Harmonia nie tkwi w liczbach[1]. O pitagorejczykach, strojach i zgodnych współbrzmieniach

Krzysztof Guczalski

I. Wprowadzenie

Według legendy 2500-letnia tradycja postrzegania związku pomiędzy muzyką a matematyką, czy też muzyką a liczbą, jako czegoś szczególnie istotnego i głębokiego ma swe źródło w przypadkowej przechadzce Pitagorasa nieopodal kuźni.

Choć 2500 lat to niewątpliwie dużo, warto zaznaczyć, że istnieje skojarzenie z muzyką jeszcze bardziej pradawne i powszechne, mianowicie to, które wiąże ją z emocjami. Skąd możemy wiedzieć, że jest ono bardziej długowieczne? Otóż skojarzenie pomiędzy muzyką a emocjami występuje właściwie we wszystkich kulturach, również w takich, które nazywamy pierwotnymi i które są oparte wyłącznie na tradycji oralnej[2]. Można zatem domniemywać, że obecne było również w kulturach prehistorycznych, kiedy muzyka rozwijała się jeszcze tylko na podstawie tradycji oralnych. Natomiast powiązanie muzyki z matematyką niewątpliwie należy do tradycji uczonych, zaawansowanych i wymaga wysokiego poziomu refleksji na temat muzyki. Jego pojawienie się było możliwe dopiero dzięki pewnym odkryciom na temat dźwięków. Otóż według wspomnianej legendy Pitagoras przechodząc obok kuźni miał zwrócić uwagę na współbrzmienia młotów kowalskich. Niektóre z nich brzmiały harmonijnie i zgodnie, niektóre wręcz przeciwnie – ostro i niezgodnie. Miało to skłonić Pitagorasa do próby odpowiedzi na pytanie, z czego wynikają różnice. Z podjętych badań wynikło słynne odkrycie, że współbrzmienia zgodne, harmonijne charakteryzują się prostymi stosunkami liczbowymi odnoszącymi

1 Powyższy tytuł jest w pewnym sensie polemiką z tytułem – ale jedynie z tytułem (por. przypis 11) – niedawno wydanej, popularnej książki pt. *Harmonia tkwi w liczbach. Muzyka i matematyka*, której autorami są Javier Arbonés i Pablo Milrud, Warszawa 2014.

2 Por. np. Piotr Podlipniak, *Uniwersalia muzyczne*, Poznań 2007, s. 179 oraz 187n.

się do obiektów wydających dźwięki – istotna była więc na przykład waga młotów kowalskich i ich proporcja, ale nie siła uderzenia. Gdy badania te zostały odniesione do strun różnej długości, a tej samej grubości i równo naprężonych, proporcje te przybrały szczególnie prostą formę: 2 : 1, 3 : 2, 4 : 3.

Jest empirycznie potwierdzonym faktem, że współbrzmienie charakteryzujące się najprostszą proporcją 2 : 1, które nazywamy oktawą, jest tak dalece zgodne, że około 75% ludzi bez profesjonalnego wykształcenia muzycznego na pytanie „Ile Pan/Pani słyszy dźwięków?" odpowiada: „Jeden". W odniesieniu do współbrzmienia, które charakteryzuje się proporcją 3 : 2 (kwinta), ciągle jeszcze około 50% osób odpowiada, że słyszy jeden dźwięk[3].

Czy faktycznie wynika stąd, że muzyka ma jakiś szczególny, głęboki związek z matematyką, liczbą i proporcjami? Zdania na ten temat są podzielone. Odpowiedź pozytywna jest jądrem tradycji pitagorejskiej, która była w kulturze europejskiej bardzo silna, najsilniejsza w średniowieczu, ale znajdująca swoich obrońców – lub przynajmniej sympatyków – w każdej epoce, także współcześnie. Sam fakt publikacji niniejszego tomu jest tego widomym przejawem. Z drugiej strony pojawiały się też głosy sceptyczne. Tytułem przykładu zestawmy jedynie dwie wyraziste wypowiedzi. Pierwsza to słynne stwierdzenie Leibniza, że „muzyka jest ukrytym ćwiczeniem arytmetycznym, w którym dusza nie wie, że liczy"[4]. W innym zaś miejscu znajdujemy jeszcze takie, bardziej rozbudowane, sformułowanie:

> Muzyka nas urzeka, mimo że jej piękno polega tylko na zgodności liczb i na rachubie, której sobie nie uprzytamniamy, a której dusza nie przestaje prowadzić, to znaczy rachubie zachodzących z pewnymi przerwami uderzeń i drgań ciał wydających dźwięki[5].

W całkiem przeciwnym duchu wypowiada się Eduard Hanslick, słynny XIX-wieczny austriacki krytyk muzyczny i być może najczęściej cytowany estetyk muzyki w dziejach. Jego rozprawa *O pięknie w muzyce*, która ukazała się w 1854 roku, jest fundamentem i początkiem nowoczesnej estetyki muzycznej. W tymże dziele Hanslick stwierdza:

---

3 Tak wynika z ustaleń Carla Stumpfa (1848-1936), niemieckiego psychologa i teoretyka muzyki. Por. np. *Encyklopedia muzyki*, red. Andrzej Chodkowski, Warszawa 1995, hasła „Konsonans", s. 460 oraz „Stopliwość dźwięków", s. 846.

4 *„Musica est exercitium arithmeticae occultum nescientis se numerare animi."* Gottfried Wilhelm Leibniz, *Epistolae ad diversos*, List 154 do Christiana Goldbacha z kwietnia 1712, [w]: *Philosophische Werke*, red. Ernst Cassirer, Lipsk 1906, t. II, s. 132.

5 Gottfried Wilhelm Leibniz, *Zasady natury i łaski oparte na rozumie*, przeł. Stanisław Cichowicz, [w]: *Wyznanie wiary filozofa*, Warszawa 1969, s. 293.

> (…) muzyczne piękno i z matematyką nie ma również nic do czynienia. (…) w utworze muzycznym, najlepszym czy najlichszym, nie znajdziemy nic obliczonego za pomocą matematyki. Twory fantazji nie są przykładami rachunkowymi[6].

II. Teoria pitagorejska

Aby przybliżyć się do rozstrzygnięcia tego sporu, trzeba uświadomić sobie, co właściwie jest jego przedmiotem. W tym celu warto wrócić do Pitagorejczyków i przypomnieć, że ich ambicje były daleko większe niż sama konstatacja prostej prawidłowości, iż np. proporcja 2 : 1 odpowiada zgodnemu współbrzmieniu. Sednem doktryny pitagorejskiej, tak jak się ją typowo przedstawia, jest teza, że liczba stanowi pierwszą zasadę bytu. Można sobie wyobrazić, że jedną z inspiracji do takiego poglądu mogły być opisane powyżej odkrycia akustyczne, które zapewne musiały wywrzeć na pitagorejczykach spore wrażenie. Bowiem po jednej stronie mamy coś, co można w dzisiejszym języku nazwać przyjemnością estetyczną – mianowicie poczucie zgodnego i harmonijnego współbrzmienia dwóch dźwięków – a więc rzecz, mogłoby się wydawać, ulotną i subiektywną. I jeśli nagle okazało się, że coś tak trudno uchwytnego udaje się wyjaśnić w sposób całkowicie jednoznaczny za pomocą prostych liczb, mogło to dodać pitagorejczykom odwagi[7], aby sądzić, że w ogóle cały świat, cały byt można wyjaśnić za pomocą liczb i proporcji. To właśnie jest sedno doktryny pitagorejskiej, które często bywa ujmowane w formie następujących, zwięzłych sentencji:

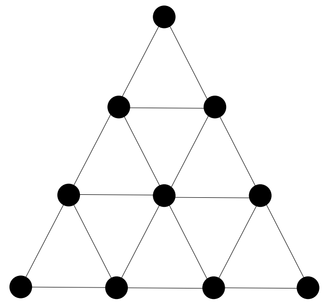
> Co jest najmądrzejsze? Liczba.
> Co jest najpiękniejsze? Harmonia.
> Czym jest cały świat? Liczbą i harmonią[8].

Pitagorejczycy mieli przy tym określone wyobrażenie, jakie liczby wyrażają ową uniwersalną harmonię. To te, które są zawarte w magicznej figurze pitagorejczyków, tzw. tetraktysie:

6 Eduard Hanslick, *O pięknie w muzyce*, przeł. Stanisław Niewiadomski, Warszawa 1903, s. 107-109.

7 A ich osiągnięcia na polu geometrii mogły być drugą, równoległą inspiracją.

8 Ta tradycyjnie powtarzana formuła nie jest chyba oryginalnym sformułowaniem Pitagorasa czy jego szkoły. Wydaje się jednak, że dobrze oddaje ona ducha tradycji pitagorejskiej. O jej popularności świadczy fakt, że można ją znaleźć na wielu popularnych portalach edukacyjnych, m.in. www.math.edu.pl/liczby, www.serwismatematyczny.pl/static/st_liczby_wstep.php.



Przykład 1. *Pitagorejski tetraktys.*

To geometrycznie ukazana liczba 10 – uznawana za doskonałą – przedstawiona jako suma: 1 + 2 + 3 + 4. Właśnie te cztery liczby – reprezentowane przez kolejne rzędy trójkąta – miały uczestniczyć we wszelkich idealnych harmoniach. W muzyce dają one proporcje 2 : 1, czyli oktawę, 3: 2, tzn. kwintę oraz 4 : 3, tzn. kwartę (pozostałe proporcje nie dają już nowych współbrzmień: 4 : 2 = 2 : 1 to znowu oktawa, 4 : 1 to interwał dwóch oktaw).

W tym przypadku ta jedność liczby i harmonii przejawia się w najbardziej egzemplaryczny, wyrazisty sposób. Co ponadto ma być za jej pomocą wyjaśnione? Po pierwsze, według doktryny pitagorejskiej, ruchy planet i gwiazd, których odległości i prędkości także miałyby podlegać zasadzie pozostawania w prostych – a przez to harmonijnych – proporcjach. Co więcej, skoro dźwięk bierze się z ruchu (to także było jednym z odkryć pitagorejczyków), tak więc i poruszające się planety oraz gwiazdy rozbrzmiewają harmonią sfer. Po drugie, również świat wewnętrzny, tzn. dusza ludzka jest swego rodzaju harmonią czy też zawiera w sobie harmonię – taka koncepcja obecna była nie tyko u pitagorejczyków: była w ogólności popularna w starożytności. Tak więc muzyka poprzez swoje idealne proporcje i harmonie z jednej strony odzwierciedla kosmos, z drugiej strony – harmonię duszy ludzkiej; jest egzemplarycznym ucieleśnieniem obu sfer i niejako łącznikiem między nimi – symbolem ich jedności: kosmosu zewnętrznego i wewnętrznego. Muzyce przypada zatem nadzwyczaj wyeksponowana, doniosła i wyróżniona pozycja. Taka jej koncepcja jest czymś bardzo pociągającym i trudno oprzeć się jej urokowi. Ale już Arystoteles z nią polemizował, a ściślej mówiąc, tylko z fragmentem tej koncepcji, mianowicie z tezą o harmonijnych współbrzmieniach wytwarzanych przez poruszające się planety (sfery niebieskie). W traktacie *O niebie* pisał:

> Gdy słońce, księżyc oraz gwiazdy tak niesłychanie liczne i olbrzymie poruszają się z nadzwyczajną szybkością, jest niemożliwe – mówią – aby nie wywoływały dźwięku (…). Opierając się (…) na fakcie, że prędkość

> gwiazd, która zależy od ich odległości, jest proporcjonalna do akordów muzycznych, twierdzą, że dźwięk wydawany przez ruch kołowy gwiazd jest harmonijny. Ponieważ jednak wydaje się czymś anormalnym, że nie słyszymy tego dźwięku, tłumaczą ten fakt tym, że dźwięk jest od naszego urodzenia tuż przy nas, wskutek czego nie odróżniamy go od jego przeciwieństwa – milczenia. (…) Takie wywody (…) wykazują wyczucie podniosłości (…); jest jednak niemożliwe, aby były uzasadnione[9].

Mówiąc o „poczuciu podniosłości" – Arystoteles docenia atrakcyjność takiej teorii, lecz mimo to nie daje się jej uwieść. W dalszym wywodzie zauważa, że dźwięki wytwarzane przez planety – ze względu na ich ogromne rozmiary i prędkości – byłyby tak głośne, iż, podobnie jak dźwięk pioruna, który potrafi roztrzaskiwać kamienie, roztrzaskałyby wszystko, co istnieje i w ogóle nic by się nie ostało. Można powiedzieć, że Ziemia pod wpływem tego strasznego, kosmicznego huku rozpadłaby się na kawałki. A ponieważ Ziemia istnieje, to znaczy, że żadnego tego rodzaju hałasu czy dźwięku nie ma.

Nie musimy rozstrzygać, czy ta argumentacja Arystotelesa jest przekonująca, choć dziś pewnie sceptycznie – także z innych powodów – odnosimy się do tezy o dźwiękach wytwarzanych przez planety. Ale wydaje się, że teoria pitagorejska może się ostać nawet wówczas, gdy pominiemy tę tezę. Możemy po prostu mówić w bardziej abstrakcyjny sposób o prostych proporcjach liczbowych i twierdzić, że są one wszędzie obecne i wyjaśniają różne dziedziny bytu, różne fragmenty uniwersum. W przypadku planet możemy rozważać proporcje pomiędzy ich prędkościami i odległościami wzajemnymi. Nie musimy wykonywać kolejnego kroku i wnioskować o dźwięku związanym z takimi proporcjami.

Natomiast całkowicie kluczowe dla podtrzymania doktryny pitagorejskiej jest twierdzenie, że wszelkie istotne przejawy bytu – w szczególności ruchy sfer niebieskich i muzyka – dają się objaśnić za pomocą prostych (idealnych) proporcji liczbowych. Jeśli chcielibyśmy pominąć to określenie i mówić wyłącznie o proporcjach liczbowych, cała teoria zostałaby pozbawiona jakiejkolwiek istotnej treści. Wszak jakieś proporcje liczbowe niczego nie wyróżniają i niczego nie wyjaśniają. Cokolwiek, nawet całkiem przypadkowe rzeczy, można pomierzyć i następnie uzyskane liczby zestawić w jakieś proporcje. W szczególności upadłoby wyjaśnienie różnicy pomiędzy konsonansami i dysonansami, pomiędzy współbrzmieniami zgodnymi i harmonijnymi *versus* ostrymi i niezgodnymi (wszak jedne i drugie charakteryzują się pewnymi proporcjami), a więc wyjaśnienie, na czym polega i jak się tłumaczy liczbowo harmonia. A to właśnie wyjaśnienie jest sednem doktryny pitagorejskiej.

---

9 Arystoteles, *O niebie*, przeł. Paweł Siwek, Warszawa 1980, s. 78-79.

III. Harmonia sfer i problemy z kalendarzem

Porzucając zatem tezę o dźwięczącej muzyce sfer, możemy teraz zadać pytanie, czy teorię pitagorejską – mówiącą o idealnej harmonii wyrażonej przez proste proporcje liczbowe – można podtrzymać w odniesieniu do sfer niebieskich (ruchów planet) i do muzyki.

Jeśli chodzi o pierwszą dziedzinę, wystarczy choćby jeden przykład, aby rozwiać wszelkie złudzenia: stosunek długości obiegu Ziemi wokół Słońca (tzn. roku astronomicznego) do długości obrotu Ziemi wokół własnej osi (tzn. jednej doby). 365 : 1 nie jest szczególnie prostą proporcją. Gdybyż jednak można liczyć przynajmniej na taką proporcję! Jak wiadomo, już w starożytności zorientowano się, że proporcja ta jest bliższa 365,25 : 1 = 1461 : 4. Z tego powodu kalendarz juliański, wprowadzony w roku 46 p.n.e. w miejsce wcześniej obowiązującego sposobu liczenia czasu, przewidywał co cztery lata rok przestępny o długości 366 dni. Łącznie więc każdy cykl czteroletni miał 1461 dni, co odpowiadało powyższej proporcji.

Jednak i ona okazała się tylko kolejnym przybliżeniem. Bowiem stopniowo, na przestrzeni wieków, kalendarz coraz bardziej oddalał się od pór roku. W XVI wieku różnica wynosiła już ponad 10 dni. W roku 1582 bullą papieża Grzegorza XIII wprowadzony został kalendarz gregoriański. Różni się on od juliańskiego wyłącznie sposobem ustalania lat przestępnych. Są to nadal lata podzielne przez 4, z wyjątkiem tych podzielnych przez 100, a niepodzielnych przez 400. Oznacza to, że – inaczej niż w kalendarzu juliańskim – nie były przestępnymi lata 1700, 1800 i 1900, natomiast przestępne pozostały lata 1600 i 2000 (oraz oczywiście wszystkie inne podzielne przez 4). Co oznacza to dla odpowiedniej proporcji? Należy teraz wziąć pod uwagę okres 400 lat. W takim okresie w kalendarzu juliańskim byłoby 100 lat przestępnych. W kalendarzu gregoriańskim – o 3 lata mniej, tzn. 97 lat. Według kalendarza juliańskiego 400 lat miałoby 100 x 1461 = 146 100 dni. Wg kalendarza gregoriańskiego o 3 dni mniej, tzn. 146 097 dni. Otrzymujemy zatem proporcję 146 097 (dni) : 400 (lat) = 365,2425 : 1.

Jednak i ta proporcja jest tylko kolejną aproksymacją. Bowiem długość roku wyrażona w dniach jest lepiej przybliżona liczbą 365,2422. Oznacza to, że według kalendarza, którego używamy, średnia długość roku jest w przybliżeniu o 0,0003 dnia tzn. o ok. 26 sekund dłuższa niż roku astronomicznego. W konsekwencji nasz kalendarz spóźnia się o 1 dzień na przestrzeni około 3300 lat (juliański spóźniał się o 1 dzień na 128 lat). Za tyle mniej więcej lat trzeba będzie znowu pominąć jeden rok przestępny (być może będzie można uznać rok 5000 za zwykły, a nie przestępny?). A proporcja w kalendarzu znowu stanie się mniej idealna, co tylko po raz kolejny pokaże, że ta proporcja, o którą naprawdę chodzi – opisująca ruchy planet (w tym wypadku Ziemi) jest całkowicie, beznadziejnie nieregularna.

IV. Muzyka – poszukiwanie stroju idealnego

Zwróćmy się teraz ku muzyce. Czy przynajmniej ona tłumaczy się prostymi proporcjami, o których mówi doktryna pitagorejska? Inaczej mówiąc, czy przynajmniej w jej zakresie można odnaleźć idealnie harmonijny system współbrzmień? Zakładamy, że konstrukcję systemu rozpoczynamy od pewnego dźwięku, którego częstotliwość[10] traktujemy jako jednostkę wyjściową i oznaczamy liczbą 1 (dla naszych rozważań jest obojętne czy ta jednostka to np. 200, 300 czy 440 herców). Wówczas dźwięk o oktawę wyższy będzie opisany przez liczbę 2 (dwukrotność częstotliwości wyjściowej), dźwięk o kwintę wyższy – przez liczbę 3/2 (trzy drugie) itd. W ten sposób liczba określająca dany dźwięk będzie jednocześnie liczbą wyrażającą jego proporcję do dźwięku wyjściowego, opisanego liczbą 1.

a) System pitagorejski

Pierwszym sposobem podejścia do konstrukcji idealnego systemu dźwiękowego jest oparcie się – wyjściowo – wyłącznie na wyróżnionych, idealnych współbrzmieniach pitagorejskich, tzn. oktawie i kwincie (kwarta jest różnicą oktawy i kwinty, tak więc oparcie się na dwóch ostatnich pozwala nam uzyskać także kwartę). Jeśli zatem konstrukcję systemu rozpoczynamy od dźwięku opisanego liczbą 1, powinniśmy do niego włączyć także harmonijnie z nim współbrzmiące dźwięki 3/2 oraz 2, położone odpowiednio o kwintę i oktawę wyżej[11]. Ograniczenie się do tych jedynie dźwięków (ewentualnie

10 Jako że dziś, jak wiadomo, rozumujemy raczej w kategoriach częstotliwości, a nie długości struny, ponieważ ta pierwsza bardziej bezpośrednio przynależy do charakterystyki samego dźwięku, a nie przedmiotów go wytwarzających.

11 To, co teraz nastąpi, to przedstawienie systemów pitagorejskiego, naturalnego, temperacji średniotonowej i równomiernej. Materiał ten przynależy w zasadzie do wiedzy encyklopedycznej, dlatego nie ma potrzeby, aby szczegółowo go dokumentować odniesieniami do literatury przedmiotu. W elementarnym zakresie informacje na ten temat można znaleźć chociażby w *Encyklopedii muzyki*, *op. cit.*, hasła „System dźwiękowy", s. 869-870 oraz „Temperacja", s. 896-897. Natomiast w wersji znacznie szerszej i szczegółowej – w książce *Systemy i skale muzyczne* Mieczysława Drobnera, Kraków 1982, zwłaszcza s. 33-37 (system pitagorejski), s. 43-46 (system naturalny), s. 69-74 (temperacja średniotonowa i równomierna). (Dziękuję za tę wskazówkę bibliograficzną anonimowemu recenzentowi mojego artykułu). Znacznie nowsza książka Marka Pilcha i Marka Toporowskiego *Dawne temperacje. Podstawy akustyczne i praktyczne wykorzystanie*, Katowice 2014 jest w większym stopniu zorientowana na praktyczne kwestie związane ze strojeniem instrumentów. Także wymieniona w przypisie 1 książka *Harmonia tkwi w liczbach. Muzyka i matematyka* przystępnie i kompetentnie przedstawia podstawowe elementy doktryny pitagorejskiej oraz wyzwanie jakie stanowią dla niej dylematy związane z różnymi systemami strojenia (por. zwłaszcza s. 25). Jednocześnie należy podkreślić, że celem niniejszego rozdziału nie jest jakikolwiek historycznie kompletny przegląd systemów i strojów w różnych kulturach muzycznych. Prezentowane przypomnienie wybranych informacji na ten temat ma służyć jedynie jako tło i podstawa dla bardziej wiarygodnego zakwestionowania doktryny pitagorejskiej w rozdziale V.

powtórzonych o oktawę wyżej lub niżej) nie wystarczy naturalnie do układania jakiejkolwiek satysfakcjonującej muzyki. Kolejne dźwięki możemy jednak uzyskać poprzez odpowiednie kombinacje – dodawanie i odejmowanie – oktaw i kwint. W szczególności możemy włączyć do systemu dźwięk o kwintę wyższy od utworzonego wcześniej dźwięku 3/2. Jaki to dźwięk? W stosunku do dźwięku wyjściowego, opisanego liczbą 1, będzie to dźwięk uzyskany przez dwukrotny krok o kwintę w górę: najpierw od 1 do 3/2, później od 3/2 do dźwięku, którego poszukujemy. Można zatem zadać ogólne pytanie, na czym polega dodawanie interwałów.

