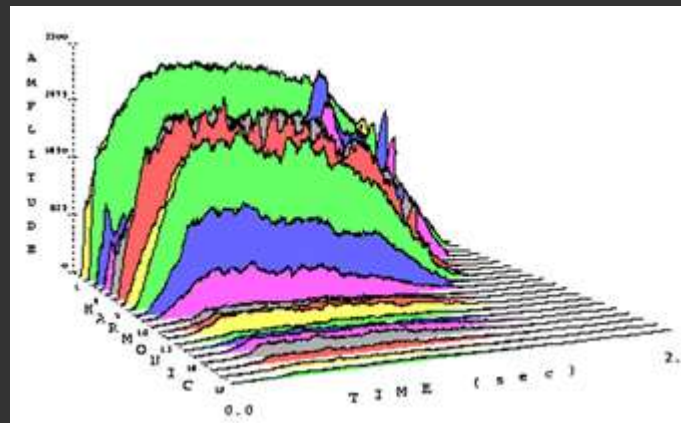


DŹWIĘK MUZYCZNY

Właściwości, analiza i resynteza addytywna



Dźwięk muzyczny

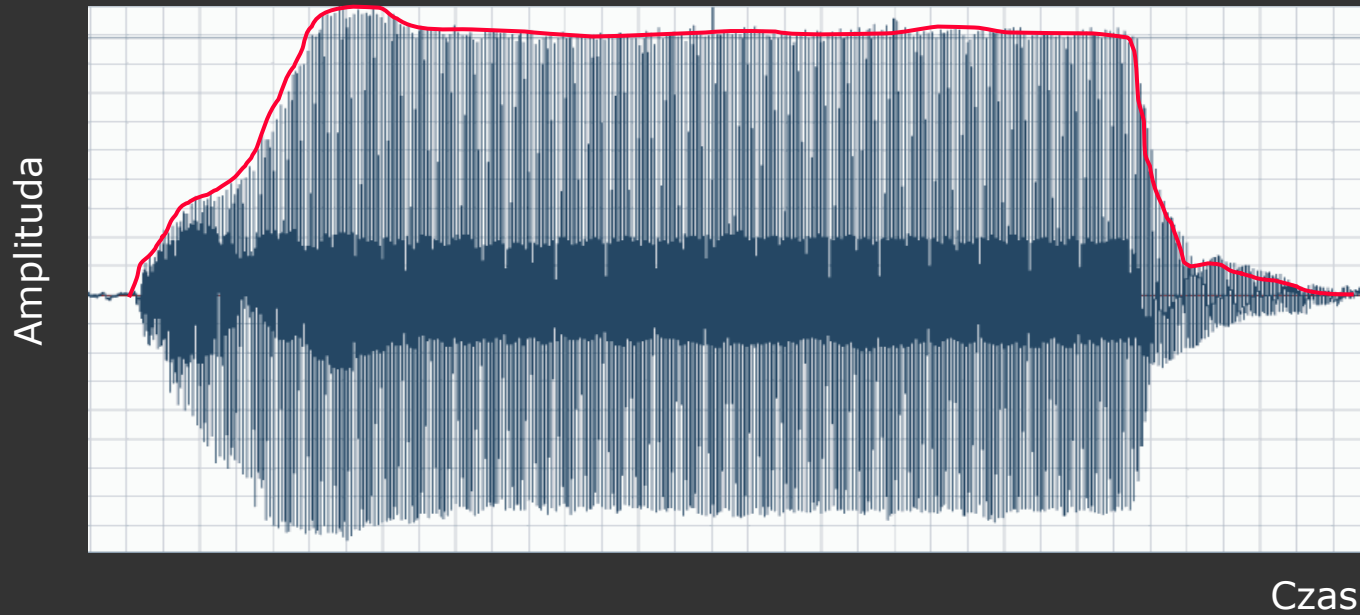
- Dźwięki wytwarzane przez instrumenty muzyczne, stosowane do tworzenia muzyki, są dźwiękami muzycznymi.
- Można wyróżnić dwa podstawowe typy dźwięków muzycznych:
 - dźwięki **melodyczne**,
 - dźwięki **rytmiczne** (np. perkusyjne).
- Dźwięki muzyczne posiadają właściwości:
 - **wysokość** (tylko dźwięki melodyczne),
 - **barwa** – brzmienie dźwięku, jego zmiany w czasie,
 - **obwiednia** – zmiany amplitudy (głośności) w czasie, czas trwania.

Syntetyczny dźwięk muzyczny

- Zadaniem **syntezy dźwięku** jest wytworzenie sygnału o właściwościach odpowiadających dźwiękowi muzycznemu.
- To nie znaczy, że musimy naśladować dźwięki rzeczywistych instrumentów.
- Mogą to być dźwięki **syntetyczne** o dowolnym brzmieniu.
Wystarczy, że posiadają one właściwości dźwięku muzycznego.
- Musimy zapewnić:
 - możliwość ustalania wysokości dźwięku (dla dźwięków melodycznych),
 - odpowiednią strukturę widma dźwięku (barwa dźwięku),
 - zmiany amplitudy w czasie (narastanie, wybrzmiewanie, itp.),
 - zmienność brzmienia dźwięku w czasie (dźwięk musi być żywy).

Obwiednia dźwięku

- **Obwiednia** (*envelope*) dźwięku obrysowuje wykres czasowy dźwięku.
- Obrazuje ona zmiany głośności dźwięku w czasie.
- Można wyróżnić fazy: narastanie, opadanie, podtrzymanie, wybrzmiewanie.



