J. Volkman. A scale for the measurement of the psychological magnitude pitch. *The Journal of the Acoustical Society of America*, 8(185), 1937.
- [47] Website: timbre. <https://asastandards.org/terms/timbre/>, March 2023.

- [48] I. Štepankova. Sensitivity of auditory perception to changes in phase spectrum. *Lekar a technika – Clinician and Technology*, 47(4):122–129, 2017.
- [49] P. Zieliński. Muzyka kwantów (I). *Foton*, 115:17–30, 2011.
- [50] P. Zieliński. Grające kwanty: wykorzystanie zjawisk dźwiękowych w nauczaniu mechaniki kwantowej. *Krakowskie Konwersatorium Fizyczne*, 8 marca, 2012.
- [51] R. Steven Turner. The ohm-seebeck dispute, hermann von helmholtz, and the origins of physiological acoustics. *The British Journal for the History of Science*, 10(1):1–24, 1977.
- [52] Website: William Brock Witherspoon, CS 29500, Digital Sound Modeling, Spring 2003. https://www.classes.cs.uchicago.edu/archive/2003/spring/29500-1/Student_work/Witherspoon/, March 2023.
- [53] Melle Jan Kromhout. The unmusical ear: Georg Simon Ohm and the Mathematical Analysis of Sound. *Isis*, 111(3):471, 2020.
- [54] Emily J. Allen, Philip C. Burton, Cheryl A. Olman, and Andrew J. Oxenham. Representations of pitch and timbre variation in human auditory cortex. *The Journal of Neuroscience*, 37(5):1284–1293, 2017.
- [55] A. Thompson and Barry N. Taylor. Logarithmic quantities and units: level, neper, bel. In *Guide for the Use of the International System of Units (SI)*. NIST Special Publication 811, November 2008.
- [56] Website: Bell. <https://encyklopedia.pwn.pl/haslo/bel;3875677.html>, March 2023.
- [57] Website: Conversion of sound units. <http://www.sengpielaudio.com/calculator-soundlevel.htm>, February 2022.
- [58] Hugo Fastl and Eberhard Zwicker. *Psychoacoustics. Facts and Models*. 2007. ISBN: 978-3-540-23159-2.

- [59] Geogr A. Gescheider. Psychophysical Scaling. *Annual Review of Psychology*, 39:169–200, 1988.
- [60] R. Sonnenschein. Summaries of the bibliographic material available in the field of otolaryngology: the functional examination of hearing: deaf-mutism and the education of te deaf annual review of the literature. *Archives of Otolaryngology*, 5:48–78, 1927.
- [61] David Cahan. *Hermann Von Helmholtz and the Foundations of Nineteenth-Century Science*. University of California Press, 1993. ISBN: 978-0-520-07353-1.
- [62] H. L. F. von Helmholtz. *Die Lehre von den Tonempfindungen als physiologische Grundlage fur die "Theorie der Musik"*. Braunschweig: Friedrich Vieweg, 1863. ISBN: 978-3-534-23890-4.
- [63] George Mather. *Foundations of Perception*. Psychology Press, 2006. ISBN: 978-1-84169-491-2.
- [64] Thomas M. Talavage, Martin I. Sereno, J. R. Melcher, P. J. Ledden, Bruce R. Rosen, and Andres M. Dale. Tonotopic organization in human auditory cortex revealed by progressions of frequency sensitivity. *Journal of Neurophysiology*, 91:1282–1296, 2004.
- [65] Charles Lightfoot Barnes. *Practical Acoustics*. Macmillan, 1897.
- [66] Richard M. Warren. *Auditory Perception. An Analysis and Synthesis*. Cambridge University Press, 3rd edition, 2008. ISBN: 978-0-521-87213-8.
- [67] Elizabeth S. Olson, Hendrikus Duifhuis, and Charles R. Steelec. Von Békésy and cochlear mechanics. *Hearing Research*, 293(1-2):31–43, 2012.
- [68] Luis Robles and Mario A. Ruggero. Mechanics of the mammalian cochlea. *Physiological Reviews*, 81(3):1305–1352, 2001.
- [69] M. Magnasco. A wave traveling over a hopf instability shapes the cochlear tuning curve. *Physical Review Letters*, 90:058101, 2003.