Jeśli rozważymy, w ogólności, trzy dźwięki o częstotliwościach odpowiednio F_1, F_2, F_3 (przy czym zakładamy, że $F_1 < F_2 < F_3$), wówczas możemy powiedzieć, że interwał pomiędzy F_1 i F_3 jest sumą interwałów od F_1 do F_2 i od F_2 do F_3. Dwa ostatnie charakteryzują się proporcjami $F_2 : F_1 = F_2/F_1$ oraz $F_3 : F_2 = F_3/F_2$. Suma tych interwałów, tzn. interwał pomiędzy F_1 i F_3 – proporcją $F_3 : F_1 = F_3/F_1$. Łatwo zauważyć, że

$$F_3/F_1 = F_3/F_2 \times F_2/F_1$$

Oznacza to, że suma dwóch interwałów, to interwał charakteryzujący się proporcją, która jest iloczynem proporcji interwałów sumowanych. Inaczej mówiąc, aby dodać dwa interwały, należy pomnożyć charakteryzujące je proporcje.

Z kolei można zadać pytanie, jaki dźwięk otrzymamy postępując o interwał charakteryzujący się proporcją *p* od dźwięku o częstotliwości F_1. Jeśli poszukiwaną częstotliwość oznaczymy F_2, wówczas proporcja pomiędzy dwoma dźwiękami to F_2/F_1, która z założenia powinna być równa *p*, tzn.

$$F_2/F_1 = p$$

A zatem:

$$F_2 = F_1 \times p$$

Tak więc aby postąpić o pewien interwał określony proporcją *p* od pewnego dźwięku wyjściowego o częstotliwości F_1, należy tę częstotliwość pomnożyć przez proporcję charakteryzującą odpowiedni interwał.

Wracając do pytania o dźwięk położony o kwintę wyżej niż 3/2, czyli o dwie kwinty wyższy niż wyjściowy dźwięk 1, stwierdzamy, że jest to dźwięk o częstotliwości opisanej przez liczbę:

$$3/2 \times 3/2 = 9/4 = 2\tfrac{1}{4}$$

Mnożenie to możemy zinterpretować na dwa równoważne sposoby opisane powyżej: albo jako dodawanie dwóch kwint, albo jako o krok o kwintę do góry od dźwięku 3/2.

Uzyskany dźwięk określony jest przez liczbę większą niż 2, a zatem jest oddalony od wyjściowego dźwięku (1) o więcej niż oktawę. Od tego dźwięku (9/4) znów wolno nam postąpić o oktawę lub kwintę w dół lub do góry. Wybieramy tę pierwszą możliwość, tzn. krok o oktawę w dół. Skoro dodawanie interwałów, to mnożenie odpowiednich proporcji, zatem odejmowanie – to dzielenie. Krok o oktawę w dół (tzn. odjęcie oktawy) będzie oznaczał podzielenie przez 2. Otrzymujemy więc:

$$9/4 : 2 = 9/8.$$

Ta proporcja nazywana jest całym tonem (a przynajmniej to jedna z możliwych wersji całego tonu) albo inaczej: to dźwięk położony o cały ton powyżej dźwięku wyjściowego, opisanego przez liczbę 1. Dźwięk ten powstał przez kombinację idealnych interwałów pitagorejskich, jako wynik dodania dwóch kwint i odjęcia oktawy. Wydaje się więc, że zasługuje na miejsce w systemie. Możemy teraz – rozpoczynając od niego – opisaną procedurę powtórzyć: postąpić znów o dwie kwinty do góry i oktawę w dół, co będzie równało się kolejnemu krokowi o cały ton w górę. W efekcie uzyskamy dźwięk:

$$9/8 \times 9/8 = 81/64$$

Proporcja ta – choć nadal wynika z dodawania i odejmowania idealnych kwint i oktaw – niewątpliwie przestaje być już prosta, co może niepokoić w kontekście pitagorejskiej zasady prostych proporcji liczbowych, które miałyby wyjaśniać świat – a w szczególności muzykę. Kolejny krok prowadzi do proporcji jeszcze bardziej skomplikowanej:

$$81/64 \times 9/8 = 729/512$$

Jednak najgorsza konsekwencja kontynuacji tej procedury wygląda następująco. Dodanie 6 kolejnych całych tonów daje dźwięk charakteryzujący się proporcją:

$$(9/8)^6 = 9^6/8^6 = 3^{12}/2^{18} = 531\ 441/262\ 144 = 2{,}0272865\ldots$$

a zatem dźwięk trochę tylko wyższy niż oktawa dźwięku wyjściowego. Jeśli przeniesiemy go o oktawę w dół otrzymujemy dźwięk:

$$(9^6/8^6)/2 = 3^{12}/2^{19} = 531\ 441/524\ 288 = 1{,}0136433\ldots$$

oddalony bardzo nieznacznie od dźwięku wyjściowego. Odległość ta, nazywana komatem pitagorejskim, jest równa w przybliżeniu 1/8 całego tonu, albo inaczej 1/4 półtonu. Taki dźwięk jest z jednej strony zbyt odległy od dźwięku wyjściowego, abyśmy nie zauważali różnicy. Gdyby tak było – gdyby np. różnica wynosiła 1/100 półtonu – moglibyśmy uzyskany właśnie dźwięk po prostu utożsamić słuchowo z dźwiękiem wyjściowym. Oczywiście takie utożsamienie byłoby z punktu widzenia zasad pitagorejskich nie do zaakceptowania, skoro nie mielibyśmy idealnej zgodności i żadnych prostych proporcji. Ale w przedstawionej sytuacji mamy problem nie tylko z idealnością, ale i z uchem ludzkim[12]: 1/8 całego tonu to coś, co słyszymy jako rozbieżność, jako dźwięk nieczysty.

Z drugiej strony jeśli dźwięk ten – i interwał dzielący go od dźwięku wyjściowego – włączymy do naszego systemu, wówczas przez kolejne jego dodawanie dostaniemy dziesiątki dźwięków oddalonych zaledwie o 1/8 tonu, których ilość będzie znacznie przekraczała nasze możliwości ich rozróżniania i zapamiętywania. O prostych proporcjach w ramach takiego systemu nie warto nawet wspominać.

Konstrukcja całości systemu, którego typowa postać została przedstawiona w poniższej tabelce, może jedynie potwierdzić te spostrzeżenia (w pierwszej linijce wpisane są tradycyjne, używane dziś nazwy dźwięków, w drugiej odpowiadające im proporcje, w trzeciej odległości pomiędzy poszczególnym stopniami skali):

c	d	e	f	g	a	h	c
1/1	9/8	81/64	4/3	3/2	27/16	243/128	2/1

9/8	9/8	256/243	9/8	9/8	9/8	256/243

Interwały kwarty i kwinty, jako podstawowe i idealne, zostały naturalnie zachowane w pierwotnej postaci. Pozostałe stopnie wynikają z omówionych powyżej kroków o jeden (9/8) i dwa (81/64) całe tony odpowiednio od dźwięku wyjściowego (1) oraz od kwinty (3/2). Proporcja 256/243 = $2^8/3^5$ to pitagorejski półton, którego wielkość wynika z różnic pomiędzy wcześniej skonstruowanymi stopniami skali. Suma dwóch takich półtonów jest nieco mniejsza niż cały ton (9/8).

12 Tutaj i dalej określenie „ucho ludzkie" jest używane jako metonimia ludzkiego słyszenia w ogólności. Naturalnie to, jak słyszymy, nie jest uzależnione wyłącznie od działania samego ucha, ale także od sposobu przetwarzania odbieranych przez nie sygnałów w układzie nerwowym, a w szczególności mózgu. W niniejszym artykule rola tych poszczególnych czynników nie będzie analizowana, możemy więc zbiorczo mówić o naszym słyszeniu jako o „uchu", przeciwstawianym jedynie obiektywnym, fizycznym cechom dźwięku, które najczęściej można jednoznacznie wyrazić liczbowo.

Rozbieżność jest równa – jak się można spodziewać – komatowi pitagorejskiemu opisanemu powyżej.

Poza skomplikowanymi proporcjami, jakimi wyrażane są niektóre stopnie skali (*e*, *a*, *h*) oraz półton, wadą takiego systemu są mocno nieczyste interwały tercji (*c* – *e*) oraz seksty (*c* – *a*) (por. dalsze wyjaśnienia w kolejnym ustępie, dotyczącym systemu naturalnego). Ponadto, gdy kontynuujemy konstrukcję tego systemu w odniesieniu do dźwięków półtonowych, dzielących wszystkie uwidocznione powyżej całe tony, dotrzemy do kwinty (*cis–as* albo *gis–es*), która okaże się znacząco nieczysta, w sposób niemożliwy do zaakceptowania: będzie różnić się od kwinty czystej o opisany powyżej komat pitagorejski wynoszący około 1/8 całego tonu, tzn. 1/4 półtonu.

b) System naturalny

Można teraz zapytać jakie jest źródło tego niepowodzenia w konstrukcji systemu dźwiękowego? Być może jest nim zbyt surowe podejście pitagorejczyków, które dopuszcza tylko oktawę i kwintę jako podstawę systemu dźwiękowego? W istocie okazuje się, że istnieją dwa kolejne interwały charakteryzujące się nadal stosunkowo prostymi proporcjami – 5 : 4 i 6 : 5 – które ciągle współbrzmią stosunkowo zgodnie. Można je rozumieć jako wynik tzw. harmonicznego podziału kwinty. Jej proporcję 3 : 2 możemy wyrazić jako 6 : 4 i następnie pomiędzy 6 i 4 wstawić liczbę pośrednią 5, uzyskując proporcję

$$6 : 5 : 4$$

Kwinta okazuje się więc sumą dwóch – oczywiście nierównych – interwałów: 6 : 5 i 5 : 4, które nazywamy odpowiednio tercją małą i tercją wielką. Ich uzupełnienia do oktawy (tzn. interwały wynikające z ich odjęcia od oktawy, a więc z podzielenia odpowiednich proporcji) to interwały 5 : 3 (2 : 6/5 = 2 × 5/6 = 5/3) oraz 8 : 5 oraz (2 : 5/4 = 2 × 4/5 = 8/5), tzn. odpowiednio seksta wielka i mała. Jak widać, w tych czterech proporcjach występują liczby 5 i 6, które nie mieszczą się w tetraktysie – magicznej figurze pitagorejczyków. Zwykle uważa się, że między innymi dlatego proporcje te nie były przez nich akceptowane. Jednocześnie są one nieporównanie prostsze niż odpowiadające im interwały w systemie pitagorejskim. Przykładowo, tercja (5/4 = 80/64) koresponduje z bardzo zbliżonym interwałem pitagorejskim złożonym z dwóch całych tonów, mianowicie

$$9/8 \times 9/8 = 81/64$$

Różnica wynosi jedynie:

$$81/64 : 80/64 = 81/80 = 1{,}0125$$

Ten z kolei niewielki interwał, nieco mniejszy niż komat pitagorejski i równy w przybliżeniu 1/9 całego tonu, jest nazywany komatem Didymosa albo komatem syntonicznym.

Być może zatem klucz do konstrukcji bardziej satysfakcjonującego systemu muzycznego leży w porzuceniu pitagorejskiego rygoryzmu i w zaakceptowaniu – jako podstawy – innych proporcji poza oktawą i kwintą. Taki ruch może się w istocie przysłużyć pitagorejskiemu dążeniu do prostych proporcji. Bo choć dodatkowe proporcje wyjściowe (5 : 4 i 6 : 5) są trochę mniej elementarne niż sama oktawa i kwinta (2 : 1 i 3 : 2), to jednocześnie, dzięki ich zaakceptowaniu, można uniknąć proporcji skrajnie złożonych, jak np. 81/64 (pitagorejska tercja) czy 27/16 (pitagorejska seksta wielka; podczas gdy teraz seksta wielka wyraża się proporcją 5/3). Podstawowy system dźwiękowy wygląda wówczas następująco:

<table>
<tr><td colspan="2">c</td><td colspan="2">d</td><td colspan="2">e</td><td colspan="2">f</td><td colspan="2">g</td><td colspan="2">a</td><td colspan="2">h</td><td colspan="2">c</td></tr>
<tr><td colspan="2">1/1</td><td colspan="2">9/8</td><td colspan="2">5/4</td><td colspan="2">4/3</td><td colspan="2">3/2</td><td colspan="2">5/3</td><td colspan="2">15/8</td><td colspan="2">2/1</td></tr>
<tr><td></td><td colspan="2">9/8</td><td colspan="2">10/9</td><td colspan="2">16/15</td><td colspan="2">9/8</td><td colspan="2">10/9</td><td colspan="2">9/8</td><td colspan="2">16/15</td><td></td></tr>
</table>

Najbardziej „nierówny" interwał – 15/8 – został wyliczony poprzez złożenie kwinty i tercji: $3/2 \times 5/4 = 15/8$. System ten, nazywany naturalnym, wygląda – w porównaniu z poprzednim – nader obiecująco. Niestety i on prędko odsłania swoje wady. Po pierwsze, pomiędzy dźwiękami d i a, gdzie oczekiwalibyśmy kwinty, mamy interwał:

$$5/3 : 9/8 = 5/3 \times 8/9 = 40/27 = 80/54$$

Natomiast kwinta to $3/2 = 81/54$. Różnica wynosi więc ponownie $81/80 = 1{,}0125$, tzn. znany nam już komat Didymosa. W sytuacji gdy kwinta jest drugim po oktawie najbardziej idealnym interwałem, jej nieczystość – w tym wypadku aż o 1/9 całego tonu – jest oczywiście dużą wadą. Skąd bierze się ta rozbieżność, skoro pomiędzy *c* i *g* mamy kwintę, a dźwięki *d* i *a* są od tamtych odległe o cały ton? Otóż okazuje się, że cały ton pomiędzy *g* i *a* jest… nieco mniejszym całym tonem. To interwał $5/3 : 3/2 = 5/3 \times 2/3 = 10/9$. Różnica pomiędzy nim, a standardowym całym tonem pitagorejskim (9/8) wynosi znowu $9/8 : 10/9 = 9/8 \times 9/10 = 81/80$ tzn. ponownie komat Didymosa (ok. 1/9 całego tonu). W ten sposób dostrzegamy drugą niedogodność systemu naturalnego: całe tony są w nim nierówne. Trzy z nich (*c-d*, *f-g*, *a-h*) mają wielkość 9/8. Dwa (*d-e* oraz *g-a*) – wielkość 10/9. Pokazuje to trzecia linijka powyższej tabelki, w której uwidocznione są interwały pomiędzy kolejnymi stopniami uzyskanej w ten sposób skali.

Półton (*e-f* oraz *h-c*) okazuje się mieć wielkość 16/15, tzn. w istocie nieco więcej niż połowa całego tonu (dzieląc duży, pitagorejski cały ton – tzn. 9/8 = 18/16 – harmonicznie, tak, jak wcześniej dzieliliśmy kwintę na tercję małą i dużą, otrzymujemy dwa nierówne półtony: mniejszy 18/17 i większy 17/16. Półton 16/15 jest naturalnie jeszcze nieco większy.)

c) Temperacja średniotonowa

Przedstawionym powyżej wadom systemu naturalnego (mocna nieczystość jednej kwinty, nierówność całych tonów) miała zaradzić tzw. temperacja[13] średniotonowa wynaleziona w XVI wieku. Jej uzasadnienie można odtworzyć w następujący sposób.

Cały ton powstaje, jak widzieliśmy przy konstrukcji systemu pitagorejskiego, jako złożenie dwóch kwint (które możemy zapisać jako *c – g – d*), przy czym rezultat trzeba odpowiednio przenieść o oktawę w dół. Tercja to w zasadzie suma dwóch całych tonów, a więc 4 kwint (*c – g – d – a – e*), znowu po odpowiednim sprowadzeniu rezultatu do właściwej oktawy. Jeśli – tak jak w systemie pitagorejskim – wszystkie 4 kwinty są naturalne (3/2), otrzymujemy interwał 81/64. Skoro w systemie naturalnym zdecydowaliśmy zastąpić go nieco mniejszą tercją czystą (80/64 = 5/4), oznacza to, że w powyższym ciągu 4 kwint nie wszystkie mogą być naturalne. Faktycznie zidentyfikowaliśmy jedną z nich (*d – a*) jako mocno nieczystą. A ponieważ jest ona jednocześnie inna (mniejsza) niż pozostałe – czyste – kwinty, więc nierówność kwint pociąga za sobą nierówność całych tonów, bo powstają one jako złożenie dwóch kwint.

Stoimy zatem przed zadaniem pozbycia się zbyt wielkiej nieczystości kwinty *d – a* oraz wyrównania wielkości kwint, co jednocześnie da gwarancję równości całych tonów. Jeśli dokonamy tego poprzez zwiększenie kwinty *d – a* do kwinty czystej, uzyskamy ponownie strój pitagorejski z tercją 81/64. Bardziej zgodne z duchem systemu naturalnego będzie zachowanie naturalnej tercji 5/4. Aby to osiągnąć musimy odpowiednio pomniejszyć kwinty w ciągu *c – g – d – a – e*. Zamiast jednak cały nadmiar ich sumy (wielkości komatu Didymosa, tzn. 81/80, ok. 1/9 całego tonu) odejmować od jednej tylko kwinty *d – a* (pozostawiając trzy pozostałe jako czyste), możemy go rozdzielić równomiernie pomiędzy wszystkie cztery kwinty, pomniejszając („temperując") każdą o jedną czwartą komatu Didymosa (ok. 1/36 całego tonu), co może być uznane za rozbieżność, która nie jest bardzo wyraźnie słyszalna i którą można tolerować[14]. W ten sposób jednocześnie: 1) zachowujemy czystość tercji, 2) rozwiązujemy problem nie

13 Tym terminem określano niewielkie, uznawane za dopuszczalne, korekty niektórych interwałów, które miały zapobiec powstawaniu bardziej poważnych niezgodności w innych miejscach stroju/systemu dźwiękowego.

14 Przykładowo w oktawie razkreślnej ta rozbieżność nie przekracza 1,6 herca różnicy częstotliwości. Natomiast różnica wielkości pełnego komatu Didymosa np. dla dźwięku a^1 przekłada się już na różnicę częstotliwości 5,5 Hz. Dlaczego ta pierwsza (poniżej 1,6 Hz) nie wywołuje poczucia ostrości czy dysonansowości, a druga (po-

dającej się zaakceptować kwinty $d - a$, 3) uzyskujemy równość kwint, a w konsekwencji także równość całych tonów, bo wynikają one z sumowania kwint. Jak wyglądają wielkości poszczególnych interwałów wynikające z takiej temperacji?

Suma czterech równych kwint powinna być równa sumie dwóch oktaw oraz czystej tercji, co wyraża się liczbą 2 × 2 × 5/4 = 5. Jeśli liczbę odpowiadającą kwincie oznaczymy *k*, uzyskujemy równość

$$k \times k \times k \times k = k^4 = 5$$

a więc

$k = \sqrt[4]{5} = 1{,}495348781\ldots$ (czwarty pierwiastek z pięciu)

Widać, że liczba ta faktycznie dobrze przybliża kwintę czystą 3/2 = 1,5[15]. Przez odpowiednie wyliczenie możemy się przekonać, że różnica wynosi faktycznie jedną czwartą komatu Didymosa. Różnica ta to

$$3/2 : \sqrt[4]{5} = 3/(2\times\sqrt[4]{5}) = r$$

Gdy ten mały interwał czterokrotnie dodamy do siebie, otrzymamy:

$$r \times r \times r \times r = r^4 = (3/2\times\sqrt[4]{5})^4 = 3^4/(2^4\times(\sqrt[4]{5})^4) = 81/(16\times5) = 81/80$$

czyli dokładnie komat Didymosa – a więc *r* jest dokładnie jego jedną czwartą. Następnie cały ton konstruujemy jako sumę dwóch kwint minus oktawę, tzn.

Cały ton $= \sqrt[4]{5} \times \sqrt[4]{5} : 2 = \sqrt[2]{5}/2$

Tak wyliczony cały ton okazuje się uśrednieniem dwóch nierównych całych tonów obecnych w systemie naturalnym (9/8 i 10/9), skąd bierze się nazwa: temperacja średniotonowa. Różniły się one o komat Didymosa, a więc interwał uśredniony powinien być o pół komatu Didymosa większy od mniejszego całego tonu (10/9) i pół komatu Didymosa mniejszy od większego całego tonu (9/8). Skoro komat Didymosa to 81/80, to jego połowa wyraża się liczbą

$$\sqrt{(81/80)} = \sqrt{81}/\sqrt{(16\times5)} = 9/4\sqrt{5}$$

wyżej 5 Hz) może już takie poczucie wywołać, wyniknie z rozdziału VIII na temat konsonansu i dysonansu.

15 Jednocześnie to przybliżenie jest niestety liczbą niewymierną. Por. dyskusję na ten temat w sekcji d).

Można się teraz przekonać, że cały ton wyrażony liczbą $\sqrt{5}/2$ jest faktycznie o tę połowę komatu Didymosa większy niż 10/9 i o połowę komatu Didymosa mniejszy niż 9/8:

$10/9 \times 9/4\sqrt{5} = 10/4\sqrt{5} = 5/2\sqrt{5} = \sqrt{5}/2$

$9/8 : 9/4\sqrt{5} = 9/8 \times 4\sqrt{5}/9 = 4\sqrt{5}/8 = \sqrt{5}/2$

Z kolei tercja wielka to suma dwóch całych tonów: $\sqrt[2]{5}/2 \times \sqrt[2]{5}/2 = 5/4$, czyli tak, jak należało się spodziewać – wszak wielkość kwinty, a poprzez to i całego tonu, została tak właśnie dobrana, aby tercja pozostała czysta. Następnie kwartę konstruujemy jako dopełnienie kwinty do oktawy tzn.

