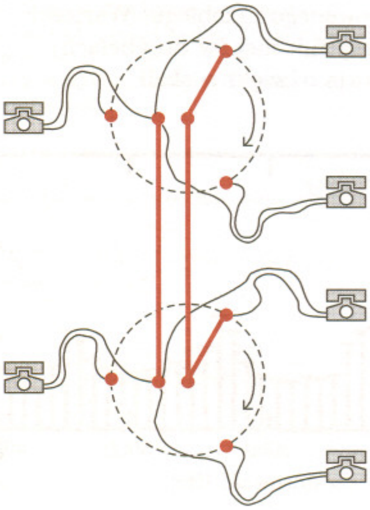
Telefony

Jacek IZYDORCZYK

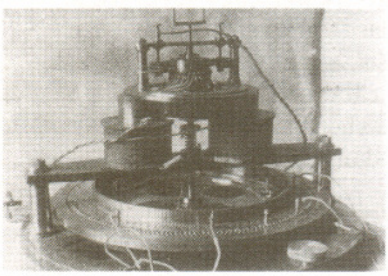
Geneza problemu próbkowania

Czy wiecie, że telefon ma już 125 lat? Chociaż aparat telefoniczny przez cały ten czas służył głównie do przenoszenia głosu ludzkiego na odległość, to jednak współczesny, cyfrowy aparat GSM jest nieporównanie mniejszy i bardziej skomplikowany niż pierwszy aparat telefoniczny skonstruowany przez A.G. Bella (1847–1922). Istota zmian zawiera się chyba w słowie „cyfrowy".



Rys. 1. Zasada działania systemu zwielokrotniania kanału transmisyjnego przez podział czasu.

Proces przetwarzania cyfrowego polega na zamianie sygnału analogowego, np. przebiegu drgań prądu elektrycznego w obwodzie mikrofonu wywołanych przez padające na membranę mikrofonu dźwięki, na sygnał cyfrowy, tzn. ciąg liczb o ograniczonej dokładności. Czy można jednak sygnał analogowy przekształcić na sygnał cyfrowy zupełnie bez strat? Czy funkcję ciągłej zmiennej t można zastąpić ciągiem liczb? Podejrzenia, że jest to możliwe, pojawiły się już 150 lat temu, gdy tylko skonstruowano telegraf. Później zaczęto budować urządzenia do przesyłania wielu rozmów telefonicznych za pomocą jednej pary przewodów. Idea była prosta. Dwa sprzężone komutatory połączone na stałe jedną parą przewodów kolejno, wiele razy na sekundę, łączyły kilka par aparatów telefonicznych (rys. 1). Pierwszy w pełni sprawny system tego typu został skonstruowany około 1900 roku (fotografia). Rozmowa telefoniczna stawała się możliwa wtedy, gdy para aparatów telefonicznych łączona była i rozłączana około 4 300 razy na sekundę. Zjawisko to sugeruje, że jeżeli wartość sygnału mowy próbkujemy (czyli sprawdzamy) co około 230 µs, to przebieg sygnału w międzyczasie jest stosunkowo „mało istotny".



Widmo

Pojedynczy ton o częstotliwości f to sygnał opisywany funkcją kosinus

$$x(t) = A\cos(2\pi ft + \phi),$$

gdzie A to amplituda dźwięku, a ϕ to przesunięcie fazowe. Korzystając ze znanej tożsamości trygonometrycznej, można to też zapisać jako

$$x(t) = a_1 \cos 2\pi ft + b_1 \sin 2\pi ft,$$

gdzie $a_1 = A\cos\phi$ oraz $b_1 = -A\sin\phi$.

Wartości współczynników a_n i b_n, które minimalizują różnicę pomiędzy sygnałem $x(t)$ i szeregiem funkcji trygonometrycznych, obliczamy według wzorów

