

DŹWIĘK MUZYCZNY

Właściwości, analiza
i resynteza addytywna

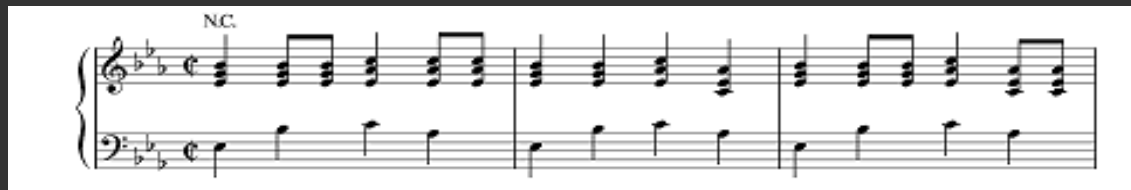
Dźwięk muzyczny

Definicja: dźwięk muzyczny jest to okresowy dźwięk wytwarzany przez instrument muzyczny (lub głos ludzki).

Dźwięk muzyczny ma następujące właściwości:

- **wysokość** (*pitch*) – umiejscowienie na skali muzycznej
- **barwa** (*timbre*) – brzmienie dźwięku
- **głośność** (*loudness*)
- **czas trwania** (*duration*)

Uwaga: głośność i barwa często zmieniają się w czasie trwania dźwięku muzycznego.

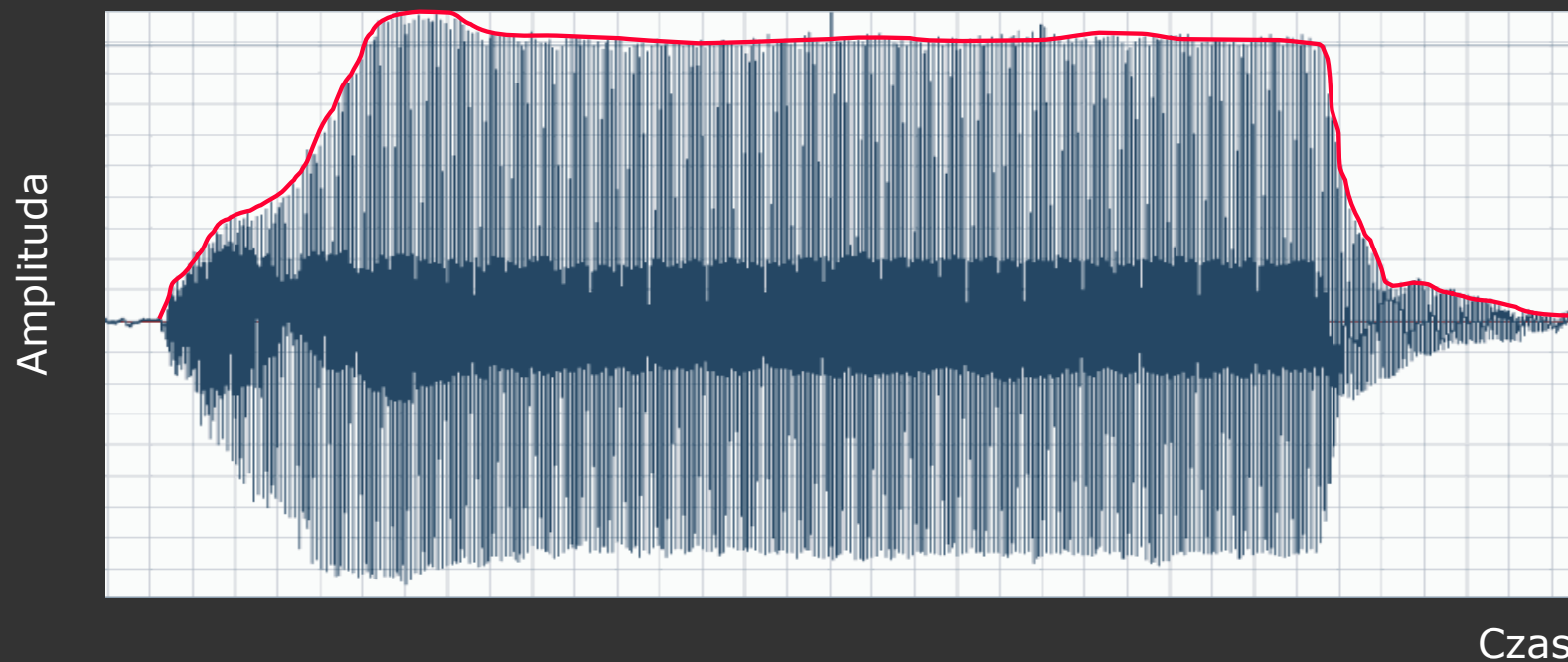


Dźwięk muzyczny

- Instrumenty: strunowe, dęte i nieliczne perkusyjne (np. ksylofon) wytwarzają dźwięki zgodne z tą definicją.
- Większość instrumentów **perkusyjnych** (np. zestaw perkusyjny) wytwarza dźwięki rytmiczne, **nie posiadające wysokości** (mają charakter szumowy).
- Zadaniem **syntezy dźwięku** jest wytworzenie sygnału o właściwościach odpowiadających dźwiękowi muzycznemu.
- To wcale nie znaczy, że musimy naśladować dźwięki rzeczywistych instrumentów. Mogą to być dźwięki **syntetyczne** o dowolnym brzmieniu, wystarczy że posiadają właściwości dźwięku muzycznego.