- [70] A. J. Hudspeth, Frank Jülicher, and Pascal Martin. A critique of the critical cochlea: Hopf—a bifurcation—is better than none. *Journal of Neurophysiology*, 104(3):1219–1229, 2010.
- [71] H. Kondhalkar and P. Mukherji. A novel algorithm for speech recognition using tonal frequency cepstral coefficients based on human cochlea frequency map. *Journal of Engineering Science and Technology*, 14:726–746, 2019.
- [72] Stanley A. Gelfand. *Essentials of Audiology*. Thieme Medical Publishers, Inc., New York, 2nd edition, 2001. ISBN: 978-1-56593-979-3.
- [73] Elizabeth S. Olson and C. Elliott Strimbu. Cochlear mechanics: new insights from vibrometry and optical coherence tomography. *Current Opinion in Physiology*, 18:56–62, December 2020.
- [74] J. Dora, D. Wojcieszak, D. Kaczmarek, M. Mazur, and A. Aksenczuk. New theory of acoustic signal detection in the inner ear – an explanation of bifilar structure of the cochlea. *Medical Hypotheses*, 140(109636), July 2020.
- [75] Marcin Majka, Paweł Sobieszczyk, Andrzej Woszczyna, Aleksander Sęk, and Piotr Zieliński. Comment on "human time-frequency acuity beats the fourier uncertainty principle".
- [76] G. S. Thekkadath and M. Spanner. Comment on "human time-frequency acuity beats the fourier uncertainty principle". *Physical Review Letters*, 114:069401, 2015.
- [77] Marcin Majka, Paweł Sobieszczyk, Robert Gębarowski, and Piotr Zieliński. Submillisecond acoustic pulses: effective pitch and weber-fechner law in discrimination of duration times. arXiv:1404.6464v2.
- [78] Marcin Majka, Paweł Sobieszczyk, Robert Gębarowski, and Piotr Zieliński. Submillisecond acoustic pulses: effective pitch and weber-fechner law in discrimination of duration times, 2014.
- [79] Hermann von Helmholtz. *On the sensations of tone as a physiological basis for the theory of music*. Aberdeen: The University Press, 4th edition, 1895.

- [80] Georg von Békésy. *Experiments in Hearing*.
- [81] Gnanasenthil Ganesh, Venkata Subramanian Srinivasan, and Sarayu Krishnamurthi. A model to demonstrate the place theory of hearing. *Advances in Physiology Education*, 40:191–193, 2016.
- [82] Aleksander Sęk. Percepcja dźwięku. *Forum Psychologiczne*, Tom 5, nr 1:5–26, 2000.
- [83] Punita G. Singh and Ira J. Hirsh. Influence of spectral locus and f0 changes on the pitch and timbre of complex tones. *The Journal of the Acoustical Society of America*, 92:2650, 1992.
- [84] W. Heisenberg. Über den anschaulichen inhalt der quantentheoretischen kinematik und mechanik. *Zeitschrift für Physik*, 43(3-4):172–198, 1946.
- [85] David J. Griffiths and F. Schroeter Darrell. *Wstęp do mechaniki kwantowej*. Wydawnictwo Naukowe PWN, 2021.
- [86] G. S. Ohm. Über die Definition des Tones, nebst daran geknüpfter Theorie der Sirene und ähnlicher tonbildender Vorrichtungen. *Annalen der Physik und Chemie*, 1843.
- [87] M. V. Laitinen, S. Dish, and V. Pulkki. Sensitivity of human auditory to changes in phase spectrum. *Journal of the Audio Engineering Society*, 61(11):860–877, 2013.
- [88] A. Miśkiewicz. *Wysokość, głośność i barwa: badanie wymiarów wrażeńiowych dźwięków muzycznych*. Akad. Muzyczna Im. Fryderyka Chopina, 2002.
- [89] Website: Innerfidelity headphone measurements explained. <https://www.stereophile.com/content/innerfidelity-headphone-measurements-explained>, April 2022.
- [90] Sho Araiba. Current diversification of behaviorism. *Perspectives on Behavior Science*, 43(1):157–175, 2019.

- [91] Simon Grondin. Timing and time perception: A review of recent behavioral and neuroscience findings and theoretical directions;. *Attention, Perception, Psychophysics*, 72(3):561–582, 2010.
- [92] Website: Creative Headset hs-600. <https://pclab.pl/art18853-4.html>, wrzesień 2020.
- [93] I. Lahdelma and T. Eerola. Cultural familiarity and musical expertise impact the pleasantness of consonance/dissonance but not its perceived tension. *Scientific Reports*, 10(1):8693, 2020.
- [94] M. Sorati and D.M. Behne. Audiovisual modulation in music perception for musicians and non-musicians. *Frontiers in Psychology*, 11(1094), 2020.
- [95] S. Chen, Y. Zhu, R. Wayland, and Y. Yang. How musical experience affects tone perception efficiency by musicians of tonal and non-tonal speakers? *Plos One*, 15/5(e0232514), 2020.
- [96] S.-K Lee, H. Lee, J. Back, K. An, Y. Yoon, K. Yum, S. Kim, and S.-U. Hwang. Prediction of tire pattern noise in early design stage based on convolutional neural network. *Applied Acoustics*, 172(15):107617, 2021.
- [97] W.K Chung, T.M Leung, C.K. Chau, and S.K. Tang. Comparing the effects of visibility of different neighborhood greenery settings on the preference ratings and noise annoyance responses to road traffic noises. *Applied Acoustics*, 169(1):107474, 2020.
- [98] K.W. Ma, C.M. Mak, and H.M. Wong. Acoustical measurements and prediction of psychoacoustic metrics with spatial variation. *Applied Acoustics*, 168:107450, 2020.
- [99] Y. Li, L. Xin, X. Liu, and X. Zhang. Dual anechoic chamber setup for over-the-air radiated testing of 5g devices. *IEEE Transactions on Antennas and Propagation*, 68(3):2469–2474, 2020.
- [100] Z. Xiong and Z. Chen. Modeling of tapered anechoic chambers. *2018 ATMA Proceedings*, 2018. 8604178.