$$a_n = \frac{2}{T}\int_{-T/2}^{T/2} x(t)\cos(2\pi nft)\mathrm{d}t,$$

$$b_n = \frac{2}{T}\int_{-T/2}^{T/2} x(t)\sin(2\pi nft)\mathrm{d}t.$$

Z czystymi tonami spotykamy się stosunkowo rzadko. Częściej mamy do czynienia z dźwiękiem o ustalonym okresie $T = 1/f$ i kształcie, który jednak trudno uznać za sinusoidę lub kosinusoidę. Można się jednak uprzeć i tak próbować dobrać amplitudy a_1 i b_1 pojedynczego tonu o częstotliwości f, aby przebieg czasowy czystego tonu był możliwie bliski sygnału $x(t)$. Gdy chcemy poprawić jakość przybliżenia, możemy użyć też tonów o częstotliwości $2f$, $3f$, $4f$, itd. W ten sposób powstaje trygonometryczny szereg Fouriera związany z okresowym sygnałem $x(t)$

$$x(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty}(a_n \cos 2\pi nft + b_n \sin 2\pi nft).$$

Jeżeli tylko moc sygnału $x(t)$ jest skończona, to moc sygnału będącego różnicą między sygnałem $x(t)$ i jego trygonometrycznym szeregiem Fouriera jest równa zeru. A zatem próbkowanie powinno być możliwe!

Wzory Eulera ($i^2 = -1$)

$$\cos t = \frac{\mathrm{e}^{it} + \mathrm{e}^{-it}}{2}, \quad \sin t = \frac{\mathrm{e}^{it} - \mathrm{e}^{-it}}{2i}.$$

Korzystając ze wzorów Eulera, trygonometryczny szereg Fouriera można przekształcić w szereg wykładniczy

$$x(t) = \sum_{n=-\infty}^{\infty} c_n e^{2\pi inft}.$$

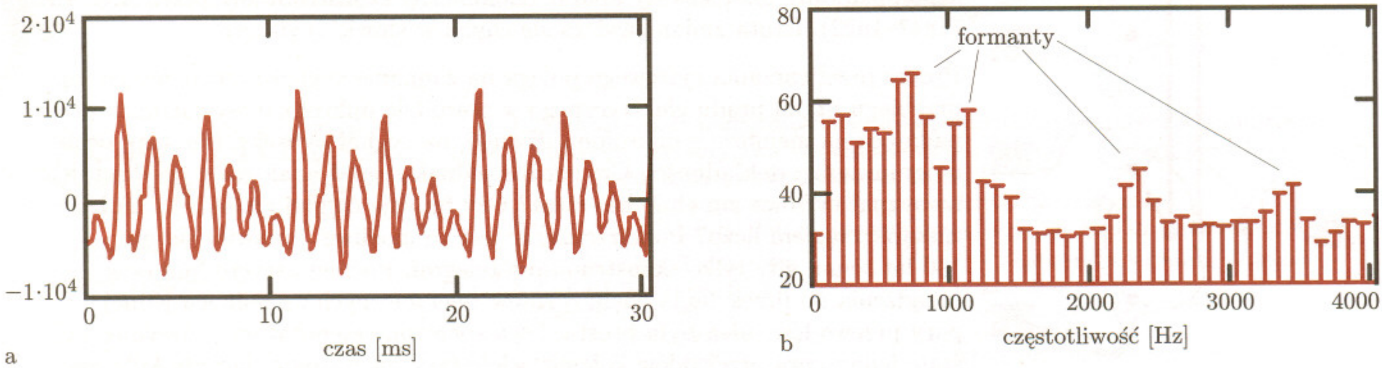
Współczynniki c_n oblicza się za pomocą wzoru

$$c_n = \frac{1}{T}\int_{-T/2}^{T/2} x(t)\mathrm{e}^{-2\pi inft}\mathrm{d}t.$$

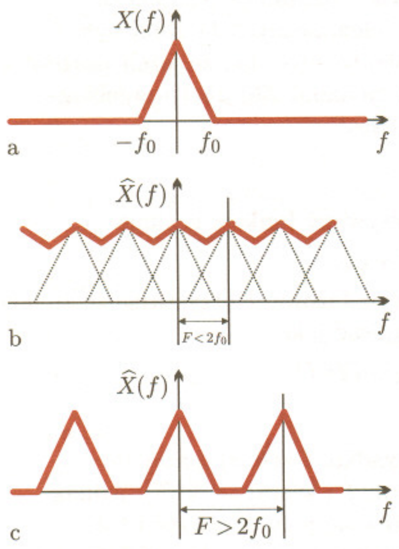
Współczynniki rozwinięcia w wykładniczy szereg Fouriera są liczbami zespolonymi i tworzą tzw. widmo sygnału. Jeżeli sygnał $x(t)$ jest sygnałem rzeczywistym, to jego widmo wykazuje symetrię, tzn. $c_{-n} = \bar{c}_n$. Moduły liczb c_n tworzą widmo amplitudowe sygnału. Argumenty liczb c_n tworzą widmo fazowe sygnału.