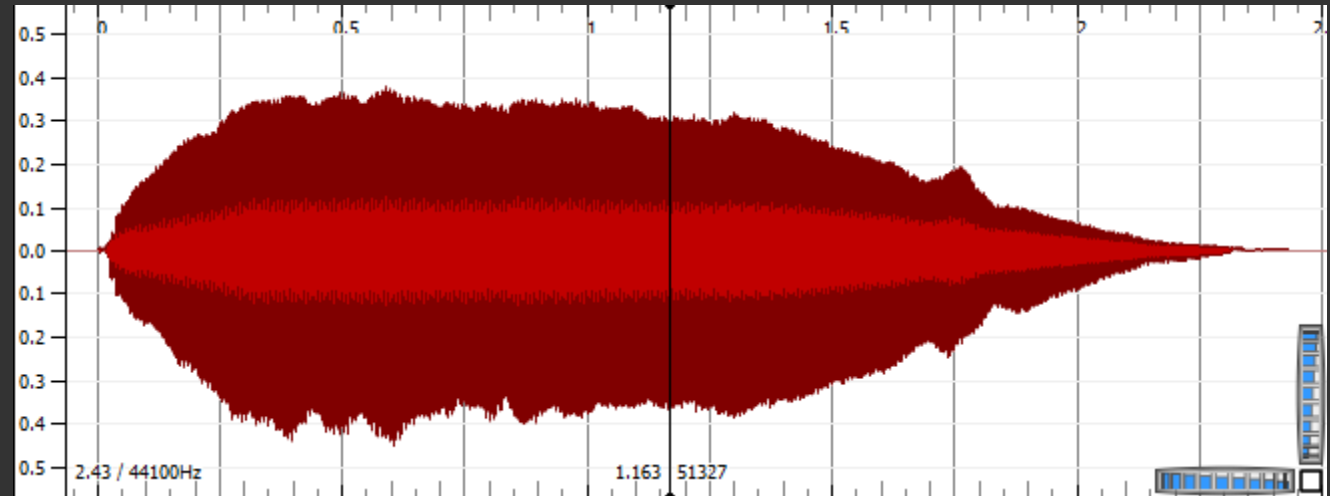
Fazy obwiedni dźwięku muzycznego

- **Fazy ataku (narastania) i opadania (*attack + decay*):**
 - budowanie się dźwięku po pobudzeniu instrumentu,
 - transjent początkowy – stan nieustalony,
 - bardzo duże zmiany barwy, duży wpływ na brzmienie.
- **Faza podtrzymania (*sustain*):**
 - stan ustalony – dźwięk jest stabilny (ale nie musi być niezmienny – np. vibrato),
 - nie musi występować (zależy od typu instrumentu).
- **Faza wybrzmiewania / zwolnienia (*release*):**
 - naturalne wygaszanie dźwięku po zaprzestaniu pobudzania instrumentu.

Przykłady obwiedni dla różnych instrumentów

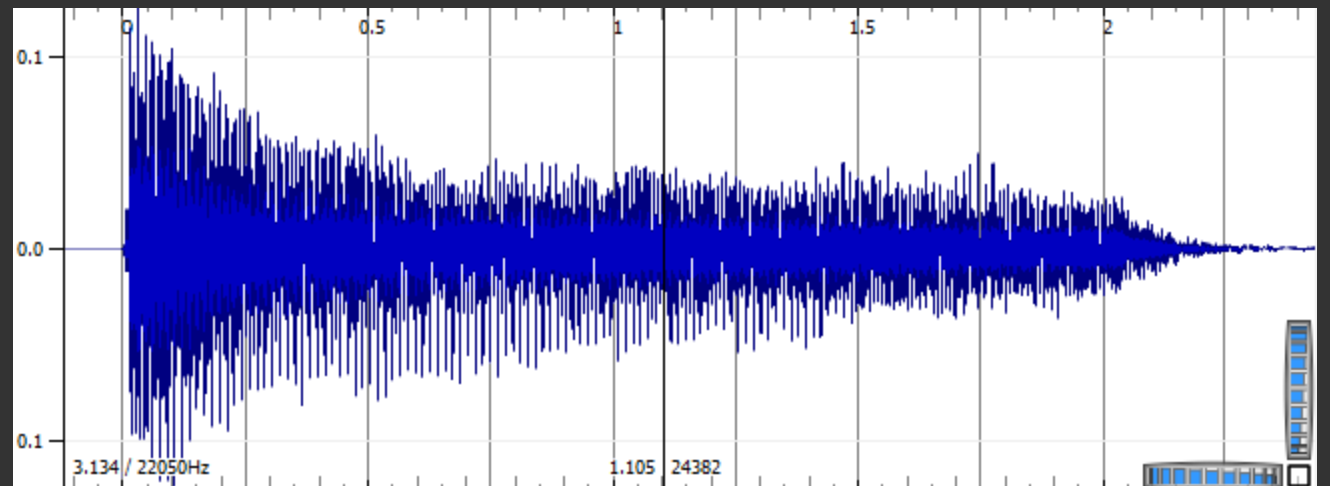
Trąbka

Dość długi atak, możliwy zanik, podtrzymanie przez muzyka, średnio długie wybrzmiewanie.



Fortepian

Bardzo krótki atak, brak podtrzymania, bardzo długie wybrzmiewanie.



Obwiednia a charakter dźwięku

- Kształt obwiedni jest różny dla dźwięków różnych typów instrumentów.
- Czasy trwania faz obwiedni zależą też od **artykulacji** - sposobu wydobywania dźwięku z instrumentu.
Np. mocniejsze szarpnięcie struny w gitarze – dłuższa faza wybrzmiewania.

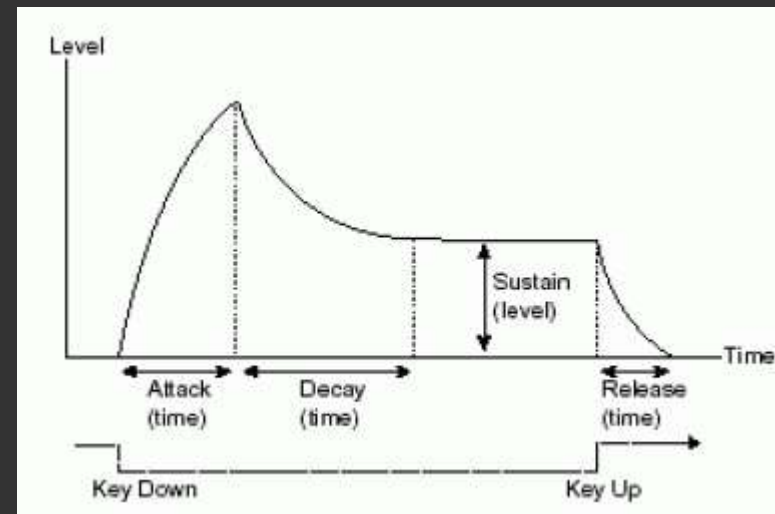
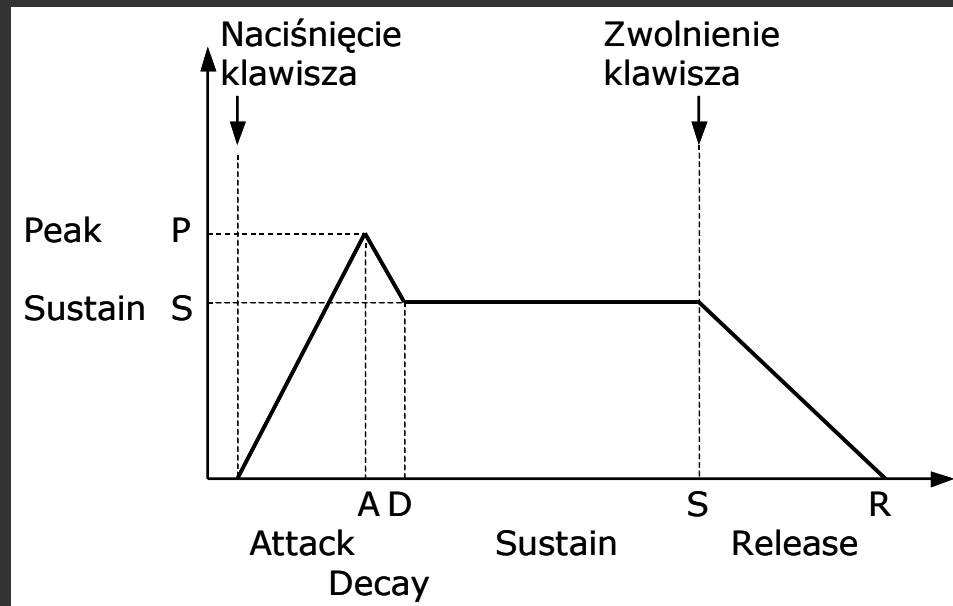
Jak uzyskać efekt obwiedni w **syntezatorze**?