- [101] R. Jacques, B. Albrecht, H.P. Schade, D. De Vries, and F. Melchior. Multichannel source directivity recording in an anechoic chamber and in a studio. *Forum Acusticum Budapest 2005: 4th European Congress on Acoustics 2005*, pages 479–484, 2005.
- [102] Website: Komora bezechowa AGH. <http://www.lat.agh.edu.pl/pomieszczenia/komora-bezechowa/>, June 2020.
- [103] Website: Head and Torso Simulator (hats), type 5128. <https://www.bksv.com/en/products/transducers/ear-simulators/head-and-torso/hats-type-5128>, June 2020.
- [104] Website: Product data. High-frequency Head and Torso Simulator Type 5128 Family. <https://www.bksv.com/-/media/literature/Product-Data/bp2573.aspx>, June 2020.
- [105] Karolina Martinson, Marcin Majka, Jarosław Rubacha, Tadeusz Kamisiński, and Piotr Zieliński. The dynamic response of the basal membrane to short acoustic pulses. In *Proceedings of 2018 Joint Conference - Acoustics, Acoustics 2018*, pages 184–187. Polish Acoustical Society, 2018.
- [106] Website: <https://www.grasacoustics.com/products/measurement-microphone-sets/product/140-46ae>. <https://www.grasacoustics.com/products/measurement-microphone-sets/product/140-46ae>, April 2023.
- [107] S. S. Stevens and J. Volkman. The relation of pitch to frequency: A revised scale. *The American Journal of Psychology*, 53(3):329–353, 1940.
- [108] D. D. Greenwood. The mel scale’s disqualifying bias and a consistency of pitch-difference equisections in 1956 with equal cochlear distances and equal frequency ratios. *Hearing Research*, 103(1-2):199–224, 1997.
- [109] Andrzej Miśkiewicz and Andrzej Rakowski. A psychophysical pitch function determined by absolute magnitude estimation and its relation to the musical pitch scale. *The Journal of the Acoustical Society of America*, 131(1):987–992, 2012.
- [110] E. Zwicker. Subdivision of the audible frequency range into critical bands (frequenzgruppen). *The Journal of the Acoustical Society of America*, 33(2):248, 1961.

- [111] William P. Shofner and George Selas. Pitch strength and Stevens's power law. *Frontiers in Psychology*, 64(3):437–450, 2002.
- [112] Brian C. J. Moore and Brian R. Glasberg. The Relationship Between Frequency Selectivity and Frequency discrimination for subjects with unilateral and bilateral cochlear impairments. *Auditory Frequency Selectivity*, 119:407–417, 1986.
- [113] Jos J. Eggermont. *Noise and the Brain*, chapter 9. Academic Press, Calgary, Alberta, Canada, 2014. ISBN: 978-0-19-983895-8.
- [114] Website: GENELEC 8020C. <https://www.genelec.com/previous-models/8020c>, March 2022.
- [115] Website: Conversion of sound units. Beyerdynamic DT 770. <https://www.soundguys.com/beyerdynamic-dt-770-studio-80ohm-review-15348/>, August 2022.
- [116] Website: Conversion of sound units. Beyerdynamic DT 880. <https://pl.manuals.plus/beyerdynamic/beyerdynamic-dt-880-250-ohm-pro-semi-open-studio-headphones-black-complete-features-instructions-manual>, August 2022.
- [117] Website: Conversion of sound units. Beyerdynamic DT 990. <https://sklepmuzyczny.pl/produkt/2532-beyerdynamic-dt-990-pro-250-ohm.html>, August 2022.

Spis rysunków

1	Widok ucha ludzkiego w przekroju [15].	18
2	Ucho środkowe i wewnętrzne. Rysunek uproszczony [17].	19
3	Schematyczne przedstawienie błony podstawnej [18].	20
4	Schematyczna mapa reprezentacji częstotliwości na błonie podstawnej pokazująca, że część błony podstawnej, która reaguje na dźwięk, zależy od danej częstotliwości dźwięku [19].	20
5	Obwiednia fali biegnącej utworzona przez ton 200 Hz. Kształt obwiedni jest opisany przez zbiór chwilowych lokalizacji (na rysunku pokazane cztery) śledzonych przez falę przemieszczającą się wzdłuż błony podstawnej [19].	21
6	Przykłady widm wywołujących wrażenie tej samej wysokości dźwięku. A: Widmo tonu prostego; B: Widmo wielotonu harmonicznego wywołującego tę samą wysokość dźwięku co widmo A; C: najwyższa harmoniczna ma najwyższą amplitudę jednak wysokość dźwięku nie zmienia się; D: Najniższa (podstawowa) harmoniczna ma amplitudę zerową lecz periodyczność sygnału jest taka sama - nie zmienia się wysokość dźwięku; E: Żaden z interwałów pomiędzy składowymi harmonicznymi nie odpowiada jego wysokości dźwięku, która jednak (dla niektórych słuchaczy) jest taka sama jak pozostałych tonów; F: Dla każdej składowej harmoniczej dodano sub-harmoniczne, dzieląc częstotliwość składowej przez kolejne liczby naturalne, następnie dla każdej sub-harmonicznej zwiększono odpowiadający przedział histogramu [45].	33
7	Badanie przeprowadzone przez autorów [4] wraz z opisem.	45
8	Efektywna Wysokość Dźwięku. Figury geometryczne bez wypełnienia dotyczą dźwięków sinusoidalnych, figury geometryczne z wypełnieniem - dźwięków czysto gaussowskich. MM, PS oraz PZ - to uczestnicy badania [78].	47