Analiza czasowa i obwiednia

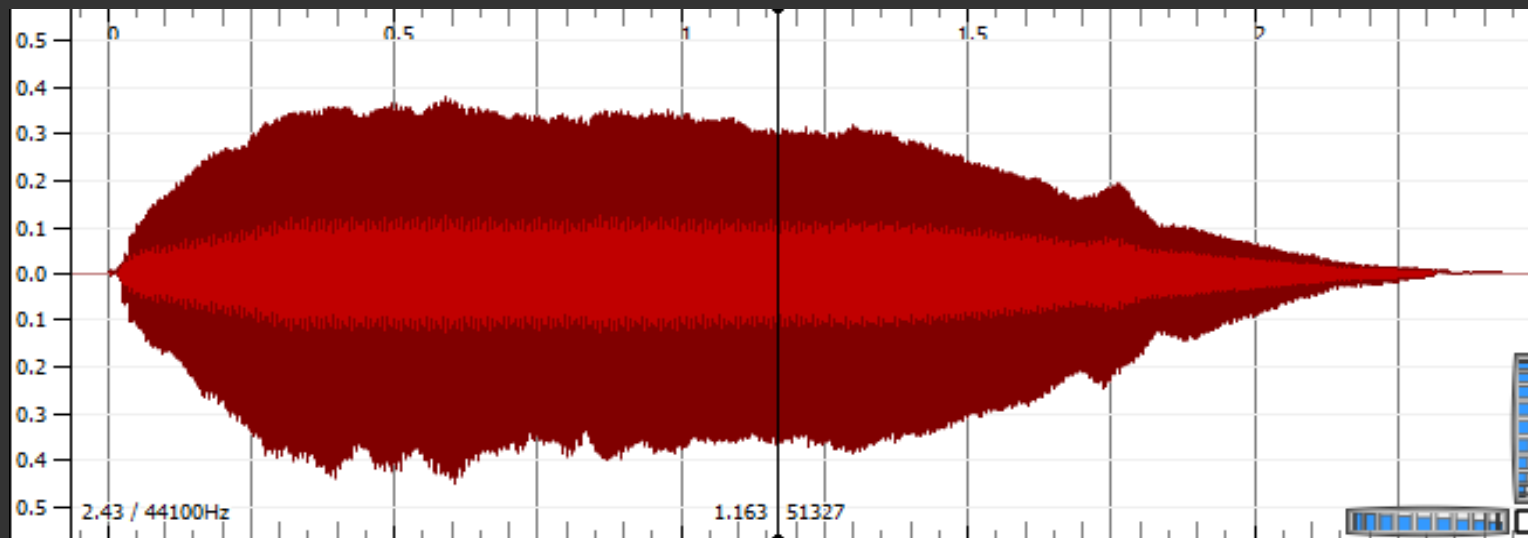
- Analiza czasowa uwidacznia zmiany **amplitudy** w czasie
- Możemy zmierzyć **czas trwania** dźwięku.
- **Obwiednia** (*envelope*) dźwięku łączy „szczyty” wykresu czasowego.
- Obwiednia informuje nas o **zmianach głośności** dźwięku w czasie.



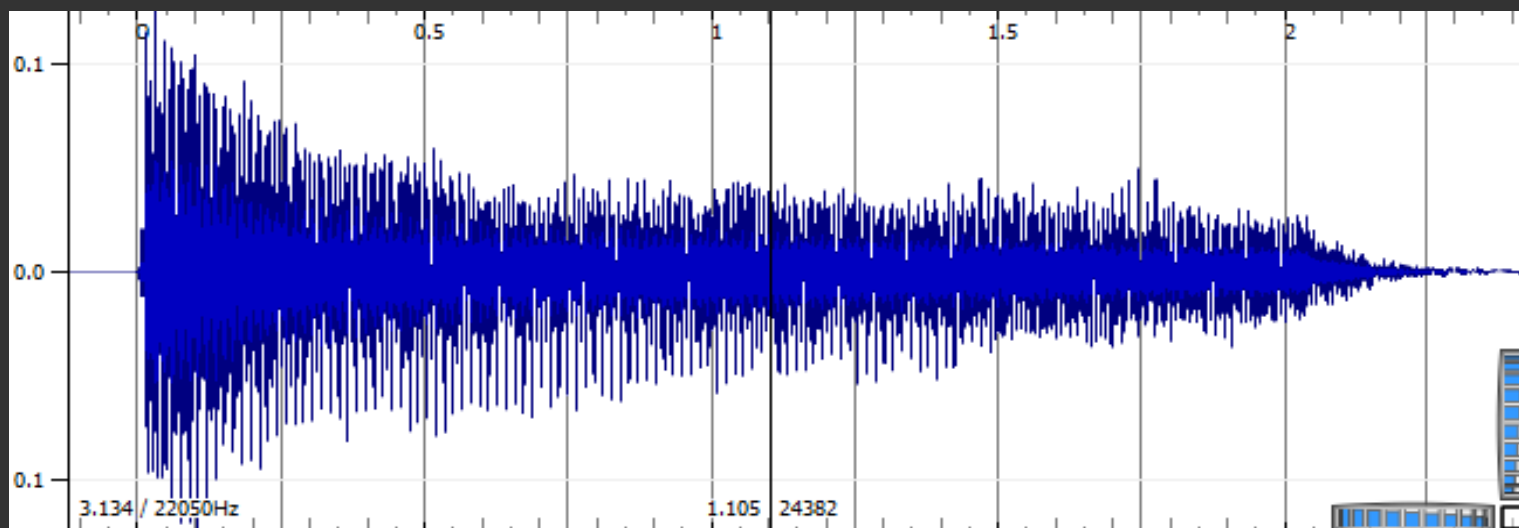
Fazy obwiedni dźwięku muzycznego

- **Faza ataku / narastania** (*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)
 - nie musi występować (zależy od typu instrumentu)
- **Faza wybrzmiewania / zwolnienia** (*release*)
 - naturalne wygaszanie dźwięku

Przykłady obwiedni dla różnych instrumentów



Trąbka



Fortepian

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 – dłuższa faza wybrzmiewania.

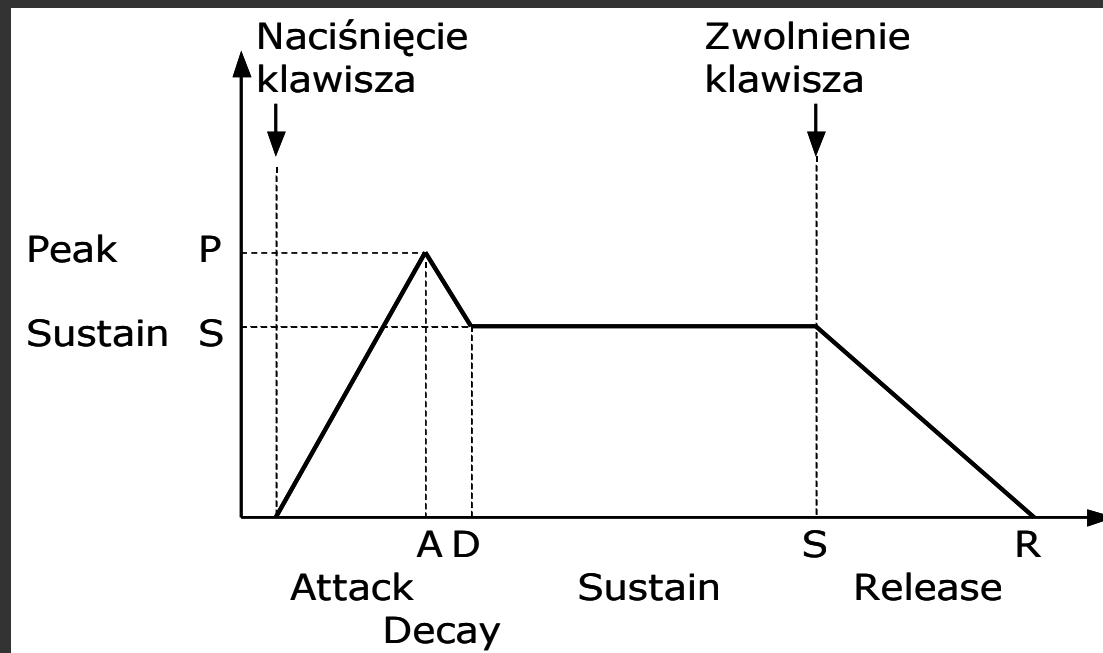
Jak uzyskać efekt obwiedni w **syntezatorze**?