Kwarta $= 2 : \sqrt[4]{5} = 2/\sqrt[4]{5} = 2 \times 5^{-1/4}$

Wreszcie seksta wielka to kwinta plus cały ton (lub kwarta plus tercja), a septyma to kwinta plus tercja (lub seksta plus cały ton):

Seksta wielka $= \sqrt[4]{5} \times \sqrt[2]{5}/2 = 5^{3/4}/2$
albo
Seksta wielka $= 2/\sqrt[4]{5} \times 5/4 = (5/\sqrt[4]{5})/2 = 5^{3/4}/2$

Septyma $= \sqrt[4]{5} \times 5/4 = 5^{5/4}/4$

co można podsumować w następującej tabelce:

c	d	e	f	g	a	h	c
1/1	$\sqrt{5}/2$	5/4	$2/\sqrt[4]{5}$	$\sqrt[4]{5}$	$5^{3/4}/2$	$5^{5/4}/4$	2/1

$\sqrt{5}/2$	$\sqrt{5}/2$	$8\times5^{-5/4}$	$\sqrt{5}/2$	$\sqrt{5}/2$	$\sqrt{5}/2$	$8\times5^{-5/4}$

Można jeszcze zauważyć, że zarówno kwarta, jak i seksta wielka są w powyższym systemie – podobnie jak kwinta – jedynie o jedną czwartą komatu Didymosa różne od odpowiednich interwałów czystych (są od nich odpowiednio większe): kwarta jest większa od 4/3, a seksta wielka – od 5/3. Można to z łatwością sprawdzić za pomocą odpowiednich obliczeń.

Jak widać, mimo zalet temperacji średniotonowej jej liczbowa forma nie wygląda zachęcająco. Z praktycznego punktu widzenia nie to jest jednak jej głównym mankamentem. Polega on raczej na tym, że kwinta temperacji średniotonowej została ustalona w taki sposób, aby zgadzać się z czystą tercją. W związku z tym, podobnie jak kwinta czysta, nie zgadza się z oktawą: złożenie 12 kwint średniotonowych ponownie wypada w pobliżu 7 oktaw, ale jest od nich tym razem nieco mniejsze (około 1/5 całego tonu). W konsekwencji – podobnie jak w przypadku systemu pitagorejskiego – kontynuacja konstrukcji systemu średniotonowego w odniesieniu do dźwięków półtonowych prowadzi do kwinty (*gis–es*), która okazuje się fatalnie nieczysta i różni się około 1/6 całego tonu od kwinty czystej – z tego powodu zasłużyła na miano wilczej.

d) Temperacja równomierna

Następnie na przestrzeni XVI, XVII i jeszcze XVIII wieku proponowano dziesiątki rozmaitych temperacji czy ich wariantów, tzn. dla różnych celów i w ramach różnych tradycji lokalnych odmiennie ustalano stroje muzyczne. Wreszcie w XVIII zaczęto szerzej stosować tzw. temperację równomierną; choć znana była już wcześniej, z różnych powodów nie cieszyła się uznaniem. Polega ona, jak wiadomo, na podziale oktawy na 12 równych półtonów – następnie z sumowania odpowiedniej liczby półtonów uzyskiwane są wszystkie inne interwały. Jaką proporcją powinien być opisany półton w temperacji równomiernej? Jeśli proporcję tę oznaczymy symbolem p, możemy powiedzieć, że zsumowanie dwunastu takich interwałów powinno dać oktawę czyli 2. Ale sumowanie interwałów to mnożenie odpowiadających im proporcji. Zatem spełnione musi być równanie:

$$p^{12} = 2$$

W konsekwencji $p = \sqrt[12]{2} = 1{,}0594631\ldots$

Zaletą takiego systemu jest, jak wiadomo, fakt, że wszystkie interwały we wszystkich pozycjach (we wszystkich miejscach skali) mają zawsze taką samą wielkość i nie musimy się obawiać pojawienia się w jakimkolwiek miejscu wilczych kwint czy innych mocno nieczystych interwałów. Kwinta, złożona z 7 półtonów, to zawsze: $p^7 = 1{,}498307\ldots$, tercja wielka, złożona z 4 półtonów, to zawsze $p^4 = 1{,}2599\ldots$ itd. Jeśli chodzi o stosunek do interwałów czystych, to kwinta równomiernie temperowana jest od kwinty czystej (3/2 = 1,5) mniejsza tylko o ok. 1/100 całego tonu (i odpowiednio kwarta równomiernie temperowana jest od kwarty czystej tylko o 1/100 całego tonu większa). Zatem wielkość najważniejszych interwałów jest bardzo dobrze przybliżona, dużo lepiej niż np. w temperacji średniotonowej. Nieco gorzej jest z tercją wielką: równomiernie temperowana jest ok. 1/14 całego tonu większa od czystej. Ale przypomnijmy, że np. tercja pitagorejska (81/64) była aż o 1/9 całego tonu większa od czystej.

Jednocześnie trzeba powiedzieć, że taki strój – nieważne jak dobrze przybliżający interwały czyste – byłby dla pitagorejczyków absolutnie nie do przyjęcia. Bowiem wszystkie interwały – poza oktawą – nie tylko nie wyrażają się prostymi proporcjami liczbowymi, ale w pewnym sensie nie wyrażają się… żadnymi proporcjami liczbowymi. To znaczy nie można ich przedstawić jako stosunku jakichkolwiek liczb całkowitych, czy też inaczej mówiąc, jako ułamków prostych. Liczby tego ostatniego rodzaju nazywane są wymiernymi. Natomiast dwunasty pierwiastek z dwóch i jego kolejne potęgi nie mogą być w ten sposób wyrażone, co dość łatwo udowodnić. W konsekwencji są zaliczane do tzw. liczb niewymiernych.

Istnienie tego rodzaju liczb stanowiło w ogólności problem dla pitagorejczyków. W pewnym sensie jest to zrozumiałe: skoro liczby – pojmowane jako liczby całkowite – i ich proporcje miały być kluczem do poznania i opisu świata, trudno zaakceptować, że istnieją rzeczy, których nie da się w ten sposób opisać. Paradoksalnie to, że istnieją wielkości, których nie można wyrazić za pomocą żadnych proporcji liczb całkowitych, wynika z twierdzenia tradycyjnie wiązanego ze szkołą, tzn. z twierdzenia Pitagorasa na temat trójkątów prostokątnych. Wynika z niego, że przekątna kwadratu o boku długości 1 ma długość równą $\sqrt{2}$ – jak łatwo pokazać, to także jest liczba niewymierna.

V. Źródło porażki – nieistnienie

Historię kolejnych prób skonstruowania idealnego systemu dźwiękowego można w pewnym sensie porównać do historii reform kalendarza. W jednym i drugim przypadku można było mieć nadzieję – zawartą w doktrynie pitagorejskiej – na odkrycie idealnej harmonii: w ruchach planet (w tym wypadku Ziemi) i w systemie muzycznym. Zamiast tego w obu przypadkach mieliśmy do czynienia tylko z historią kolejnych przybliżeń i mniej lub bardziej satysfakcjonujących kompromisów. Można teraz zadać pytanie, z czego wynika to niepowodzenie. W przypadku ruchu Ziemi można by powiedzieć – w duchu platońskim – że mamy po prostu do czynienia z niesforną, niedoskonałą empirią, z dziedziną tego, co badane za pomocą mało wiarygodnych zmysłów. W sferze idealnej, sferze „prawdziwych ruchów, (…) prawdziwej prędkości i powolności istotnej, w ich prawdziwej liczbie", którą „można tylko myślą ująć"[16], moglibyśmy liczyć na bardziej harmonijny i doskonały ład.

Czy jednak niepowodzenie w budowie doskonałego systemu muzycznego na tym samym polega? W żadnym razie. Nie wynika ono z niedoskonałości tego, co postrzegane zmysłowo, nie bierze się np. stąd, że realnych, fizycznych instrumentów nie da się – z powodu ich ułomnej materialności – w taki czy inny sposób nastroić. Opisane powyżej poszukiwania systemu muzycznego prowadzone były właśnie za pomocą samego rozumu, a więc w sposób, który powinien zaaprobować Platon. Opierały się wyłącznie na pewnych wyliczeniach matematycznych. Oczywiście istnieją problemy matematyczne, które są jeszcze nierozstrzygnięte i czekają na rozwiązanie. Czy kwestia ide-

16 Platon, *Państwo*, ks. VII, 529D, przeł. Władysław Witwicki, Kęty 2003, s. 237.

alnego systemu dźwiękowego jest tego rodzaju otwartym problemem? Czy po prostu jeszcze nie dość długo, nie dość uporczywie próbowaliśmy skonstruować taki system? Niestety proste wyliczenie pokazuje, że takiego idealnego systemu w muzyce zbudować w ogóle nie można, że jego istnienie jest po prostu niemożliwe, a nie tylko dotychczas nieodkryte.

Jeżeli mianowicie dopuścimy w naszym systemie oktawę i kwintę oraz kombinacje tych interwałów, to musielibyśmy oczekiwać, że można uzyskać idealną zgodność pomiędzy pewną liczbą kwint a pewną liczbą oktaw. Skoro kwinta to 3/2, a oktawa to 2, musielibyśmy oczekiwać, że dla pewnych liczb całkowitych a i b zachodzi równość:

$$(3/2)^a = 2^b$$

Odpowiednio ją przekształcając otrzymujemy:

$$3^a/2^a = 2^b$$
$$3^a = 2^a \times 2^b = 2^{a+b}$$

Oznacza to, że pewna potęga trójki musiałaby być równa pewnej potędze dwójki. Jest to oczywiście niemożliwe, choćby dlatego, że pierwsza jest zawsze liczbą nieparzystą, druga – zawsze liczną parzystą. W związku z tym jest jasne, że idealnego systemu opartego na oktawie i kwincie nie da się zbudować.

Można powiedzieć, że pierwsze niepowodzenie – związane z systemem pitagorejskim – było przypadkiem szczególnym tej właśnie, ogólnej niemożności. To, co w istocie tam czyniliśmy, to składanie kwint i oktaw: posuwaliśmy się o dwie kwinty do góry i jedną oktawę w dół. Po sześciu powtórzeniach tego zabiegu uzyskaliśmy złożenie 12 kwint w górę i 6 oktaw w dół. Następnie wynik (531 441/262 144 = 2,0272865…) przenieśliśmy jeszcze o oktawę w dół (dodając siódmą do wcześniejszych 6 oktaw) i otrzymaliśmy mały interwał, który nazwaliśmy komatem pitagorejskim. Ten interwał, to w istocie różnica pomiędzy sumą 12 kwint ($(3/2)^{12} = 3^{12}/2^{12}$ = 531 441/4 096 = 129,74633789…) i 7 oktaw (2^7 = 128). To, że równość jest nieosiągalna, wynika z powyższego, ogólnego rozumowania.

Konstrukcję systemu równomiernie temperowanego można opisać jako próbę uzgodnienia pojawiającej się tutaj nierówności. Skoro suma 12 kwint przekracza 7 oktaw jedynie o komat pitagorejski (mniej niż 1/8 całego tonu), zatem pomniejszenie kwinty o 1/12 (jedną dwunastą) komatu pitagorejskiego (ok. 1/100 całego tonu[17]) spowodowałoby, że 12 kwint dawałoby w sumie dokładnie 7 oktaw. (Naturalnie najbar-

17 1/12 z 1/8 to dokładnie 1/96. Ale komat pitagorejski to nieco mniej niż 1/8 całego tonu, więc przybliżenie jego jednej dwunastej jako 1/100 całego tonu jest uprawnione.

dziej idealne współbrzmienie – oktawa – nie podlega negocjacjom.) Zamiast proporcji 3/2 poszukujemy zatem odrobinę mniejszej proporcji k, której dwunastokrotne złożenie da nam dokładnie 7 oktaw, tzn. będzie zachodziła równość:

$$k^{12} = 2^7$$

Znaczy to, że

$$k = \sqrt[12]{2^7} = (\sqrt[12]{12})^7$$

Okazuje się, że kwinta w ten sposób uzyskana jest równa kwincie równomiernie temperowanej. Opierając się na takiej kwincie możemy skonstruować resztę systemu: kwartę, jako różnicę pomiędzy oktawą a kwintą, cały ton, jako różnicę między kwintą a kwartą itd. Okaże się, że system ten to po prostu system równomiernie temperowany, uzyskany w nieco inny sposób.

Wracając do kwestii dowodu niemożności konstrukcji idealnego systemu dźwiękowego można dodać: jeśli poza oktawą i kwintą zaraz na wstępie dopuścimy jeszcze inne proporcje, np. odpowiadające tercji i sekście, system będzie dopuszczał konstrukcję jeszcze większej ilości interwałów i o zgodność (idealne zharmonizowanie) wszystkich będzie jeszcze trudniej. Co pozostaje? Wyzbyć się kwinty i zostawić tylko oktawę. Wtedy uzyskiwalibyśmy tylko złożenia oktaw: 2 : 1, 4 : 1, 8 : 1 itd. W ten sposób zachowalibyśmy czystość systemu. Harmonia i zgodność proporcji byłaby idealna, muzyka oparta na takim systemie byłaby skrajnie niesatysfakcjonująca. Należy powątpiewać, że o takiej muzyce mogli myśleć pitagorejczycy, gdy formułowali tezę, że odzwierciedla ona muzykę sfer i całość bytu.

VI. Konsonans – poszukiwanie proporcji idealnej

Czy w związku z tym nie istnieje żaden szczególny, głęboki, tajemny związek pomiędzy muzyką a matematyką, muzyką a liczbą? Poszukując odpowiedzi, można zadać jeszcze następujące pytanie: czy muzyka nie odróżnia się pod tym względem choćby od innych dziedzin sztuki: literatury i malarstwa?

Barwy, które są tworzywem tego ostatniego, mogą być, podobnie jak dźwięki, opisane jako fala, tym razem elektromagnetyczna a nie mechaniczna. Zatem przysługują jej takie atrybuty jak długość czy częstotliwość. Możemy więc, podobnie jak czyniliśmy to w odniesieniu do dźwięków, zestawiać pewne kolory i pytać o proporcję pomiędzy odpowiadającymi im długościami fali czy częstotliwościami. Przykładowo można wskazać, że pomiędzy pewnym odcieniem czerwieni o długości fali 705 nanometrów a pewnym odcieniem błękitu o długości fali 470 nanometrów zachodzi proporcja 705 : 470 = 3 : 2, tzn. taka, jaka charakteryzuje kwintę. Jednak rozważania tego rodzaju nie są zwykle przeprowadzane. W szczególności z faktu, że pewne pary barw charakteryzują się prostymi proporcjami liczbowymi, nie wyprowadza się wniosku o szczególnym pokrewieństwie pomiędzy malarstwem a liczbą. Skąd bierze się ta asymetria pomiędzy muzyką a malarstwem?

Odpowiedź jest dość oczywista: w stosunkach barw nie występuje tak wyrazisty jak w muzyce fenomen konsonansu i dysonansu. Proporcja 3 : 2 bardzo jednoznacznie wyróżnia niektóre współbrzmienia jako zgodne, przeciwstawiane innym, ostrym, a w żaden sposób nie wyróżnia par barwnych. Możemy oczywiście spotkać się ze stwierdzeniami, że pewne kolory dobrze ze sobą współgrają, a inne „się gryzą". Ale określenia takie są niewątpliwie dalece subiektywne i rozbieżne. O tak szerokim, powszechnym konsensusie, z jakim mamy do czynienia w odniesieniu do oktawy i kwinty, czy też nawet w odniesieniu do przeciwstawienia np. tercji i półtonu, w ogóle nie może być mowy.

Może więc specyficzny fenomen konsonansu i dysonansu i jego powiązanie z pewnymi proporcjami liczbowymi, sam w sobie (tzn. niezależnie od kwestii nieistnienia w pełni zgodnego systemu dźwiękowego) jest czymś, co wyróżnia muzykę i wskazuje na jej szczególny związek z liczbą? Innymi słowy, może nawet jeśli harmonia globalna, systemowa, nie daje się w idealny sposób ustalić liczbowo, to przynajmniej harmonia każdorazowego interwału (konsonansu) tkwi w liczbie, mianowicie w jego prostej proporcji? Właściwie bardzo prędko trzeba ten pitagorejski pogląd odrzucić. Przykładowo możemy wskazać proporcję

$$5141 : 3427 = 1{,}500145900204\ldots$$

która jest nie tylko w stosunku do 3 : 2 = 1,5, ale także w stosunku do 9 : 8 czy np. 17 : 12, bardzo skomplikowana i „nieharmonijna". Ale określa ona interwał, który różni się od kwinty o mniej niż jedną… tysięczną całego tonu. Oczywiście nie mamy żadnych szans, aby percepcyjnie wychwycić różnicę i słyszymy ten interwał po prostu jako czystą, konsonującą kwintę. Natomiast dużo prostsze proporcje jak 9 : 8 = 1,125 (cały ton) czy 17 : 12 (tryton), to ostre dysonanse. Jak widać również rozwinięcie dziesiętne całego tonu jest stosunkowo proste, a trochę tylko mniejszego interwału wyrażonego proporcją 10 : 11 = 1,1 – tak samo ostro dysonującego – jest jeszcze prostsze. Natomiast rozwinięcie dziesiętne kwinty równomiernie temperowanej, będące liczbą niewymierną, jest nie tylko nieskończone (podobnie jak proporcji 5141 : 3427), ale jeszcze w dodatku nieokresowe (nie wykazuje żadnej powtarzalności w rozwinięciu dziesiętnym). A kwinta ta jest podobnie zgodna i konsonująca jak tylko o jedną setną tonu większa kwinta czysta.

Skoro zatem proste proporcje liczbowe nie tłumaczą zjawiska konsonansu i nie wyjaśniają różnicy pomiędzy nim a dysonansem, trzeba zadać pytanie, czy dysponujemy alternatywnym wyjaśnieniem tej różnicy, które w jakiś inny sposób wskazywałoby na szczególny związek pomiędzy harmonią a liczbą.

VII. Konsonans a składowe harmoniczne

Kolejne wyjaśnienia zjawiska konsonansu, jakie pojawiły się w XVII i XVIII wieku, odwoływały się do obecności w dźwiękach tzw. składowych harmonicznych, zwanych też alikwotami. Należy krótko przypomnieć to pojęcie. Jak wiadomo, wysokość dźwięku jest określona przez jego częstotliwość, która wynika z częstotliwości drgań obiektu wytwarzającego dźwięk. Okazuje się jednak, że każdy fizyczny obiekt drgający emituje nie tylko dźwięk o pewnej częstotliwości podstawowej, ale jednocześnie dźwięki o innych, zazwyczaj wyższych częstotliwościach, które są nazywane składowymi dźwięku. W wypadku typowych dźwięków używanych w muzyce, wytwarzanych przez drgające struny lub słup powietrza w instrumentach dętych, te dodatkowe częstotliwości są najczęściej całkowitymi wielokrotnościami częstotliwości podstawowej i wówczas nazywane są składowymi harmonicznymi. Jeśli więc przykładowo częstotliwość podstawowa wynosi 100 herców (100 Hz, 100 drgań na sekundę), wówczas jednocześnie emitowane są dźwięki o częstotliwościach 200, 300, 400, 500, 600… herców. Dźwięk podstawowy (w tym wypadku 100 Hz) nazywamy pierwszą składową harmoniczną, jego dwukrotność (200 Hz) – drugą itd. W dalszym ciągu będziemy zajmować się już tylko takim przypadkiem, gdzie wszystkie składowe są harmoniczne, tzn. są całkowitymi wielokrotnościami częstotliwości podstawowej (składowe nieharmoniczne są zawarte np. w dźwiękach dzwonów czy niektórych instrumentów perkusyjnych; takie dźwięki nie będą więc przedmiotem naszych rozważań).

Natężenie[18] wyższych składowych harmonicznych (począwszy od drugiej) może pozostawać w różnych relacjach do natężenia dźwięku podstawowego (tzn. pierwszej składowej). Określając jego natężenie jako 100%, natężenie kolejnych składowych może być wyrażone liczbami: 50%, 25%, 10%, 5%. Ale też np. 60%, 10%, 40%, 5%. Ilość możliwości jest teoretycznie nieograniczona, choć oczywiście najczęściej natężenie pierwszej składowej (dźwięku podstawowego) jest największe. Rozkład i natężenie składowych harmonicznych dźwięku jest najistotniejszym czynnikiem, który decyduje o jego barwie i odróżnia np. dźwięk skrzypiec od dźwięku fletu. W tym ostatnim siła

18 Natężenie to, mówiąc w uproszczeniu, obiektywna, mierzalna fizycznie siła dźwięku. Parametr ten jest główną – choć nie jedyną – cechą dźwięku, która decyduje o jego subiektywnie odczuwanej głośności.

wyższych składowych jest stosunkowo niewielka, co przekłada się na jego gładkie brzmienie, które ktoś mógłby nazwać metalicznym i pustym – w przeciwieństwie do pełniejszego i bardziej soczystego dźwięku skrzypiec, o większym udziale składowych[19]. Istnieją źródła dźwięku o jeszcze mniejszym udziale składowych niż flet: np. niektóre rodzaje piszczałek organowych lub kamerton. Ale dźwięk posiadający wyłącznie częstotliwość podstawową, bez żadnych innych składowych, zwany tonem prostym, praktycznie nie występuje „w naturze", tzn. nie może zostać wytworzony przez fizyczny obiekt drgający. Tony proste można natomiast uzyskać za pomocą elektrycznego obwodu drgającego (generatora elektroakustycznego), tak jak dzieje się to w elektronicznych syntezatorach dźwięku.

Świadomość istnienia składowych harmonicznych pojawiła się stosunkowo wcześnie – właściwie jednocześnie z ustaleniem, że wysokość dźwięku jest bezpośrednio zależna od jego częstotliwości, a tylko pośrednio od właściwości obiektów drgających: długości, grubości czy naprężenia struny, długości piszczałki itd. Powiązanie wysokości dźwięku z jego częstotliwością zwykle przypisuje się Isaacowi Beeckmanowi (1588–1637), Galileuszowi (1564–1642) oraz Marinowi Mersenne (1588–1648). Ten ostatni badał też zjawisko składowych harmonicznych, które na gruncie praktycznym było już wcześniej znane renesansowym konstruktorom organów czy ludwisarzom[20]. Zaś jeszcze bardziej zaawansowane ustalenia na temat składowych harmonicznych są zasługą Josepha Sauveura (1653-1716), który uchodzi czasem za twórcę akustyki jako nowoczesnej dziedziny naukowej.