9	Cosinusoida z obwiednią gaussowską o malejącym parametrze szerokości σ . W dziedzinie czasu i częstości. Szerokość krytyczną σ_c zdefiniowano jako wartość przy której maksimum widma dochodzi do zera.	51
10	Wartość oczekiwana pulsacji w widmie sygnału cosinusoidalnego z obwiednią gaussowską o czasie trwania (szerokości) σ	52
11	Iloczyn niepewności $\Delta\omega \cdot \Delta t = 2\pi \cdot \Delta f \cdot \Delta t$ jako funkcja czasu trwania impulsu σ wg wzoru (19 - 21).	53
12	Wartość oczekiwana dla sygnałów cosinusoidalnego (linia zielona) i sinusoidalnego (linia pomarańczowa) uzyskana przy użyciu widma mocy.	54
13	Iloczyn niepewności czasu i pulsacji w funkcji σ dla sygnałów cosinusoidalnych (linia zielona) i sinusoidalnych (linia pomarańczowa).	55
14	Wartość oczekiwana pulsacji z widm sygnału cosinusoidalnego (linia zielona) i sinusoidalnego (linia pomarańczowa) z obwiednią gaussowską o szerokości σ otrzymana z użyciem widma mocy	56
15	Iloczyny nieokreśloności czasu i pulsacji gaussowskich sygnałów cosinusoidalnych i sinusoidalnych otrzymane z użyciem widm mocy.	57
16	Poglądowy rysunek przedstawiający pomiar dźwięku rejestrowanego przez mikrofon oraz odbiór tego samego dźwięku przez człowieka. Geometria głowy, rozstaw uszu oraz kształt tułowia powodują wiele odbić i opóźnień [89].	66
17	Schemat wpływu różnych czynników związanych z geometrią głowy, ucha i przewodu słuchowego na wzmocnienie lub osłabienie odbioru tonów prostych w zależności od ich częstotliwości [89].	67
18	Interfejs programu testującego użytego podczas badań behawioralnych przeprowadzonych w Politechnice Krakowskiej.	68
19	Odpowiedzi uzyskane podczas testu behawioralnego dla wybranych uczestników (uczestnicy 1, 2, 4, 5 - niemuzycy, 3 - muzyk, którego odpowiedź przedstawiona jest w tabeli 3).	71
20	Słuchawki Creative, HS-600 wraz z mikrofonem [92].	72
21	Stanowisko pomiarowe.	73

22	Odpowiedzi uczestników badania behawioralnego dla grupy muzyków, która dokonywała rozróżnienia pomiędzy dwoma dźwiękami na podstawie wysokości dźwięku.	74
23	Odpowiedzi uczestników badania behawioralnego dla grupy muzyków, która dokonywała rozróżnienia pomiędzy dwoma dźwiękami na podstawie barwy dźwięku.	74
24	Odpowiedzi uczestników badania behawioralnego dla grupy niemuzyków - pierwsza zauważalna różnica.	75
25	Wykres przedstawiający zebrane średnie z rysunków 22, 23 oraz 24.	75
26	Wykres przedstawiający średnie z odpowiedzi muzyków dot. wysokości dźwięku oraz dopasowaną prostą $y = 0,039(0,003) \cdot x + 0,019(0,001)$ (dopasowanie liniowe a)).	76
27	Wykres przedstawiający średnie z odpowiedzi muzyków dot. wysokości dźwięku oraz dopasowaną prostą dla pierwszych 17 punktów: $y = 0,052(0,004) \cdot x + 0,0171(0,0001)$ (dopasowanie liniowe b)).	76
28	Wykres przedstawiający średnie z odpowiedzi muzyków dot. wysokości dźwięku oraz dopasowaną funkcją kwadratową $y = -0,062(0,01) \cdot x^2 - 0,075(0,09) \cdot x + 0,016(0,001)$ (dopasowanie 2 st.).	77
29	Zestawienie wszystkich dopasowań do średniej najmniejszej zauważalnej różnicy czasu trwania impulsu $\Delta\sigma_{DL}$ w funkcji czasu trwania σ stwierdzonej przez muzyków na podstawie wysokości dźwięku.	77
30	Wykres przedstawiający średnie z odpowiedzi muzyków dot. barwy dźwięku oraz dopasowaną funkcją liniową $y = 0,031(0,004) \cdot x + 0,026(0,002)$ (dopasowanie liniowe a)).	78
31	Wykres przedstawiający średnie z odpowiedzi muzyków dot. barwy dźwięku oraz dopasowaną funkcją liniową dla pierwszych 17 punktów. $y = 0,055(0,006) \cdot x + 0,022(0,001)$ (dopasowanie liniowe b)).	78
32	Wykres przedstawiający średnie z odpowiedzi muzyków dot. barwy dźwięku oraz dopasowaną funkcję kwadratową $y = -0,13(0,02) \cdot x^2 - 0,10(0,01) \cdot x + 0,019(0,001)$ (dopasowanie 2 st.).	79
33	Zestawienie wszystkich dopasowań do średniej najmniejszej zauważalnej różnicy czasu trwania impulsu $\Delta\sigma_{DL}$ w funkcji czasu trwania σ stwierdzonej przez muzyków na podstawie barwy dźwięku.	79