- **Generator obwiedni** wytwarza sygnał o odpowiednim kształcie – tzw. obwiednia **ADSR**.
- Sygnał ten steruje **wzmocnieniem** końcowego wzmacniacza, dzięki temu uzyskujemy **zmiany głośności** zgodnie z zadaną funkcją.

Obwiednia ADSR

Parametry funkcji obwiedni ADSR:

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

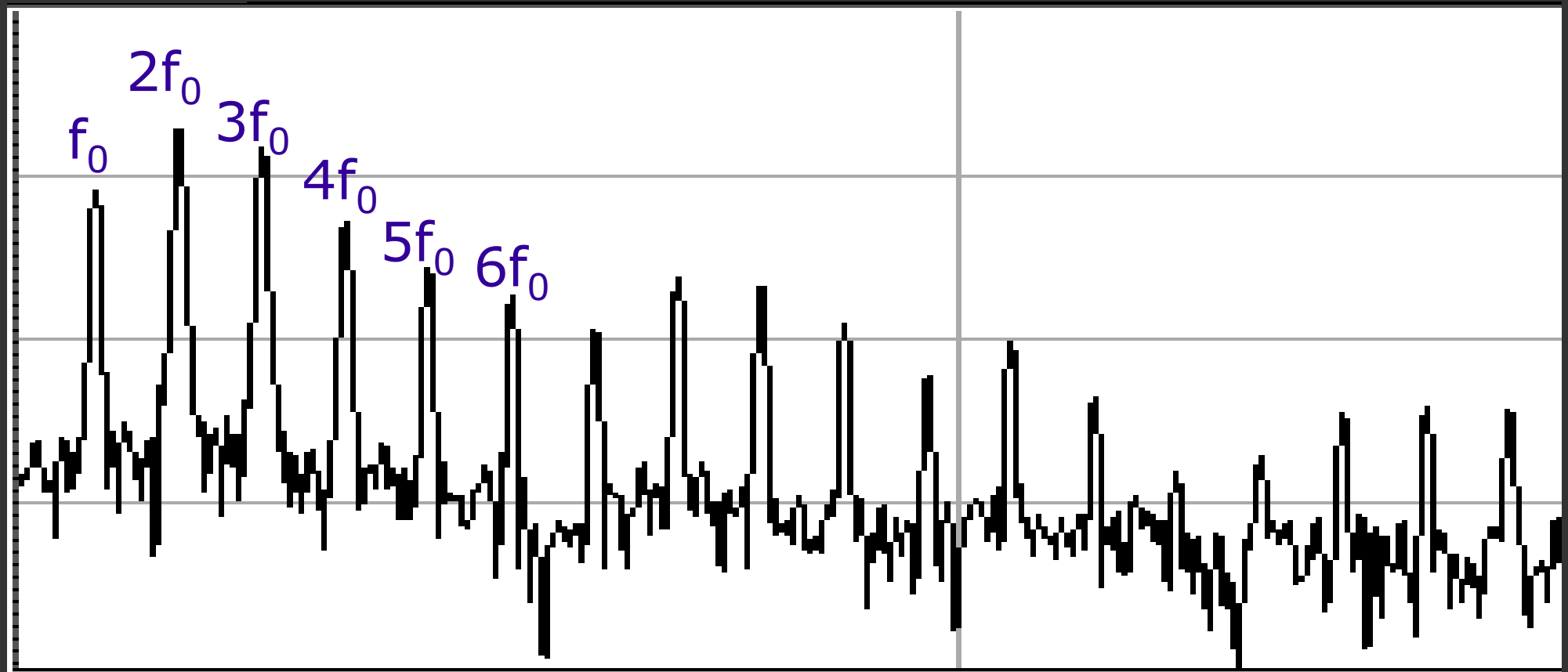


Analiza częstotliwościowa

- Właściwości dźwięku w dziedzinie częstotliwości.
- Widmo częstotliwościowe dźwięku muzycznego decyduje o jego **wysokości** i **barwie**.
- Analiza Fouriera: dowolny sygnał okresowy można przedstawić w postaci sumy sygnałów sinusoidalnych o odpowiednich amplitudach i częstotliwościach.
- Jak obliczamy:
 - „wycinamy” fragment sygnału (idealnie: okres) za pomocą funkcji okna
 - obliczamy transformatę Fouriera (FFT)
 - wynik: amplituda widma w funkcji częstotliwości.