Wiedzę o składowych harmonicznych wykorzystał Jean-Philippe Rameau (1683-1764) do naturalnego ugruntowania zasad harmonii[21]. Nie zajmował się co prawda wprost wyjaśnieniem zjawiska konsonansu. Raczej interesowało go naturalistyczne uzasadnienie trójdźwięków durowego i molowego i uznał, że istnienie składowych harmonicznych takiego uzasadnienia dostarcza. Jeśli mianowicie

---

19 Dość trudno jest opisać słownie barwy instrumentów. Niektóre określenia mogą być uznane za subiektywne i ukrycie wartościujące. W żadnym razie nie chciałbym obstawać przy użytych tutaj określeniach. Niektórym dźwięk fletu może wydawać się miękki i łagodny, a skrzypiec – ostry. W znaczącym stopniu barwa dźwięku zależy też od rejestru: jest inna dla niskich, a inna dla wysokich dźwięków danego instrumentu.

20 Por. Burdette Green, David Butler, *From acoustics to* Tonpsychologie, [w]: *Cambridge History of Western Music Theory,* red. Thomas Christensen, Cambridge 2008, s. 248-250.

21 Swoją teorię wyłożył po raz pierwszy w roku 1722 w *Traité de l'harmonie réduite à ses principes naturels (Traktat o harmonii sprowadzonej do jej naturalnych zasad)*, a następnie udoskonalał i modyfikował w szeregu rozpraw i esejów publikowanych aż do roku 1762. Nie trzeba dodawać, że w powyższym zdaniu termin „harmonia" odnosi się do reguł budowy i następstwa akordów, tzn. harmonii w techniczno-muzycznym sensie (harmonika modalna, harmonika dur-moll itp.), która zasadniczo nie jest tematem niniejszego artykułu.

oznaczymy, tak jak poprzednio, częstotliwość dowolnie wybranego dźwięku liczbą 1, wówczas jego sześć pierwszych składowych to tony o częstotliwościach:

1 – 2 – 3 – 4 – 5 – 6

A więc składowe czwarta, piąta oraz szósta, tzn. 4 – 5 – 6 tworzą między sobą właśnie trójdźwięk durowy. Wszak proporcja 5 : 4 to tercja wielka, 6 : 5 – to tercja mała, a 6 : 4 = 3 : 2 to kwinta. Tak więc trójdźwięk durowy rozbrzmiewa automatycznie w każdym dźwięku – można powiedzieć, że wytwarza go sama natura.

Przechodząc do kwestii konsonansu możemy powiedzieć: ponieważ trójdźwięk durowy uznajemy za akord konsonujący, jego powyższe uzasadnienie można uogólnić i odnieść do wszelkich konsonansów. Można mianowicie powiedzieć, że konsonanse to te współbrzmienia, które wytwarza sama natura: które w naturalny sposób występują pomiędzy różnymi (łącznie z pierwszą) składowymi harmonicznymi dźwięku. A więc konsonansami byłyby współbrzmienia określone przez proporcje 2 : 1 (oktawa), 3 : 1 (duodecyma, tzn. oktawa plus kwinta), 4 : 1 (dwie oktawy), 5 : 1 (dwie oktawy plus tercja wielka), 6 : 1 (dwie oktawy plus kwinta). W dalszej kolejności: 3 : 2 (kwinta), 4 : 2 (oktawa), 5 : 2 (oktawa plus tercja wielka), 6 : 2 = 3 : 1 (oktawa plus kwinta). Z kolei: 4 : 3 (kwarta), 5 : 3 (seksta wielka), 6 : 3 (oktawa) itd.

Trójdźwięk molowy nie poddaje się, według Rameau, takiemu uzasadnieniu jak durowy. Tłumaczy się on raczej w następujący sposób: to akord tego rodzaju, że wszystkie jego dźwięki mają wspólną składową harmoniczną. Jeśli zapiszemy schematycznie taki trójdźwięk jako sekwencję częstotliwości 10 – 12 – 15 (pierwszy interwał 12 : 10 = 6 : 5 to tercja mała; drugi, 15 : 12 = 5 : 4, to tercja wielka), wówczas ton o częstotliwości wyrażonej liczbą 60 jest szóstą składową pierwszego dźwięku (6 × 10 = 60), piątą – drugiego (5 × 12 = 60) oraz czwartą – trzeciego (4 × 15 = 60).

To uzasadnienie naturalności trójdźwięku molowego także można uogólnić na wszelkie konsonanse. Mianowicie można powiedzieć, że konsonans to współbrzmienie dwóch dźwięków, których pewne składowe się pokrywają. Przykładowo jeśli rozważymy dowolny dźwięk (jego częstotliwość jak zwykle oznaczamy liczbą 1) i jego składowe harmoniczne:

1 – 2 – 3 – 4 – 5 – 6 – 7 – 8...

wówczas składowe dźwięku o oktawę wyższego wyglądają następująco:

2 – 4 – 6 – 8...

Okazuje się, że wszystkie są składowymi dźwięku wyjściowego (o oktawę niższego). Dźwięk o oktawę wyższy jest więc już niejako w całości w nim zawarty. Gdy rozbrzmiewa, nie pojawiają się żadne nowe tony, a jedynie niektóre ulegają wzmocnieniu. W ten sposób tłumaczy się nadzwyczajna zgodność oktawy – tak daleka, że jak pamiętamy z wprowadzenia, około 75% osób bez wykształcenia muzycznego sądzi, że słyszy jeden dźwięk. Z kolei dla interwału kwinty, 3 : 2, uzyskujemy następujące składowe:

2 – 4 – 6 – 8 – 10 – 12 – 14 – 16 – 18…

3 – 6 – 9 – 12 – 15 – 18…

Jak widać co trzecia składowa niższego dźwięku (trzecia, szósta, dziewiąta) pokrywa się z co drugą (drugą, czwartą, szóstą) składową wyższego. Zgodność nie jest już tak idealna, jak w przypadku oktawy, ale nadal znacząca. Dla kwarty pokrywać się będzie co czwarta składowa niższego dźwięku z co trzecią wyższego, a więc zgodność będzie znowu nieco mniejsza niż kwinty.

Takie uzasadnienie pozwala więc nie tylko wyjaśnić konsonansowość pewnych współbrzmień, ale także jego stopień: im więcej wspólnych składowych harmonicznych tym bardziej zgodny, doskonalszy konsonans. Cechy tej nie miało wyjaśnienie poprzednie: zarówno oktawa (2 : 1) jak i kwinta (3 : 2), kwarta itd. są obecne w szeregu harmonicznym każdego dźwięku i jako takie są naturalne („stworzone przez naturę"), nie widać więc powodu, dlaczego któryś miałby być lepszy niż inny.

Można teraz zadać pytanie, na ile zadowalające są te wyjaśnienia? Na wstępie warto zauważyć, że pod względem zakresu wyznaczają one w zasadzie taką samą klasę konsonansów jak pierwotne wyjaśnienie pitagorejskie – jeśli potraktujemy je nieco mniej restryktywnie niż pitagorejczycy, tzn. ograniczymy się do określenia „proste proporcje liczbowe", nie zastrzegając, że muszą ograniczać się do liczb od 1 do 4 obecnych w tetraktysie. Inaczej mówiąc, dokładnie te same interwały okażą się konsonansami w świetle wszystkich trzech wyjaśnień. Interwał jest określony prostą proporcją liczbową (wyjaśnienie pitagorejskie) dokładnie wtedy, gdy jest wyrażony proporcją pomiędzy elementami szeregu składowych harmonicznych, bo szereg ten to właśnie początkowe liczby całkowite: 1, 2, 3, 4… W odniesieniu zaś do wyjaśnienia drugiego, możemy rozumować następująco. Jeśli interwał jest określony proporcją n : k, gdzie n i k są liczbami całkowitymi nie mającymi żadnych wspólnych dzielników (tzn. ułamka n/k nie można już uprościć), wówczas pierwsza wspólna składowa harmoniczna to n × k (a kolejne to 2 × n × k, 3 × n × k itd.) Jest to więc n-ta składowa częstotliwości k i k-ta składowa częstotliwości n. Im mniejsze są zatem liczby n i k, tym wcześniej wspólna składowa wystąpi w szeregu harmonicznym każdego dźwięku i tym więcej składowych będzie wspólnych. Widzieliśmy to już na przykładzie oktawy, kwinty i kwarty powyżej.

Wydaje się więc, że modele konsonansu odwołujące się do istnienia składowych harmonicznych są w pewnym sensie równoważne z wyjaśnieniem pitagorejskim. Czy oznacza to, że niczego nowego one nie wnoszą? Wręcz przeciwnie: można powiedzieć, że dostarczają właśnie pewnego uzasadnienia odwołującego się do natury dźwięków dla tezy pitagorejskiej, która wynikała z obserwacji empirycznych, ale pozostawała w gruncie rzeczy niewyjaśniona (jeśli abstrahować od pewnych metafizycznych spekulacji). Można wszak było zapytać: dlaczego proste proporcje wiążą się z konsonansami? Modele odwołujące się do składowych harmonicznych próbują udzielić odpowiedzi na to pytanie. Niestety ich zbyt wielkie pokrewieństwo z tezą pitagorejską oznacza, że odnoszą się do nich te same zarzuty.

Rozpoczynając od modelu pierwszego: w szeregu harmonicznym występują również następujące proporcje: 7 : 5, 9 : 5, 9 : 8. Ta pierwsza jest dość bliska trytonu, druga to jedna z wersji septymy małej (por. wyjaśnienia w przypisach do tabeli 1 w następnym rozdziale), trzecia to cały ton. Wszystkie są oczywiście dysonansami. Z drugiej strony żaden interwał temperowany nie występuje w szeregu harmonicznym (tak jak zostało wyjaśnione wcześniej, wszystkie są określone liczbami niewymiernymi, których nie można wyrazić za pomocą żadnej proporcji liczb całkowitych), a przynajmniej kwintę i kwartę temperowaną, różne zaledwie o 1/100 tonu od ich czystych odpowiedników, należy uznać za konsonanse, z pewnością znacznie doskonalsze niż wymienione właśnie proporcje 7 : 5, 9 : 5, 9 : 8. Proporcja 5141 : 3427 teoretycznie mogłaby występować w szeregu harmonicznym. Praktycznie jednak żadne dostępne nam obiekty drgające nie wytwarzają tak wysokich składowych. A gdyby nawet je wytwarzały, ucho ludzkie nie mogłoby ich usłyszeć. Można więc w zasadzie powiedzieć, że proporcja 5141 : 3427 nie występuje w szeregu harmonicznym. A mimo to, jak pamiętamy, daje niemal idealny konsonans.

Teoretycy XVII i XVIII-wieczni – w szczególności Rameau – oczywiście zdawali sobie sprawę z tych niepożądanych własności szeregu harmonicznego, który oprócz kwinty, kwarty itd. zawiera w sobie również tryton, septymę i cały ton. Dlatego ograniczali liczbę rozważanych i uznawanych za istotne składowych – tyle że do 6, a nie do 4, jak niegdyś pitagorejczycy – co odpowiadało traktowaniu w XVII i XVIII wieku tercji (5 : 4 i 6 : 5) oraz seksty (5 : 3) jako konsonansów[22]. Ograniczenie takie jest jednak czysto arbitralne i odzwierciedla jedynie obowiązującą wówczas tradycję niewiele w istocie tłumacząc.

Podobnie przedstawia się sprawa z drugim wyjaśnieniem. Każde najdrobniejsze rozstrojenie konsonansu (nawet takie, którego nie je-

22 Por. B. Green, D. Butler, *op.cit.*, s. 251 i 253.

steśmy w stanie usłyszeć) powoduje, że składowe przestają się pokrywać, a konsonans nadal takowym pozostaje. W proporcji 5141 : 3427 wspólna byłaby dopiero 3427-dma składowa pierwszego dźwięku i 5141-sza składowa drugiego – czyli praktycznie żadna. W kwincie temperowanej nawet czysto teoretycznie żadne składowe nie są wspólne. Natomiast w interwale 7 : 5 wspólna jest piąta składowa jednego dźwięku i siódma drugiego, następnie dziesiąta i czternasta itd.; w proporcji 9 : 8 – ósma jednego i dziewiąta drugiego. Ta zgodność w żaden sposób nie generuje konsonansowości wymienionych interwałów i nie zapobiega ich dysonansowości.

VIII. Hermann Helmholtz – dudnienia jako źródło dysonansu

Takie (niesatysfakcjonujące) wyjaśnienie konsonansu – odwołujące się do pokrywania się niektórych składowych – jest czasem przedstawiane jako stanowisko Hermanna Helmholtza (1821-1894)[23], wszechstronnego niemieckiego uczonego, który dokonał przełomowych odkryć w dziedzinie fizyki, chemii, fizjologii, percepcji wzrokowej i słuchowej czy medycyny. W interesującym nas tutaj dziele *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik*[24] (*Teoria percepcji dźwięku jako fizjologiczna podstawa teorii muzyki*) faktycznie znajdują się sformułowania, które zdają się wyrażać taką właśnie koncepcję (por. poniżej s. 83). Jednak Helmholtz jest w istocie autorem daleko bardziej satysfakcjonującego wyjaśnienia konsonansu, które unika zasadniczych mankamentów wcześniejszych koncepcji i tłumaczy, dlaczego konsonansem jest kwinta równomiernie temperowana a także proporcja 5141 : 3427, a nie jest nim wyrażony dużo prostszą proporcją cały ton.

Koncepcja ta odwołuje się do zjawiska dudnienia, które najłatwiej opisać w odniesieniu do tonów prostych, tzn. dźwięków o jednej tylko częstotliwości, nie zawierających dalszych składowych. Gdy dwa tony proste o nieznacznie różniących się częstotliwościach rozbrzmiewają jednocześnie, ich drgania nakładają się i wytwarzają dźwięk wypadkowy, którego częstotliwość jest średnią arytmetyczną częstotliwości obu dźwięków, a jego głośność pulsuje z częstotliwością równą różni-

23 Por. np. Ulrich Michels, *Atlas Muzyki*, t. 1, przeł. Piotr Maculewicz, Warszawa 2002, s. 21.

24 Hermann Helmholtz, *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik*, Braunschweig 1863. Faksymile tego pierwszego wydania jest dostępne pod adresem: http://vlp.mpiwg-berlin.mpg.de/library/data/lit3483. Dzieło to było następnie w kolejnych wydaniach poprawiane i uzupełniane. W niniejszym artykule będę odwoływał się do VI wydania (1913), które z kolei jest dostępne w wygodniejszej wersji tekstowej pod adresem https://archive.org/details/dielehrevondento028665mbp.

cy częstotliwości dźwięków wyjściowych. Przykładowo dwa dźwięki o częstotliwościach 200 i 202 Hz wytworzą jeden dźwięk o częstotliwości 201 Hz, którego głośność będzie pulsować z częstotliwością 2 Hz – tzn. będzie narastać i cichnąć dwa razy na sekundę. Zjawisko to jest najłatwiej postrzegalne, gdy różnica częstotliwości dźwięków wyjściowych jest niewielka i narastanie oraz zanikanie głośności nie jest zbyt prędkie. Poza tym dźwięk wydaje nam się całkiem gładki i jednolity, jedynie jego głośność się zmienia – podlega rytmicznemu pulsowaniu. Warto podkreślić, że nie jest to złudzenie słuchowe, tzn. wrażenie jednego dźwięku o pośredniej wysokości i pulsującej głośności nie jest jedynie wytworem naszego umysłu. Powstawanie takiego dźwięku wypadkowego jest obiektywnym zjawiskiem fizycznym, które możemy zarówno uzasadnić na podstawie odpowiedniej teorii drgań i fal, jak i w sposób empiryczny zmierzyć np. za pomocą oscyloskopu.

Gdy różnica ta wzrasta – wynosi 3, 5 czy nawet 10 Hz – ciągle jeszcze słyszymy coraz szybsze pulsowanie głośności dźwięku wypadkowego. Stopniowo jednak staje się ono za szybkie dla naszego aparatu percepcyjnego i zamiast niego zaczynamy słyszeć coś, co można by nazwać warkotem, który staje się coraz bardziej ostry i nieprzyjemny. A wreszcie po prostu słyszymy dźwięk, który odbieramy jako ostry i szorstki. Używając wcześniejszej terminologii można powiedzieć, że tony proste w tym momencie mocno dysonują. Gdy w dalszym ciągu zwiększamy odległość pomiędzy nimi i częstotliwość dudnienia (pulsowania głośności) dalej wzrasta, ta szorstkość stopniowo słabnie, aż wreszcie całkiem zanika, tzn. dźwięki zaczynają tworzyć konsonans. Dalsze zwiększanie odległości między dźwiękami (tonami prostymi!) niczego już zasadniczo nie zmienia, tzn. pozostają one konsonansem.

Dla wyjaśnienia zjawiska konsonansu i dysonansu niebagatelne jest ustalenie, w którym momencie nieprzyjemne odczucie ostrości (szorstkości) współbrzmienia dwóch tonów prostych jest najsilniejsze. A także – w którym momencie to odczucie ostrości w ogóle się pojawia i następnie kiedy zanika. W omówieniach teorii Helmholtza zwykle podaje się, że uznał on, iż maksimum ostrości współbrzmienia występuje przy różnicy częstotliwości równej 33 Hz albo w przedziale 30–40 Hz[25]. Podczas gdy stwierdzenia takie faktycznie pojawiają się u Helmholtza[26], w żadnym razie nie stanowią jego ostatniego słowa. Zobaczymy później, że kwestia ta będzie miała w miarę istotne znaczenie dla całości koncepcji i jej oceny. W tej chwili, akceptując chwilowo założenie, że ostrość współbrzmienia tonów prostych jest naj-

25 33 Hz pojawiają się np. w *Cambridge History of Western Music Theory*, red. Thomas Christensen, Cambridge, 2008, s. 261. Z kolei przedział 30–40 Hz wymieniają R. Plomp, W. J. M. Levelt, *Tonal Consonance and Critical Bandwidth*, „Journal of the Acoustical Society of America”, 38/1965, s. 549 oraz David J. Benson, *Music. A Mathematical Offering*, Cambridge, 2007, s. 149.

26 H. Helmholtz, *op. cit.* wyd. VI, s. 309 oraz s. 286.

większa w przedziale 30–40 Hz, przedstawmy koncepcję Helmholtza – tym razem już w odniesieniu do naturalnych dźwięków występujących w muzyce, posiadających składowe harmoniczne. Otóż stwierdza on, że stopień dysonansowości współbrzmienia dwóch dźwięków jest rezultatem nagromadzenia ostrych dudnień – tych w sąsiedztwie 33 Hz – pomiędzy ich składowymi harmonicznymi, włączając w to pierwszą składową, tzn. dźwięk podstawowy. Interwały, w których takie ostre dudnienia pomiędzy składowymi nie występują lub są odpowiednio słabe, odbieramy jako konsonanse.

Ilustruje to poniższa tabela. Wszystkie interwały są w niej wyliczone od dźwięku o częstotliwości 300 Hz. Dźwięk ten został wybrany z typowego zakresu skali (oktawa razkreślna), w taki sposób, aby jednocześnie wyrażał się prostą liczbą. Dźwięk ten leży w bezpośrednim sąsiedztwie d^1, którego częstotliwość w stroju równomiernie temperowanym to 293,66 Hz. W kolumnie „Dalsze składowe harmoniczne" dla dźwięku 300 Hz, od którego liczymy wszystkie interwały, wpisano ich osiem, do 2400 Hz włącznie. Dla wyższych dźwięków uwzględniono te kolejne składowe, które mogą trafić w pobliże tych ośmiu. Inaczej mówiąc, wyliczania zaprzestawano, gdy któraś kolejna składowa znacząco przekraczała 2400 Hz. Wreszcie w ostatniej kolumnie wpisano te odległości pomiędzy składowymi, które dla danego interwału znajdują się najbliżej krytycznej odległości 33 Hz – uznanej za produkującą najostrzejsze, dysonujące współbrzmienie. Jeśli takich wielkości było więcej niż jedna, podane zostały w kolejności, w jakiej występują w szeregu składowych, rozpoczynając od odległości pomiędzy najniższymi składowymi.

Objaśnienia do tabeli:

1) Tryton powinien być w zasadzie interwałem pośrodku pomiędzy kwartą a kwintą. Podział całego tonu (9/8), jaki dzieli kwartę i kwintę, na połowę daje tryton temperowany, który jest jednocześnie dokładnie połową oktawy; a więc przewrotem (uzupełnieniem do oktawy) trytonu jest znowu tryton. Cały ton pomiędzy kwartą a kwintą można też podzielić na półtony harmonicznie, za pomocą sekwencji 16 – 17 – 18. Dodając pierwszy półton (17/16) do kwarty, albo odejmując drugi (18/17) od kwinty dostaniemy właśnie interwał 17/12. Natomiast tryton wyrażony jako 7 : 5 to raczej koncesja na rzecz operowania prostymi proporcjami. Jest on, jak widać, bardziej oddalony od dwóch pozostałych. Dodatkowo można jeszcze utworzyć tryton będący przewrotem tego ostatniego, który zatem wyraża się proporcją 10 : 7 = 1,42857… i który z kolei można by nazwać trytonem dużym.

2) Septyma mała jest częściej określana jako 16/9, co stanowi przewrót (uzupełnienie do oktawy) regularnego (pitagorejskiego, wielkiego) całego tonu 9/8. Ale można ją też wyliczyć jako przewrót mniejszego całego tonu (10/9) i wtedy uzyskujemy prostszą proporcję 9/5 – choć jednocześnie bardziej oddaloną od równomiernie temperowanej.