34	Wykres przedstawiający średnie z odpowiedzi niemuzyków oraz dopasowaną funkcją liniową $y = 0,045(0,007) \cdot x + 0,026(0,003)$ (dopasowanie liniowe a)).	80
35	Wykres przedstawiający średnie z odpowiedzi niemuzyków dla pierwszych 17 punktów oraz dopasowaną funkcję liniową $y = 0,094(0,005) \cdot x + 0,017(0,001)$ (dopasowanie liniowe b)).	80
36	Wykres przedstawiający średnie z odpowiedzi niemuzyków oraz dopasowaną funkcję kwadratową $y = -0,21(0,3) \cdot x^2 - 0,17(0,02) \cdot x + 0,013(0,002)$ (dopasowanie 2 st.).	81
37	Zestawienie wszystkich dopasowań do średniej najmniejszej zauważalnej różnicy czasu trwania impulsu $\Delta\sigma_{DL}$ w funkcji czasu trwania σ stwierdzonej przez niemuzyków.	81
38	Sztuczna głowa: Head and Torso Simulator (HATS), typ 5128 [103].	83
39	Przekrój przez prawe ucho sztucznej głowy w symulatorach HATS. Zawiera małżowinę, kanał uszny oraz symulator ucha niedroźnego [104].	84
40	Konfiguracja 1, mikrofon - głośnik [źródło własne].	85
41	Konfiguracja 2, sztuczna głowa - głośnik [źródło własne].	86
42	Konfiguracja 3, sztuczna głowa - słuchawki [źródło własne].	86
43	Sygnal gaussowski (po lewej) oraz sygnal cosinusoidalny (po prawej)	87
44	Przebiegi impulsów generowanych w środowisku MatLab oraz odpowiedzi głośnika na zaproponowane wymuszenie: a) generowany sygnal cosinusoidalny z obwiednią gaussowską, b) odpowiedź głośnika na wymuszenie sygnałem a), c) generowany sygnal gaussowski, d) odpowiedź głośnika na wymuszenie sygnałem c). Czas trwania wszystkich impulsów wynosi 13,60 ms [105].	87
45	Wartości częstotliwości uzyskane z widma mocy dla słuchawek Creative wyliczone jako maksimum widma mocy oraz wartość oczekiwana	89
46	Mikrofon G.R.A.S 46AE użyty podczas badań w komorze bezechowej.	91
47	Maksymalna wartość dla każdego σ z zakresu 1 – 25 zarejestrowana przez mikrofon.	91
48	Przybliżenie (zoom) na widmo mocy dźwięku $n = 9$, ($\sigma = 0,204ms$).	93

51	Wysokość dźwięku w skali MIDI obliczona z wartości oczekiwanej częstotliwości widm otrzymanych dla sztucznej głowy jako funkcja czasu trwania impulsu σ . Linia przerywana odpowiada wartości 69, tj. klawiszowi a razkreślne (a^1).	93
49	Widma impulsów gaussowskich otrzymane z mikrofonu rejestrującego krótkie impulsy z głośnika dla σ z zakresu od 1 do 12.	94
50	Wartość oczekiwana częstości.	95
52	Podwójnie logarytmiczny wykres zależności przedstawionej na rysunku 50a. Dopasowana prosta ma wzór: $-0,99(0,01) \cdot x + 2,08(0,01)$. Nachylenie bardzo bliskie jedności wskazuje na proste prawo potęgowe z wykładnikiem bliskim -1	96
53	Wykres przedstawiający prawo Webera-Fechnera dla sztucznej głowy bez czynnika dp_{lim}	97
54	Prawo psychofizyczne z dokładnością do dp_{lim} otrzymane z danych z rys. 50a wg wzoru (54), wraz z dopasowaniami liniowymi dla dwóch segmentów. Segment dla małych σ : $\Delta\sigma_{DL} = 0,000532(6,7 \cdot 10^{-6})\sigma + 0,0000151(6,0 \cdot 10^{-7})(ms)$, segment dla $\sigma > 0,25ms$: $\Delta\sigma_{DL} = 0,00111(0,00001)\sigma - 0,000112(4 \cdot 10^{-6})(ms)$	97
55	$F_r(\sigma)$ obliczone dla słuchawek Creative wg wzoru (67).	98
56	Zależność najmniejszej zauważalnej różnicy czasu trwania czasu impulsu $\Delta\sigma_{DL}(\sigma)$ od czasu trwania σ otrzymana na podstawie wzoru (67) i (62), z dokładnością do czynnika $\Delta T_{lim(\sigma)}$ przyjętego za stały, otrzymane z widm mocy zarejestrowanych przez sztuczną głowę przy użyciu tych samych słuchawek co w badaniach behawioralnych (Creative).	99
57	Logarytmiczno-logarytmiczny wykres wysokości dźwięku przypisanej przez autorów pracy [3] sygnałom gaussowskim o różnym czasie trwania. Czas trwania impulsu jest tu oznaczony Δt , zaś częstość f_{EP} została obliczona na podstawie zgłaszanych przez testowanych interwałów muzycznych względem ustalonego dźwięku odniesienia.	105