Widmo typowego dźwięku muzycznego

Amplituda widma



Częstotliwość

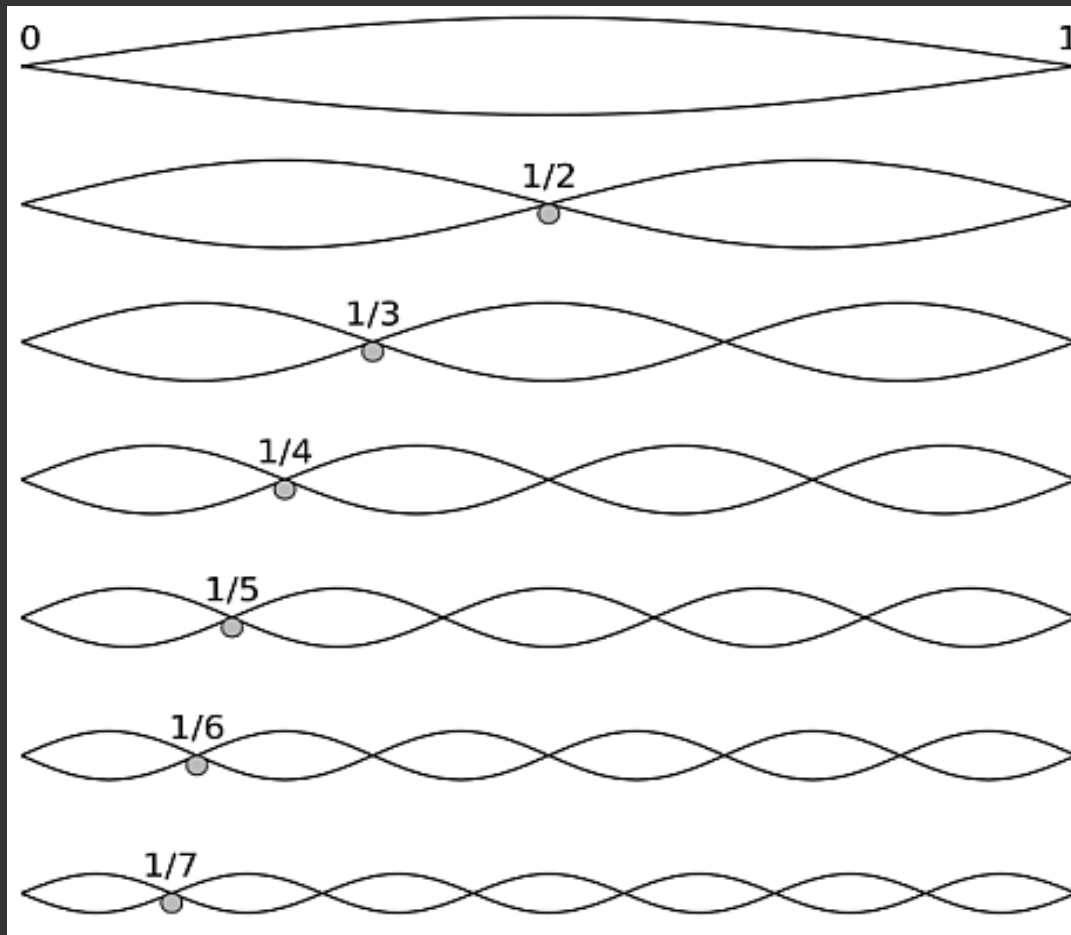
Widmo dźwięków muzycznych

Co zauważamy?

- Widmo jest **prążkowe** – ma wyraźne maksima.
- Prążki są położone w równych odstępach od siebie – tworzą **szereg harmoniczny**.
- Częstotliwość pierwszego prążka w szeregu to **częstotliwość podstawowa** dźwięku (f_0).
- Wyższe prążki to **harmoniczne**: pierwsza ($2f_0$), druga ($3f_0$), trzecia ($4f_0$), itd.
- Widmo zawiera też zwykle prążki nieharmoniczne (o znacznie niższej amplitudzie) oraz szum.

Widmo dźwięków muzycznych

Dlaczego właśnie takie widmo?



Widmo dźwięków muzycznych

PROSZĘ ZAPAMIĘTAĆ!!!

- Częstotliwość podstawowa dźwięku muzycznego wyznacza jego **wysokość**.
- Rozkład amplitud wszystkich prążków widma decyduje o **barwie** (brzmieniu) dźwięku.

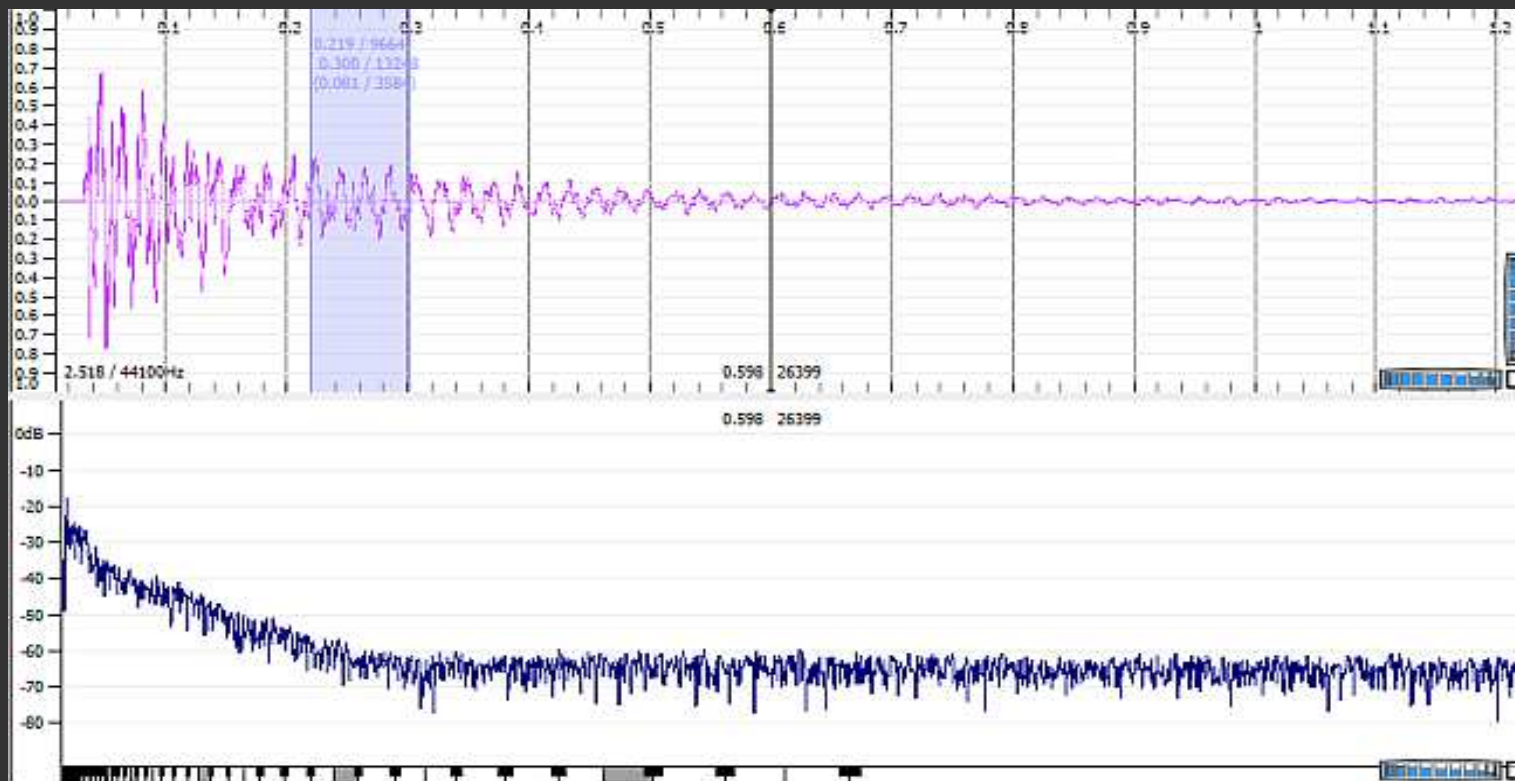
Widmo dźwięków muzycznych

Aby nie było zbyt prosto:

- Nie zawsze pierwszy prążek w widmie to częstotliwość podstawowa (ma być pierwszy prążek w szeregu h.).
- W widmie niektórych dźwięków instrumentów muzycznych może brakować „parzystych” prążków – to wciąż jest widmo harmoniczne.
- Niektóre instrumenty (np. dzwony) mają widmo prążkowe nieharmoniczne.
- Instrumenty perkusyjne mają widmo typu szumowego, bez prążków (dlatego nie mają wysokości).

