

2. Cyfrowe reprezentacje sygnału fonicznego

Treść niniejszego rozdziału została opracowana przy założeniu, że Czytelnik jest zaznajomiony z podstawami teorii sygnałów dyskretnych. Podstawowe zagadnienia, związane z próbkowaniem, kwantyzacją i transformacjami, które są przytaczane, zostały uzupełnione wskazaniem niektórych źródeł i metod redukowania zniekształceń, które występują w cyfrowych torach fonicznych w związku z omawianymi operacjami. Bardziej złożone problemy, np. związane z próbkowaniem nierównomiernym, konwersją częstotliwości próbkowania przy wykorzystaniu filtrów polifazowych i in. zostały omówione w oparciu o praktycznie stosowane algorytmy.

Zagadnienia związane z budową i zasadą działania cyfrowych torów fonicznych i wpływem ich ograniczeń konstrukcyjnych na jakość dźwięku zostaną przedyskutowane w roz. 3. Materiał przedstawiony w rozdziałach 2 i 3 ułatwi późniejsze zapoznanie Czytelnika ze współczesnymi koncepcjami i zastosowaniami metod przetwarzania, rejestracji i syntezy dźwięku, które są tematem rozdziałów 4 do 8.

2.1 Próbkowanie sygnału fonicznego

2.1.1 Sygnał dyskretny w dziedzinie czasu

Cyfrowe operacje na ograniczonym pasmowo, analogowym sygnale fonicznym, który będzie oznaczany jako $f(t)$, rozpoczynają się w praktyce od przeprowadzenia jego próbkowania za pomocą ciągu impulsów próbkujących $s(t)$, w wyniku czego otrzymuje się ciąg próbek $f_s(t)$ dany zależnością:

$$f_s(t) = f(t) \cdot s(t) \quad (2-1)$$

Gdyby możliwe było pobieranie próbek w nieskończenie krótkim czasie, to można byłoby przyjąć, że impuls próbkujący, który występuje z okresem T_s ma postać funkcji delta $\delta(t)$. Wówczas sygnał $f_s(t)$ wyraża się zależnością:

4 Cyfrowe reprezentacje sygnału fonicznego

$$f_s(t) = f(t) \sum_{n=-\infty}^{\infty} \delta(t - nT_s) \quad (2-2)$$

gdzie: n – indeks czasu (numer chwili czasu)

Ze względu na fakt, że transformatą Fouriera ciągu impulsów $\delta(t)$ w dziedzinie czasu jest ciąg impulsów jednostkowych w dziedzinie widma, widmo sygnału próbkowanego $F_s(j\omega)$ jest określone następującą zależnością (iloczyn postaci czasowych zamienia się w spłot w dziedzinie widma):

$$F_s(j\omega) = \frac{1}{2\pi} F(j\omega) * \omega_s \sum_{k=-\infty}^{\infty} \delta(\omega - k\omega_s) \quad (2-3)$$

gdzie: k – indeks częstotliwości ($k = 0, \pm 1, \pm 2, \dots$)

$F(j\omega)$ – widmo sygnału $f(t)$,

ω_s – jest pulsacją próbkowania,

znak “*” - reprezentuje operację spłotu

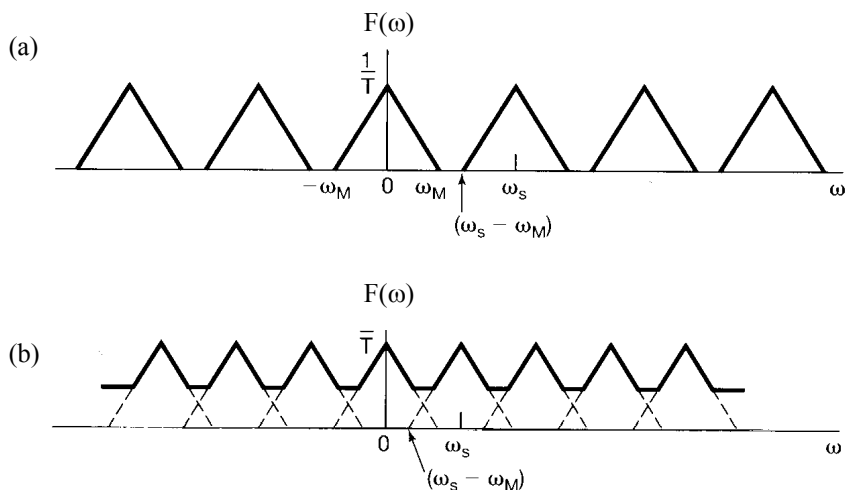
Bardziej użyteczny od postaci (2-3) jest następujący zapis tej zależności:

$$\frac{1}{\omega_s} \sum_{k=-\infty}^{\infty} [F(j\omega) * \delta(\omega - k\omega_s)] = \frac{1}{\omega_s} \sum_{k=-\infty}^{\infty} F[j(\omega - k\omega_s)], \quad (2-4)$$

która pozwala bezpośrednio zobrazować widmo spróbkowanego sygnału (jak na rys. 2-1). Podstawowemu widmu, które rozciąga się w zakresie pulsacji od $-\omega_M$ do ω_M , towarzyszą jego repliki wokół wielokrotności pulsacji próbkowania. Wynika to z faktu, że dokonanie obrotu wskazów na płaszczyźnie zespolonej o kąt 2π daje za każdym razem identyczną sytuację wyjściową do analizy widmowej i w efekcie powoduje wyznaczenie takiej samej postaci widma.