58	Próba odtworzenia częstości odpowiadającej efektywnej wysokości dźwięku na podstawie linowego dopasowania b) do danych otrzymanych przez niemuzyków wg równania (88).	108
59	Próba odtworzenia częstości odpowiadającej efektywnej wysokości dźwięku na podstawie dopasowania 2 st. do danych otrzymanych przez niemuzyków, wg równania (90).	109
60	Próba odtworzenia częstości odpowiadającej efektywnej wysokości dźwięku na podstawie linowego dopasowania b) do danych otrzymanych przez muzyków na podstawie różnic w wysokości dźwięku porównywanych impulsów wg wzoru (88).	110
61	Próba odtworzenia częstości odpowiadającej efektywnej wysokości dźwięku na podstawie dopasowania drugiego stopnia do danych otrzymanych przez muzyków na podstawie różnic w wysokości dźwięku porównywanych impulsów wg wzoru (90)	111
62	Porównanie prawa psychofizycznego dla czasu trwania impulsu otrzymanego dla niemuzyków (zob. rys. 35) z wyprowadzonym na podstawie wzoru (54) analogiem tego prawa dla sztucznej głowy przy założeniu jednakowego nachylenia liniowych dopasowań do otrzymanych wyników.	112
63	Porównanie prawa psychofizycznego dla czasu trwania impulsu otrzymanego dla muzyków stwierdzających różnice trwania impulsów na podstawie wysokości dźwięku (zob. rys. 26) z wyprowadzonym na podstawie wzoru (54) analogiem tego prawa dla sztucznej głowy przy założeniu jednakowego nachylenia liniowych dopasowań do otrzymanych wyników.	113
64	Maksymalna wartość dla każdego σ z zakresu 1 – 25 zarejestrowana przez sztuczną głowę	122
65	Widma impulsów gaussowskich otrzymane dla sztucznej głowy (HATS) słuchającej impulsów z głośnika dla σ z zakresu od 1 do 12. Próbkowanie impulsów: 44100Hz.	123
66	Widma impulsów gaussowskich otrzymane dla sztucznej głowy (HATS) słuchającej impulsów z głośnika dla σ z zakresu od 13 do 24. Próbkowanie impulsów: 44100Hz.	124

67	Widma impulsów gaussowskich otrzymane dla sztucznej głowy (HATS) słuchającej impulsów z głośnika dla wybranych σ . Próbkowanie impulsów: 44100Hz.	125
68	Słuchawki Bayerdynamic DT770, [115].	126
69	Słuchawki Bayerdynamic DT880, źródło: [116].	127
70	Słuchawki Bayerdynamic DT990, źródło [117].	127
71	Maksymalna wartość dla każdego σ z zakresu 1 – 50 zarejestrowana przez HATS słuchającą sygnałów z 4 różnych typów słuchawek. . .	128
72	Wartość oczekiwana dla każdego σ z zakresu 1 – 25 zarejestrowana przez HATS słuchającą sygnałów z 4 różnych typów słuchawek. . .	129
73	Maksymalna wartość dla każdego σ z zakresu 1 – 100 zarejestrowana przez sztuczną głowę słuchającą sygnałów ze słuchawek Creative. . .	130
74	Widma impulsów gaussowskich otrzymane dla sztucznej głowy (HATS) słuchającej impulsów dla σ z zakresu od 1 do 12 dla słuchawek Creative, model Hs-600. Próbkowanie impulsów: 44100Hz.	131
75	Widma impulsów gaussowskich otrzymane dla sztucznej głowy (HATS) słuchającej impulsów dla σ z zakresu od 13 do 24 dla słuchawek Creative, model Hs-600. Próbkowanie impulsów: 44100Hz.	132
76	Widma impulsów gaussowskich otrzymane dla sztucznej głowy (HATS) słuchającej impulsów ze słuchawek DT770 σ z zakresu od 1 do 12. Częstotliwość próbkowania 44100 Hz.	133
77	Widma impulsów gaussowskich otrzymane dla sztucznej głowy (HATS) słuchającej impulsów ze słuchawek DT770 σ z zakresu od 13 do 24. Częstotliwość próbkowania 44100 Hz.	134
78	Widma impulsów gaussowskich otrzymane dla sztucznej głowy (HATS) słuchającej impulsów ze słuchawek DT880 σ z zakresu od 1 do 12. Częstotliwość próbkowania 44100 Hz.	135
79	Widma impulsów gaussowskich otrzymane dla sztucznej głowy (HATS) słuchającej impulsów ze słuchawek DT880 σ z zakresu od 13 do 24. Częstotliwość próbkowania 44100 Hz.	136
80	Widma impulsów gaussowskich otrzymane dla sztucznej głowy (HATS) słuchającej impulsów ze słuchawek DT990 σ z zakresu od 1 do 12. Próbkowanie impulsów: 44100Hz.	137