Dźwięki perkusyjne

Dźwięki typu perkusyjnego syntetyzuje się przez filtrowanie szumu i nałożenie odpowiedniej obwiedni.



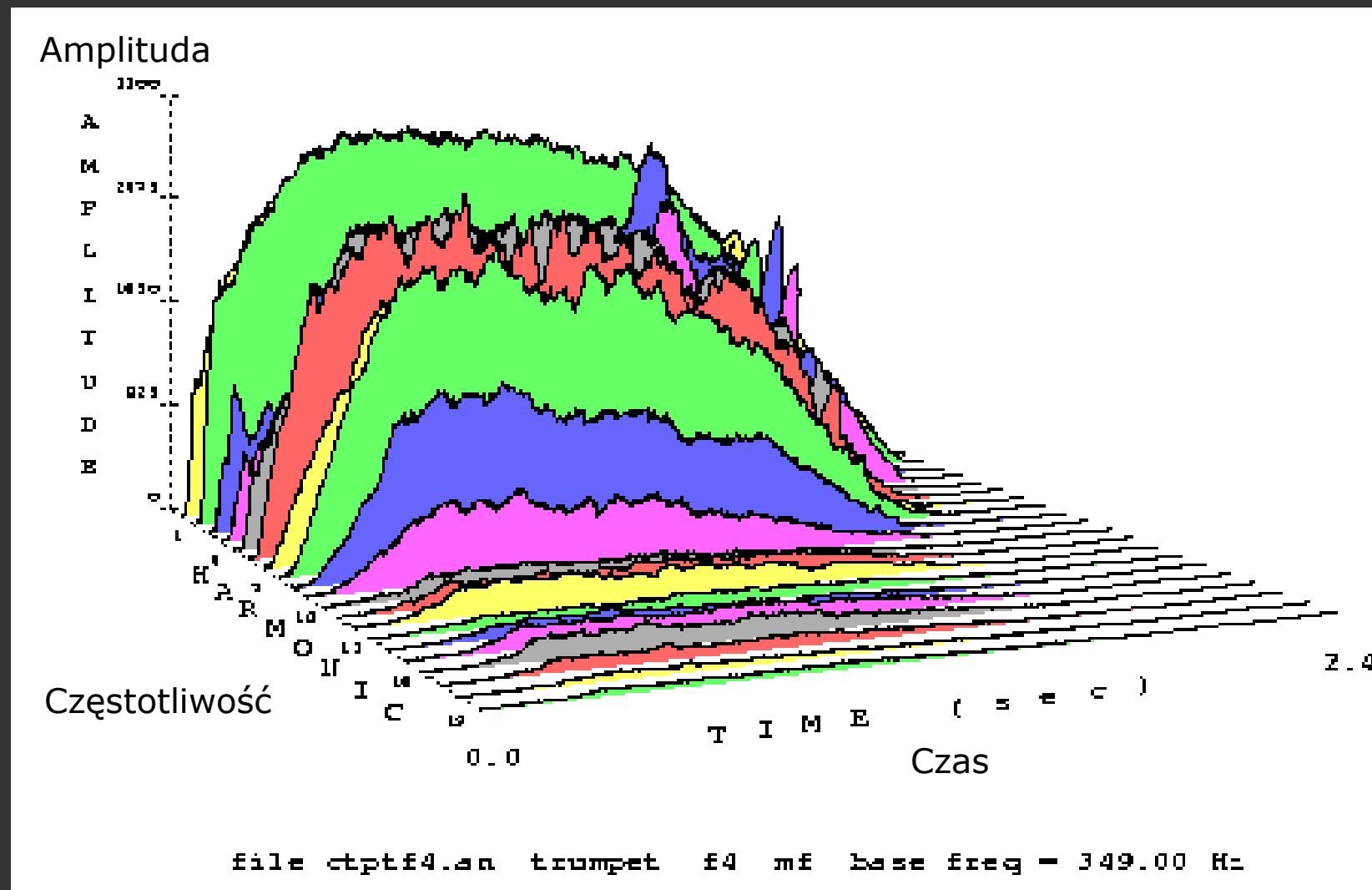
Synteza addytywna

- Wiemy że można rozłożyć dźwięk muzyczny na „sumę sinusów” w porządku harmonicznym.
- Możemy też zrobić na odwrót: zsumować „sinusy”:
 - o **częstotliwościach** wynikających z szeregu harmonicznego, o odpowiedniej cz. podstawowej,
 - o **amplitudach** tak dobranych, aby uzyskać pożądany kształt widma.
- Jest to **addytywna synteza dźwięku** (*additive synthesis*), od łac. *additio*: dodawanie. Metoda nie była praktycznie wykorzystywana w instrumentach muzycznych.

Zmienność widma dźwięku muzycznego

- Podany wcześniej wykres widma został „złapany” w wybranym momencie.
- Jeżeli widmo przez cały czas trwania dźwięku 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.
- Artykulacja (sposób gry) ma bardzo duży wpływ na zmiany barwy, szczególnie w fazie ataku.
- Aby dźwięk był „żywy” i ciekawy, musimy uwzględnić ten aspekt w syntezie.