- **Generator obwiedni** wytwarza sygnał sterujący o odpowiednim kształcie.
- Sygnał ten steruje **wzmocnieniem** wyjściowym Instrumentu.
- Dzięki temu uzyskujemy **zmiany głośności** zgodnie z zadanym kształtem obwiedni.
- Początek nuty (wciśnięcie klawisza) – włączenie (*trigger*) fazy ataku.
- Koniec nuty (zwolnienie klawisza) – rozpoczęcie fazy wybrzmiewania.

Obwiednia ADSR

Klasyczna, czteroodcinkowa obwiednia ADSR w syntezatorze:

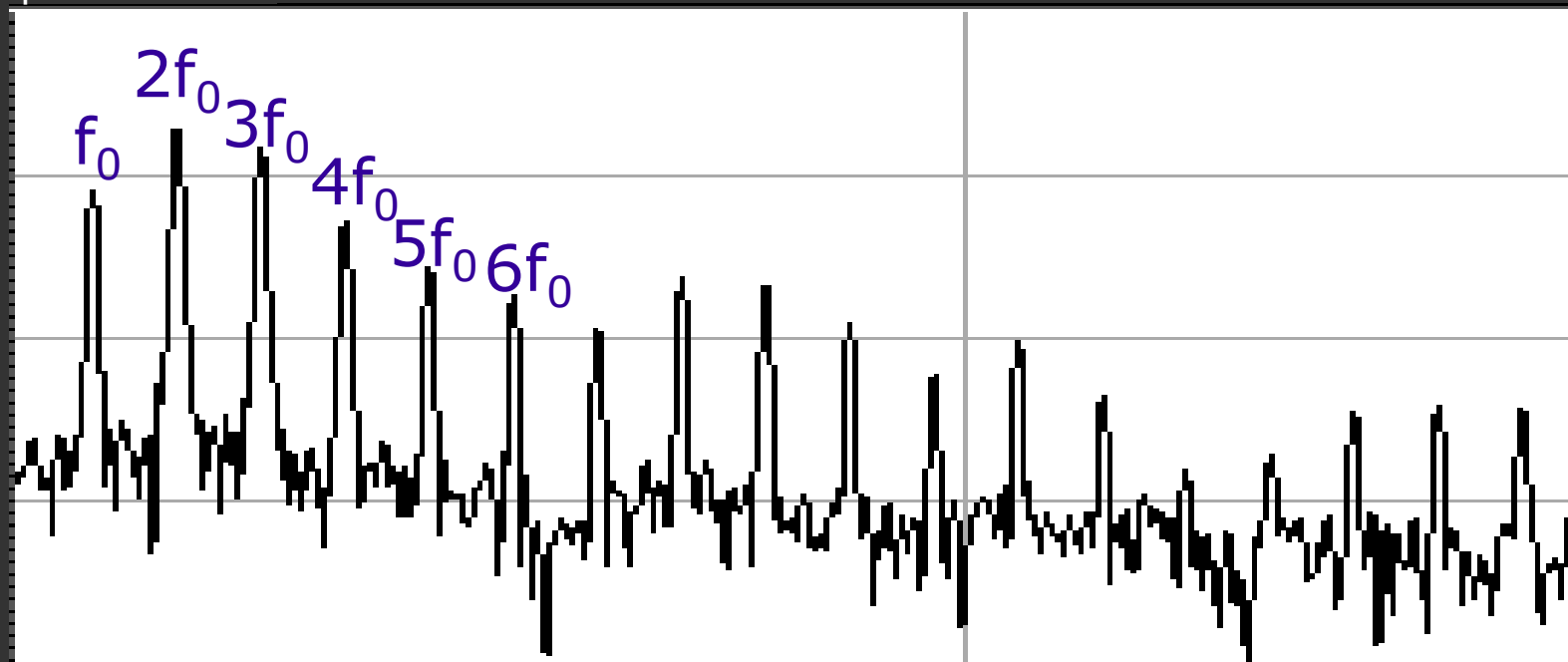
- **A**: czas trwania fazy ataku (*attack*)
- **D**: czas trwania fazy opadania (*decay*)
- **S**: poziom stanu podtrzymania (*sustain*)
- **R**: czas trwania fazy wybrzmiewania (*release*)



Widmo dźwięku muzycznego

- Struktura widmowa dźwięku decyduje o jego **barwie (brzmieniu)** oraz o **wysokości** dźwięku.
- Prawie wszystkie melodyczne dźwięki muzyczne mają widmo **harmoniczne**. Zawiera składowe o częstotliwościach $N \cdot f_0$ ($N = 1, 2, 3\dots$).