2.1.2 Błędy w procesie próbkowania

Jak wynika z rys. 2-1b, gdy nie jest spełniony warunek Nyquista, następuje nakładanie się widm (ang. *aliasing*). Zjawisko to można interpretować zarówno na podstawie bezpośredniej postaci czasowej sygnału, jak i biorąc pod uwagę jego reprezentację operatorową w postaci pary wskazów wirujących w przeciwnych kierunkach na płaszczyźnie zespolonej. Jeżeli dla uproszczenia wzięty zostanie pod uwagę sygnał harmoniczny, to łatwo zauważyć, że każda połówka okresu tego sygnału powinna być reprezentowana przez więcej, niż jedną próbkę. Jeżeli



Rys. 2-1. Widmo, będące wynikiem próbkowania sygnału analogowego o szerokości pasma równej ω_M : (a) przypadek $\omega_s > 2\omega_M$; (b) przypadek $\omega_s < 2\omega_M$

częstotliwość takiego sygnału będzie zbyt duża w stosunku do częstotliwości próbkowania, to przebieg próbkowany zostanie zinterpretowany jako sygnał o niższej częstotliwości, niż sygnał oryginalny. W przypadku analizy tej sytuacji na płaszczyźnie zespolonej można zauważyć, że pulsacje o wartości $\omega > 0.5 \omega_s$ można interpretować, jako pulsacje ujemne. Istotne jest przy tym, że jeżeli próbkowany sygnał będzie zawierał pulsacje o wartościach powyżej połowy pulsacji próbkowania, to zostaną one utożsamione z pulsacjami ujemnymi, co spowoduje błędy w wyznaczeniu widma (nakładanie się dodatnich i ujemnych połówek jego obrazów).

Próbkowanie sygnałów fonicznych realizuje się praktyce w oparciu o koncepcję tzw. próbkowania podpasmowego. Koncepcja próbkowania podpasmowego zakłada, że sygnał musi być próbkowany z częstotliwością wyższą od podwojonej górnej częstotliwości granicznej pasma sygnału użytecznego, nie zaś od częstotliwości najwyższej składowej występującej w widmie sygnału. Postępując w ten sposób zakłada się, że założone pasmo (np. pasmo 20 Hz - 20 kHz, reprezentujące zakres częstotliwości słyszalnych) zawiera wszystkie potrzebne składowe. Konsekwencją tego założenia jest konieczność ograniczenia pasma za pomocą filtracji dolnoprzepustowej na wypadek, gdyby w rzeczywistym sygnale występowały jednak składowe o częstotliwościach wyższych (są one obecne np. w widmie wielu instrumentów muzycznych).

Z punktu widzenia praktycznej realizacji układów próbkująco-pamiętających (rozdział 3) istotne jest, że w rzeczywistych warunkach czas pobierania próbki nie może być zerowy ($\tau \neq 0$). Zatem nieskończony ciąg impulsów typu delta w zależności (2-3) musi zostać zastąpiony przez falę prostokątną o (niezerowym) współczynniku wypełnienia, o wartości:

6 Cyfrowe reprezentacje sygnału fonicznego

$w = \frac{\tau}{T_s}$. Określony w ten sposób sygnał prostokątny posiada widmo prążkowe o obwiedni amplitudowej określonej zależnością:

$$S_n = A_w \frac{\sin(\pi k w)}{\pi k w} = A_w \text{sinc}(\pi k w) \quad (2-5)$$