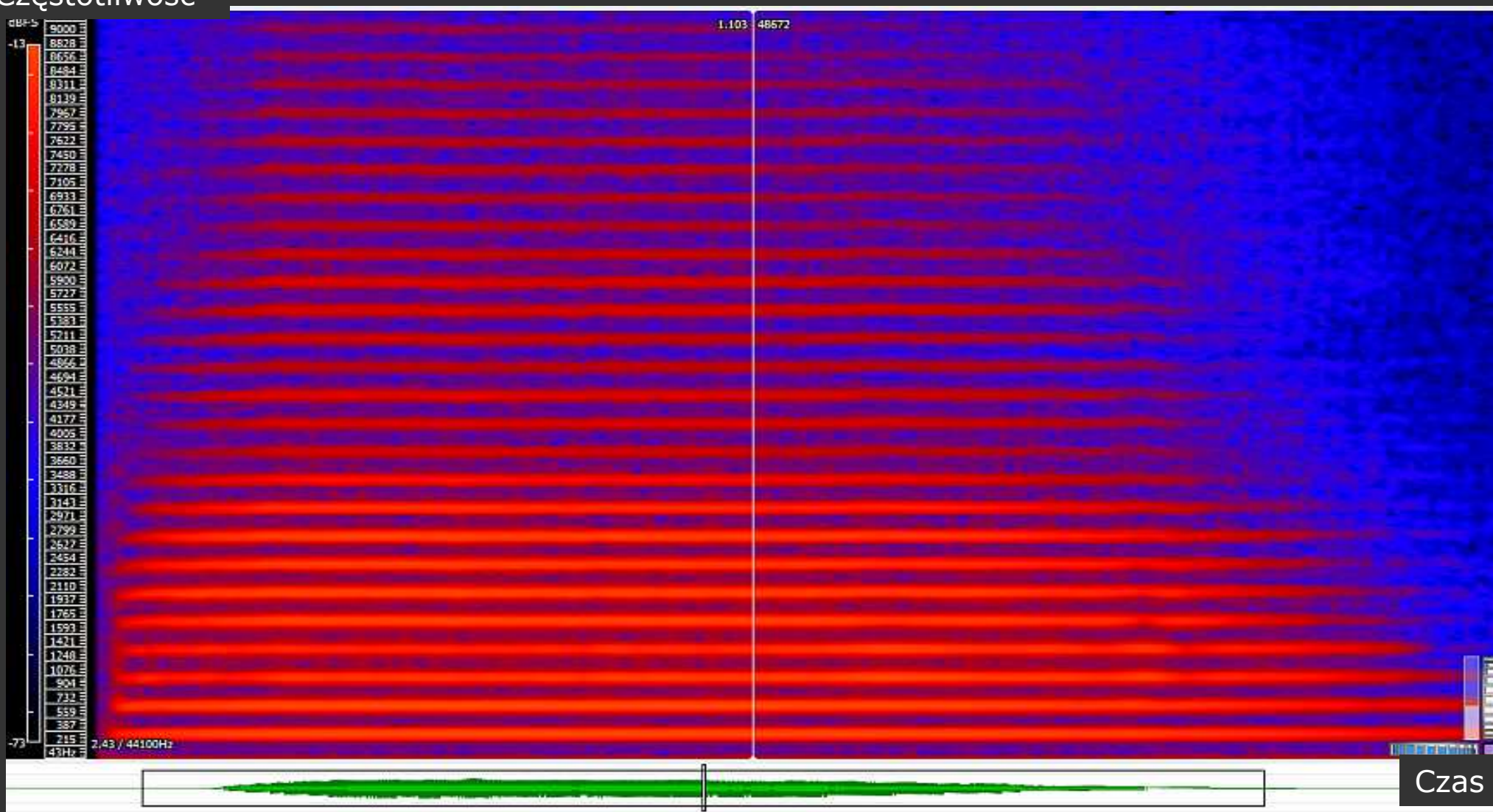
Wykres „waterfall”



Spektrogram - wykres zmienności widma

Widmo 3D: czas – częstotliwość – amplituda widma (kolor)

Częstotliwość



Parametry dźwięku a brzmienie

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

- Różny kształt obwiedni czasowej.
- Różna struktura widma statycznego.
- Różna zmienność widma w trakcie trwania.

Jak zatem zróżnicować brzmienie dźwięku w syntezie?

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

Zmienność widma w syntezie addytywnej

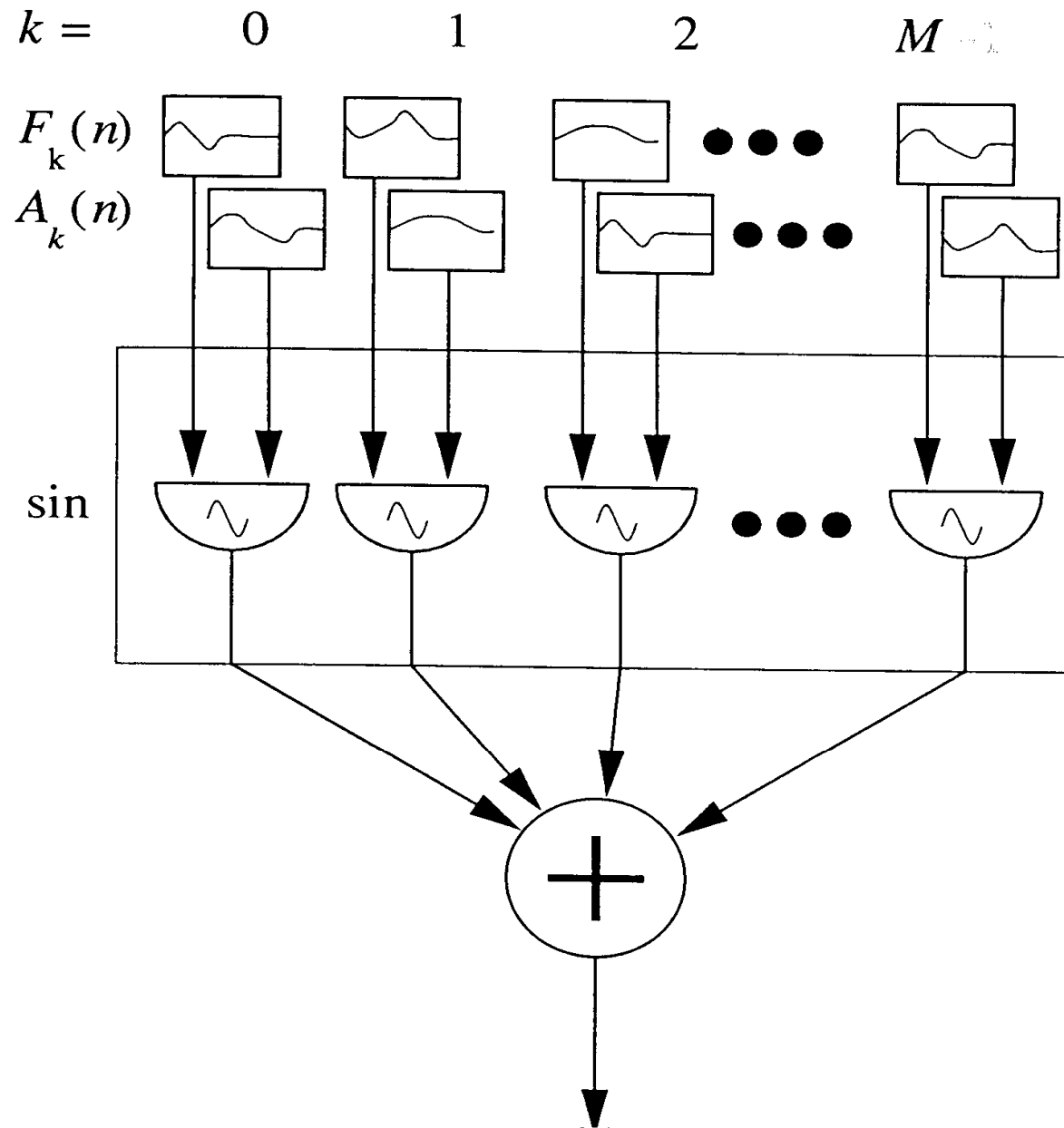
Wracamy do syntezy addytywnej. Jak zapewnić zmienność widma? Parametry nie mogą być stałe, lecz muszą zmieniać się w czasie. Potrzebujemy funkcji czasowych:

- **amplitudy** każdej harmonicznej $A_k(t)$
- **odchyłki częstotliwości** (faza) każdej harmonicznej $\Delta f_k(t)$

Jeżeli ktoś bardzo lubi wzory 😊:

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

Schemat syntezy addytywnej



Najprostsza
synteza jaką
da się zrobić.

Problem praktyczny

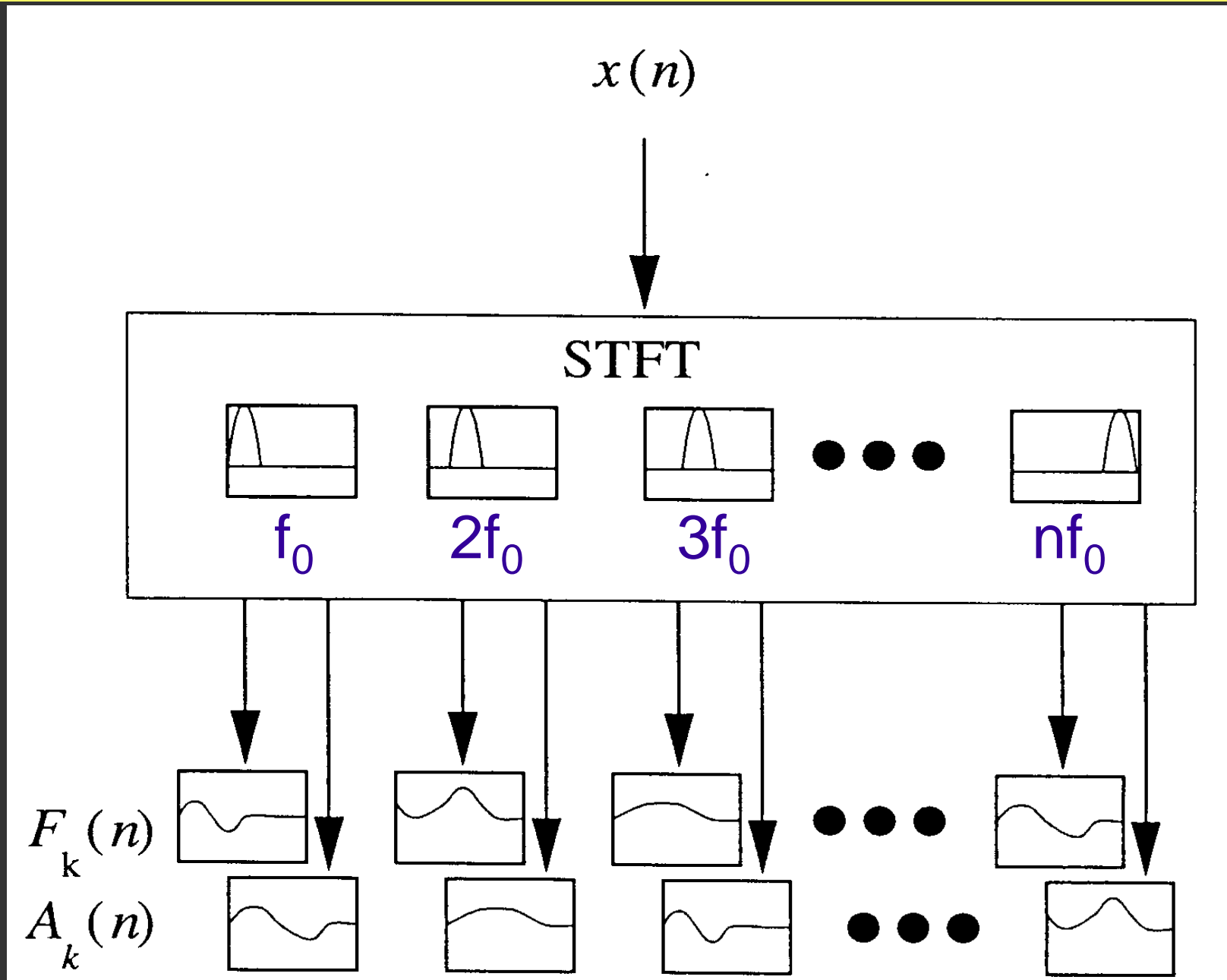
Skąd wziąć funkcje potrzebne do syntezy addytywnej?

- Można projektować „ręcznie” – mocno niepraktyczne (próbowano – *Fairlight*, nikt tego nie chciał robić).
- Ale można też „wyciągnąć” potrzebne parametry poprzez **analizę** nagranych dźwięków muzycznych.
- Odtworzenie dźwięku syntetycznego poprzez **resynteze addytywną**.
- Można też zrobić resynteze poprzez odwrotną transformację Fouriera IFFT (prościej).
- Metoda nie stosowana w praktyce, ponieważ sampling daje zbliżony efekt znacznie mniejszym kosztem.