Dla głosu męskiego częstotliwość $f = 1/T$, zwana częstotliwością tonu krtaniowego, jest równa ok. 100 Hz. Dla głosów żeńskich częstotliwość tonu krtaniowego jest odpowiednio wyższa i może dochodzić do 200 Hz – żeńskie głosy są wyższe od męskich.

W ludzkiej mowie można znaleźć fragmenty bardzo zbliżone do sygnału okresowego. Przykładowo, na rysunku 2a przedstawiony jest przebieg sygnału powstałego w wyniku rejestracji głoski „a". Okres T, z jakim sygnał powtarza się, odpowiada okresowi drgań wiązadeł głosowych. Na rysunku 2b przedstawione jest widmo amplitudowe zarejestrowanego przebiegu. Wartości współczynników c_n przedstawione są w skali logarytmicznej (w decybelach) między innymi dlatego, że narządy słuchu odbierają dźwięki w skali logarytmicznej.



Rys. 2. Fragment przebiegu czasowego głoski „a" i jego widmo.



Rys. 3. Przykład nałożenia widm.

Jeżeli okres F jest mniejszy niż $2f_0$, przesunięte względem siebie wersje widma sygnału nakładają się i na podstawie przebiegu funkcji $\widehat{X}(f)$ (rys. 3b) nie da się odtworzyć widma sygnału $X(f)$. Jeżeli okres F jest większy lub równy $2f_0$, to przedział obejmuje obszar, w którym widmo $X(f)$ jest różne od zera, przy czym przebieg funkcji $\widehat{X}(f)$ jest w tym przedziale identyczny z przebiegiem widma $X(f)$ (rys. 3c).

Mowa ludzka oraz otaczające nas dźwięki zawierają jednak także fragmenty nieokresowe, a nawet takie, które wykazują wiele cech losowości. W takim przypadku dźwięk nie jest już sumą tonów tylko o częstotliwościach f, $2f$, $3f$, itd., gdyż pojawiają się w nim dowolne częstotliwości, przy czym jedne w większym, inne w mniejszym natężeniu. Zamiast sumowania mamy więc całkowanie (bo bierzemy continuum częstotliwości, a nie tylko ich przeliczalną ilość), a zamiast współczynników a_n i b_n mamy całą funkcję $X(f)$, która określa, z jakim natężeniem poszczególne częstotliwości wchodzą w skład całego sygnału. Funkcję tę nazywamy widmem sygnału. Sygnał dany jest zatem wzorem

$$x(t) = \int_{-\infty}^{\infty} X(f) \mathrm{e}^{ift} \mathrm{d}f.$$

Cechą charakterystyczną systemów telekomunikacyjnych jest to, że widmo przekazywanych sygnałów jest różne od zera tylko w ograniczonym zakresie częstotliwości. Na przykład ludzie porozumiewają się za pomocą mowy, której widmo z pewnością kończy się na częstotliwości 22 000 Hz, a w praktyce wystarcza im mowa, której widmo ograniczono do 4 000 Hz, bo w tym obszarze zawarta jest cała wypowiadana informacja tekstowa, wszystkie informacje prozodyczne (np. akcentowanie) oraz wiele danych pozwalających na rozpoznanie mówcy. Ta właściwość głosu ludzkiego została wykorzystana w telefonii, która przesyła sygnały o częstotliwości nie przekraczającej 3 500 Hz.

Twierdzenie Shannona o próbkowaniu

Załóżmy, że mamy do czynienia z sygnałem $x(t)$, którego widmo $X(f)$ jest różne od zera jedynie w przedziale $\langle -f_0, f_0 \rangle$. Chcielibyśmy rozwinąć widmo w szereg Fouriera. Ale zrobić to można tylko z funkcją okresową. Tworzymy więc z widma funkcję o okresie F

$$\widehat{X}(f) = \sum_{n=-\infty}^{+\infty} X(f - nF),$$

gdzie F jest nie mniejsze niż $2f_0$. W tym przypadku współczynniki rozwinięcia funkcji w wykładniczy szereg Fouriera opisują jednocześnie przebieg widma $X(f)$. Można nietrudno wyliczyć, że są one równe

$$c_n = \frac{1}{F} x\left(-\frac{n}{F}\right).$$

Ze wzoru na współczynniki mamy bowiem

$$c_n = \frac{1}{F} \int_{-f_0}^{f_0} X(f) \mathrm{e}^{-2\pi i \frac{n}{F} f} \mathrm{d}f = \frac{1}{F} x\left(-\frac{n}{F}\right).$$

Aby odtworzyć sygnał $x(t)$, należy pobierać próbki tego sygnału w odstępach czasu $t = 1/F$. Wielkość F ma znaczenie częstotliwości próbkowania sygnału

Claude Elwood Shannon, ur. 1916, amerykański matematyk i inżynier.
Władimir A. Kotielnikow, ur. 1908, radiotechnik rosyjski.
I. Somey, radiotechnik japoński, twierdzenie o próbkowaniu sformułował w książce pt. „Hakei Denso" (Transmisja sygnałów).
Harry Nyquist, 1889–1976, amerykański elektrotechnik.

i, jak zauważyliśmy wcześniej, powinna być przynajmniej dwa razy większa od największej częstotliwości występującej w widmie sygnału próbkowanego. Jest to treść twierdzenia Shannona–Kotielnikowa–Someya o próbkowaniu. Minimalną częstotliwość próbkowania sygnału nazywa się częstotliwością Nyquista.

Ciąg kolejnych próbek sygnału analogowego nazywa się sygnałem dyskretnym. Wartości próbek sygnału zapisuje się zwykle z ograniczoną dokładnością w binarnym systemie zapisu liczb. Taki sygnał nazywa się sygnałem cyfrowym.

Wiemy bowiem, że $x(t)$ jest dany wzorem

$$\frac{1}{F}\sum_{n=-\infty}^{\infty} x\left(\frac{n}{F}\right)\int_{-F/2}^{F/2} e^{2\pi i f(t-n/F)}\,df = = \sum_{n=-\infty}^{\infty} x\left(\frac{n}{F}\right)\frac{\sin\pi(Ft-n)}{\pi(Ft-n)}.$$

Pozostaje problem odtworzenia oryginalnego sygnału analogowego na podstawie sygnału dyskretnego (cyfrowego). Wartości próbek sygnału $x(t)$ pobrane z częstotliwością Nyquista lub wyższą wyznaczają wartości współczynników rozwinięcia c_n okresowej wersji widma w wykładniczy szereg Fouriera.
Po wyzerowaniu funkcji poza przedziałem otrzymuje się widmo sygnału analogowego. Ten ostatni uzyskuje się w standardowy sposób z widma, co prowadzi do wzoru

$$x(t) = \sum_{n=-\infty}^{\infty} x\left(\frac{n}{F}\right)\frac{\sin\pi(Ft-n)}{\pi(Ft-n)}.$$

Aproksymacja sygnału pomiędzy próbkami odbywa się zatem z wykorzystaniem funkcji typu $\sin t/t$. Funkcja ta dla $t \to 0$ dąży do wartości 1 i zeruje się w punktach $k\pi$, k – całkowite. W ten sposób odtworzony sygnał jest równy dokładnie pobranej próbce w każdej z chwil $t = n/F$. Wartość sygnału pomiędzy próbkami odtwarzana jest na podstawie wszystkich próbek wcześniejszych i późniejszych. Im odleglejsza próbka, tym mniejszy jej wpływ na przebieg sygnału. W praktyce do aproksymacji stosuje się funkcję ograniczoną do skończonego zakresu wartości argumentu t.

Dźwięki zapisywane są na płycie kompaktowej w postaci próbek pobieranych z częstotliwością $F = 44{,}1$ kHz. Zgodnie z twierdzeniem o próbkowaniu zapewnia to poprawną rejestrację dźwięków, których widmo rozciąga się do częstotliwości 22 050 Hz, czyli poza obszar słyszalności większości ludzi. Każda próbka sygnału zapisana jest w systemie dwójkowym z wykorzystaniem szesnastu cyfr binarnych – bitów.
Na płycie zapisywany jest dźwięk stereofoniczny, tzn. dwa osobne sygnały: jeden pochodzący z kanału lewego, a drugi pochodzący z kanału prawego.
W rezultacie strumień bitów odczytywany podczas odtwarzania muzyki wynosi

$$2\cdot 44\,100\ \frac{\text{próbek}}{\text{s}}\cdot 16\ \frac{\text{bitów}}{\text{próbkę}} = = 1\,411\,200\ \frac{\text{bitów}}{\text{s}}.$$

Połączenie abonenta z centralą telefoniczną w znakomitej większości ma charakter analogowy. Tymczasem cyfrowa centrala telefoniczna przenosi rozmowy abonentów, faksy, sygnały analogowe generowane przez modemy itd. w postaci sygnału cyfrowego. Sygnał analogowy po stronie centrali przechodzi przez filtr dolnoprzepustowy, który ogranicza widmo sygnału do 3,5 kHz. Następnie sygnał jest próbkowany z częstotliwością 8 kHz. Każda próbka sygnału zapisywana jest z wykorzystaniem ośmiu bitów. Generowany strumień danych wynosi

$$8\,000\ \frac{\text{próbek}}{\text{s}}\cdot 8\ \frac{\text{bitów}}{\text{próbkę}} = 64\,000\ \frac{\text{bitów}}{\text{s}}.$$

Muzyka zapisana w postaci sygnału stereofonicznego to sygnał, którego widmo rozciąga się w zakresie od 16 Hz do 50 kHz. Obraz telewizyjny przekazywany jest w postaci sygnału, którego widmo nie przekracza 10 MHz.

W przypadku telefonii cyfrowej ISDN zamiana sygnału analogowego na cyfrowy następuje wewnątrz aparatu telefonicznego. Transmisja między centralą a abonentem ma charakter cyfrowy. Od abonenta do centrali płynie strumień danych 64 000 bitów/s. Centrala nadsyła taki sam strumień niosący głos abonenta zdalnego. Łącze cyfrowe można tak skonfigurować, aby przesyłać w jednym kierunku strumień 128 000 bitów/s z wykorzystaniem obu kanałów jednocześnie. Taki tryb pracy łącza przydaje się podczas pracy z Internetem.
Cyfrowy aparat telefonii komórkowej GSM przesyła do centrali cyfrowy sygnał mowy łączem radiowym. Sygnał ten jeszcze wewnątrz aparatu poddaje się procesowi kompresji, tzn. usuwa się z sygnału te wszystkie elementy, które są przez zmysł słuchu słabo odbierane. W rezultacie odtworzony u abonenta zdalnego sygnał mowy może dość istotnie różnić się od sygnału zarejestrowanego. Tyle tylko, że różnice te są słabo słyszalne. Natomiast łączem radiowym przesyła się dzięki temu strumień danych równy 13 000 bitów/s lub w nowszych wersjach systemu zaledwie 6 500 bitów/s.

Aparat głosowy człowieka ma charakter akustycznej rury rezonansowej. Podczas artykulacji mowy kluczową rolę odgrywają dwie najniższe częstotliwości rezonansowe tej „rury" nazywane formantami. Częstotliwości te mieszczą się w zakresie od 100 Hz do 700 Hz (pierwszy formant) i od 100 Hz do 2 000 Hz (drugi formant). To właśnie dlatego, kiedy na początku wieku prowadzono prace nad komutowaniem wielu rozmów telefonicznych po jednej parze przewodów, stwierdzono, że możliwą do przyjęcia jakość połączenia uzyskuje się, gdy każda para aparatów łączona jest 4 300 razy na sekundę. Odpowiada to poprawnemu odtwarzaniu sygnałów o częstotliwości do 2 150 Hz.

⋆⋆⋆

Przywykliśmy, że telefon komórkowy otrzymujemy od operatora za symboliczną złotówkę. Jest zwykle taki mały i niepozorny. Wielkością przypomina otoczaki znajdowane na dnie potoku. Ale w rzeczywistości jest rezultatem pracy i dociekliwości całej rzeszy ludzi genialnych, którzy przez wieki starali się dociec istoty Przyrody. Jednak nawet ten telefon nie jest ostatnim słowem techniki. Może więc warto wziąć udział w przygodzie z telekomunikacją?