Amplituda widma



Częstotliwość

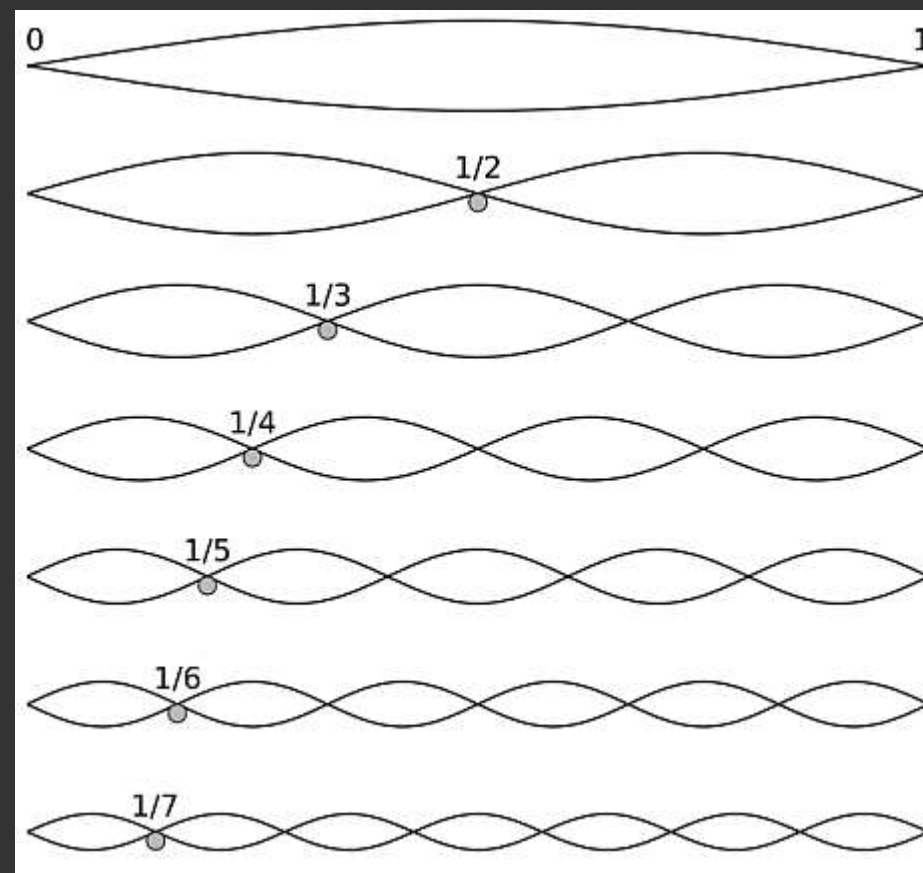
Widmo dźwięku muzycznego

- Mamy wyraźne, izolowane składowe widmowe.
- Składowa **podstawowa** (*fundamental*) – f_0 , pierwsza składowa, która rozpoczyna szereg harmoniczny.
- Składowe na wielokrotnościach f_0 ($2f_0$, $3f_0$, ...) – **harmoniczne** (*partials, overtones, harmonics*, w polskiej terminologii muzycznej: alikwoty).
- Obecny jest szum, mogą pojawić się niewielkie składowe nieharmoniczne.
- Rozkład amplitud składowych widmowych oraz ich zmienność w czasie decydują o **barwie dźwięku** (*timbre*), czyli o tym jak brzmi dany instrument.

Widmo dźwięków muzycznych

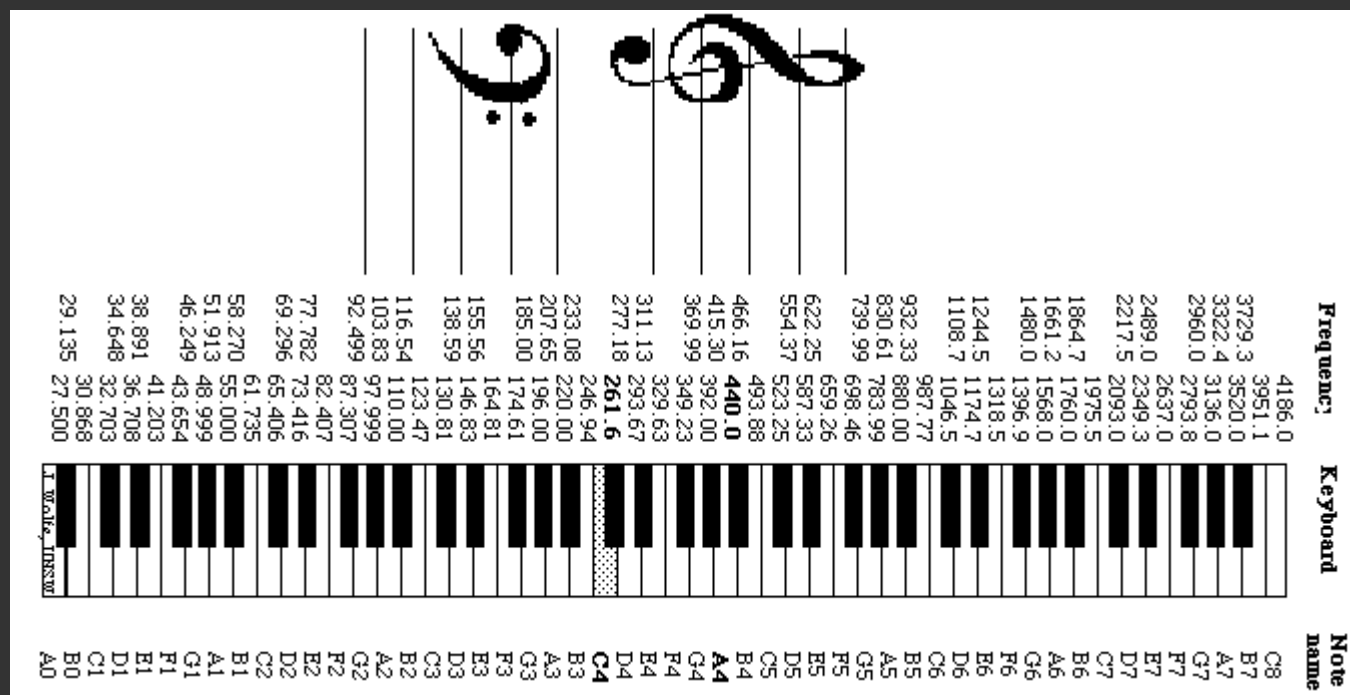
Dlaczego właśnie takie widmo?

- W instrumencie (np. na strunie) powstaje **fala stojąca**.
- Długość fali decyduje o **częstotliwości podstawowej** drgań (f_0).
- Wyższe **mody** o mniejszej długości fali ($1/2$, $1/3$, ...) tworzą **harmoniczne** ($2f_0$, $3f_0$, ...).
- Wynikowe widmo jest **ważoną sumą** wszystkich fal stojących (modów).
- Amplituda modu decyduje o amplitudzie danej harmonicznej.