gdzie: A – amplituda impulsów

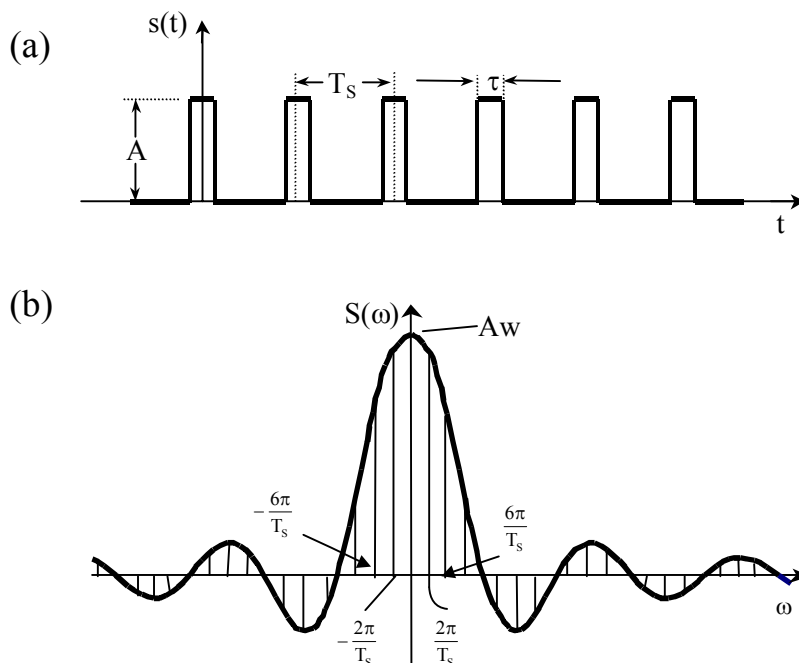
Biorąc pod uwagę zależność (2-5) można zauważyć, że widmo spróbkowanego w taki sposób sygnału będzie z pewnością odbiegało swym kształtem od widma określonego splotem (2-3). Próbkowanie przy użyciu fali impulsów prostokątnych jest zatem powodem występowania zniekształceń w wynikowym sygnale cyfrowym. Zniekształcone widmo można wyznaczyć, podobnie jak poprzednio, na podstawie twierdzenia o splotcie:

$$\begin{aligned} F_s(j\omega) &= \frac{1}{2\pi} F(j\omega) * 2\pi A_w \sum_{k=-\infty}^{\infty} \frac{\sin \pi k w}{\pi k w} \delta(\omega - k\omega_s) = \\ &= A_w \sum_{k=-\infty}^{\infty} \frac{\sin \pi k w}{\pi k w} F[j(\omega - k\omega_s)] \end{aligned} \quad (2-6)$$

Skutki niezerowego czasu pobierania próbek są, niestety, trudne do wyeliminowania w praktycznie realizowalnym torze cyfrowym, ponieważ nawet gdyby czas akwizycji próbek zmierzał do zera, to i tak konieczne jest „zamrożenie” sygnału w fazie pamiętania, co umożliwia prawidłowe zadziałanie (nadażenie) kwantyzera, będącego następnym w kolejności elementem toru cyfrowego. W związku z tym, wymnożenie rzeczywistego sygnału, który może zmieniać się w okresie T_s (patrz rys. 2-2) przez impuls („schodek”) o niezmiennej w określonym czasie wysokości prowadzi do następującej postaci widma sygnału spróbkowanego:

$$F_s(j\omega) = A_w \text{sinc} \frac{1}{2} \omega \tau \sum_{k=-\infty}^{\infty} F[j(\omega - k\omega_s)] \quad (2-7)$$

gdzie τ - oznacza w tym przypadku łączny czas pobierania próbki i jej pamiętania, będący parametrem układu próbkująco-pamiętającego,
k – indeks częstotliwości, jak wyżej