Interwał (Temperacja oznacza temperację równomierną)	Proporcja	Wielkość wyrażona liczbowo	Częstotl. dźwięku podstawowego (pierwsza składowa)	Dalsze składowe harmoniczne	Krytyczne odległości pomiędzy składowymi
Pryma	1 : 1	1	300	600, 900, 1200, 1500, 1800, 2100, 2400	
Półton temp.		1,0595	317,85	636, 954, 1271, 1589, 1907, 2225, 2543	18, 36
Półton naturalny	16:15	1,0667	320	640, 960, 1280, 1600, 1920, 2240, 2560	20, 40
Cały ton temp.		1,1225	336,75	674, 1010, 1347, 1684, 2021, 2357, 2694	37, 43
Cały ton	9 : 8	1,125	337,5	675, 1013, 1350; 1688, 2025, 2363, 2700	37, 37
Tercja mała temp.		1,1892	356,76	714, 1070, 1427, 1784, 2141, 2497	57, 16, 41
Tercja mała	6 : 5	1,2	360	720, 1080, 1440, 1800, 2160, 2520	60
Tercja wielka	5 : 4	1,25	375	750, 1125, 1500, 1875, 2250, 2625	75
Tercja wielka temp.		1,2599	377,97	756, 1134, 1512, 1890, 2268, 2646	66, 12
Tercja wielka pitagor.	81:64	1,2656	379,68	759, 1139, 1519, 1898, 2278, 2658	61, 19
Kwarta	4 : 3	1,3333	400	800, 1200, 1600, 2000, 2400, 2800	100
Kwarta temper.		1,3348	400,44	801, 1201, 1602, 2002, 2403	99, 1, 98, 3
Tryton mały[1)]	7 : 5	1,4	420	840, 1260, 1680, 2100, 2520	60
Tryton temper.		1,4142	424,26	849, 1273, 1697, 2121, 2546	51, 21
Tryton naturalny[1)]	17:12	1,4167	425	850, 1275, 1700, 2125, 2550	50, 25
Kwinta temper.		1,4983	449,5	899, 1348, 1798, 2247, 2697	1, 148, 2
Kwinta	3 : 2	1,5	450	900, 1350, 1800, 2250, 2700	150
Seksta mała temp.		1,5874	476,22	952, 1429, 1905, 2381, 2857	52, 19
Seksta mała	8 : 5	1,6	480	960, 1440, 1920, 2400, 2880	60
Seksta wielka	5 : 3	1,6667	500	1000, 1500, 2000, 2500, 3000	100
Seksta wielka temp.		1,6818	504,54	1009, 1514, 2018, 2523	14, 82
Septyma m.[2)] mniejsza	16 : 9	1,7778	533,34	1067, 1600, 2133, 2667	67, 33
Septyma mała temp.		1,7818	534,54	1069, 1604, 2138, 2673	65, 38
Septyma m.[2)] większa	9 : 5	1,8	540	1080, 1620, 2160, 2700	60, 60
Septyma wielka	15 : 8	1,875	562,5	1125, 1688, 2250, 2813	38, 75
Septyma wlk. temp.		1,8877	566,31	1133, 1699, 2265, 2832	34, 67
Oktawa	2 : 1	2	600	1200, 1800, 2400	300

Tabela 1.
Krytyczne *różnice* częstotliwości pomiędzy składowymi harmonicznymi interwałów

Na pierwszy rzut oka tabela ta całkiem nieźle zgadza się z tym, co zwykle sądzimy na temat dysonansowości i konsonansowości poszczególnych interwałów. Największą – i bezpiecznie oddaloną od krytycznych 33 Hz – odległość pomiędzy składowymi ma oktawa (300 Hz), następnie kwinta (150 Hz), kwarta (100 Hz), seksta wielka (100 Hz) i tercja wielka (75 Hz). To nie stanowi jednak szczególnego osiągnięcia tej koncepcji, bo i z wcześniejszych wyjaśnień wynikała taka właśnie, zgodna z empirią, hierarchia konsonansów. Ale teoria Helmholtza, zilustrowana powyższą tabelą, ponadto po raz pierwszy zgadza się z inną empiryczną obserwacją: że niewielkie rozstrojenie (zmiana) poszczególnych interwałów, które prowadzi często do drastycznego skomplikowania proporcji, nie ma istotnego znaczenia dla ich konsonansowości czy też dysonansowości. W kwincie równomiernie temperowanej odległość pomiędzy składowymi zmniejsza się do 148 Hz – ze 150 Hz dla kwinty czystej – co nie ma żadnego znaczenia, bo jedna i druga odległość jest już daleko poza zakresem, w którym postrzegana jest ostrość współbrzmienia tonów prostych. Dodatkowo pojawia się różnica 1 Hz i 2 Hz pomiędzy pewnymi składowymi. To z kolei odległości, które skutkują pulsowaniem głośności (jeden lub dwa razy na sekundę), ale daleko im do jakiejkolwiek ostrości. Takie pulsowanie głośności mogłoby co prawda zostać uznane za zakłócenie jednolitości i gładkości współbrzmienia, a więc za zakłócenie idealności konsonansu. W odniesieniu do współbrzmienia generowanego przez tony proste (bez wyższych składowych) można by się z tym zgodzić: głośność narasta do pewnego maksimum a następnie spada praktycznie do zera i znowu narasta – trudno uznać to za niezakłócone, jednolite trwanie harmonijnego współbrzmienia (choć pewnie nie nazwalibyśmy tego również dysonansem – wszak kojarzymy z nimpewną ostrość współbrzmienia, której brak w tym wypadku).

Jednak w przypadku kwinty temperowanej pomiędzy dźwiękami naturalnymi (ze składowymi) mamy do czynienia z następującą konfiguracją: współbrzmienie drugiej składowej kwinty (899 Hz) i trzeciej składowej prymy (900 Hz) pulsuje raz na sekundę; współbrzmienie czwartej składowej kwinty (1798 Hz) i szóstej prymy (1800 Hz) – dwa razy na sekundę; wreszcie szóstej składowej kwinty (2697 Hz) i dziewiątej prymy (2700 Hz) – trzy razy na sekundę. Już te trzy pulsowania głośności w różnym tempie, nakładając się na siebie, częściowo się równoważą i zmniejszają wyrazistość każdego z nich. Jednocześnie rozbrzmiewają w stałej głośności: pierwsza, druga, czwarta, piąta, siódma i ósma składowe prymy (300, 600, 1200, 1500, 2100, 2400 Hz) oraz pierwsza, trzecia i piąta składowe kwinty (449,5, 1348, 2247 Hz). Pierwsze składowe obu dźwięków (300 Hz i 449,5 Hz) są zazwyczaj najsilniejsze. W konsekwencji pulsowanie głośności wynikające z bliskości pomiędzy niektórymi wyższymi składowymi nakłada się na jednolitą, stałą głośność dużo większej ilości składowych, może więc skutkować zmianami sumarycznej głośności całkowitego współbrzmienia sięgającymi zaledwie paru procent – z tego powodu praktycznie go nie słyszymy. Mimo to Helmholtz stwierdza:

> (…) jedynie interwały nastrojone według prostych proporcji liczbowych dają spokojne współbrzmienie, podczas gdy już całkiem małe odchylenia od tego matematycznego stroju zdradzają się przez niespokojnie falujące dudnienia[27].

Całkowity brak jakichkolwiek słyszalnych dudnień wymaga, aby składowe dokładnie się pokrywały albo były oddalone na bezpieczną odległość, jak np. 150 Hz w powyższym przykładzie. I faktycznie nieco dalej Helmholtz powiada: „Konsonanse scharakteryzowaliśmy poprzez to, że któreś dwie składowe obu dźwięków się pokrywają”[28]. tak jakby przychylając się do wcześniej przedstawionej koncepcji, którą uznaliśmy za niesatysfakcjonującą. Uzupełnienie Helmholtza polegałoby jedynie na dodatkowym wyjaśnieniu, że takie pokrywanie się konieczne jest dla uniknięcia słyszalnych dudnień, które zakłóciłyby konsonans. Taka uzupełniona koncepcja – wskazująca na zjawisko dudnienia – byłaby i tak lepsza niż samo odwołanie się do zgodności niektórych składowych. Wyjaśnia ona mianowicie, dlaczego np. czysty cały ton (9/8) nie jest konsonansem mimo stosunkowo prostej proporcji: otóż dokładne pokrywanie się odpowiednio ósmej i dziewiątej składowej zapobiega co prawda powstawaniu dudnień pomiędzy nimi, ale wcześniejsze składowe (w szczególności pierwsze, tzn. tony podstawowe) nie pozostają w bezpiecznej odległości i ostre dudnienia powstają pomiędzy nimi. Jednak nadal koncepcja taka nie rozróżniałaby między całym tonem a np. kwintą temperowaną – oba uznawałaby za dysonanse. Jest jednak jasne, że Helmholtz widzi taką różnicę. Stwierdza np.:

> …dwa c^1, które różnią się o jedną dziesiątą półtonu, wytwarzają tylko około jedno dudnienie na sekundę, co można zauważyć jedynie przy uważnej obserwacji i co przynajmniej nie daje żadnej ostrości (szorstkości) współbrzmienia…[29]

Helmholtz mówi tutaj o dudnieniu występującym pomiędzy tonami podstawowymi obu dźwięków, tzn. pomiędzy typowo najsilniejszymi składowymi – i o nich powiada, że można je zauważyć jedynie przy uważnej obserwacji (z pewnością z tego względu, że na to dudnienie nakładają się następne: dwa na sekundę drugich składowych,

27 *Ibidem*, s. 301. Wszystkie cytaty podaję we własnym przekładzie.

28 *Ibidem*, s. 310.

29 *Ibidem*, s. 286.

trzy na sekundę – trzecich itd.). Tym trudniej naturalnie dostrzec dudnienie pomiędzy drugą a trzecią składową dźwięków składających się na kwintę temperowaną, przykryte dużą ilością niedudniących składowych. A nawet jeśli czyjeś wrażliwe i wyszkolone ucho potrafi tego dokonać, nie jest to z pewnością, jak zauważa Helmholtz, nieprzyjemne wrażenie szorstkości (ostrości) charakterystyczne dla szybszych dudnień. W innym miejscu Helmholtz mówi też:

> …słabo rozstrojony konsonans brzmi niemal tak samo dobrze jak czysty, i lepiej niż mocniej rozstrojony, podczas gdy proporcje liczbowe właśnie dla słabo rozstrojonego konsonansu z reguły są najbardziej skomplikowane…[30]

Jest zatem jasne, że sugestia, jakoby koncepcja konsonansu Helmholtza ograniczała się do wskazania na pokrywanie się składowych harmonicznych, musi zostać uznana za błędną. Jego być może najbardziej jednoznaczne i wyraziste określenie dysonansu i konsonansu brzmi następująco:

> Gdy współbrzmią dwa dźwięki muzyczne, zazwyczaj powstają zaburzenia ich współbrzmienia spowodowane poprzez dudnienia wytwarzanie przez ich składowe, tak że (…) współbrzmienie staje się ostre (szorstkie). Nazywamy to dysonansem.
>
> Ale istnieją pewne proporcje pomiędzy częstotliwościami, przy których występują wyjątki od tej reguły: nie powstają żadne dudnienia albo dochodzą one tak słabo do ucha, że nie powodują nieprzyjemnego zakłócenia współbrzmienia; te wyjątki nazywamy konsonansami[31].

Ponieważ zaś, jak się niebawem przekonamy (por. poniżej Tabela 2), sytuacja, gdy nie powstają żadne dudnienia[32] niemal nie występuje, z konsonansem mamy do czynienia praktycznie wtedy, gdy dudnienia są wystarczająco słabo słyszalne. A zatem powyższe

30 *Ibidem*, s. 378.

31 *Ibidem*, s. 320. Pisząc „nie powstają żadne dudnienia" Helmholtz musiał mieć oczywiście na myśli „żadne potencjalnie słyszalne". Bowiem w ogólności dudnienia powstają zawsze. Pomiędzy tonami prostymi o częstotliwościach 300 i 600 Hz występuje dudnienie o częstotliwości 300 Hz, które jednak leży całkowicie poza zakresem postrzegalności. Zakres ten obejmuje dudnienia powolne – te które możemy policzyć – oraz trochę szybsze, które postrzegamy jako ostrość (szorstkość) współbrzmienia.

32 Tzn. żadne potencjalnie słyszalne (por. poprzedni przypis). Gdyby chodziło o dowolne dudnienia, wówczas można by całkiem kategorycznie powiedzieć, że zawsze powstają. Ale pomiędzy tonami prostymi, gdy są wystarczająco oddalone, nie występuje żadne postrzegalne dudnienie. Gdy jednak rozważamy dźwięki naturalne, ze składowymi harmonicznymi, nie tylko jakiekolwiek, ale także potencjalnie postrzegalne dudnienia występują niemal zawsze.

określenie Helmholtza nie przesądza jeszcze, co jest konsonansem, a co dysonansem. Dostarcza tylko ogólnej formuły, która wymaga dookreślenia, jakie dudnienia uznajemy za wystarczająco słabe, aby nie burzyły zgodności współbrzmienia. Granicę tę można wyznaczyć w różny sposób. W szczególności, w najbardziej restryktywnej wersji, możemy uznać, że nawet powolne i łagodne dudnienia, jakie występują w kwincie temperowanej, uznajemy za zaburzające jej konsonansowość. Mimo niektórych sformułowań Helmholtza nie wydaje się – sądząc na podstawie innych, przywołanych powyżej wypowiedzi – aby taka była jego intencja, tzn. aby chciał uznać kwintę temperowaną za dysonans[33]. Ale w gruncie rzeczy nie musimy tego rozstrzygać. Jego wyjaśnienie konsonansu przedstawione w ostatnim cytacie, odwołujące się do braku lub słabości słyszalnych dudnień, jest niezależne od rozstrzygnięcia, jak wyznaczamy granicę pomiędzy silnymi (istotnymi) i słabymi (nieistotnymi) dudnieniami. Możemy nie zgadzać się z Helmholtzem w sprawie wyznaczenia tej granicy, a jednocześnie w pełni akceptować jego ogólną koncepcję i jej moc wyjaśniającą.

Wracając do komentarza na temat przedstawionej tabeli: to, co powyżej o kwincie temperowanej, można powiedzieć także o temperowanej kwarcie. Odpowiednie liczby dla temperowanej tercji, bardziej oddalonej od tercji czystej niż w przypadku kwinty i kwarty, wskazują już na możliwość zauważalnego jej pogorszenia: zamiast bezpiecznej odległości 75 Hz pojawia się 66 Hz, a następnie 12 Hz, jeszcze w miarę odległe od krytycznego przedziału 30–40 Hz, ale już przekraczające granicę powolnego i łagodnego pulsowania głośności[34]. Widać jednak, że pogorszenie tercji pitagorejskiej jest znacząco większe: pojawia się odległość 19 Hz, która już całkiem niebezpiecznie zbliża się do krytycznego przedziału 30–40 Hz. Znowu: to, co o tercji wielkiej, można powiedzieć o sekście wielkiej. Kolejna obserwacja dotyczy całego tonu, półtonu oraz septymy wielkiej: widać, że nie ma istotnej różnicy pomiędzy ich wersją naturalną a tem-

33 Uznanie kwinty temperowanej za konsonans naturalnie nie musi przesądzać o akceptacji stroju równomiernie temperowanego, co jest kwestią o znacznie szerszych implikacjach.

34 W tym momencie możemy też podać wyjaśnienie, zapowiedziane w przypisie 14, dlaczego kwinta temperacji średniotonowej, mniejsza o jedną czwartą komatu Didymosa od kwinty czystej, mogła być akceptowana, podczas gdy kwinta *d* – *a* stroju naturalnego, mniejsza o komat Didymosa od kwinty czystej – już nie. To ostatnie zmniejszenie odniesione do kwinty 300 – 450 Hz to interwał 300 – 444,44 Hz. Powstaje tutaj dudnienie pomiędzy trzecią składową prymy (900 Hz) a drugą kwinty (888,88 Hz) o częstotliwości ponad 11 Hz, podobne jak w tercji wielkiej temperowanej, ale pojawiające się pomiędzy niższymi składowymi (więc potencjalnie wyraźniej słyszalne) i dotyczące bardziej konsonującego interwału kwinty, a więc wprowadzające relatywnie większe zaburzenie. Z drugiej strony kwinta średniotonowa to np. interwał 300 – 448,6 Hz. Dudnienie w nim obecne (pomiędzy 900 a 897,2 Hz) ma częstotliwość 2,8 Hz, nie powoduje więc żadnej ostrości.

perowaną: wszystkie są ostrymi dysonansami. Podobnie więc jak w przypadku kwinty i kwarty, gdzie zamiana prostych proporcji na skrajnie skomplikowane w zasadzie nie zaszkodziła ich konsonansowości, tak w wypadku tych ostatnich interwałów zamiana niewymiernych proporcji wersji temperowanych na dużo prostsze wersji naturalnych w żadnym stopniu nie zmniejsza ich dysonansowości. Wszystko zgadza się z tym, co już dawno wiadomo o tych interwałach na podstawie empirii.

Staje się jasne, że konsonansowość czy też dysonansowość niewiele ma wspólnego z proporcjami – raczej leży w umiejscowieniu interwałów w pewnych przedziałach (zakresach), wyznaczonych przez wrażliwość ucha. Właściwie było to już wcześniej oczywiste – i jeszcze bardziej wyraziste – w odniesieniu do tonów prostych: ich bliskość – całkiem niezależna od prostoty czy też skomplikowania proporcji – skutkuje słyszalnymi dudnieniami, najpierw powolnymi i łagodnymi, potem coraz szybszymi i ostrzejszymi, a po osiągnięciu pewnego maksimum znowu coraz słabiej postrzegalnymi. Wreszcie zgodne współbrzmienie tonów prostych pojawia się po przekroczeniu pewnej odległości (na razie wstrzymujemy się z rozstrzygnięciem, jak ten próg określić: czy jako 60 Hz – jak mogłaby sugerować powyższa tabela – czy jakoś inaczej), poza którą ich konsonansowość już się w zasadzie nie zmienia – jest podobna dla trytonu, kwinty, septymy, oktawy czy nony – podczas gdy stosunek ich częstotliwości przebiega wszelkie możliwe wartości, wyrażone zarówno prostymi jak i bardzo skomplikowanymi proporcjami. Słowem-kluczem jest w tym wypadku nie proporcja, a odległość: odpowiednio duża daje konsonans, stosunkowa mała – dysonans, przy czym maksimum jego ostrości jest wyznaczone wyłącznie przez wrażliwość naszego ucha.

Po omówieniu sukcesów powyższej tabeli, przejdźmy do znaków zapytania. Zgodne z oczekiwaniami jest porównywalne miejsce tercji małej i seksty małej na skali konsonansowości (60 Hz to dla każdej z nich najmniejsza odległość między składowymi). Jednak niepokojący jest fakt, że podobną pozycję zdają się zajmować: jedna wersja trytonu i jedna wersja septymy małej. Po drugie, porównanie różnych wersji septymy małej również budzi wątpliwości. Największa jej wersja (9/5), a więc najbliższa ostro dysonującej septymy wielkiej, miałaby być znacznie mniej dysonująca (podobnie zgodna jak tercja mała) niż dwie pozostałe wersje septymy małej, odleglejsze od septymy wielkiej, a bliższe seksty. Wydaje się to intuicyjnie niewiarygodne.

Jak można odpowiedzieć na takie wątpliwości? Po pierwsze, należy wskazać na uproszczenia w konstrukcji powyższej tabeli. To, co zostało w niej wyeksponowane, to odległości pomiędzy składowymi najbliższe krytycznego przedziału 30–40 Hz. Nie została jednak wyraźnie odnotowana liczba tych zbliżeń pomiędzy różnymi parami składowych. Co więcej, i co ważniejsze, nie zostało wyraziście uwidocznione, pomiędzy którymi składowymi mają miejsce wspomniane zbliżenia. Wszak czym innym jest krytyczna bliskość np. pomiędzy drugą składową jednego dźwięku a tonem podstawowym drugiego, a czym innym taka sama odległość np. pomiędzy piąta a siódmą składową, które mogą być już na tyle słabe, że ostrość ich współbrzmienia może niemal nie dochodzić do głosu. Zwracając uwagę na ten czynnik można rozbroić paradoks rzekomo mniej dysonującej septymy małej o proporcji 9 : 5 w stosunku do dwóch pozostałych. Wskazane tam częstotliwości dudnień 33 Hz dla jednej wersji septymy oraz 38 Hz dla drugiej, to odległości pomiędzy czwartą składową septymy i siódmą prymy. Zwłaszcza ze względu na słabość tej ostatniej dudnienia te są prawdopodobnie tak słabo słyszalne, że praktycznie nie mają znaczenia dla oceny dysonansowości tych interwałów. Natomiast 60 Hz w przypadku septymy o proporcji 9 : 5 to odległość pomiędzy jej tonem podstawowym a drugą składową prymy – i to współbrzmienie z pewnością nie jest bez znaczenia. Analogiczne odległości dla pozostałych dwóch septym małych to 67 i 65 Hz. Jeśli jednak położymy nacisk na to współbrzmienie, okaże się, że wszystkie wersje septymy małej są podobnymi konsonansami jak tercja mała (lub nawet lepszymi) – i trudno taki werdykt zaakceptować. W istocie ta niezgodność z empirią wskazuje na ogólniejszą trudność w koncepcji Helmholtza, tak jak została ona przedstawiona powyżej. Zobaczymy, że sedno problemu leży w określeniu maksimum ostrości współbrzmienia dla tonów prostych, które zostało powyżej przyjęte – zgodnie z jedną z sugestii Helmholtza – jako 33 Hz. W tym momencie warto zatem przejść do ogólnych zarzutów wobec tej koncepcji i ocenić ich wagę.

IX. Kapryśność ucha

Pierwszy zarzut wobec teorii Helmholtza oraz jednocześnie odpowiedź na niego zostały przekonująco sformułowane przez Plompa i Levelta:

> Podstawowy powód, dla którego Helmholtza wyjaśnienie konsonansu (…) zostało odrzucone przez wielu badaczy, był taki, że w ich opinii stopień konsonansowości lub dysonansowości interwału nie zmienia się, gdy usuniemy składowe harmoniczne (…). Analiza obserwacji, na których opierała się ta opinia pokazuje, że osoby, które miały oceniać interwały były – bez wyjątku – wykształcone muzycznie. Nie było to uznawane za trudność ale, wręcz przeciwnie, za istotny warunek, aby otrzymać właściwe odpowiedzi. Sam Stumpf, być może najważniejszy krytyk teorii opartej na dudnieniach, może być przedstawiony jako dobry przykład. (…) Dla niego ocena konkretnego interwału była równoznaczna z odgadnięciem jego nazwy muzycznej i ta wiedza całkowicie

> determinowała stopień konsonansowości, który przypisywał temu interwałowi. Z tego powodu uznawał interwały takie jak 8 : 15 [septyma wielka] i 7 : 10 [tryton] jako dysonansowe, także w przypadkach bez słyszalnych składowych (…) Najwyraźniej takie podejście było dla niego (i wielu innych) tak oczywiste, że nie zdawał sobie sprawy, iż jego wyniki nie miały nic wspólnego ze źródłem konsonansu i dysonansu, ale muszą być uznane jedynie za demonstrację sukcesu jego wykształcenia muzycznego[35].