Wysokość dźwięku

- Wysokość dźwięku (*pitch*) pozwala umieścić dźwięk na skali muzycznej.
- Wysokość dźwięku zależy od częstotliwości podstawowej dźwięku - od częstotliwości pierwszej składowej, która tworzy szereg harmoniczny.
- Częstotliwość drgań fali jest powiązana z wysokością dźwięku poprzez strój muzyczny.
- W syntezatorach standardowo stosuje się strój: $a^1 = 440 \text{ Hz}$, ale możliwe są też inne stroje stosowane w muzyce.



Widmo dźwięków muzycznych

Aby nie było zbyt prosto.

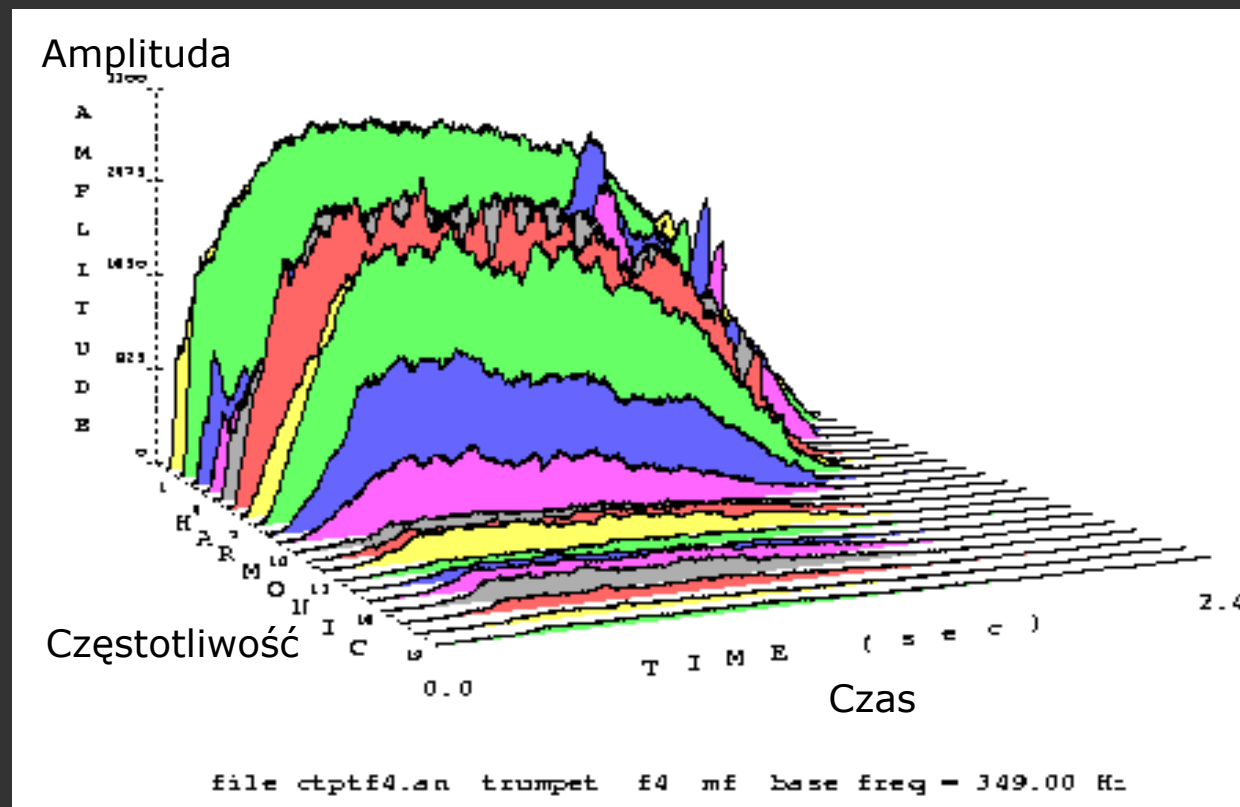
- W widmie dźwięków niektórych instrumentów muzycznych (np. klarnetu) może brakować „parzystych” składowych – to wciąż jest widmo harmoniczne.
- Niektóre instrumenty (np. dzwony) mają widmo nieznacznie nieharmoniczne.
- Jeżeli będzie brakowało składowej f_0 , to i tak ta nieistniejąca składowa decyduje o wysokości dźwięku (bo wyznacza szereg harmoniczny).
- Jeżeli pojawi się podharmoniczna (*subharmonic*) na częstotliwości $\frac{1}{2}f_0$, to nie zmieni ona wysokości dźwięku, ponieważ nie rozpoczyna ona szeregu harmonicznego (będzie słyszalna jako osobny dźwięk).

Dźwięki instrumentów perkusyjnych

- Większość instrumentów perkusyjnych jest używana jako instrumenty **rytmiczne** – wyznaczają rytm muzyki.
- W instrumentach perkusyjnych, drgania (a przez to i fale stojące) powstają **w dwóch wymiarach** (w strunie i słupie powietrza – tylko w jednym).
- Wynikowy dźwięk jest kombinacją wielu modów (X, Y).
- Składowe wytwarzane przez te mody nie są zazwyczaj w porządku harmonicznym. Jest tych składowych tak dużo, że często tworzą **szum**.
- Niektóre instrumenty (np. ksylofon, dzwony) mogą wytwarzać dźwięk w przybliżeniu harmoniczny – tu można wyznaczyć wysokość dźwięku.
- Dla większości instrumentów perkusyjnych (np. zestaw perkusyjny): **brak szeregu harmonicznego**, więc nie można mówić o wysokości dźwięku, nawet jeżeli występuje maksimum widma dające wrażenie „wyżej/niżej”.