81	Widma impulsów gaussowskich otrzymane dla sztucznej głowy (HATS) słuchającej impulsów ze słuchawek DT990 σ z zakresu od 13 do 24. Próbkowanie impulsów: 44100Hz.	138
82	Widma impulsów gaussowskich otrzymane dla sztucznej głowy (HATS) słuchającej impulsów ze słuchawek Creative σ z zakresu od 1 do 12. Częstotliwość próbkowania dźwięków 192000 Hz	140
83	Widma impulsów gaussowskich otrzymane dla sztucznej głowy (HATS) słuchającej impulsów ze słuchawek Creative σ z zakresu od 13 do 24. Częstotliwość próbkowania dźwięków 192000 Hz	141
84	Widma impulsów gaussowskich otrzymane dla sztucznej głowy (HATS) słuchającej impulsów ze słuchawek DT770 σ z zakresu od 1 do 12. Częstotliwość próbkowania dźwięków 192000 Hz.	142
85	Widma impulsów gaussowskich otrzymane dla sztucznej głowy (HATS) słuchającej impulsów ze słuchawek DT770 σ z zakresu od 13 do 24. Częstotliwość próbkowania dźwięków 192000 Hz.	143
86	Widma impulsów gaussowskich otrzymane dla sztucznej głowy (HATS) słuchającej impulsów ze słuchawek DT990 σ z zakresu od 1 do 12. Częstotliwość próbkowania dźwięków 192000 Hz.	144
87	Widma impulsów gaussowskich otrzymane dla sztucznej głowy (HATS) słuchającej impulsów ze słuchawek DT990 σ z zakresu od 13 do 24. Częstotliwość próbkowania dźwięków 192000 Hz.	145

Spis tablic

1	Struktura narządu słuchu [8].	16
2	Lista behawioryzmów i ich charakterystyka [90].	65
3	Przykładowa odpowiedź uczestnika badań behawioralnych (nr 3 na wykresie 19). W kolumnie pierwszej przedstawiony jest numer dźwięku referencyjnego, w kolumnie drugiej numer pary dla której zarejestrowano różnicę w wysokości dźwięku, oraz w kolumnie trzeciej numer pary dla której zarejestrowana została różnica w barwie dźwięku.	70
4	Tabela przedstawiająca pierwsze 3 wyraźne maksima dla dźwięków o podanych σ zmierzonych za pomocą mikrofonu rejestrującego sygnał z głośnika.	92
5	Przykłady rozwiązań równań dla funkcji $f(\sigma)$ sugerowanych przez wyniki obecnych badań behawioralnych	102
6	Przykłady rozwiązań równań dla funkcji $f(\sigma)$ sugerowanych przez wyniki obecnych badań behawioralnych	102
7	Używane w muzyce nazwy dźwięków odpowiadających przedziałowi częstości przedstawionych na rys. 50a.	106
8	Tabela przedstawiająca pierwsze 4 wyraźne maksima dla dźwięków o $\sigma = 2, \sigma = 5, \sigma = 10$ i $\sigma = 15$ zmierzonych w konfiguracji: HATS - głośnik.	122

Osiągnięcia naukowe

Publikacje

- K. Martinson, P. Zieliński, „Statistical Properties of Tonal and Atonal Music Pieces”, Acta Physica Polonica B Proceedings Supplement, Vol. 13 (2020), No 4, pp. 823-828.
- K. Martinson, J. Rubacha, T. Kamisiński, P. Zieliński, „The Perception of Short-time Acoustic Pulses”, Proceedings of inter.noise 2019 noise control for a better environment, no. 1500, Madrid, 16-19 czerwca 2019
- K. Martinson, M. Majka, J. Rubacha, T. Kamisiński, „Prawo Webera-Fechnera w ocenie czasu trwania krótkich impulsów dźwiękowych przez narząd słuchu”, Zagadnienia aktualnie poruszane przez młodych naukowców 11 Warszawa, Creativetime ISBN: 978-83-63058-72-2
- K. Martinson, P. Zieliński, „The Statistical Properties of Tonal and Atonal Music Pieces”, Proceedings of the 14th International Symposium on CMMR, p. 815-822, Marseille, France, 14-18 października 2019
- K. Martinson, M. Majka, J. Rubacha, T. Kamisiński, „Percepcja ultrakrótkich impulsów dźwiękowych: modele błony podstawnej”, Zagadnienia aktualnie poruszane przez młodych naukowców 13 Warszawa, Creativetime, s. 17-21, ISBN: 978-83-63058-83-8
- M. Majka, K. Martinson, T. Kamisiński, P. Zieliński, „Duration discrimination of the ultrashort acoustic pulses”, Advances in Acoustic 2018 Gdańsk 2018, s. 68-72, ISBN 978-83-951890-0-5
- K. Martinson, M. Majka, J. Rubacha, T. Kamisiński, P. Zieliński, „The dynamic response of the basal membrane to short acoustic pulses”, Proceedings of Joint Conference Acoustics 2018, ISBN: 978-1-5386-7114-6

Nagrody

- I-INCE Young Professional Grant, 23-26 sierpnia 2020, Seul, Korea.
- Pierwsza nagroda za najlepszą prezentację plakatową podczas 45tego Zjazdu Fizyków Polskich w Krakowie, 13-18 września 2019.
- Nagroda za najlepsze dydaktyczne wystąpienie podczas konferencji „Dynamical Properties of Solids XXXVI International Symposium Cracow” za wystąpienie zatytułowane: „Weber-Fechner law in short time acoustic pulses”, 27-31 sierpnia 2017.