Inaczej mówiąc, dla Stumpfa interwały takie jak tryton czy septyma wielka z definicji były dysonansami. A jeśli tak, to ewentualne usunięcie składowych harmonicznych niczego nie zmieniało, bo odległość pomiędzy tonami prostymi dalej pozostawała trytonem czy septymą – a więc z definicji dysonansem – niezależnie od jakiejkolwiek ostrości współbrzmienia lub jej braku.

Poważniejsza jest druga trudność, która dochodziła do głosu już w problemach z interpretacją niektórych interwałów w powyższej tabeli. Jeśli mianowicie przeprowadzimy podobną analizę w odniesieniu do innych dźwięków wyjściowych niż wybrane tam 300 Hz, uzyskamy całkiem odmienne, sprzeczne z empirią wnioski. Możemy zilustrować to na przykładzie dźwięków 100 Hz i 1000 Hz. Pierwszy jest o duodecymę (tzn. oktawę plus kwintę) niższy od 300 Hz (okolice d^1), a więc znajduje się w okolicach dźwięku G wielkie. Drugi jest o tercdecymę (oktawę plus sekstę wielką) wyższy od 300 Hz, leży więc w okolicach h^2. Są to więc ciągle dźwięki ze stosunkowo typowo używanego zakresu skali. (Niemal wszystkie instrumenty muzyczne zawierają w swoim rejestrze przynajmniej jeden z nich.) Bez ponownego konstruowania całej tabeli dla tych dźwięków możemy powiedzieć tylko, że w świetle przedstawionej koncepcji niemal wszystkie interwały od 100 Hz w górę byłyby dysonansami. Jedynie dla kwinty (100 – 150 Hz) najmniejsza odległość pomiędzy składowymi (w istocie pomiędzy tonami podstawowymi) wynosiłaby 50 Hz, co zresztą oznaczałoby, że kwinta ta jest i tak mniej zgodnym konsonansem niż np. tercja mała, tryton (7/5), seksta mała i septyma mała (9/5) od dźwięku 300 Hz – dla wszystkich bowiem najmniejsza odległość między składowymi wynosi 60 Hz. Kwarta 100 – 133 Hz oraz seksta wielka 100 – 167 Hz byłyby najostrzejszymi dysonansami, z odległością pomiędzy składowymi równą 33 Hz. Takim zresztą dysonansem musiałaby być także nieco niższa kwinta 66 – 99 Hz.

Z drugiej strony wszystkie interwały powyżej 1000 Hz byłyby konsonansami. Najmniejsza odległość pomiędzy składowymi półtonu temperowanego i naturalnego wynosi odpowiednio 60 i 67 Hz, które to odległości we wcześniejszej tabeli (dla dźwięku 300 Hz) wiązały się np. z tercją i sekstą małą. Natomiast dla całych tonów i septym wiel-

35 R. Plomp, W. J. M. Levelt, *op. cit.*, s. 550.

kich od dźwięku 1000 Hz w górę odpowiednie liczby wynoszą zawsze ponad 100 Hz. Byłyby to więc konsonanse przynajmniej tak zgodne jak kwarta 300 – 400 Hz i seksta wielka 300 – 500 Hz we wcześniejszej tabeli.

Naturalnie Helmholtz nie mógł nie zdawać sobie sprawy z tego rodzaju niewiarygodnych konsekwencji przyjęcia, że maksimum ostrości (szorstkości) współbrzmienia dla tonów prostych występuje zawsze przy różnicy 33 Hz czy też w przedziale 30–40 Hz oraz że pewna różnica częstotliwości (np. 100 Hz) zawsze gwarantuje konsonans. W istocie porównywał on współbrzmienia w różnych rejestrach skali i konkludował:

> Ostrość współbrzmienia zależy więc w złożony sposób od wielkości interwału i od liczby dudnień[36].

Liczba dudnień wynika z różnicy częstotliwości, wielkość interwału – z ich proporcji (ilorazu). Jedna i druga wielkość (różnica i iloraz) nie zmieniają się jednocześnie w tym samym stopniu. Wyznaczone stałą proporcją 3 : 2 kwinty 100 – 150, 200 – 300, 300 – 450, 400 – 600, 600 – 900 Hz charakteryzują się wzrastającą różnicą częstotliwości: 50, 100, 150, 200, 300 Hz. Z drugiej strony interwały określone kolejno częstotliwościami 50 – 100, 100 – 150, 150 – 200, 200 – 250, 250 – 300, 400 – 450 Hz wszystkie charakteryzują się różnicą częstotliwości 50 Hz (i generowałyby właśnie takie dudnienie), ale są to kolejno: oktawa, kwinta, kwarta, tercja wielka, tercja mała, cały ton. Dla ustalonego interwału (np. kwinty) różnica częstotliwości pomiędzy tworzącymi go dźwiękami zmienia się wraz z rejestrem. Z kolei dla stałej różnicy częstotliwości wraz z rejestrem zmienia się wielkość interwału.

Określenie, że maksymalna ostrość współbrzmienia tonów prostych występuje przy dudnieniu 33 Hz, sugeruje oczywiście, że ostrość ta uzależniona jest od różnicy częstotliwości. Ale wynikałoby stąd, że wszystkie interwały w ostatniej sekwencji – a w szczególności oktawa i cały ton – charakteryzują się takim samym stopniem konsonansowości (czy też dysonansowości). Konstatacja, że tak nie jest, skłoniła Helmholtza do zacytowanej powyżej konkluzji. Ale nie określił on bliżej i bardziej szczegółowo tej złożonej zależności ostrości współbrzmienia od dwóch wskazanych parametrów i okazjonalnie powracał do swojego przybliżonego szacunku 33 Hz (np. na s. 309 i 316).

To, czego nie dopełnił Helmholtz, starali się nadrobić Plomp i Levelt[37]. Przeprowadzili oni wśród osób niewykształconych muzycznie (a więc niepotrafiących rozpoznać słyszanych interwałów

---

36 H. Helmholtz, *op. cit.*, s. 287.

37 R. Plomp, W. J. M. Levelt, *op. cit.*, s. 552–556.

i w konsekwencji nie kierujących się w ocenie ich określeniem) badanie postrzegania dysonansowości/konsonansowości współbrzmienia tonów prostych, dla pięciu częstotliwości wyjściowych: 125, 250, 500, 1000 i 2000 Hz. Z uzyskanych wyników zdaje się wypływać wniosek, że postrzeganie dysonansowości zależne jest w pierwszym rzędzie od wielkości interwału, tzn. proporcji, a nie od różnicy częstotliwości.

Ściśle rzecz biorąc do określenia maksimum ostrości dysonansu autorzy używają odniesienia do tzw. szerokości pasma krytycznego (*critical bandwidth*). Terminu tego, który odsyła do fizjologicznych mechanizmów słyszenia, nie musimy tutaj wyjaśniać[38]. Wystarczy nam informacja, że w pewnym zakresie częstotliwości szerokość ta jest nieco mniejsza od tercji małej[39]. Autorzy stwierdzają, że dla współbrzmień tonów prostych maksimum dysonansowości występuje w przybliżeniu w jednej czwartej szerokości tego pasma, natomiast ostrość współbrzmienia zanika, gdy interwał osiąga tę szerokość. Skoro tercja mała to 6/5 = 1,2, a tercja mała temperowana to ok. 1,189, możemy pewnie przyjąć, że szerokość pasma krytycznego – która ma być nieznacznie mniejsza – to interwał wyznaczony w przybliżeniu liczbą 1,18. Wówczas każdy większy interwał pomiędzy tonami prostymi byłby już całkiem pozbawiony ostrości, natomiast jej maksimum występowałoby dla jednej czwartej takiego interwału, tzn., dla $\sqrt[4]{1{,}18} = 1{,}042$. To nieco mniej niż półton (ca. 1,06) – można orientacyjnie powiedzieć, że to około ¾ półtonu, skoro tercja mała to trzy półtony.

Oznacza to, że zamiast uprzedniej różnicy częstotliwości wskazującej na maksymalną dysonansowość (33 Hz), jako jej kryterium uzyskujemy teraz proporcję (iloraz) częstotliwości równą 1,042. Dla interwału 780–813 Hz obie miary pokrywają się, bo 813/780 = 1,042 a jednocześnie 813 – 780 = 33 Hz. Dla wyższych dźwięków zgodnie z obecną miarą ostrości współbrzmienia (1,042) maksimum będzie występować dla większej różnicy częstotliwości, tzn. dla szybszych dudnień; dla niższych dźwięków – przy niższej różnicy i wolniejszych dudnieniach.

38 Omówienie tego pojęcia można znaleźć np. [w:] Juan G. Roederer, *The Physics and Psychophysics of Music*, wyd. IV, Nowy Jork 2008, s. 34-42.

39 Ten szacunek nie obowiązuje jednak dla wszelkich poziomów częstotliwości, tzn. dla wszystkich rejestrów. Patrz dalsze wyjaśnienia poniżej.

X. Szacowanie dysonansów

Dysponując taką nową miarą dysonansowości możemy sporządzić podobną tabelę jak poprzednio, tym razem wpisując w nią nie różnice częstotliwości pomiędzy składowymi, ale proporcje pomiędzy nimi. W związku z tym podanie konkretnych częstotliwości poszczególnych dźwięków nie jest istotne. Wystarcza ponownie oznaczenie częstotliwości niższego dźwięku każdego interwału w sposób abstrakcyjny liczbą 1. (Dzięki temu nasza tabela odnosi się do interwałów na dowolnej wysokości, w dowolnym rejestrze.) Wówczas częstotliwość wyższego dźwięku to np. 1,25 dla tercji czystej, 1,2599… – dla równomiernie temperowanej, 1,5 dla kwinty itd. Wielokrotności tak oznaczonych częstotliwości to składowe harmoniczne. A więc dla dolnego dźwięku (prymy) będą to liczby: 1 (ton podstawowy) – 2 – 3 – 4 – 5…, dla jej tercji wielkiej: 1,25 – 2,5 – 3,75 – 5 – 6, 25…, dla jej kwinty: 1,5 – 3 – 4,5 – 6 – 7,5… itp. Z kolei możemy wyliczyć proporcje pomiędzy tak ustalonymi składowymi. Np. proporcja pomiędzy czwartą składową prymy, a trzecią składową jej tercji wielkiej to 4 : 3,75 = 1,0666…, pomiędzy piątą składową prymy a trzecią kwinty to 5 : 4,5 = 1,111… itp.

W poniższej tabeli nie zostały uwidocznione składowe harmoniczne, ale od razu odpowiednie proporcje pomiędzy nimi. Przy czym, nieco inaczej niż w tabeli poprzedniej, przedstawione zostały proporcje pomiędzy każdą kolejną składową prymy (od pierwszej do ósmej), a najbliższą (wyższą lub niższą) składową górnego dźwięku interwału. Najbliższą dlatego, że właśnie relatywna bliskość pomiędzy tonami prostymi jest tym, co generuje dysonansowość. Jeśli więc rozważamy np. piątą składową prymy (jej częstotliwość wynosi 5) i składowe kwinty, wówczas pierwsza (1,5) lub piąta (7,5) są całkiem nieinteresujące, bo nie generują żadnego dysonansu. Najbliższe – w stosunku do częstotliwości określonej jako 5 – składowe kwinty to trzecia (4,5) i czwarta (6). Proporcja z pierwszą to 5 : 4,5 = 1,111… (mały cały ton), z drugą – 6 : 5 = 1,2 (tercja mała). Ponieważ ta pierwsza jest mniejsza i generuje większą ostrość, ją właśnie wpisaliśmy do tabelki w kolumnie 5, zawierającej najbliższe współbrzmienia z piątą składową prymy.

Interwał	Proporcja	Wielkość wyrażona liczbowo	Krytyczny (najbliższy) interwał z kolejnymi składowymi prymy							
			1	2	3	4	5	6	7	8
Półton temp.		1,0595	1,060	1,060	1,060	1,060	1,060	1,060	1,060	1,060
Półton naturalny	16:15	1,0667	1,067	1,067	1,067	1,067	1,067	1,067	1,067	1,067
Cały ton temp.		1,1225	1,123	1,123	1,123	1,123	1,114	1,069	1,039	1,018
Cały ton	9 : 8	1,125	1,125	1,125	1,125	1,125	1,111	1,067	1,037	1,016
Tercja mała temp.		1,1892	1,189	1,189	1,189	1,121	1,051	1,009	1,019	1,041
Tercja mała	6 : 5	1,2	1,200	1,200	1,200	1,111	1,042	1,000	1,029	1,050
Tercja wielka	5 : 4	1,25	1,250	1,250	1,200	1,067	1,000	1,042	1,071	1,067
Tercja wielka temp.		1,2599	1,260	1,260	1,191	1,058	1,008	1,050	1,080	1,058
Tercja wlk. pitagor.	81:64	1,2656	1,266	1,266	1,185	1,054	1,012	1,055	1,085	1,054
Kwarta	4 : 3	1,3333	1,333	1,333	1,125	1,000	1,067	1,111	1,050	1,000
Kwarta temper.		1,3348	1,335	1,335	1,124	1,001	1,068	1,112	1,049	1,001
Tryton mały	7 : 5	1,4	1,400	1,400	1,071	1,050	1,120	1,071	1,000	1,050
Tryton temper.		1,4142	1,414	1,414	1,061	1,061	1,131	1,061	1,010	1,061
Tryton naturalny	17:12	1,4167	1,417	1,411	1,059	1,063	1,133	1,059	1,012	1,063
Kwinta temper.		1,4983	1,498	1,335	1,001	1,124	1,112	1,001	1,070	1,068
Kwinta	3 : 2	1,5	1,500	1,333	1,000	1,125	1,111	1,000	1,071	1,067
Seksta mała temp.		1,5874	1,587	1,260	1,058	1,191	1,050	1,058	1,102	1,008
Seksta mała	8 : 5	1,6	1,600	1,250	1,067	1,200	1,042	1,067	1,094	1,000
Seksta wielka	5 : 3	1,6667	1,667	1,200	1,111	1,200	1,000	1,111	1,050	1,042
Seksta wielka temp.		1,6818	1,682	1,189	1,121	1,189	1,009	1,121	1,041	1,051
Septyma m. mniejsza	16 : 9	1,7778	1,778	1,125	1,185	1,125	1,067	1,125	1,016	1,111
Septyma mała temp.		1,7818	1,782	1,123	1,188	1,123	1,069	1,123	1,018	1,114
Septyma m. większa	9 : 5	1,8	1,800	1,111	1,200	1,111	1,080	1,111	1,029	1,111
Septyma wielka	15 : 8	1,875	1,875	1,067	1,250	1,067	1,125	1,067	1,071	1,067
Septyma wlk. temp.		1,8877	1,888	1,059	1,258	1,059	1,133	1,059	1,079	1,059
Oktawa	2 : 1	2	2,000	1,000	1,333	1,500	1,200	1,333	1,143	1,250

Tabela 2.
Krytyczne *proporcje* częstotliwości pomiędzy składowymi harmonicznymi interwałów

Naturalnie powyższa tabela nie uwzględnia składowych aż do ósmej dla wyższych dźwięków wszystkich interwałów. Przykładowo dla seksty wielkiej interwał 1,042 w ósmej kolumnie powstaje pomiędzy ósmą składową prymy a piątą składową seksty. Szósta, siódma i ósma składowa górnego dźwięku seksty pozostają nieuwzględnione. Mogłyby one wchodzić w krytyczną bliskość dopiero z 10–13 składową prymy, które na potrzeby naszego obecnego rozważania uznaliśmy za wystarczająco słabe, aby można je pominąć.

Jak należy wykorzystać powyższą tabelę do oceny relatywnej dysonansowości poszczególnych interwałów? Jak widać pomiędzy składowymi wielokrotnie występują interwały z przedziału 1,02 – 1,09, tzn. w okolicach półtonu (1,06), które jednocześnie poprzez swoją bliskość do interwału 1,042, uznanego za najostrzejszy, należą do najbardziej dysonansowych. Wszystkie zostały zaznaczone kolorem pomarańczowym. Występuje też spora liczba interwałów w przedziale 1,10 – 1,14, tzn. w okolicach całego tonu (1,125) – oznaczono je kolorem szarym. Ten przedział leży mniej więcej w połowie odległości pomiędzy uznanym za najostrzejszy interwałem 1,042, a proporcją 1,18 (pobliże tercji małej), gdzie dysonansowość (dla tonów prostych, jakimi są składowe) już całkowicie zanika. Zatem – według badań Plompa i Levelta – ostrość położonych tutaj współbrzmień można uznać za około połowę mniejszą niż tych w okolicach półtonu. Kolorem czarnym zaznaczono szczególne, bardzo małe odległości (w zakresie 1,008–1,010, tzn. 1/15–1/12 tonu) pomiędzy niektórymi składowymi interwałów temperowanych (tercji, seksty, trytonu). Są one bardzo oddalone od najostrzejszej proporcji 1,042 i produkują jedynie powolne dudnienia o co najwyżej nieznacznej szorstkości. Wreszcie takież odległości pomiędzy składowymi temperowanej kwarty i kwinty (1,001 tzn. ok. 1/100 tonu) pozostały niezaznaczone – podobnie jak nie produkujące już żadnej szorstkości interwały powyżej 1,18.

Jest jasne, że o dysonansowości wyjściowych interwałów pomiędzy dźwiękami naturalnymi nie można wnioskować wyłącznie na podstawie samej obecności ostrych współbrzmień pomiędzy ich składowymi, bo współbrzmienia takie występują w zasadzie wszędzie. Należy zatem w jakiś sposób ocenić ilość i jakość takich współbrzmień w poszczególnych interwałach. Z pewnością należy uwzględnić miejsce ich pojawienia się, zakładając w pierwszym przybliżeniu, że niższe składowe są silniejsze, a więc występująca między nimi dysonansowość – istotniejsza. Następnie należy zadać pytanie jaką wagę przypisujemy – mówiąc umownie – „półtonom” i „całym tonom”[40].

40 Taki podział występujących mniej lub bardziej ostrych współbrzmień (wszystkich poniżej 1,18) nie jest bynajmniej arbitralny. Łącznie jest ich w powyższej tabeli 130 (pomijamy najmniejsze z nich, charakterystyczne dla interwałów temperowanych): 85 z nich leży w przedziale 1,016 – 1,08, a 42 w przedziale 1,11-1,143. W szerokim paśmie rozdzielającym te dwa przedziały (pomiędzy 1,08 a 1,11) leżą zaledwie 3. Zatem zakresy wyraźnie ostrzejszych półtonów (1,016 – 1,08) i znacznie mniej

Uznając – za Plompem i Leveltem – półtony (pobliże interwału 1,042) za znacznie ostrzejsze niż całe tony, możemy im przyznać decydującą rolę w całkowitej dysonansowości interwałów, a całym tonom jedynie rolę pomocniczą. Wówczas interwały można by podzielić na następujące grupy:

I. bez półtonów: oktawa
II. z półtonami od 7 składowej wzwyż: kwinta, seksta wielka
III. z półtonami od 6 składowej: cały ton
IV. z półtonami od 5 składowej: tercja mała, kwarta, septyma mała
V. z półtonami od 4 składowej: tercja wielka
VI. z półtonami od 3 składowej: seksta mała, tryton
VII. z półtonami od 2 składowej: septyma wielka
VIII. z półtonami od 1 składowej: półton

Ponieważ te same określenia stosują się jednakowo do wszystkich wariantów każdego z interwałów (tzn. do wersji czystych, temperowanych, mniejszych, większych itd.), warianty te nie zostały wyszczególnione. W grupach z więcej niż jednym interwałem (II, IV, VI) kolejność ustalona została na podstawie występowania całych tonów. A więc kwinta została umieszczona przed sekstą wielką, ponieważ pierwszy cały ton występuje w niej na czwartej składowej prymy, a w sekście wielkiej – już na trzeciej. To samo dotyczy grupy IV, natomiast w grupie VI seksta mała została umieszczona przed trytonem, ponieważ w tym ostatnim kolejny półton (na 4 składowej) występuje wcześniej niż w sekście małej.

Taka hierarchia jest naturalnie nie do przyjęcia ze względu na pozycję całego tonu: jedynie za oktawą, kwintą i sekstą wielką, a przed wszystkimi innymi interwałami. Mniej rażąca jest pozycja septymy małej, ale jej miejsce przed tercją wielką też jest naturalnie nie do zaakceptowania. Gdyby można te dwa interwały przenieść w dół skali, pozostała kolejność nadal byłaby dość osobliwa (seksta wielka – tercja mała – kwarta – tercja wielka), ale już nie tak bardzo bulwersująca. Należy zatem spróbować innego podejścia, być może początkowo traktując całe tony i półtony na równi, a różnice pomiędzy nimi wykorzystując dopiero w drugiej kolejności. Wówczas możliwe byłoby ułożenie następującej hierarchii:

ostrych całych tonów (1,11 – 1,143) są dobrze odseparowane. Przydzielenie zaledwie trzech współbrzmień leżących pomiędzy 1,08 a 1,11 do jednej lub drugiej klasy (powyżej granica została ustalona jako 1,09) jest już naturalnie częściowo arbitralne, ale ze względu na ich znikomą ilość (zresztą dwa spośród trzech występują w interwałach obocznych: 1,085 w tercji pitagorejskiej, 1,102 w sekście małej temperowanej) nie będzie miało istotnego wpływu na całość obrazu.

Interwały z dysonansami (półtonami lub całymi tonami):

I. począwszy od 7 składowej: oktawa
II. począwszy od 4 składowej: kwinta, tercja mała, tercja wielka
III. począwszy od 3 składowej: seksta wielka, kwarta, seksta mała, tryton
IV. począwszy od 2 składowej: septyma mała, septyma wielka
V. począwszy od 1 składowej: cały ton, półton.