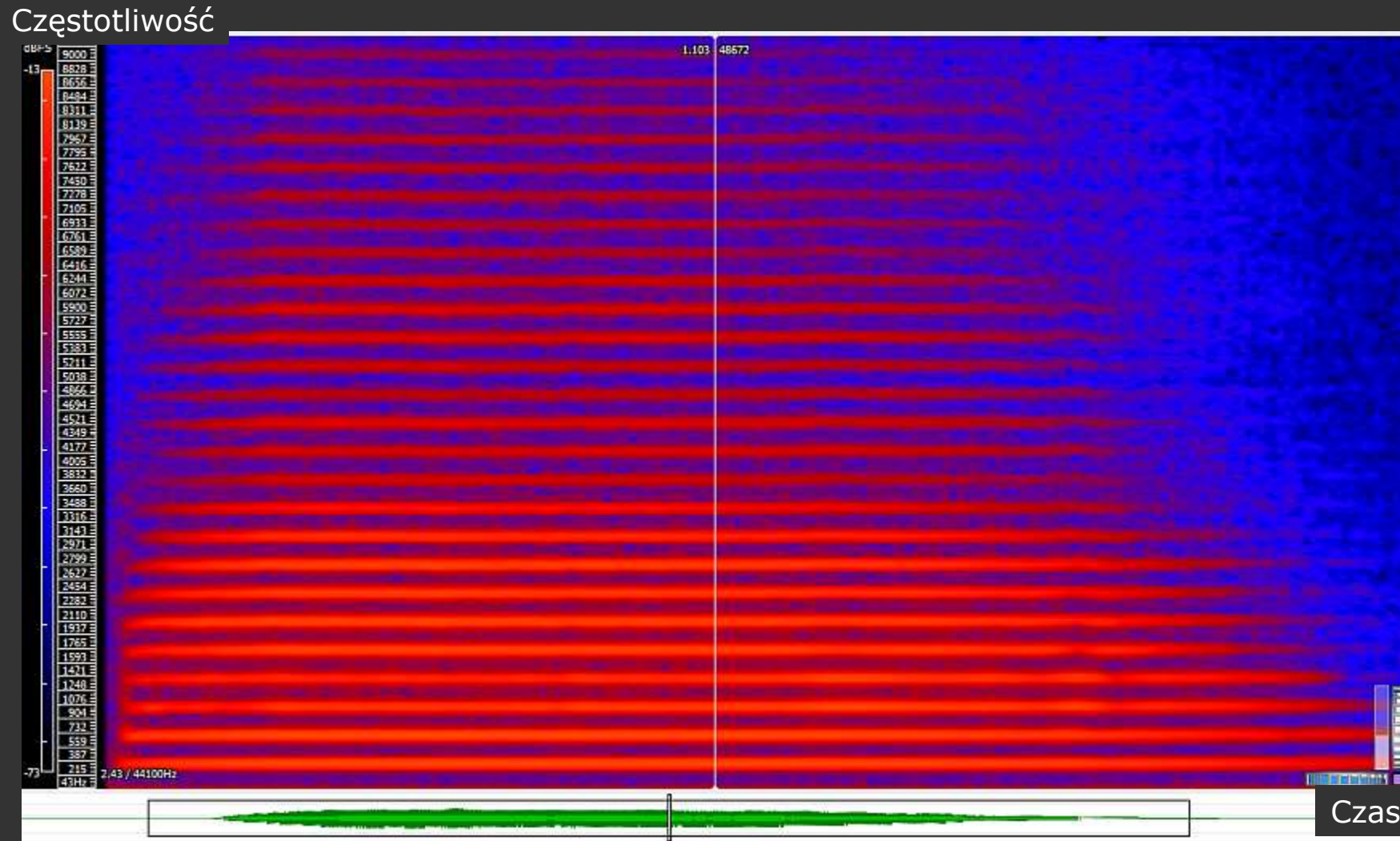
Zmienność widma dźwięku muzycznego

- Jeżeli widmo dźwięku muzycznego przez cały czas jego trwania będzie takie samo, uzyskany dźwięk będzie „martwy”, o bardzo nieciekawym brzmieniu.
- Widmo (a przez to barwa) dźwięku naturalnych instrumentów **zmienia się** w trakcie trwania dźwięku, np. wskutek ekspresji muzyka.
- Aby dźwięk był „żywy” i ciekawy, musimy uwzględnić ten aspekt w syntezie. Musimy zmieniać parametry syntezy odpowiedzialne za barwę dźwięku w trakcie jego generowania, szczególnie w fazie ataku.



Spektrogram - wykres zmienności widma

Widmo 3D: czas – częstotliwość – amplituda widma (kolor). Widoczne fazy dźwięku.



Opisywanie dźwięku

- Dźwięk **niski / wysoki**:
 - zależy od częstotliwości podstawowej (większa = wyższy dźwięk),
 - nie zależy od barwy dźwięku.
- Dźwięk **ciemny / jasny**:
 - opisuje **barwę** dźwięku (charakter brzmienia),
 - **jasny**: szerokie pasmo, wiele silnych wyższych składowych („dużo góry”),
 - **ciemny**: przytłumiony, wąskie pasmo, brak wyższych składowych,
 - nie zależy od wysokości dźwięku.

Częsty błąd studentów: opisywanie dźwięku bardziej jasnego jako „wyższy”.

Opisywanie dźwięku

- Dźwięk **żywy, ciepły, dynamiczny**:
 - zmieniający swój charakter w trakcie grania,
 - zmiany: barwy, wysokości (wibrato), głośności,
 - np. analogowe generatory i filtry – wynik niedoskonałości elementów analogowych, tutaj pożądanych.
- Dźwięk **zimny, statyczny, „syntetyczny”**:
 - brak istotnych zmian brzmienia w trakcie odgrywania dźwięku,
 - np. cyfrowe generatory – stały kształt sygnału, idealna precyzja,
 - nieciekawe, nudne brzmienie,
 - aby ożywić dźwięk, wprowadza się np. **modulację** parametrów syntezy.

Parametry dźwięku a brzmienie

Podsumujmy: dlaczego dźwięki dwóch instrumentów o tej samej wysokości brzmią inaczej?

- Różny kształt obwiedni czasowej (ADSR).
- Różna struktura widma statycznego (amplitudy harmoniczných).
- Różna zmienność widma w trakcie trwania.

Jak zatem różnicować brzmienie dźwięku w syntezie?

- Ustawić odpowiedni kształt obwiedni (proste).
- Ukształtować odpowiednio widmo statyczne (dość proste).
- Zapewnić zmienność widma w czasie, aby dźwięk stał się żywy (znacznie trudniejsze).