Wystąpienia konferencyjne

- 26 marca 2022, wystąpienie pt.: „Percepcja krótkich sygnałów dźwiękowych” zorganizowane przez Polskie Towarzystwo Fizyczne - oddział w Białymstoku, Białystok, Polska.
- 9 grudnia 2021, wystąpienie pt.: „Fizyczne i fizjologiczne aspekty percepcji skrajnie krótkich sygnałów dźwiękowych”, Zakład Fizyki Biomedycznej, Uniwersytet Warszawski, Warszawa, Polska.
- 7-11 września 2020, wystąpienie pt.: „Absolute Treshold Needed to Define the Frequency of the Sound” podczas „5th Grandmaster Early-Career Workshop in Physics”, Praga, Czechy.
- 23-26 sierpnia 2020, wystąpienie pt.: „The Behavioural Test and Artificial Head’s Responses Involving the Short Sounds with the Gaussian Envelope” podczas „INTER-NOISE 2020”, „The 49th International Congress and Exhibition on Noise Control Engineering” w Seulu: e-congress, Korea.
- 14-18 października 2019, prezentacja plakatu pt.: „The Statistical Properties of Tonal and Atonal Music Pieces” podczas „14th International Symposium on Computer Music Multidisciplinary Research (CMMR)”, Marsylia, Francja.

- 13-18 września 2019, prezentacja plakatu pt.: „Statystyczne własności czasowych przebiegów parametrów fizycznych w muzyce”, podczas „45 Zjazdu Fizyków Polskich” w Krakowie.
- 1-7 września 2019, wystąpienie pt.: „Statistical Methods as Applied to Time Series of Physical Parameters in Music” podczas „4th Grandmaster Early-Career Workshop in Physics”, Split, Chorwacja.
- 16-19 czerwca 2019, wystąpienie pt.: „The perception of Short-time Acoustic Pulses” podczas INTER-NOISE 2019, „the 48th International Congress and Exhibition on Noise Control Engineering” w Madrycie, Hiszpania.
- 16th maja 2019, wystąpienie pt.: „Prawo Zipfa na przykładzie utworów muzycznych: tonalnego i atonalnego”, podczas „Krakowsko-Katowickiego Seminarium”, Kraków, Polska.
- 5-7 października 2018, Udział w „II Konferencji Doktorantów PAN”, Raszyn-Falenty, Warszawa.
- 11-14 września 2018, wystąpienie pt.: „The dynamic response of the basal membrane to short acoustic pulses”, podczas „Acoustic 2018”, Ustka, Polska
- 29th maja 2018 - wystąpienie podczas Seminarium Oddziału Fizyki i Materii Skondensowanej pt.: „Fizyczne i fizjologiczne aspekty percepcji krótkich sygnałów akustycznych: badania behawioralne i eksperymenty w komorze bezechowej”, Kraków, Polska.
- 13-16 maja 2018 - prezentacja plakatu pt.: „The physics and physiology of perception of short acoustic pulses” podczas „The Mediterranean Society of Otolology and Audiology (MSOA) 16th International Meeting in Jerusalem”, Jerozolima, Izrael.
- 21 kwietnia 2018 wystąpienie pt.: „Percepcja ultrakrótkich impulsów dźwiękowych: modele błony podstawnej”, podczas „VI edycji konferencji Dokonania Naukowe Doktorantów - Creative Time”, Kraków, Polska.

- 27-31 sierpnia 2017 udział w konferencji: „Dynamical Properties of Solids XXXVI International Symposium Cracow” oraz wystąpienie pt.: „Weber-Fechner law in short time acoustic pulses”, Kraków, Polska.
- 3-6 lipca 2017 Udział w konferencji „Multiscale phenomena in molecular matter (MULTIS)” oraz prezentacja plakatu pt.: „Time scale effects in perception of acoustic signals”, Kraków, Polska.
- 2 czerwca 2017, udział w „XII Rzeszowskiej Konferencji Młodych Fizyków” na Uniwersytecie Rzeszowskim oraz wystąpienie pt.: „Zachowanie Prawa Webera-Fechnera w ocenie czasu trwania dźwięków milisekundowych”, Rzeszów, Polska
- 20 kwietnia 2017, prezentacja plakatu pt.: „Prawo Webera-Fechnera w percepcji bardzo krótkich impulsów dźwiękowych” podczas „XIV Katowicko-Krakowskiego seminarium Fizyki Fazy Skondensowanej” na Uniwersytecie Jagiellońskim, Kraków, Polska.
- 22 listopada 2016, wystąpienie podczas Seminarium Oddziału Fizyki i Materii Skondensowanej pt.: „Porównanie statystycznych i korelacyjnych własności utworów solowych tonalnych i atonalnych”, Kraków, Polska.