W ramach każdej grupy interwały zostały uszeregowane na podstawie rozróżnienia pomiędzy całymi tonami a półtonami i kolejności ich występowania na następnych składowych. Tę hierarchię z pewnością łatwiej zaakceptować, choć czymś, co kłóci się z tradycyjnymi wyobrażeniami, jest pozycja tercji małej przed wielką (jak się później okaże, nie jest to przypadkowe), a przede wszystkim niska pozycja kwarty, którą zwykle umieszcza się zaraz po kwincie, a przed tercjami i sekstami. Gdy porównamy ją w powyższej tabeli np. z tercją wielką, widać, że pierwszy dysonans w kwarcie następuje już na trzeciej składowej, a w tercji wielkiej – dopiero na czwartej, co właśnie było powodem umieszczenia tej ostatniej wyżej w hierarchii. Ale jednocześnie dysonans na czwartej składowej tercji wielkiej jest ostrzejszy niż na trzeciej składowej kwarty. Podobne porównanie pomiędzy tercją wielką a sekstą wielką przemawia jeszcze bardziej na korzyść tej ostatniej, a wbrew powyższej hierarchii. Kolejna więc próba oszacowania dysonansowości interwałów mogłaby polegać na odpowiednim „odważeniu" wpływu tych dwóch czynników: ostrości dysonansów pomiędzy składowymi i ich pozycji w szeregu, a także ich łącznej ilości. Naturalnym sposobem, aby to uczynić, byłoby przyznanie każdemu z czynników odpowiedniej wagi wyrażonej liczbowo, a następnie zsumowanie rezultatu. Ponieważ krytyczne odległości pomiędzy składowymi fortunnie dzielą się na dwie wyraziste klasy: półtonów i całych tonów[41], możemy, dla uproszczenia obliczeń, przypisać całym tonom stopień dysonansowości 1, a ostrzejszym półtonom – wyższy stopień dysonansowości 2. Małym odległościom (1,008–1,010, tzn. poniżej 1/12 tonu), które pojawiają się w interwałach temperowanych, przypisujemy jedynie stopień dysonansowości 0,5[42]. Następnie mogliby-

41 Pewnie wynika to stąd, że rozważane interwały są w mniejszym lub większym stopniu wielokrotnościami półtonu. Gdybyśmy rozważali kroki interwałowe np. co 1/10 tonu, wówczas rozkład dysonansów pomiędzy składowymi w sposób bardziej ciągły przechodziłby od półtonów do całych tonów i trudniej byłoby ustalić jakąś niearbitralną granicę między tymi grupami.

42 Dodatkowo interwałom bliskim tamtym, pojawiającym się w tercji wielkiej pitagorejskiej oraz trytonie naturalnym (1,012 tzn. ok. 1/10 tonu) nie przypisujemy pełnej dysonansowości półtonu (tzn. 2), a jedynie – całego tonu (1). Zostało to uwidocznione w powyższej tabeli przez kolor szary zamiast pomarańczowego.

śmy zsumować te liczby dla każdego interwału i uzyskać jego prosty indeks dysonansowości. Przykładowo:

dla tercji małej wynosiłby on: $1 + 2 + 2 + 2 = 7$
dla kwarty: $1 + 2 + 1 + 2 = 6$

Dla kwinty i seksty wielkiej – również 6. Takie postępowanie nie uwzględnia jednak w ogóle pozycji dysonansu w szeregu składowych: tak samo wycenia jego wystąpienie np. na drugiej jak i na ósmej składowej. W konsekwencji nie rozróżnia pomiędzy kwintą, kwartą a sekstą wielką. Aby uwzględnić również ten czynnik można przypisać wystąpieniom dysonansów na kolejnych składowych odpowiednią wagę, np. 1, 1/2, 1/3, 1/4, 1/5 itd., ponieważ typowo siła kolejnych składowych jest coraz mniejsza. Tzn. przy sumowaniu stopni dysonansowości będziemy je najpierw mnożyć przez odpowiedni czynnik: 1/3 jeśli dysonans występuje na 3 składowej, 1/6 – jeśli występuje na szóstej itd. W ten sposób wystąpienie takiego samego dysonansu (powiedzmy półtonu) na trzeciej składowej ($1/3 \times 2 = 2/3$) znacznie silniej wpłynie na zsumowany indeks dysonansowości niż jego wystąpienie na szóstej składowej ($1/6 \times 2 = 1/3$). Uzyskany w ten sposób indeks dysonansowości nazwiemy ważonym.

Naturalnie przyjęcie akurat takich mnożników jest decyzją nieobojętną, bo niejako przypisuje hipotetycznemu instrumentowi pewną barwę. Jednak każda próba rozróżnienia pomiędzy występowaniem dysonansu na różnych składowych wiązałaby się z pewną decyzją w tym względzie. Możemy powiedzieć, że po prostu testujemy teoretycznie wiarygodność przedstawianej koncepcji konsonansu dokonując takiego akurat wyboru, który może być potraktowany jako pierwsze przybliżenie tego, co typowe. Jednocześnie można mieć nadzieję, że niewielkie zmiany wybranych mnożników miałyby niewielki wpływ na relatywną dysonansowość interwałów i w konsekwencji nie wpłynęłyby na ich hierarchię. Obliczenia wykonane według powyższych założeń dają następującą hierarchię interwałów ustaloną na podstawie indeksu ważonego dla 8 składowych. Gwiazdki w ostatniej kolumnie oznaczają interwały, których pozycja według indeksu ważonego dla 6 składowych byłaby niższa.

Interwał	Indeks dysonansowości dla 8 składowych		Indeks dysonansowości dla 6 składowych		
	Ważony	Prosty	Ważony	Prosty	
Oktawa	0,1430	1	0,0000	0	
Kwinta	0,9860	6	0,4500	2	
Kwinta temperowana	0,9860	6	0,4500	2	
Seksta wielka	1,0357	6	0,4997	2	
Seksta wielka temperowana	1,1357	6,5	0,5997	2,5	
Kwarta	1,1857	6	0,8997	4	*
Kwarta temperowana	1,1857	6	0,8997	4	*
Tercja mała	1,1860	7	0,6500	3	
Tercja mała temperowa	1,2693	7,5	0,7333	3,5	
Tercja wielka	1,3693	8	0,8333	4	
Tercja wielka temperowana	1,4693	8,5	0,9333	4,5	
Seksta mała	1,5423	7	1,3993	6	*
Tercja wielka pitagorejska	1,5693	9	1,0333	5	
Seksta mała temperowana	1,6048	7,5	1,3993	6	*
Septyma mała mniejsza	1,7277	8	1,3167	5	
Septyma mała temperowana	1,7277	8	1,3167	5	
Septyma mała większa	1,7277	8	1,3167	5	
Tryton mały	1,9493	9	1,6993	7	
Tryton temperowany	2,0208	9,5	1,6993	7	
Tryton naturalny	2,0923	10	1,6993	7	
Septyma wielka	2,5693	11	2,0333	7	
Septyma wielka temperowana	2,5693	11	2,0333	7	
Cały ton	3,1523	11	2,6163	7	
Cały ton temperowany	3,1523	11	2,6163	7	
Półton naturalny	5,4353	16	4,8993	12	
Półton temperowany	5,4353	16	4,8993	12	

Tabela 3.
Hierarchia konsonansowości/dysonansowości interwałów według indeksu ważonego dla 8 składowych

Jak widać interwały temperowane znajdują się w bezpośrednim sąsiedztwie interwałów czystych, co oznacza, że wybór jednej lub drugiej wersji nie wpływa na pozycję interwału w hierarchii. Możemy więc, dla uzyskania większej przejrzystości, ograniczyć się do interwałów czystych:

Interwał	Indeks dysonansowości dla 8 składowych		Indeks dysonansowości dla 6 składowych		
	Ważony	Prosty	Ważony	Prosty	
Oktawa	0,1430	1	0,0000	0	
Kwinta	0,9860	6	0,4500	2	
Seksta wielka	1,0357	6	0,4997	2	
Kwarta	1,1857	6	0,8997	4	*
Tercja mała	1,1860	7	0,6500	3	
Tercja wielka	1,3693	8	0,8333	4	
Seksta mała	1,5423	7	1,3993	6	*
Septyma mała	1,7277	8	1,3167	5	
Tryton mały	1,9493	9	1,6993	7	
Tryton naturalny	2,0923	10	1,6993	7	
Septyma wielka	2,5693	11	2,0333	7	
Cały ton	3,1523	11	2,6163	7	
Półton	5,4353	16	4,8993	12	

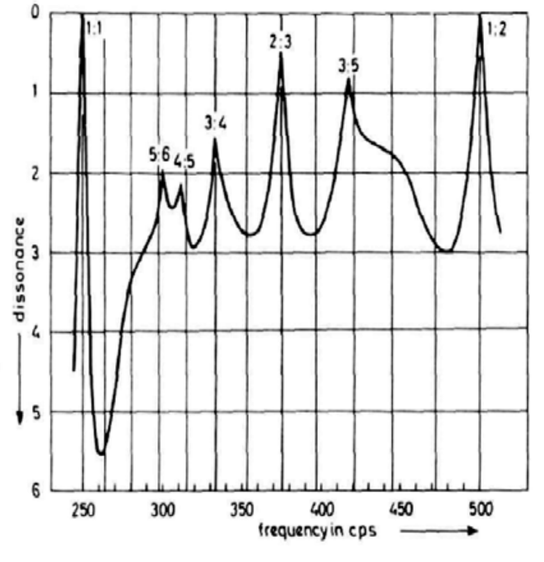
Tabela 4.
Hierarchia konsonansowości/dysonansowości interwałów czystych

Hierarchia ta, uzyskana przy mocno upraszczających założeniach, prezentuje się zaskakująco dobrze. Jedyne łatwo widoczne punkty rozbieżności ze średnią typowych ujęć to wysokie pozycje seksty wielkiej oraz tercji małej: przed tercją wielką i praktycznie na równi z kwartą. Jak się jednak okaże, pozycje te w żadnym razie nie są przypadkowe, a w pewnym sensie jak najbardziej poprawne.

Co więcej, nawet wartość indeksu prostego (kolumna 3) w zasadzie zgadza z hierarchią ustaloną według indeksu ważonego. Jedyna rozbieżność polega na zamianie pozycji tercji wielkiej i seksty małej, co łatwo zrozumieć: wszystkie dysonanse występujące w tej pierwszej pojawiają się na późniejszych składowych niż w przypadku seksty małej, czego indeks uproszczony w ogóle nie uwzględnia i dopiero indeks ważony wykazuje wynikającą stąd wyższość tercji wielkiej. (Naturalnie sam w sobie indeks prosty jest niewystarczający, skoro nie rozróżnia np. pomiędzy kwintą, kwartą i sekstą wielką czy pomiędzy tercją wielką a septymą małą.)

Widać także, że hierarchia powyższa w zasadzie rozwiązuje trudność związaną z wcześniejszą tabelą 1, opartą na różnicach częstotliwości. Pojawiła się tam sugestia, że tercja mała, tryton, seksta mała i septyma mała są interwałami porównywalnymi co do ich konsonansowości/ dysonansowości, z czym naturalnie trudno było się zgodzić. Hierarchia oparta na stosunkach częstotliwości wykazuje wyraźne różnice pomiędzy nimi. Zresztą aby dostrzec przewagę tercji małej nad pozostałymi trzema interwałami nie trzeba się nawet uciekać do obliczeń – wystarczy spojrzeć na tabelę 2 i rozkład dysonansów w każdym z wymienionych interwałów.

Jak można się spodziewać, podobnych obliczeń – opartych na dodawaniu wartości dysonansów pomiędzy składowymi – dokonali także Plomp i Levelt. Wynik przedstawili za pomocą następującego wykresu:[43]



Przedstawia on zmienność poziomu konsonansowości dla interwału przebiegającego w sposób ciągły wszystkie wielkości: od unisonu (250–250 Hz) do oktawy (250–500 Hz). Pionowe linie oznaczają interwały równomiernie temperowane. Jak widać wykres uwidacznia ich praktyczną identyczność w wypadku kwinty i kwarty i niewielką rozbieżność dla tercji i seksty. Z wykresu tego można wyczytać następującą hierarchię interwałów:

43 R. Plomp, W. J. M. Levelt, *op. cit.*, s. 556.

Oktawa, kwinta, seksta wielka, kwarta, septyma mała (sic!), tercja mała, tercja wielka, seksta mała – tryton (praktycznie nierozróżnialne), septyma wielka, cały ton, półton.

Pozycja septymy małej jest naturalnie nie do zaakceptowania. Jak wyglądały obliczenia autorów prowadzące do takich rezultatów? Przede wszystkim uwzględnili oni tylko 6 składowych. To jednak nie wydaje się tłumaczyć szczególnej pozycji septymy małej. W powyższej tabeli 3 zostały przedstawione, w czwartej i piątej kolumnie, rezultaty obliczeń dla sześciu składowych. W uproszczonej tabeli 4, obejmującej tylko interwały czyste, widać, że w stosunku do wyliczeń opartych na ośmiu składowych kolejność musiałaby zostać zamieniona tylko w dwóch przypadkach: kwarta musiałaby zostać zdegradowana poniżej obu tercji, a seksta mała – poniżej septymy małej. To zrozumiałe, bo zarówno kwarta jak i seksta mała mają unison na ósmej składowej, więc zyskują na jej uwzględnieniu, a tracą na jej pominięciu. U Plompa i Levelta kwarta ciągle jednak jest przed tercjami, przed którymi jest, zaskakująco, także septyma mała. Dalsza różnica w metodologii obliczeń autorów polegała na tym, że nie opierali się na szacunkowej ocenie ostrości współbrzmień między składowymi (1 – dla całych tonów, 2 – dla półtonów), ale dokładnie wyznaczali ich dysonansowość według odpowiedniej skali. Ich wyliczenia były więc pod tym względem bardziej precyzyjne. Natomiast najwyraźniej nie stosowali żadnego algorytmu uwzględniającego pozycję poszczególnych dysonansów w szeregu składowych. Innymi słowy wyliczali to, co powyżej nazwaliśmy indeksem prostym. Ale jak widać w powyższej tabeli 4, dla sześciu składowych indeks ten daje w zasadzie identyczne rezultaty jak indeks ważony[44] i nadal odmienne niż Plompa i Levelta[45]. Ponieważ autorzy nie zaprezentowali swoich obliczeń, trudno ustalić, co zadecydowało o takich właśnie wynikach. Warto natomiast zwrócić uwagę na jeszcze jedną ich cechę: otóż wykres pokazuje ogromne oddalenie w stopniu dysonansowości pomiędzy całym tonem a półtonem. To samo zresztą wynika z wyliczeń przedstawionych w tabeli 3, podczas gdy zwykle sądzi się, że np. półton *c*^{1}–*cis*^{1} i cały ton *c*^{1}–*d*^{1} są porównywalne pod względem ostrości (a wykres Plompa i Levelta odnosi się właśnie do interwałów z niższym dźwiękiem 250 Hz, a więc pomiędzy *h* i *c*^{1}) lub różnią się co najwyżej nieznacznie. Na pierwszy rzut oka wydaje się stanowić to znacznie mniejszy problem niż nie-

44 Wartości indeksu prostego dla 6 składowych w piątej kolumnie nie zgadzają się z kolejnością uszeregowania interwałów. Ale nie jest to rozbieżność z indeksem ważonym dla 6 składowych, a z indeksem dla 8 składowych. Gdyby interwały uszeregować zgodnie z indeksem ważonym dla 6 składowych, wówczas również wartości indeksu prostego będą się zgadzać z taką kolejnością.

45 Przy okazji warto zauważyć, że ich hierarchia niemal pokrywa się z pierwszą, najbardziej przybliżoną, która była rozważana powyżej i odrzucona. Różnica polega w zasadzie jedynie na nienaturalnie wysokiej pozycji całego tonu, która nie występuje u Plompa i Levelta.

naturalnie wysoka pozycja septymy małej. A jednak ta rozbieżność wskazuje na pewne istotne ograniczenie całego dotychczas przedstawionego szacunku stopnia dysonansowości.

XI. Dysonujące kwinty – zgodne całe tony

W ogólności należy przypomnieć, że inaczej niż wcześniejsza tabela 1, skonstruowana dla dźwięku wyjściowego 300 Hz (co najwyżej z nadzieją, że dla innych dźwięków wyniki mogą być podobne[46]), tabela 2 odnosi się do interwałów położonych na dowolnej wysokości i pokazuje odległości interwałowe pomiędzy składowymi, tzn. proporcje częstotliwości, które są zawsze identyczne we wszystkich rejestrach – wynika to z samych obliczeń arytmetycznych. Gdyby więc nasze postrzeganie ostrości interwałów pomiędzy tonami prostymi było zależne wyłącznie od ich wielkości (i niezależne od rejestru), oznaczałoby to, że i percepcja konsonansów i dysonansów pomiędzy dźwiękami złożonymi byłaby taka sama we wszystkich rejestrach. Byłaby to oczywiście ogromna zaleta takiego modelu: jedna tabela stosowałaby się do wszystkich interwałów na wszelkich wysokościach, w przeciwieństwie do wcześniejszej wersji odwołującej się do różnic częstotliwości, gdzie w zasadzie dla każdej wysokości trzeba by konstruować osobną tabelę i przeprowadzać nowe obliczenia. Jednocześnie przypuszczenie, że konsonanse i dysonanse są zawsze w zasadzie takie same, może nie budzić na pierwszy rzut oka wątpliwości. Wszak gdy mówimy, że kwinta jest konsonansem, a cały ton dysonansem, to milcząco zakładamy, że jest to niezależne od rejestru. Co więcej, jesteśmy może nawet skłonni sądzić, że kwinta jest zawsze tak samo zgodnym konsonansem, a cały ton – tak samo ostrym dysonansem. Wyjaśnienie pitagorejskie z pewnością przyczynia się do umocnienia takiego domniemania. Wszak jeśli konsonansowość kwinty wynika z prostej proporcji 3 : 2, to nie ulega wątpliwości, że proporcja ta jest taka sama dla każdej kwinty, w każdym rejestrze.

A jednak muzykom dobrze wiadomo, że konsonanse i dysonanse nie wszędzie są takie same i każdy laik wyposażony w poprawnie nastrojoną klawiaturę może się o tym łatwo przekonać. Tercja $c^1 – e^1$ i kwarta $c^1 – f^1$ brzmią całkiem dobrze w oktawie razkreślnej, pogarszają się w oktawie małej, a w oktawie wielkiej tercja ta brzmi wręcz jak dysonans. Nawet kwinta $C – G$ jest w oktawie wielkiej wyraźnie gorszym współbrzmieniem. To zjawisko pogarszania się konsonansowości było od dawna znane i powodowało unikanie mniej doskonałych interwałów w niższych rejestrach: ograniczano się tam do oktaw i co najwyżej kwint. Nieco rzadziej wspomina się o analogicznym zjawisku w odniesieniu do wyższego rejestru: tutaj z kolei jakość niektó-

46 Jak pamiętamy nadzieja ta okazała się uzasadniona co najwyżej dla pobliskich częstotliwości, a całkiem błędna – dla znacząco oddalonych.

rych współbrzmień wyraźnie się poprawia. W szczególności cały ton *c¹–d¹*, który jest niemal tak ostry jak półton *c¹–cis¹*, dwie, a zwłaszcza trzy oktawy wyżej zdecydowanie traci na ostrości i staje się znacznie mniej dysonujący niż sąsiadujący półton. To samo można powiedzieć o tercji małej: *c¹–es¹* brzmi wyraźnie gorzej niż tercja wielka *c¹–e¹*. Oktawę wyżej ich zgodność praktycznie się wyrównuje, a kolejne dwie oktawy wyżej tercja mała staje się wręcz lepszym konsonansem. Nie tylko więc stopień konsonansowości zmienia się wraz z rejestrem, ale co więcej dzieje się to nierównomiernie dla różnych interwałów – tercja mała najwyraźniej poprawia się szybciej niż wielka – co w efekcie prowadzi do odwrócenia hierarchii.

Okazuje się, że przedstawiona teoria konsonansu może wyjaśnić te zjawiska – choć wydaje się, że sami Plomp i Levelt nie wyciągnęli w pełni właściwych wniosków z przeprowadzonych badań. Określili oni – jak wspomniano powyżej – miejsce zaniku ostrości współbrzmienia tonów prostych jako szerokość pasma krytycznego (*critical bandwidth*), a miejsce maksimum ostrości – jako jedną czwartą tej szerokości. Zaś szerokość ta miała być – to kluczowy punkt – nieznacznie mniejsza niż tercja mała. Właśnie przy takim założeniu robione były wyliczenia stopnia konsonansowości interwałów pomiędzy dźwiękami złożonymi uwidocznione w tabelach 3 i 4.

Jednak takie określenie szerokości pasma krytycznego – jako nieco mniej niż tercja mała – stosuje się bez zastrzeżeń jedynie do częstotliwości większych niż 1000 Hz. W literaturze przedmiotu najczęściej pojawia się stwierdzenie, że począwszy od częstotliwości ok. 500 Hz w dół ta szerokość zaczyna się szybko zwiększać[47]. Jednak z wykresów zamieszczonych u Roederera[48] oraz Plompa i Levelta[49] można odczytać, że zmiana rozpoczyna się w istocie około 1000 Hz: już wtedy szerokość pasma krytycznego zaczyna znacząco przybliżać się do tercji małej. Na wysokości 500 Hz jest bliższa tercji wielkiej, a na wysokości 250 Hz – kwinty.

Oznaczałoby to, że dopiero dla takiego interwału zanikałaby ostrość współbrzmienia pomiędzy tonami prostymi. Z tego jednak wynikałoby, że np. tercja wielka 200–250 Hz (tzn. w przybliżeniu as – c¹) powinna być już stosunkowo ostrym dysonansem (dysonans występowałby tutaj pomiędzy pierwszymi składowymi obu, byłby więc dominujący). Wiadomo, że tak nie jest. Okazuje się, że faktyczne badania empiryczne Plompa i Levelta dla interwałów o średniej częstotliwości 250 Hz[50] wskazują, iż ostrość współbrzmienia między

47 Por. np. R. Plomp, W. J. M. Levelt, *op. cit.*, s. 556.

48 Juan G. Roederer, *op.cit.*, s. 40.

49 R. Plomp, W. J. M. Levelt, *op. cit.*, s. 555.

50 *Ibidem*, wykres na s. 553. Interwały o średniej częstotliwości 250 Hz to np. interwały 249–251 Hz, 245–255 Hz, 230–270 Hz itd. Określenie to wyznacza więc rejestr, w jakim znajduje się interwał, ale nie jego wielkość.

tonami prostymi zanika już właśnie w okolicach tercji wielkiej, a więc znacznie poniżej szerokości pasma krytycznego, która miałaby tutaj wynosić niemal kwintę. Nie ulega jednak wątpliwości, że granica zaniku dysonansu zwiększa się wyraźnie wraz z obniżaniem się rejestru. Na wysokości 500 Hz przekracza wyraźnie tercję małą, na wysokości 250 Hz sięga tercji wielkiej. Oznacza to, że tercja mała jednoznacznie wkracza tutaj w zakres umiarkowanej ostrości współbrzmienia.