Synteza addytywna

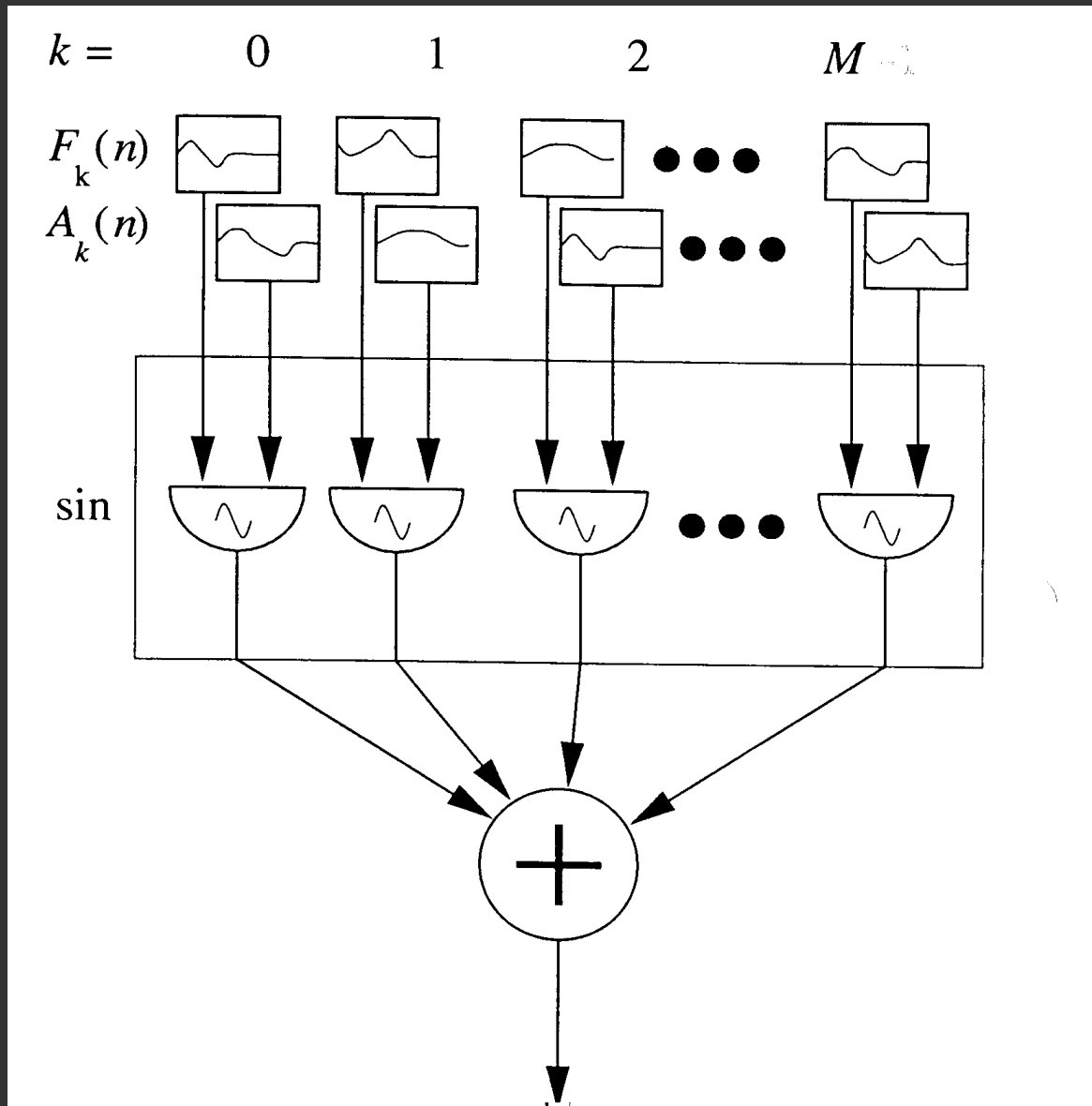
- Wiemy, że harmoniczny dźwięk muzyczny można rozłożyć na ważoną sumę „sinusów” w porządku harmonicznym.
- Możemy też zrobić na odwrót: zsumować „sinusy”:
 - o **częstotliwościach** wynikających z szeregu harmonicznego, o pożądanej częstotliwości podstawowej (wysokości),
 - o **amplitudach** tak dobranych, aby uzyskać pożądany kształt widma.
- Jest to **addytywna synteza dźwięku** (*additive synthesis*).
- Ze względu na skomplikowanie (bardzo duża liczba parametrów), metoda ma praktyczne zastosowanie tylko w cyfrowych instrumentach, szczególnie programowych.

Zmienność widma w syntezie addytywnej

- Jak zapewnić zmienność widma w syntezie addytywnej?
- Parametry syntezy nie mogą być stałe, lecz muszą zmieniać się w czasie.
- Potrzebujemy funkcji zależnych od czasu, które kontrolują:
 - **amplitudę** każdej składowej $A_k(t)$,
 - **odchyłki częstotliwości** każdej składowej $\Delta f_k(t)$
- Wynik syntezy addytywnej: suma składowych (sinusów) o zmiennej w czasie amplitudzie i nieznacznie zmiennej częstotliwości.

$$y(n) = \sum_{k=1}^M A_k(n) \sin(2\pi n(k \cdot f_0 + \Delta f_k(n)))$$

Schemat syntezy addytywnej



Najprostsza koncepcyjnie synteza, jaką da się zrobić.

Problem praktyczny

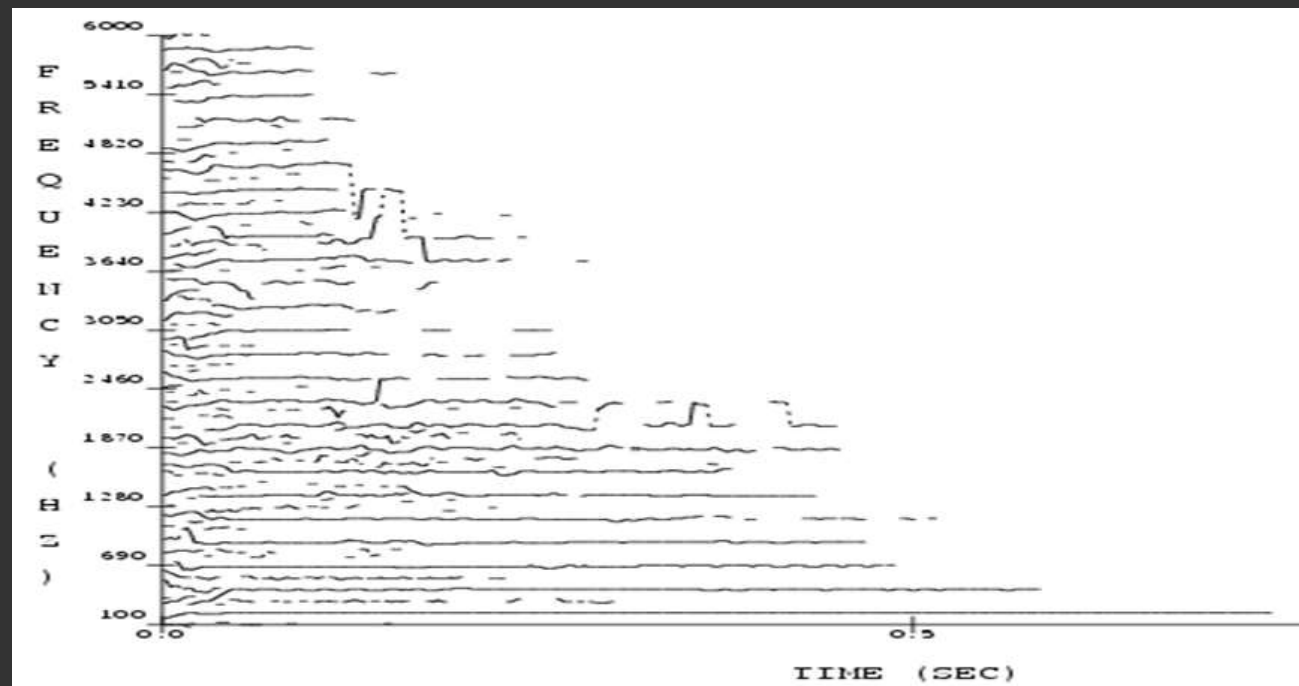
Skąd wziąć funkcje potrzebne do syntezy addytywnej?