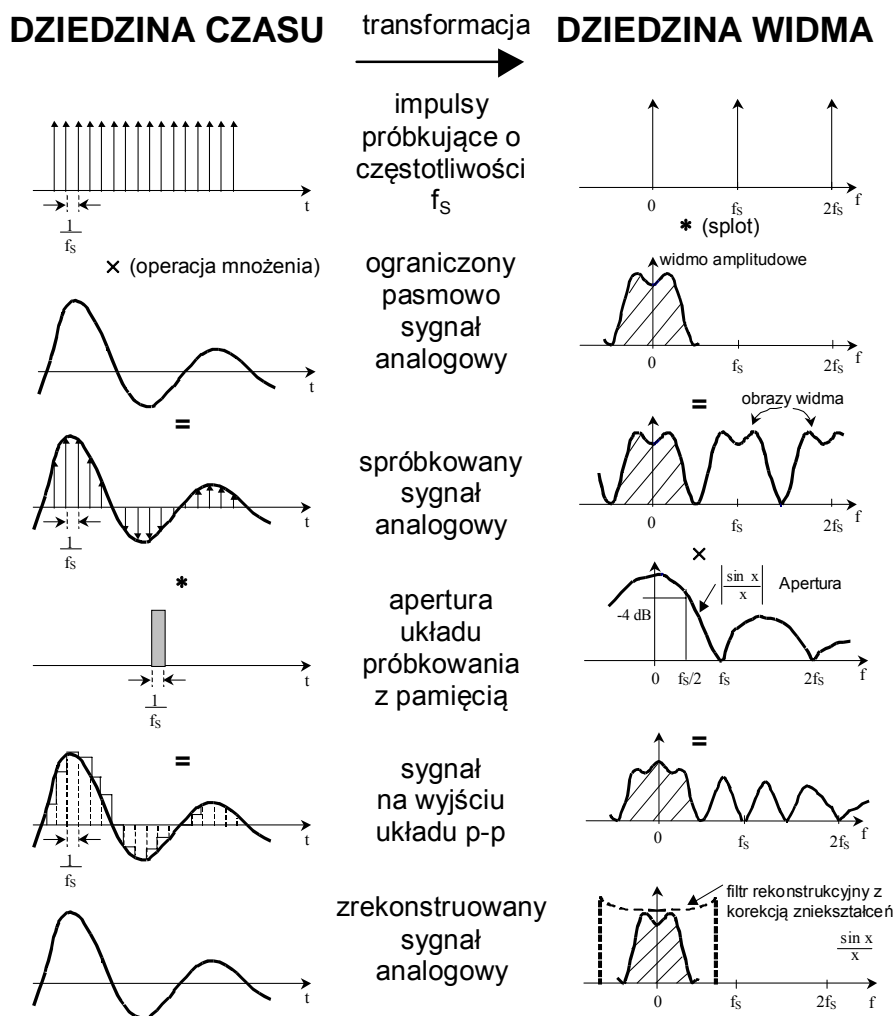


Rys. 2-2. Próbkowanie przy użyciu impulsów o szerokości $\tau > 0$; (a) przebieg próbkujący; (b) widmo przebiegu próbkującego

Jak wynika z zależności (2-7), składowe widma spróbkowanego sygnału maleją wraz ze wzrostem pulsacji, o czym decyduje czynnik:

$$A\omega \operatorname{sinc} \frac{1}{2} \omega \tau = A \frac{\tau}{T_s} \operatorname{sinc} \frac{1}{2} \omega \tau.$$

Przebiegi ilustrujące powstawanie związanych z tym zniekształceń apertury zostaną zilustrowane w roz. 3. W praktyce, sposobem zmniejszenia tego typu zniekształceń staje się zatem minimalizacja czasu τ przetwarzania sygnału przez konwerter analogowo-cyfrowy, gdyż okres próbkowania T_s jest na ogół wielkością określoną przez standard próbkowania. Na późniejszym etapie rekonstrukcji sygnału można przeciwdziałać skutkom zniekształceń poprzez wprowadzenie filtracji korygującej zniekształcenia typu $\operatorname{sinc}(\omega\tau/2)$ – patrz rys. 2-3. Zagadnienie to zostanie poruszone dokładniej w par. 2.1.3.



Rys. 2-3. Graficzna ilustracja skutków niezerowego czasu pobierania próbek. Po lewej - przebiegi w postaci czasowej, po prawej - odpowiadające im widma. Pomiędzy rysunkami sąsiadującymi w pionie zaznaczono symbolicznie operacje przeprowadzane na sygnałach

W pewnych przypadkach występuje niestabilność (nieperiodyczność) częstotliwości próbkowania. W praktycznych układach proces próbkowania jest synchronizowany zegarem kwarcowym, jednak występują pewne sytuacje (np. stany przejściowe związane z konwersją formatów - patrz par. 2.1.6), w których trzeba uwzględnić wzrost szumów powodowanych przez to zjawisko. Oszacowania poziomu powstających szumów można

dokonać zakładając dla uproszczenia, że sygnał ma charakter harmoniczny. Wówczas fluktuacje okresu próbkowania będą powodowały konieczność obliczenia poprawki kąta fazowego wskazu wirującego na płaszczyźnie zespolonej. Wynikająca stąd chwilowo zmienna pulsacja może być zapisana jako: $\omega(t_i + \delta_i)$, gdzie δ_i jest błędem i -tej próbki. Jeżeli założyć, że błąd pulsacji posiada rozkład równomierny w zakresie $-\delta$ do $+\delta$, to można wykazać (wyprowadzenie nie będzie tu przytaczane), że moc szumów powodowanych przez fluktuacje okresu próbkowania jest określona zależnością:

$$P_{sf} = \frac{1}{2\delta} \int_{-\delta}^{+\delta} [2 - 2 \cos(\omega\delta_i)] d\delta_i \cong \frac{(\omega\delta)^2}{3} \quad (2-8)$$

Wyznaczony na tej podstawie stosunek wartości skutecznej sygnału harmonicznego do szumu jest łatwy do obliczenia i wynosi:

$$\text{SNR}_{sf} = \frac{3}{2(\omega\delta)^2} = \frac{3T_s^2}{8\pi^2\delta^2} \quad (2-9)$$

Można oszacować na podstawie zależności (2-9), że wzrost stosunku δ/T_s o rząd wielkości spowoduje pogorszenie stosunku sygnału do szumu o 20 dB.