Podobnie rozszerza się zakres maksymalnej ostrości współbrzmienia. Podczas gdy w okolicach częstotliwości 1000 Hz rozciąga się on pomiędzy 1,015 a 1,083[51], czyli dokładnie w pobliżu półtonu (1,06), w okolicach częstotliwości 500 Hz i 250 Hz[52] sięga raczej od półtonu do całego tonu. Na tych wysokościach nie ma zatem powodu, aby cały ton uznawać za słabiej dysonujący niż półton.

Jakie wynikają stąd wnioski dla oceny konsonansowości i dysonansowości interwałów pomiędzy dźwiękami naturalnymi? Okazuje się, że wcześniejsze wyliczenia stosują się bez ograniczeń tylko powyżej 1000 Hz. To dźwięk pomiędzy h^2 i c^3, a więc ustalona powyżej hierarchia interwałów (tabela 3 i 4) obowiązuje dopiero począwszy od oktawy trzykreślnej wzwyż. Jednak właśnie poniżej tej wysokości rozpoczyna się rejestr najbardziej muzycznie istotny: typowy zakres sopranów, altów, tenorów i wreszcie basów[53]. Ale w tym zakresie, wraz z obniżaniem się rejestru, stopniowo zwiększa się zarówno granica zaniku dysonansu (powyżej tercji małej) jak i maksymalnej ostrości (powyżej półtonu) pomiędzy tonami prostymi, co naturalnie wpływa na konsonansowość/dysonansowość interwałów pomiędzy dźwiękami naturalnymi. Okazuje się więc, że stałość konsonansowości i dysonansowości rozpoczyna się tam (powyżej 1000 Hz), gdzie niemal kończy się najistotniejszy rejestr muzyczny. A w ramach tego ostatniego dysonansowość interwałów zwiększa się wraz z obniżaniem się wysokości.

Możemy teraz ustaloną wcześniej hierarchię (tabela 3 i 4) wziąć za punkt wyjścia i rozważyć, co będzie się działo wraz z obniżaniem rejestru poniżej 1000 Hz. Ponieważ granica zaniku dysonansu najpierw przekroczy tercję małą, więc w pierwszym rzędzie ucierpią te interwały, w których odległość taka występuje pomiędzy składowymi. To oczywiście przede wszystkim sama tercja mała, z takim właśnie interwałem na pierwszej, drugiej i trzeciej składowej. Z tego powodu spadnie ona poniżej tercji wielkiej, a także poniżej kwarty, jeżeli była powyżej niej w hierarchii wyliczonej na podstawie tylko pierwszych sześciu składowych. W kwarcie bowiem tercja mała nie występuje po-

51 *Ibidem*, wykres na s. 554.

52 *Ibidem*, wykresy na s. 553.

53 Typowo określa się zakres sopranów chóralnych jako $c^1 - a^2$. Jedynie sopran solowy sięga do początków oktawy trzykreślnej. Natomiast basy sięgają w dół do niższych partii oktawy wielkiej.

między składowymi. W drugiej kolejności pogorszy się seksta wielka, w której tercja mała występuje na drugiej i czwartej składowej – także więc i seksta wielka w którymś momencie okaże się gorszym konsonansem niż kwarta, a być może także niż tercja wielka, w której tercja mała występuje dopiero na trzeciej składowej.

Z drugiej strony istotne będzie zwiększenie ostrości całego tonu pomiędzy tonami prostymi i jego zrównanie w tym względzie z półtonem, co w tabeli 2 można by zilustrować zmianą w odpowiednich miejscach koloru szarego na pomarańczowy. W pierwszym rzędzie wpłynie to naturalnie na sam interwał całego tonu, choć trzeba zwrócić uwagę, że nie od razu na wszystkich jego pierwszych pięciu składowych powinien pojawić się kolor pomarańczowy. Jeśli rozważymy np. wysokość 500 Hz, wówczas pierwsze składowe całego tonu mogą być uznane za podobnie dysonujące jak półton. Ale drugie składowe znajdują się w okolicach wysokości 1000 Hz, a tam cały ton jest jeszcze znacznie mniej ostry niż półton.

Kolejnym interwałem, którego ostrość się zwiększy, będzie septyma mała, w której cały ton występuje na drugiej, czwartej, szóstej i ósmej składowej. Jeśli w hierarchii wyliczonej na podstawie tylko pierwszych 6 składowych była ona nieco powyżej seksty małej, teraz spadnie poniżej niej, bowiem w sekście małej cały ton, który mógłby ulec pogorszeniu, występuje dopiero na siódmej składowej. Ostrość septymy małej będzie stopniowo zbliżać się do septymy wielkiej i przewyższy ostrość trytonu, w którym cały ton występuje jedynie na piątej składowej. Dodatkowo zwiększenie ostrości całego tonu może wpłynąć na relatywne pogorszenie seksty wielkiej (cały ton na trzeciej składowej) w stosunku do tercji wielkiej, w której cały ton nie występuje. Podsumowując, zmiany będą wyglądały tak:

1. spadek tercji małej i seksty wielkiej poniżej kwarty i tercji wielkiej;
2. spadek septymy małej poniżej trytonu;
3. zbliżenie się całego tonu do półtonu.

Okazuje się zatem, że gdy z wysokości 1000 Hz zstąpimy do najbardziej typowego rejestru oktawy razkreślnej, hierarchia interwałów wyznaczona przez rozważany model teoretyczny przyjmie następującą postać:

Oktawa, Kwinta, Kwarta, Tercja wielka, Seksta wielka, Tercja mała, Seksta mała, Tryton, Septyma mała, Septyma wielka, Cały ton, Półton

przy czym ostrość całego tonu będzie porównywalna z półtonem. Całość właściwie idealnie pokrywa się ze standardowym postrzeganiem relatywnej konsonansowości odpowiednich interwałów. Jednocześnie model ten wyjaśnia, dlaczego tercja mała wyprzedza wielką w wyższych rejestrach i dlaczego cały ton staje się tam znacznie łagodniejszy niż półton. Okazuje się więc, że nieco osobliwa hierarchia przedstawiona w tabelach 3 i 4 jest w pełni poprawna – tyle, że odnosi się do częstotliwości powyżej 1000 Hz, tzn. do oktawy trzykreślnej i wyższych. A typowe hierarchie konsonansów konstruowane empirycznie, takie jak powyższa, opierały się naturalnie na dźwiękach z okolic oktawy razkreślnej.

Wykres zaoferowany przez Plompa i Levelta również odnosił się do oktawy razkreślnej: dolny dźwięk wszystkich interwałów to 250 Hz, czyli wysokość pomiędzy *h* a c^1. Dlaczego zatem tercja mała ciągle jest tam lepsza niż wielka, seksta wielka lepsza niż kwarta, a cały ton dużo lepszy niż półton? Wydaje się, że pewnie z powodu prostego dodawania miary dysonansowości pomiędzy poszczególnymi składowymi (indeks prosty), bez uwzględniania pozycji poszczególnych dysonansów w szeregu składowych. Przykładowo gdy rozważamy interwał całego tonu 250–281,25 Hz, cały ton pomiędzy jego pierwszymi i drugimi składowymi (tzn. 250–281,25 Hz i 500 – 562,5 Hz) jest podobnie ostry jak półton. Ale już na kolejnych składowych (w okolicach 750 Hz, 1000 Hz, 1250 Hz) jest znacząco mniej ostry. Tak więc prosta suma dysonansowości pomiędzy składowymi dla półtonu (por. tabela 2) będzie ciągle wyraźnie większa niż dla całego tonu, zwłaszcza jeśli weźmie się pod uwagę tylko 6 pierwszych składowych. Jest jednak jasne, że w takiej sytuacji ostrość współbrzmienia pomiędzy najsilniejszymi pierwszymi składowymi okaże się dominująca i cały ton będzie co najwyżej nieznacznie mniej dysonujący niż półton, co poprawnie pokazałby ważony indeks dysonansowości.

Z kolei dla tercji małej (250–300 Hz) taki właśnie interwał pomiędzy pierwszymi i drugimi składowymi wykazuje już pewną, choć na tej wysokości jeszcze niewielką, ostrość, której tercja wielka jest tutaj pozbawiona. Przy prostym dodawaniu najwyraźniej ostrzejsze współbrzmienia tercji wielkiej na czwartej i szóstej składowej ciągle jeszcze rachunkowo przeważają i wykres Plompa i Levelta przedstawia ją jako interwał nadal gorszy niż tercja mała. W rzeczywistości jednak decydujący wpływ na jakość współbrzmienia mają pierwsze, silniejsze składowe, a tam tercja mała jest na tej wysokości znacząco gorsza – indeks ważony pokazałby więc właściwie jej rzeczywistą pozycję w hierarchii. Podobne rozważania pozwoliłyby pewnie wyjaśnić inne osobliwości hierarchii wynikającej z wykresu Plompa i Levelta: pozycję seksty wielkiej przed kwartą, nienaturalnie wysoką pozycję septymy małej, oraz zrównanie seksty małej z trytonem.

XII. Harmonia tkwi w uchu ludzkim[54]

Jak zatem należy podsumować powyższe rozważania? Źródłowe zjawisko dysonansu między tonami prostymi bierze się z wytwarzanych przez nie szybkich dudnień, które w pewnym zakresie postrzegamy jako ostrość współbrzmienia. Pojawia się ona – najpierw bardzo słabo – powyżej kilku dudnień na sekundę, osiąga maksimum – w zależności od rejestru – pomiędzy 3/8 tonu a całym tonem, a następnie zanika pomiędzy tercją małą a wielką (lub jeszcze większym interwałem w niskich rejestrach). Powyżej tej granicy współbrzmienie tonów prostych jest już konsonansem. Gdyby natura wytwarzała przede wszystkim tony proste, wówczas konsonans byłby powszechny, a dysonans ograniczony tylko do wąskich interwałów o szerokości nieprzekraczającej tercji małej lub wielkiej. Jednak dźwięki naturalne są złożone ze składowych, które eksportują dysonującą bliskość do wyższych rejestrów, a przez to – do szerokich interwałów, które byłyby konsonansami, gdyby składały się z tonów prostych. W ogólności, gdy wybierzemy przygodny interwał (nie ograniczając się do stopni jakiejkolwiek skali) jest wysoce prawdopodobne, że składowe tworzących go dźwięków – które wszak następują w coraz mniejszych odstępach interwałowych (druga oktawę wyżej nad pierwszą, trzecia o kwintę wyżej nad drugą, czwarta znowu o kwartę wyżej, piąta – o tercję itd.) – wielokrotnie zbliżą się do siebie na krytyczną odległość, produkując dysonans. Jedynie w wyjątkowych przypadkach, gdy stosunek częstotliwości wyraża się prostą proporcją (lub jest takiej proporcji bardzo bliski), relacja pomiędzy niektórymi składowymi z krytycznej bliskości zamienia się w tożsamość (lub niemal tożsamość), która rozbraja dysonansowość i przemienia ją w całkowitą zgodność (lub prawie zgodność). W ten sposób ilość potencjalnie niezgodnych składowych znacząco się zmniejsza, co redukuje poziom dysonansowości całości współbrzmienia i czyni z niego relatywny konsonans. W ogólności zatem – jeśli nie ograniczamy się do stopni jakiejś skali – dysonans jest normą, a konsonans szczęśliwym wyjątkiem[55]. To właśnie miał na myśli Helmholtz stwierdzając w przytoczonym powyżej cytacie:

54 Por. przypis 12 powyżej.

55 Rozważmy dla przykładu kwartę i kwintę o wspólnym dźwięku dolnym (np. *c*–*f* i *c*–*g*). Wówczas odległość pomiędzy dźwiękami górnymi wynosi cały ton. Jak wynika z przypisów 14 i 34 kwintę możemy rozstroić o jedną czwartą komatu Didymosa, tzn. o 1/36 całego tonu, bez wyraźnego zaburzania jej konsonansowości. Załóżmy, że byłaby to nawet 1/20 całego tonu. I że przy rozstrojeniu o 1/10 całego tonu otrzymamy już dysonans. Jeśli to samo założymy o kwarcie, wówczas dla interwałów c–x, gdzie x jest pewnym dźwiękiem z przedziału pomiędzy *f* i *g*, uzyskamy (jeszcze) konsonans na przestrzeni 2/10 całego tonu (1/10 powyżej kwarty [*f*] i 1/10 poniżej kwinty [*g*]), natomiast na pozostałej przestrzeni 8/10 całego tonu – otrzymamy dysonanse. Będą one więc zajmowały 4 razy większą część przedziału pomiędzy *f* i *g* niż konsonanse.

> Gdy współbrzmią dwa dźwięki muzyczne, zazwyczaj powstają zaburzenia ich współbrzmienia (…). Nazywamy to dysonansem.
> Ale istnieją pewne proporcje pomiędzy częstotliwościami, przy których występują wyjątki od tej reguły (…) te wyjątki nazywamy konsonansami[56].

Wyjątki występują właśnie przy prostych proporcjach. Czy w takim razie nie wracamy jednak do punktu wyjścia? Tzn. do stwierdzenia, że harmonia z nich właśnie wynika? Przede wszystkim, jak pamiętamy, proste proporcje nie muszą być ścisłe, a tym samym przestają być proste. Może jednak sam fakt, że konsonanse grupują się, jeśli nie dokładnie w punktach prostych proporcji, to przynajmniej blisko wokół nich, jest wystarczająco godny uwagi i wskazuje na szczególną zgodność pomiędzy harmonią a liczbą?

Które jednak proporcje są proste? 6 : 5 i 5 : 4 powyżej 500 Hz są proste. Ale poniżej 200 Hz stają się – zwłaszcza 6 : 5 – już skomplikowane. Poniżej 100 Hz (tzn. w dolnej części oktawy wielkiej) nawet 3 : 2 (kwinta) zaczyna być „skomplikowana”. Z drugiej strony skomplikowany cały ton (9 : 8) w wysokich rejestrach stopniowo zmierza ku prostocie. Wszystko jest naturalnie zależne od omówionych powyżej progów postrzegania ostrości przez ucho. Ale przecież progi te mogłyby wyglądać nieco inaczej. Gdybyśmy byli nietoperzami lub żółwiami – z pewnością prezentowałyby się całkiem odmiennie. Gdyby np. granica zaniku ostrości pomiędzy tonami prostymi zawsze przewyższała kwintę (tak zresztą jest w bardzo niskim rejestrze), wówczas wszystkie współbrzmienia pomiędzy dźwiękami naturalnymi byłyby dysonansami: interwały nie większe od trytonu miałyby dysonans pomiędzy pierwszymi składowymi, interwały szersze – pomiędzy drugą składową niższego, a pierwszą wyższego dźwięku. Jedynie oktawa brzmiałaby nieco lepiej, ale i tu mielibyśmy dysonansowe współbrzmienia na wszystkich składowych począwszy od trzeciej (por. tabela 2 powyżej): kolejno kwartę, kwintę, tercję małą, znowu kwartę. Złe brzmienie oktawy byłoby zresztą skorelowane z faktem, że wszystkie dźwięki naturalne same w sobie brzmiałyby ostro: wszak w każdym takim dźwięku interwał kwinty występuje pomiędzy drugą a trzecią składową, kwarty – pomiędzy trzecią a czwartą itd. Jedyne dostępne uchu ludzkiemu konsonanse występowałyby pomiędzy tonami prostymi odległymi o więcej niż kwintę. Ale konsonanse takie nie miałyby nic wspólnego z prostymi proporcjami, a jedynie z wielkością interwału – przekraczające granicę ostrości seksta, septyma, oktawa, nona oraz wszystkie interwały pośrednie byłyby konsonansami.

To jednak nie koniec alternatywnych możliwości. Możemy wyobrazić sobie, że zakres słyszalnych częstotliwości mógłby być znacznie węższy niż jest w istocie: że obejmowałby jedynie septymę wielką

56 H. Helmholtz, *op. cit.*, s. 320. Por. pełny cytat powyżej, w rozdziale VIII, s. 84.

(powiedzmy $c^1 – h^1$)[57]. Nie jest to bynajmniej całkiem absurdalne wyobrażenie. Wszak w odniesieniu do innego naszego zmysłu, mianowicie wzroku, tak właśnie jest: zakres światła widzialnego obejmuje orientacyjnie długości fali pomiędzy 390 a 700 nanometrów, czyli mniej niż „oktawę" (700/390 = 1,79, co w przybliżeniu odpowiada septymie małej). Takie wyobrażenie w odniesieniu do słyszenia nie musiałoby bynajmniej podważać możliwości istnienia muzyki. Wszak wiele prostych melodii mieści się w granicach septymy. Może zresztą w takiej sytuacji nasza rozdzielczość słuchu byłaby znacznie wyższa i powszechnie posługiwalibyśmy się skalą ćwierćtonową lub jeszcze bardziej rozdrobnioną, co naturalnie poszerzałoby możliwości tworzenia zróżnicowanej muzyki. Jednak w takiej sytuacji nie słyszelibyśmy wyższych składowych żadnych dźwięków: wszak druga składowa jest zawsze o oktawę wyższa niż ton podstawowy. Inaczej mówiąc, wszystkie dźwięki naturalne słyszelibyśmy w zasadzie jak tony proste. Tak więc konsonans byłby znowu tylko i wyłącznie kwestią odległości pomiędzy dźwiękami – i nie miałby żadnego związku z proporcjami.

Można by pewnie nakreślić jeszcze wiele alternatywnych modeli słyszenia i być może okazałoby się, że większość z nich została w naturze – w świecie zwierzęcym – zrealizowana. W wielu z nich nie występowałoby w ogóle rozróżnienie na konsonanse i dysonanse (wrażliwość na częstotliwość dudnień między dźwiękami także nie jest czymś, co koniecznie musimy zakładać). A tam, gdzie by się pojawiało, nie musiałoby się w żaden sposób wiązać z proporcjami, o czym poświadczają modele przedstawionych powyżej.

Warto w tym momencie raz jeszcze powrócić do sprawy idealnego stroju muzycznego, omawianej w rozdziale IV. W rozdziale V kwestia ta została przedstawiona jako problem matematyczny: uzgodnienia pewnej ilości kwint z pewną ilością oktaw, co na gruncie czysto rachunkowym okazało się niemożliwe. Mogło się zatem wydawać, że kwestia ta nie miała związku z wrażliwością ucha. Faktycznie, gdy problem zostanie sformułowany jako pytanie o możliwość uzgodnienia proporcji 2 : 1 z proporcją 3 : 2, staje się on czysto rachunkowy. Teraz jednak okazuje się, że u podstaw wyróżnienia takich proporcji i sformułowania problemu ich uzgodnienia stoją właśnie szczególne cechy ludzkiego słyszenia i wrażliwość ucha. Gdyby konsonans był kwestią odległości (niezgodność dla małych interwałów, zgodność dla wszystkich większych) a nie proporcji, problem właściwego stro-

57 W rzeczywistości zakres dźwięków słyszalnych bywa zwykle określany jako 20 do 20 000 Hz, tzn. obejmuje około 10 oktaw (20 000/20 = 1000; jednocześnie 2^{10} = 1024). Z wiekiem górna granica tego zakresu zwykle szybko opada: do 15 000 Hz dla trzydziestolatków i 5 000 Hz dla sześćdziesięciolatków. Jednak nie oznacza to, że sześćdziesięciolatkowie słyszą „cztery razy mniej" niż ludzie młodzi: wszak spadek z 20 000 Hz do 5 000 Hz to jedynie dwie oktawy. Zakres od 20 do 5 000 Hz obejmuje ciągle jeszcze około 8 oktaw.

ju – a więc harmonii systemu muzycznego jako całości – jawiłby się całkiem inaczej. Tak więc również ten problem okazuje się u swych podstaw uzależniony od specyfiki ludzkiego słyszenia.

Fakt, że konsonanse rzeczywiście mają pewien – choć nie całkiem bezpośredni i oczywisty – związek z prostymi proporcjami, właśnie z tej specyfiki wynika. Pitagorejczycy nie mogli naturalnie wiedzieć o takiej zależności, której ustalenie – w pierwszym przybliżeniu – udało się dopiero w XIX wieku. Dlatego natrafiając na związek pomiędzy konsonansami a prostymi proporcjami mogli mieć wrażenie, że odkryli ogólne prawo dotyczące całości kosmosu i uniwersalnych praw harmonii. Jednak dziś wyobrażenie, że coś mającego tak specyficzny związek z naszym aparatem słuchowym miałoby odzwierciedlać strukturę kosmosu byłoby doprawdy zadziwiającą uzurpacją. Harmonia – przynajmniej muzyczna – nie tkwi w liczbach. Harmonia tkwi w uchu ludzkim.

Abstract

Harmony does not lie in numbers. Of the Pythagoreans, temperaments and harmonious chords

Krzysztof Guczalski

The conviction of a profound link between music and numbers or music and mathematics has been present in the European tradition at least since the times of the Pythagoreans. Its origins are usually sought in the discovery that simple numerical proportions correspond to the most harmonious combinations of notes. During the Middle Ages, music – understood as an abstract science of numerical proportions and their resultant harmonies – was treated alongside arithmetic, geometry and astronomy as one of the mathematical disciplines that formed the so-called Quadrivium. In later times, however, that conviction was sometimes contested, including by Eduard Hanslick.

In this article, I attempt to resolve that dilemma. Since one of the key arguments in favour of the existence of a special link between music and mathematics is the above-mentioned discovery made by the Pythagoreans and the supposedly pursuant numerical grounding of harmony in music, this article will deal primarily with a critical analysis of the latter conviction.

To that end, I will consider the possibility of constructing an ideal musical system in which all chords are expressed through simple numerical proportions. It turns out that the existence of such a system is impossible from a mathematical point of view. Next to be considered is the question as to whether at least the harmoniousness of individual chords can be explained through simple numerical proportions. That hypothesis also proves untenable, leading one to ponder the source of the difference between harmonious and disharmonious chords. Put forward in answer to that question will be Helmholtz's explanation (and its later refinements), according to which dissonance stems from beating.

Analysis of that explanation, illustrated with concrete examples, and rectification of the inaccuracies that frequently appear in its reception lead to the conclusion that the difference between consonances and dissonances results from specific properties of human hearing, and not from numerical proportions.