Analiza PV

- **PV** – metoda wokodera fazowego (*phase vocoder*)
- Bank filtrów wąskopasmowych nastrojonych na częstotliwość podstawową i harmoniczne
- Filtry mierzą **energię** każdego pasma – z tego powstają **funkcje** amplitudowe
- Metoda zgrubna i mało dokładna (mała rozdzielczość), słaba jakość dźwięku po resyntezie
- Musimy znać częstotliwość podstawową
- Metoda nie sprawdza się gdy częstotliwość składowych zmienia się – zaburzenia w fazie ataku

Analiza PV



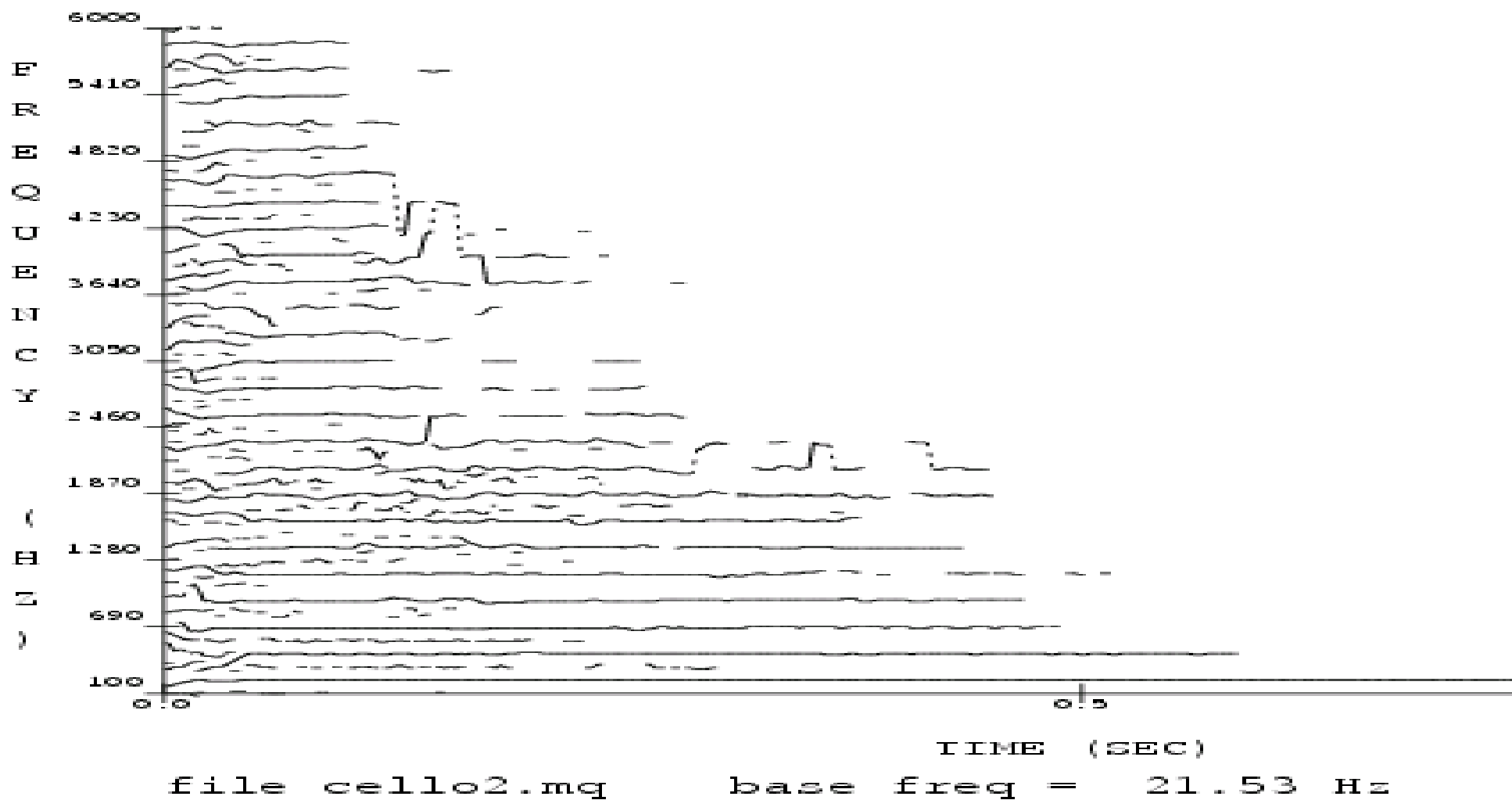
Analiza MQ

MQ – analiza McAulay – Quatieri

- Analiza cyfrowa – FFT w krótkich oknach czasowych
- W każdym oknie znajdujemy lokalne **maksima** widma i zaznaczamy je na wykresie czas-częstotliwość
- Maksima występujące w sposób ciągły tworzą **ścieżki**
- Jako **funkcje** do resyntezy wybieramy ścieżki o określonej minimalnej długości i poziomie
- Resynteza – zwykle przez IFFT
- Metoda znacznie bardziej dokładna, większa rozdzielczość niż PV, dużo lepsza jakość resyntezy

Analiza MQ

Przykład wyniku analizy – ścieżki



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 (które występują w samplingu)
- **rozciąganie/skracanie** dźwięku – również łatwiejsze i bez większych zniekształceń (nie trzeba zapętlać dźwięku)
- usuwanie niepotrzebnych składowych i szumu
- (potencjalnie) dowolne modyfikacje struktury widmowej dźwięku

Instrumenty addytywne (przykłady)

Kurzweil K150 (1986)

- cyfrowa synteza addytywna
- 240 generatorów
- kontrolowanie każdej harmoniczej (komputer)



Kawai K5000 (1996)

- zaawansowana stacja robocza
- s. addytywna + sampling
- trudna obsługa



Synteza addytywna - podsumowanie

Zalety:

- można naśladować dźwięki prawdziwych instrumentów
- możliwość bezpośredniego wpływania na widmo
- łatwe operacje zmiany częstotliwości/czasu
- prosta koncepcyjnie metoda

Wady:

- nie nadaje się do szukania nowych brzmień
- skomplikowane tworzenie brzmień – dużo parametrów
- sampling uzyskuje podobny efekt niższym „kosztem”

Synteza dźwięków muzycznych

Na koniec: jak można wykonać sztucznie dźwięk muzyczny?

- Złożyć widmo z sinusów (s. *addytywna* – to już znamy).
- Ukształtować „surowy” sygnał harmoniczny (s. *subtraktywna*).
- Wytworzyć sygnał algorytmem matematycznym, np. przez modulację (*FM*).
- Odtworzyć nagrany dźwięk (*sampling*).
- Zbudować model instrumentu, który wygeneruje pożądany dźwięk (s. *falowodowa*).

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.
- Vintage Synth Explorer: www.vintagesynth.com
- Wikipedia (wersja angielska)