2.1.3 Rekonstruowanie sygnału na podstawie próbek

Aby zrekonstruować sygnał w postaci ciągłej na podstawie próbek, wykorzystując widmo sygnału spróbkowanego (przedstawione na rys. 2-1a), teoretycznie wystarczyłoby odfiltrować spróbkowany sygnał za pomocą filtru dolnoprzepustowego o charakterystyce $H(j\omega)$ równej 1 w paśmie od $-\omega_c$ do ω_c i 0 poza tym pasmem. Operacja ta spowodowałaby interpolację przebiegu czasowego sygnału w obszarze pomiędzy próbkami za pomocą funkcji typu sinc x . Idealny filtr o potrzebnej w tym wypadku charakterystyce nie jest jednak realizowalny praktycznie. W związku z tym, zachodzi konieczność rozwiązania problemu rekonstrukcji w inny sposób. Bardzo prosty sposób interpolacji bywa często używany w sytuacji, gdy wartość próbki jest utrzymywana na jednakowym poziomie przez znaczną część okresu próbkowania T_s . Sytuacja ta występuje często w praktycznych układach, ze względu na wykorzystanie układów próbkująco-pamiętających o krótkim czasie pobierania próbki i długim czasie jej pamiętania. Można założyć, że w tej sytuacji przebiegiem interpolującym sygnał rzeczywisty o

10 Cyfrowe reprezentacje sygnału fonicznego

postaci ciągłej jest przebieg schodkowy (o szerokości schodków τ), uzyskiwany na wyjściu układu próbkująco-pamiętającego. Jak wynika z zależności (2-5), filtr potrzebny do przeprowadzenia rekonstrukcji zostanie w ten sposób zastąpiony przez wycinek funkcji sinc ($\pi n \frac{\tau}{T_s}$). Na rys. 2-4b

pokazano, jak powinna wyglądać charakterystyka filtru korygującego, który skompensuje zniekształcenia powstające w paśmie użytecznym. Przy stosowaniu tej metody, proces interpolacji jest wspomagany przez użycie dodatkowego filtru dolnoprzepustowego o częstotliwości granicznej $\omega_s/2$ i o takim (stosunkowo łagodnym) nachyleniu zbocza, które pozwoli na dodatkowe stłumienie składowych powyżej częstotliwości $\omega_s/2$.

Opisany powyżej, uproszczony sposób dokonywania interpolacji przebiegu powoduje jednak powstawanie dość znacznych zniekształceń w rekonstruowanym sygnale. Dlatego w bardziej zaawansowanych rozwiązaniach torów cyfrowych wykorzystywana jest inna metoda interpolacji sygnału pomiędzy dostępnymi próbkami, które ponadto nie muszą być pamiętane w ciągu dłuższej części okresu próbkowania. Metoda ta opiera się na wykorzystaniu faktu, że widmo funkcji sinc x jest oknem prostokątnym. Wykorzystując właściwości splotu i biorąc pod uwagę zależność (2-2), zrekonstruowany sygnał $f_r(t)$, który jest równy z założenia sygnałowi oryginalnemu (na zasadzie identyczności widm), można przedstawić w następującej formie:

$$f_r(t) = T_s f_s(t) * \frac{\omega_M}{\pi} \text{sinc}(\omega_M t) = \sum_{n=-\infty}^{\infty} f(nT_s) \text{sinc}[\omega_M(t - nT_s)] \equiv f(t) \quad (2-10)$$

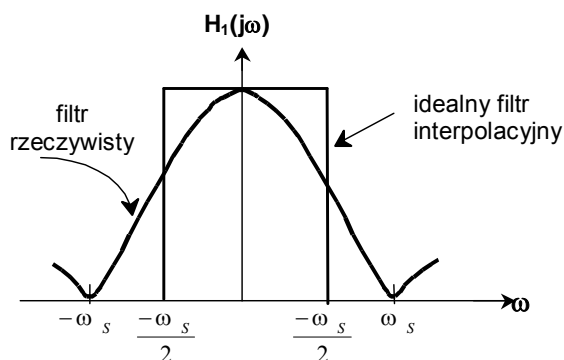
gdzie: ω_M - górna graniczna pulsacja pasma sygnału

Uzyskana zależność została zilustrowana na rys. 2-5, na przykładzie rekonstruowania funkcji harmonicznej.

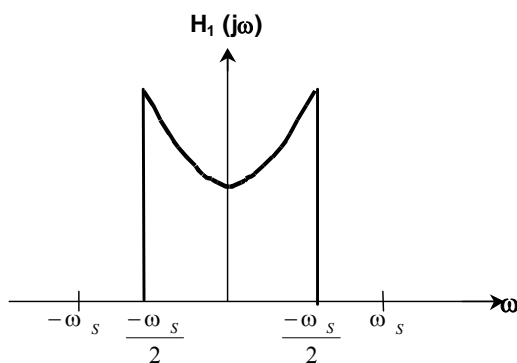
Jak wynika z zależności (2-10), interpolacja dokonywana tą metodą przebiega na drodze splatania (dodawania przemnożonych przez wartość bieżącej próbki i przesuniętych o kolejne okresy próbkowania kopii funkcji sinc ω_M). Metoda rekonstrukcji z wykorzystaniem splotu z funkcją typu sinc x została zilustrowana graficznie na rys. 2-6. Bardziej szczegółowy opis procesu interpolacji z wykorzystaniem funkcji typu sinc x można znaleźć w następnych paragrafach.