- Można projektować widmo „ręcznie” – trudne i niepraktyczne.
- Ale można też „wyciągnąć” potrzebne parametry poprzez **analizę** nagranych dźwięków muzycznych.
- Odtworzenie dźwięku syntetycznego poprzez **resyntezę addytywną**.
- Zamiast generować i sumować sinusy, można prościej zrobić resyntezę poprzez zbudowanie widma i odwrotne przekształcenie Fouriera (IFFT).
- Metoda wymaga w praktyce implementacji cyfrowej.
- Synteza addytywna jest rzadko stosowana w instrumentach, ponieważ sampling daje zbliżony efekt znacznie mniejszym kosztem.

Analiza MQ

MQ – analiza McAulay–Quatieri

- Analiza cyfrowa – FFT w krótkich oknach czasowych (ramkach).
- W każdym oknie znajdujemy lokalne **maksima** widma.
- Maksima występujące w kolejnych oknach tworzą ciągłe **ścieżki**.
- Jako **funkcje** do resyntezy wybieramy ścieżki o określonej minimalnej długości i minimalnym poziomie.
- Resynteza – zwykle przez IFFT.



Korzyści z resyntezy addytywnej

Rozkładamy dźwięk na parametry, a potem składamy go z powrotem.
Co nam to daje?

Odpowiedź: możliwość wygodnych operacji na parametrach:

- **transpozycja** – łatwa zmiana wysokości dźwięku, bez zniekształceń czasowo-częstotliwościowych (skracania/wydłużania dźwięku), które występują w samplingu,
- **rozciąganie/skracanie** dźwięku – również łatwiejsze i bez większych zniekształceń (nie trzeba zapętlać dźwięku),
- usuwanie niepotrzebnych składowych i szumu,
- dowolne modyfikacje struktury widmowej dźwięku: dodawanie/usuwanie składowych, miksowanie dźwięków, dodawanie efektów, itp.

Generatory wielotonów (addytywne)

- **Generatory wielotonów** (*multitone generator*) są czasem nazywane też generatorami addytywnymi.
- Stosowane w niektórych syntezatorach (głównie programowych).
- Wytwarzają harmoniczne **wielotony** (*multitone*) poprzez sumowanie harmoniczných (sinusów) z różnymi amplitudami.
- Nie jest to synteza addytywna: amplitudy harmoniczných są stałe w trakcie generowania dźwięku.
- Wymagana jest filtracja i modulacja, aby ukształtować brzmienie dźwięku.

Synteza addytywna - podsumowanie

Zalety:

- można naśladować dźwięki prawdziwych instrumentów
- możliwość bezpośredniego wpływania na widmo (na barwę dźwięku)
- łatwe modyfikacje skali częstotliwości i czasu
- prosta koncepcyjnie metoda

Wady:

- ograniczona do sygnałów harmoniczných
- słabo nadaje się do tworzenia nowych brzmień
- skomplikowana kontrola brzmienia – dużo parametrów
- sampling uzyskuje podobny efekt niższym „kosztem”

Typy dźwięków w synteźatorach

Najważniejsze typy dźwięków tworzonych przez synteźatory.

- *Lead, Key* – wiodące dźwięki tworzące główny motyw muzyczny, dość krótkie, o wyrazistym brzmieniu.
- *Pad* – rozciągnięte w czasie dźwięki o łagodnym brzmieniu, używane jako podkład muzyczny w tle (wypełniacz).
- *Bass* – niskie dźwięki, krótkie, mocno akcentowane, stosowane do tworzenia rytmicznych partii basowych (*bassline*).
- *Drum* – dźwięki rytmiczne, zwykle o nieustalanej wysokości, służące do wytwarzania perkusyjnego rytmu.
- *Effect* – niemuzyczne dźwięki stosowane jako efekty brzmieniowe.

Synteza dźwięków muzycznych

Na koniec: jak można wykonać sztucznie dźwięk muzyczny?

- Złożyć widmo z sinusów (synteza *addytywna*).
- Ukształtować „surowy” sygnał harmoniczny (s. *subtraktywna* i *tablicowa*).
- Wytworzyć sygnał algorytmem matematycznym, np. przez modulację (*FM*).
- Nie robić syntezy, tylko odtworzyć nagrany dźwięk (*sampling*).
- Zbudować model instrumentu, który wygeneruje pożądany dźwięk (synteza metodą modelowania fizycznego).

Znamy już syntezę addytywną. Pozostałe podejścia omówimy na kolejnych wykładach.

Literatura

- SPEAR - Sinusoidal Partial Editing Analysis and Resynthesis: <http://www.klingbeil.com/spear/>
- M.K. Klingbeil: *Spectral Analysis, Editing and Resynthesis: Methods and Applications*. Columbia Univ. 2009 (dostępne ze strony SPEAR)
- R.J. McAulay, T.F. Quatieri: *Speech Analysis/Synthesis Based on A Sinusoidal Representation*. IEEE Trans. on Acoustics, Speech, and Signal Processing, vol. ASSP-34, no. 4, Aug. 1986, pp. 744-754.
- M. Russ: *Sound Synthesis and Sampling*. Focal Press, Oxford 1996.
- Wikipedia (wersja angielska)

Materiały wyłącznie do użytku wewnętrznego dla studentów przedmiotu *Elektroniczne instrumenty muzyczne*, prowadzonego przez Katedrę Systemów Multimedialnych Politechniki Gdańskiej. Wykorzystywanie do innych celów oraz publikowanie i rozpowszechnianie zabronione.

This presentation is intended for internal use only, for students of Multimedia Systems Department, Gdansk University of Technology, attending the „Electronic musical instruments” course. Other uses, including publication and distribution, are strictly prohibited.