Nieprzyczynowość (spowodowana symetrią względem osi $t=0$) funkcji $\frac{\omega_M}{\pi} \text{sinc}(\omega_m t)$, która reprezentuje postać czasową idealnego filtra dolnoprzepustowego powoduje, że filtracja określona zależnością (2-10) nie może być bezpośrednio zrealizowana praktycznie, dlatego zastępuje się tę funkcję przez funkcję, która jest określona dla $t \geq 0$. W wyniku uzyskuje się jednak funkcję niesymetryczną, co pociąga za sobą powstanie nieliniowości charakterystyki fazowej. Nieliniowość ta może być następnie kompensowana przy użyciu filtrów wszechprzepustowych (patrz rozdział 3). Jest rzeczą korzystną, że zniekształcenia fazowe, o których mowa, nie są

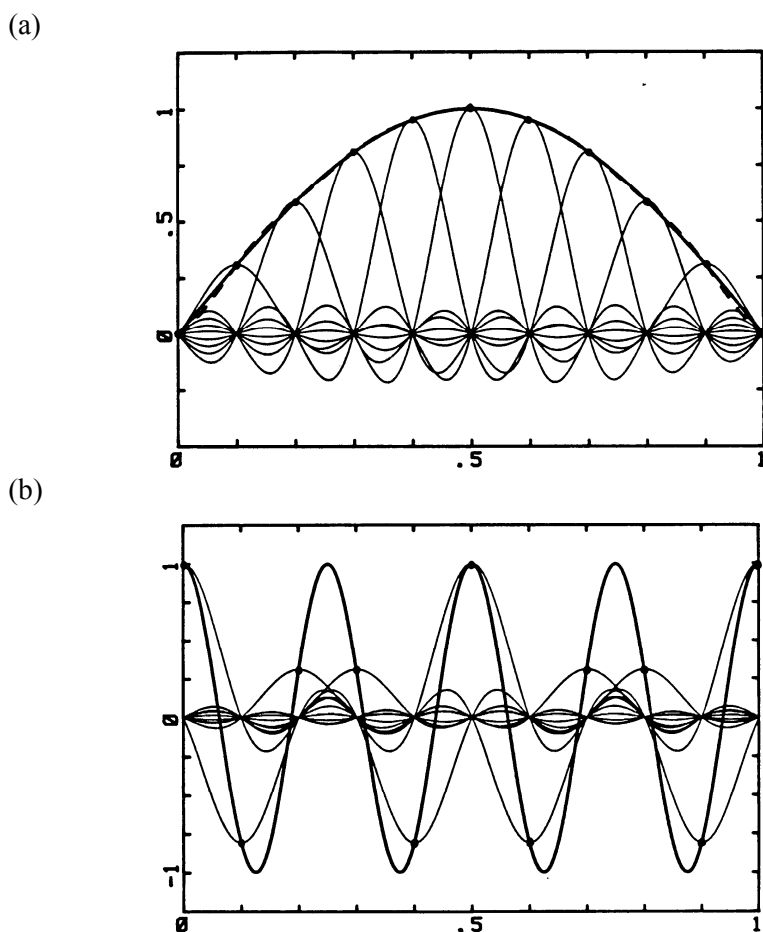
(a)



(b)

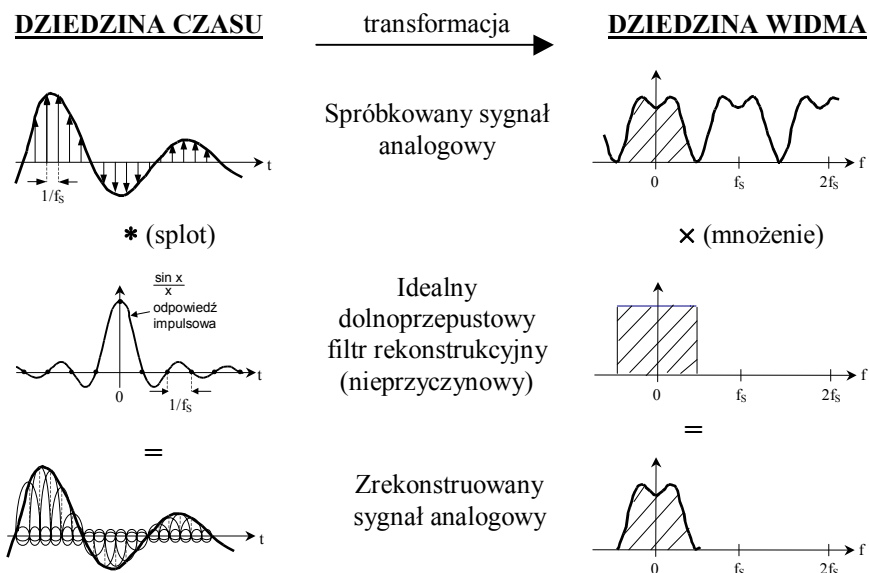


Rys. 2-4. Porównanie charakterystyk amplitudowych filtrów dolnoprzepustowych:
 (a) opartego na wycinku funkcji typu sinc x oraz (b) filtru korygującego



Rys. 2-5. Rekonstrukcja przebiegów na podstawie próbek: (a) sygnał o niższej częstotliwości (2,2 kHz); (b) sygnał o wyższej częstotliwości (17,6 kHz).
Częstotliwość próbkowania wynosiła 44,1 kHz

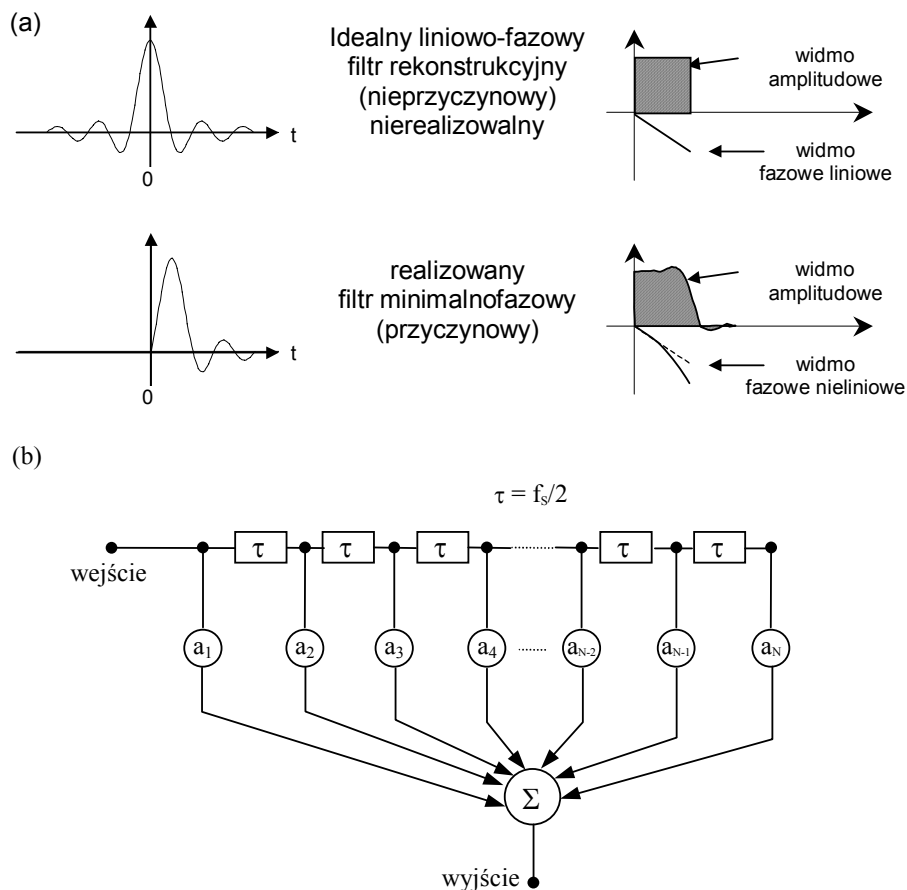
na ogół wyraźnie słyszalne, zwłaszcza dla sygnałów o wyższych częstotliwościach. Do realizacji interpolacji w torach fonicznych o wysokiej jakości wykorzystuje się filtr cyfrowy, który dokonuje mnożenia próbek sygnału przez współczynniki odpowiadające wartościom funkcji interpolującej. Zagadnienie to wyjaśniono na rys. 2-7. Liczba sekcji opóźniająco-mnożących tego filtra rzutuje na jakość interpolacji, wpływając jednocześnie na złożoność obliczeniową procesu interpolacji.



Rys. 2-6. Ilustracja procesu interpolacji sygnału na podstawie próbek -przypadek idealny. Po lewej - przebiegi w postaci czasowej, po prawej - odpowiadające im widma. Pomiędzy rysunkami sąsiadującymi w pionie zaznaczono operacje przeprowadzane na sygnałach

Skończony czas działania układów wewnętrznych konwertera cyfrowo-analogowego, który poprzedza w łańcuchu cyfrowym filtr rekonstrukcyjny jest powodem, dla którego napięcie wyjściowe konwertera nie zmienia się skokowo, tak jak zakładano w dotychczasowych rozważaniach. W związku z tym, na jego wyjściu powstają charakterystyczne zniekształcenia (opóźnienie narastania odpowiedzi, efekt Gibbsa, zniekształcenia nieliniowe), które mogą być słyszalne. Aby zapobiec tego typu zniekształceniom w torze cyfrowym, bezpośrednio za konwerterem cyfrowo-analogowym stosuje się dodatkowy układ próbkująco-pamiętający, najczęściej o takiej samej konstrukcji, jak układ wykorzystywany na wejściu toru. Układ ten pełni rolę bramki impulsowej. Na wyjściu tego układu otrzymuje się “zregenerowane” impulsy o kształcie prostokątnym, które są wykorzystywane w procesie interpolacji opisanym powyżej.

Obok opisanych rozwiązań istnieje wiele koncepcji filtrów specjalnych, np. filtrów o zmiennym opóźnieniu, które mogą być używane do interpolacji w procesie rekonstruowania sygnału na podstawie próbek. Wybrane algorytmy interpolacji zostaną przedstawione w par. 2.1.5.1 i 2.1.5.2.



Rys. 2-7. Ilustracja procesu interpolacji sygnału na podstawie próbek: (a) przypadek idealny i przypadek rzeczywisty; (b) filtr interpolacyjny

2.1.4 Próbkowanie i interpolacja sygnałów dyskretno-czasowych

Operacje określone w tytule niniejszego paragrafu występują stosunkowo często w cyfrowych układach fonicznych m.in. ze względu na potrzebę zwiększenia stosunku sygnału do szumu w paśmie użytecznym oraz wymagania związane z konstrukcją filtrów antyzakładowych. Sygnał dyskretno-czasowy będzie w tym miejscu rozumiany jako sygnał określony w dyskretnych chwilach czasu, przyjmujący wartości amplitud z dziedziny ciągłej. Bardzo ważnym zastosowaniem tej techniki jest także konwersja standardów, ze względu na towarzyszącą jej potrzebę dokonywania zmian

częstotliwości próbkowania sygnałów cyfrowych (par. 2.1.6). Jeszcze jednym zastosowaniem jest synteza dźwięku (roz. 8), w związku z którą występuje konieczność dokonywania zmian wysokości dźwięków zarejestrowanych w pamięci instrumentu muzycznego w czasie gry.

Zagadnienie próbkowania sygnału dyskretnego zostało zilustrowane w sposób schematyczny na rys. 2-8. Możliwe są przy tym dwa przypadki: pierwszy, gdy nowa częstotliwość próbkująca jest mniejsza od częstotliwości próbkowania sygnału (wykresy (a) i (b)) oraz drugi przypadek, gdy jest ona większa (wykresy (d) i (e)). W pierwszym przypadku uzyskuje się sygnał w postaci próbek, które są równe próbkom oryginalnego sygnału w chwilach czasu będących wielokrotnościami oryginalnego okresu próbkowania T_S . W pozostałych chwilach czasu sygnał będzie przyjmował wartości zerowe. Tak więc, $f_S(t)$ jest zdefiniowany przez wartości $f(t)$ dla $t = n \cdot N \cdot T_S$ oraz przez wartości 0 dla pozostałych t , gdzie N jest liczbą naturalną określającą nowy sposób próbkowania ciągu, zaś n - jest ciągiem liczb całkowitych. Drugi przypadek zostanie rozpatrzony później.

Dla obu przypadków łącznie można wyznaczyć postać czasową $f_S(t)$ w analogiczny sposób, jak miało to miejsce dla zależności (2-1) i (2-2). W związku z tym otrzymuje się:

$$f_S(t) = f(t) \cdot s(t) = \sum_{n=-\infty}^{\infty} f(t) \cdot \delta\left(t - n \frac{N}{M} T_S\right) \quad (2-11)$$

gdzie: $s(t)$ – sygnał próbkujący, N – współczynnik decymacji,
 M – współczynnik interpolacji

Należy zauważyć, że pierwszy omawiany przypadek oznacza, że $M=1$ oraz $N>1$, zaś drugi przypadek jest związany z wartościami $M>1$ oraz $N=1$. Zatem pierwszy przypadek upraszcza się do zapisu:

$$f_S(t) = \sum_{n=-\infty}^{\infty} f(t) \cdot \delta(t - nNT_S) \quad (2-12)$$

Oznacza to, że sygnał $f_S(t) = f(t)$ dla całkowitych krotności $N \cdot T_S$. Pozostałe wartości sygnału $f(t)$ zostają pominięte. Dla ściślejszego opisu należałoby uwzględnić przesunięcie fazowe w stosunku do sygnału $f(t)$, co prowadzi do następującego wzoru: