

AKUSTOELEKTRONIKA

OD PODSTAW

Adam Kawalec
Mateusz Pasternak
e-skrypt WAT

Spis treści

1. Fale – pojęcia podstawowe	3
1.1. Fale w ciałach stałych	18
1.2. Sprężyste własności materiałów anizotropowych	21
1.3. Materiały piezosprężyste	28
2. Akustyczne fale powierzchniowe	31
2.1. Fala Rayleigha w piezoelektrykach	33
2.1.1. Warunki brzegowe	36
2.1.2. Rozwiązanie ogólne	36
2.1.3. Współczynnik sprzężenia piezoelektrycznego	37
2.1.4. Temperaturowy współczynnik czasu opóźnienia	38
2.1.5. Kąt przepływu mocy	38
2.2. Wybrane rodzaje akustycznych fal powierzchniowych	39
3. Synteza filtrów z AFP	43
3.1. Model funkcji δ	43
3.2. Teoria widmowa	45
3.3. Model schematu zastępczego	49
3.3.1. Zależności energetyczne	52
3.3.2. Dopasowanie energetyczne	53
3.4. Efekty drugiego rzędu	55
3.4.1. Odbicia międzyelektrodowe	55
3.4.2. Efekt trzeciego echa (potrójnego przejścia)	57
3.4.3. Inne efekty drugiego rzędu	58
3.5. Metody kształtowania przetworników	59
3.6. Przetworniki jednokierunkowe	61
3.6.1. Elementy analizy przetworników SPUDT	65
4. Urządzenia z AFP	68
4.1. Banki filtrów	68
4.2. Linie opóźniające	73
4.3. Linie dyspersyjne	75
4.4. Rezonatory	77
4.4.1. Struktury odbijające	77
4.4.2. Podstawowe konfiguracje rezonatorów	79
4.5. Konwoluty	81
4.6. Procesory Fouriera	86
4.7. Korelatory	90
4.8. Czujniki	91
4.9. Silniki	94
4.10. Układy zdalnej identyfikacji	96
4.11. Urządzenia na podłożach modyfikowanych	97
4.12. Przykłady zastosowań	104
5. Technologia urządzeń z AFP	107
5.1. Wytwarzanie kryształów	107
5.2. Wytwarzanie struktur z AFP	110
Dodatek	112
Analiza harmoniczna sygnałów	112
Transformacja Fouriera przebiegu okresowego	114
Dyskretna transformacja Fouriera	120
Szybka transformacja Fouriera (FFT)	122
Podstawy akustoelektronicznej obróbki sygnałów	125

1. Fale – pojęcia podstawowe

Duża część bodźców odbieranych przez zmysły organizmów żywych i w szczególności przez zmysły człowieka dociera do odpowiednich receptorów za pośrednictwem fal. Najważniejszy u ludzi zmysł wzroku jest z punktu widzenia fizyki zwykłym (i niezbyt dobrym) detektorem fal elektromagnetycznych, zaś drugi co do swej wagi zmysł słuchu, detektorem fal sprężystych (także nie należący do najlepszych).

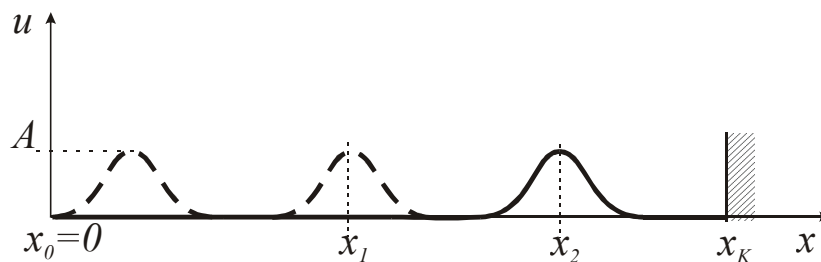
Wspólną cechą wszystkich zjawisk falowych (mechanicznych, elektromagnetycznych i in.) jest przenoszenie energii w czasoprzestrzeni. Dzięki temu fale znalazły i zapewne znajdują jeszcze wiele ważnych zastosowań technicznych. Na podstawie samej obserwacji fal można sformułować następującą definicję:

Fala = ruch zaburzenia w ośrodku ciągłym

Wnikliwy obserwator zauważyć też może, że przenoszona poprzez fale energia ulega ciągłym i cyklicznym przemianom z jednej postaci w drugą. W przypadku fal mechanicznych są to cykliczne przemiany energii potencjalnej w kinetyczną i na odwrót, przy czym przemiana taka jest możliwa tylko wtedy gdy odkształcony ośrodek dąży do przywrócenia stanu w jakim znajdował się przed odkształceniem.

Odrębną naturę mają fale elektromagnetyczne, których ruch polega na wzajemnej indukcji pól elektrycznego i magnetycznego nie wymagających (w sensie fizyki klasycznej) do swego istnienia żadnego ośrodka materialnego.

W celu ściślejszego opisu ruchu falowego wykonajmy następujący eksperyment. Przymocujmy jeden koniec liny w określonym miejscu, a drugi cyklicznie poruszajmy. Wytworzone w ten sposób zaburzenie będzie się poruszać wzdłuż liny przy czym sama lina po przejściu takiego zaburzenia zasadniczo nie zmieni swojego położenia. Przypiszmy linie kartezjański układ współrzędnych z końcem liny zamocowanym w x_k i początkiem w początku układu jak na rys 1.

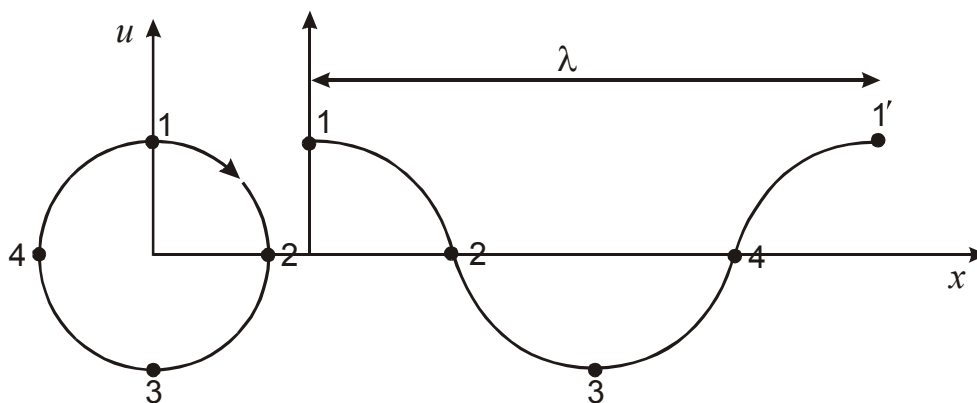


Rys. 1. Doświadczenie z liną

Jeśli swobodny koniec liny (lewy) porusza się wzdłuż osi u ruchem harmonicznym:

$$u = A \cos \omega t \quad (1)$$

to każdy punkt liny także poruszał się będzie takim ruchem, przy czym ruch kolejnych punktów będzie przesunięty w czasie (opóźniony w fazie) w stosunku do ruchu początku liny. Ilustrację ruchu harmonicznego przedstawia rys. 2.



Rys 2. Ruch harmoniczny. Jest on analogiczny do ruchu obrotowego koła. Czas potrzebny na zmianę fazy fali do położenia 1 do 1' (analogicznie do jednego pełnego obrotu koła) nazywa się okresem T , liczba cykli fali (w analogii do liczby obrotów koła) w jednostce czasu to częstotliwość f , zaś wyrażona w radianach zmiana fazy fali w jednostce czasu (kąt obrotu koła w jednostce czasu) to pulsacja ω . Z definicji zatem $\omega = 2\pi f = \frac{2\pi}{T}$. Dystans jaki pokonuje fala w czasie trwania jednego okresu nazywam się długością fali λ . Maksymalna wartość u (równa promieniowi koła) jest nazywana amplitudą fali A

Dla dowolnego punktu x opóźnienie fazowe względem początku linii proporcjonalne będzie do przebytej przez falę drogi. Jeśli współczynnik proporcjonalności oznaczymy przez k to wychylenie linii w dowolnym punkcie wyniesie:

$$u = A \cos(\omega t - kx). \quad (2)$$

Bezwzględna odległość dwóch punktów, których faza różni się o 2π (360°) nazywa się długością fali λ . Zatem:

$$\lambda = |x_1 - x_2|. \quad (3)$$

Ponieważ dla tych punktów różnica faz wynosi 2π zatem można zapisać:

$$(\omega t - kx_2) - (\omega t - kx_1) = 2\pi, \quad (4)$$

z czego po podstawieniu (3) uzyskać można:

$$k\lambda = 2\pi \text{ lub } k = \frac{2\pi}{\lambda} \quad (5)$$

Obliczony tu współczynnik proporcjonalności nosi nazwę liczby falowej.

Jeśli w równaniu (2) uwzględnimy związek $\omega = \frac{2\pi}{T}$ to otrzymamy:

$$u = A \cos 2\pi \left(\frac{t}{T} - \frac{x}{\lambda} \right). \quad (6)$$

Związek (6) jest równaniem monochromatycznej¹ fali biegnącej. Pozwala on wyliczyć wychylenie u w dowolnym punkcie x .

¹ tzn. sinusoidalnej o stałej amplitudzie i częstotliwości.

Jeśli w czasie jednego okresu T fala pokonuje dystans λ to musi się ona poruszać (propagować) z prędkością $v = \frac{\lambda}{T} = f\lambda = \frac{\omega\lambda}{2\pi} = \frac{\omega}{k}$, stąd $k = \frac{\omega}{v}$. Ściślej, jest to prędkość poruszania się dowolnej fazy fali stąd też nazywa się ona prędkością fazową.

Po wprowadzeniu tego pojęcia równanie (6) przyjmie postać:

$$u = A\cos\left(\omega t - \frac{\omega}{v}x\right) = A\cos\omega\left(t - \frac{x}{v}\right). \quad (7)$$

Ponieważ liczba falowa zależna jest od prędkości fali, która to wielkość jest wektorem zamiast liczby falowej lepiej posługiwać się wektorem falowym \mathbf{k} , co jest o tyle korzystniejsze, że wektor ten zawiera informację o kierunku ruchu fali. Liczba falowa jest wtedy długością wektora falowego.

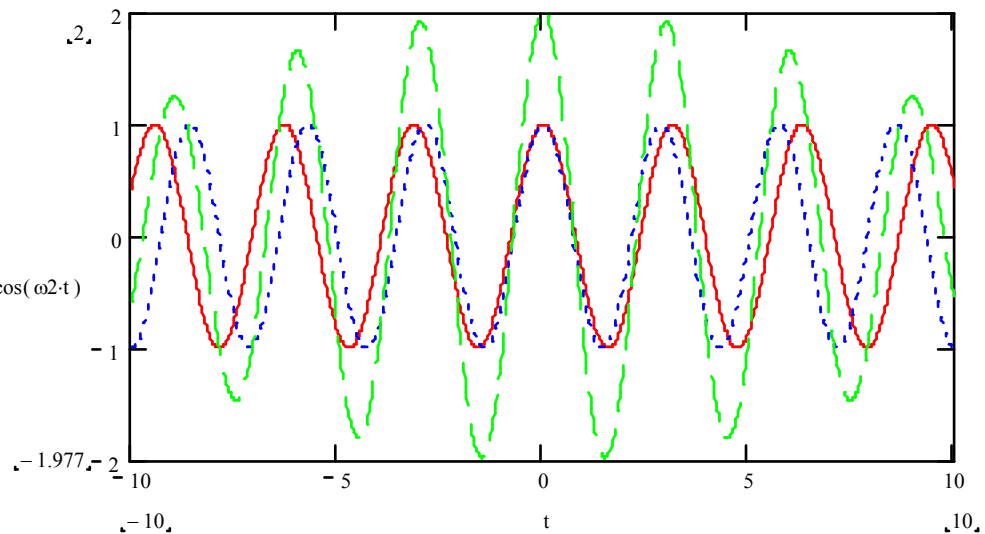
Stosując wektor falowy równanie fali monochromatycznej propagującej się w kierunku \mathbf{k} przyjmie postać:

$$u = A\cos(\omega t - \mathbf{k}\mathbf{r}) \quad (8)$$

gdzie \mathbf{r} jest wektorem położenia (wektorem wodzącym). Dla fali propagującej się wzdłuż osi x iloczyn wektorowy przechodzi w odpowiedni iloczyn skalarny $\mathbf{r}\mathbf{k} = xk$

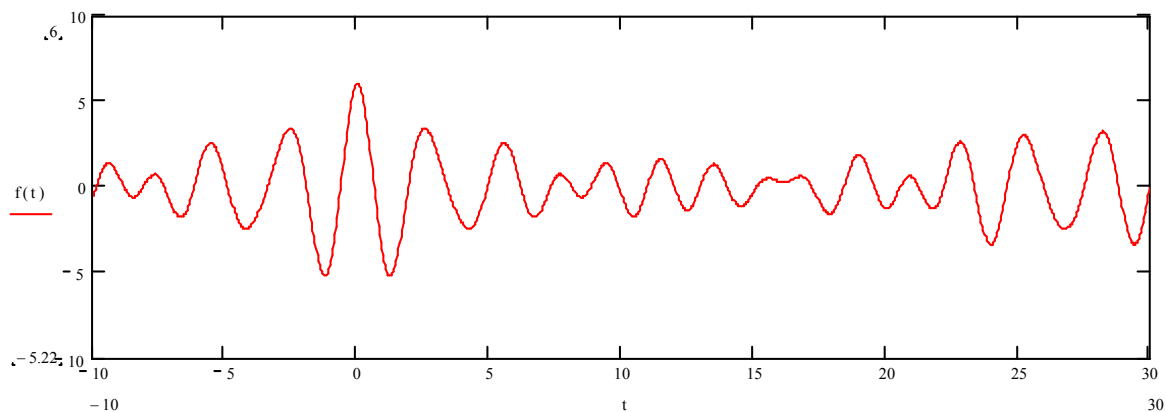
Fala monochromatyczna jest pewną idealizacją rzeczywistych procesów falowych. Na ogół bowiem fale występujące w przyrodzie lub też generowane sztucznie są superpozycjami (sumami czy też złożeniami) wielu fal monochromatycznych.

Rys. 3 prezentuje dwie fale o 10 % różnicy częstotliwości się o oraz ich sumę.



Rys. 3. Dwie fale i ich superpozycja

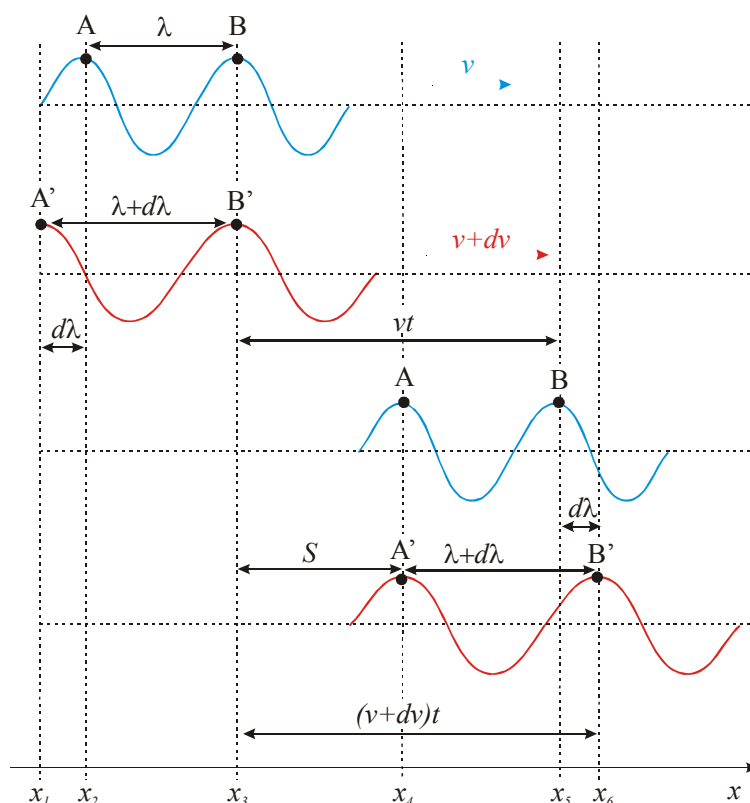
Zasada składania fal jest jak widać prosta. Maksymalna amplituda złożenia występuje w miejscach gdzie fazy fal składowych są zgodne (interferencja konstruktywna), minimalna gdzie przeciwne (interferencja destruktywna). Sens terminów konstruktywny i destruktywny widać lepiej na rys. 4.



Rys. 4. Superpozycja 6 fal. Wyraźnie widać wzmocnienia i wygaszenia. Amplitudy fal składowych, podobnie jak na rys. 3. są równe i wynoszą 1

W przypadku złożenia fal nie można posługiwać się pojęciem prędkości fazowej ponieważ fazy poszczególnych fal składowych mogą poruszać się z różnymi prędkościami. Ośrodki, w których to ma miejsce są nazywane ośrodkami dyspersyjnymi².

W takim ośrodku konieczne staje się określenie prędkości propagacji grupy fal (lub inaczej paczki falowej). Pomocny w wyznaczeniu tej prędkości będzie rys. 5, na którym złożenie (grupę) dwóch fal rozłożono na fale składowe.



Rys. 5. Diagram pomocny przy wyprowadzaniu relacji na prędkość grupy fal

Zaobserwujemy na rys. 5 ruch punktów A i B fali niebieskiej i A' oraz B' fali czerwonej. Jak widać w czasie kiedy punkty A i A' znajdują się w miejscach odpowiednio x_2 i x_1 punkty B

² Dyspersja jest to zależność parametrów ośrodka od częstotliwości, w szczególności mianem tym określa się zależność prędkości fali od częstotliwości w danym ośrodku.

i B' znajdują się we wspólnym miejscu x_3 . W tym miejscu amplituda fali wypadkowej będzie maksymalna. Po pewnym czasie, jako że fala czerwona jest szybsza punkt A' „dogoni” punkt A i znajdą się razem w miejscu x_4 . I w tym miejscu przypadnie też maksymalna amplituda złożenia obydwu fal. Punkty B i B' zajmą zaś pozycje odpowiednio x_5 i x_6 . Oczywiście ponieważ fala niebieska ma prędkość v to punkt B przemieści się w tym czasie o dystans vt , zaś punkt B' o dystans $(v+dv)t$. S to dystans pomiędzy najbliższymi maksimumami złożenia obydwu fal. Prędkość grupowa jest właśnie prędkością przemieszczania się tego maksimum. Stąd:

$$v_g = \frac{S}{t}. \quad (9)$$

Z rysunku wynika, że $S = vt - \lambda$ zatem:

$$v_g = v - \frac{\lambda}{t} \quad (10)$$

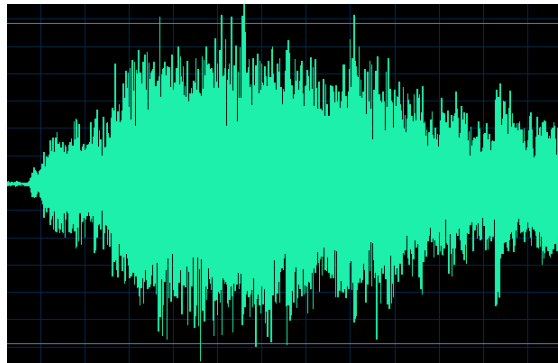
Łatwo też zauważyć, że $d\lambda = (v + dv)t - vt = dv \cdot t$ czyli $t = \frac{d\lambda}{dv}$, co po podstawieniu do wyrażenia (10) daje:

$$v_g = v - \lambda \frac{dv}{d\lambda} \quad (11)$$

Zależność (11) jest związkiem między prędkością grupową a prędkością fazową. Jeśli prędkość fazowa nie zależy od długości fali (a tym samym od częstotliwości), co ma miejsce w ośrodkach bezdyspersyjnych, prędkość grupowa jest równa prędkości fazowej.

Przy wyprowadzaniu zależności na prędkość grupową wykorzystano superpozycję dwóch fal. W praktyce jednak fale wykorzystywane są do przesyłania informacji, co jest równoznaczne z ich modulacją tj. nałożeniem na falę, którą zwykle nazywamy nośną, pewnej informacji. Należy zaznaczyć, że informacje też są swego rodzaju falami.

Rys. 6 przedstawia fragment paczki falowej zmodulowanej sygnałem informacyjnym (paczka falową zawierającą informację).



Rys. 6. Sygnał złożony

Jak łatwo zauważyć, przedstawiony na rys. 6. sygnał złożony składa się z bardzo dużej liczby fal monochromatycznych. W praktyce uznaje się, że jest ich nieskończenie wiele.

Jeśli w charakterze zmiennej wybierzemy liczbę falową pakiet taki można ściśle opisać następująco:

$$u(x, t) = \int_{-\infty}^{\infty} A(k) e^{i(\omega t - kx)} dk. \quad (12)$$

Wielkość $A(k)$ nazywana jest spektralną gęstością amplitudy pakietu falowego.

Uwaga.

Wykorzystując wzór Eulera dla liczb zespolonych $e^{i\varphi} = \cos \varphi + i \sin \varphi$ ($i = \sqrt{-1}$) można zapisać drgania harmoniczne w postaci wykładniczej, szczególnie dogodnej przy różniczkowaniu.

$$s = A \sin(\omega t + \varphi_0) = A \cos(\omega t + \varphi_1) = A e^{i(\omega t + \varphi_1)} = \tilde{A} e^{i\omega t}$$

gdzie: $\varphi_1 = \varphi_0 - \frac{\pi}{2}$. Wielkość $\tilde{A} = A e^{i\varphi_1}$ nazywana jest amplitudą zespoloną.

Z tego względu wyrażenie (12) może mieć też postać:

$$u(x, t) = \int_{-\infty}^{\infty} A(k) \cos(\omega t - kx) dk$$

W ośrodkach bezdyspersyjnych każda fala składowa propaguje się z tą samą prędkością i do dowolnego punktu x wszystkie fale dotrą w tym samym czasie (z jednakowym opóźnieniem):

$$t = \frac{x}{v}. \quad (13)$$

W takich ośrodkach paczka falowa nie ulega zniekształceniom. Zmiana fazy fali po przejściu drogi S jest proporcjonalna do jej częstotliwości.

$$\varphi = -\omega \frac{S}{v}, \quad (14)$$

a czas opóźnienia całej paczki falowej wynosi:

$$\tau_g = -\frac{d\varphi}{d\omega} = \frac{S}{v} \quad (15)$$

i nie zależy od częstotliwości.

W ośrodkach dyspersyjnych sytuacja jest odmienna. Częstotliwość nośna sygnałów złożonych f_0 jest zwykle dużo większa od częstotliwości sygnału modulującego, a gęstość widmowa amplitudy jest funkcją o wartościach ulokowanych w otoczeniu liczby falowej (lub inaczej w pobliżu ω_0). W takim przypadku można posłużyć się następującym przybliżeniem:

$$\omega(k) \approx \omega(k_0) + \left(\frac{d\omega}{dk} \right)_{k_0} (k - k_0). \quad (16)$$

Wielkość $\left(\frac{d\omega}{dk} \right)_{k_0} = v_g$.

Wstawiając (16) do (12) otrzymuje się następujący związek:

$$u(x, t) = e^{i[\omega(k_0) - k_0 v_g]t} \int_{-\infty}^{\infty} A(k) e^{-ik(x - v_g t)} dk, \quad (17)$$

co w skrócie można zapisać w postaci:

$$u(x, t) = e^{i\Omega_0 t} F(x - v_g t). \quad (18)$$

gdzie: $u(x, 0) = F(x)$ oraz $\Omega_0 = \omega(k_0) - k_0 v_g$.

Z wyrażenia (18) wynika, że w czasie t grupa fal przemieszcza się o odległość $S = v_g t$, gdzie

$$v_g = \left(\frac{d\omega}{dk} \right)_{k_0}.$$

Wyrażenie na prędkość grupy fal można znaleźć także tzw. metodą stacjonarnej fazy. Rozumowanie jest następujące:

Ponieważ w ośrodku dyspersyjnym każda fala składowa może mieć inną prędkość cały pakiet fal może ulegać zniekształceniom. Oczywiście składowe o częstotliwościach wyższych będą się propagowały wolniej niż składowe o częstotliwościach niższych. Maksima takiej grupy fal mogą więc ze sobą interferować. Aby tego uniknąć konieczna jest fazowa zgodność składowych z wektorami falowymi położonymi w pobliżu k_0 . Warunek ten można zapisać w formie:

$$\varphi(k) = \omega(k)t - kx = \text{const}, \quad (19a)$$

a po zróżniczkowaniu:

$$\left(\frac{d\varphi}{dk} \right)_{k_0} = \left(\frac{d\omega}{dk} \right)_{k_0} t - x = 0, \quad (19b)$$

stąd

$$v_g = \left(\frac{d\omega}{dk} \right)_{k_0}.$$

Jak widać uzyskuje się tą drogą wynik identyczny.

Jeśli ośrodek nie jest stratny tzn. amplitudy fal składowych nie maleją w czasie propagacji, prędkość grupowa jest wtedy równa prędkości przenoszenia energii paczki falowej.

Korzystając z relacji dla falowodu (czyli przenoszącego fale ośrodka ograniczonego) w postaci:

$$\omega = vk = v \sqrt{k_G^2 + \left(\frac{n\pi}{l} \right)^2} \quad (20)$$

można wyrazić prędkość grupową w falowodzie następująco:

$$v_G = \frac{d\omega}{dk_G} = vk_G \sqrt{k_G^2 + \frac{n^2\pi^2}{l^2}} \quad (21)$$

v jest prędkością fazową poza falowodem.

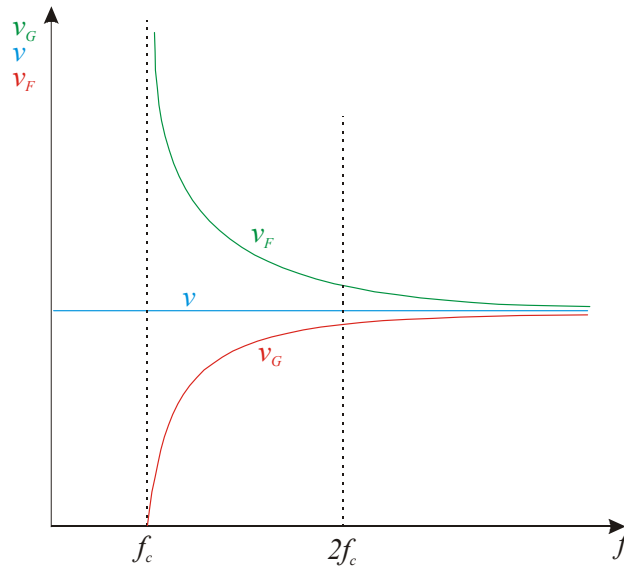
Prędkość fazowa w falowodzie wynosi :

$$v_F = \frac{\omega}{k_G} = \frac{v}{k_G} \sqrt{k_G^2 + \frac{n^2\pi^2}{l^2}}. \quad (22)$$

Z porównania (21) i (22) wynika że

$$v_G v_F = v^2 \quad (23)$$

Przebieg tych prędkości w funkcji częstotliwości przedstawia rys. 7.



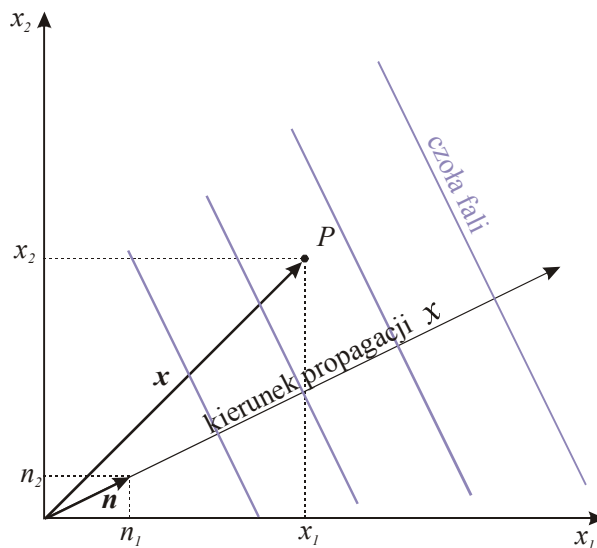
Rys. 7. Zależność prędkości w falowodzie od częstotliwości

Jak widać przy dostatecznie wielkich częstotliwościach prędkości fazowa i grupowa w falowodzie oraz prędkość w wolnej przestrzeni mają zbliżone wartości (fala jest na tyle krótka, że falowód „odczuwa” podobnie jak wolną przestrzeń). Przy zbliżaniu się do częstotliwości krytycznej f_c czyli wartości częstotliwości, poniżej której fala nie może propagować się w falowodzie (jest zbyt długa), prędkość grupowa maleje do zera, a fazowa rośnie do nieskończoności. Nie ma w tym nic sprzecznego ze Szczególną Teorią Względności, która nakłada ograniczenia na maksymalną prędkość rozchodzenia się sygnałów, ponieważ informacja przemieszcza się z prędkością grupową, a nie fazową.

Prędkość fazowa jest bowiem prędkością poruszania się obserwatora widzącego falę w tej samej fazie tzn. dla którego:

$$\omega\left(t - \frac{x}{v}\right) = const.$$

Rozpatrzmy stan drgania w pewnym punkcie P , który opisany jest wektorem \mathbf{x} łączącym P z początkiem kartezjańskiego układu współrzędnych jak na rys. 8.



Rys. 8. Przemieszczanie się czoła fali (frontu falowego albo in. miejsc ekwifazowych czyli miejsc o stałej fazie)

W dowolnym punkcie P określonym przez wektor \mathbf{x} drganie związane z falą propagującą się w kierunku jednostkowego wektora \mathbf{n} ma postać:

$$\mathbf{u}(\mathbf{x}, t) = F\left(t - \frac{\mathbf{n}\mathbf{x}}{v}\right) \quad (24)$$

Rzut punktu P na kierunek propagacji fali reprezentuje rzut wektora \mathbf{x} na kierunek \mathbf{n} wtedy:

$$u = F\left(t - \frac{\mathbf{n}\mathbf{x}}{v}\right) = F\left(t - \frac{n_1x_1 + n_2x_2}{v}\right) \quad (25)$$

Równanie (25) opisuje falę na płaszczyźnie. W celu zapisania tego wyrażenia w przestrzeni trójwymiarowej trzeba wprowadzić trzecią współrzędną punktu x_3 oraz trzecią składową wektora \mathbf{n} tj. n_3 . W trzech wymiarach uzyskamy:

$$u = F\left(t - \frac{\mathbf{n}\mathbf{x}}{v}\right) = F\left(t - \frac{n_1x_1 + n_2x_2 + n_3x_3}{v}\right) \quad (26)$$

Dla harmonicznej fali płaskiej (powierzchnie ekwifazowe są płaszczyznami) otrzymamy związek:

$$u = A \cos \omega \left(t - \frac{\mathbf{n}\mathbf{x}}{v}\right) = A \cos(\omega t - \mathbf{k}\mathbf{x}), \quad (27)$$

gdzie wektor falowy $\mathbf{k} = \frac{\omega \mathbf{n}}{v} = \frac{2\pi \mathbf{n}}{\lambda} = k\mathbf{n}$.

Interferencja

O interferencji była już mowa przy wyjaśnianiu sensu fizycznego prędkości grupowej. Przyjrzyjmy się dokładnie temu zjawisku. Ma ono miejsce gdy w danym punkcie przestrzeni punkt materialny doznaje wychyleń będącego sumą wychyleń docierających do tego punktu fal (tzw. zasada superpozycji).

Rozważmy dwie identyczne fale docierające do wspólnego punktu po przebyciu różnych dróg:

$$u_1 = A \cos \omega \left(t - \frac{x_1}{v}\right) \text{ i } u_2 = A \cos \omega \left(t - \frac{x_2}{v}\right) \quad (28)$$

Ich suma wyniesie:

$$u_1 + u_2 = A \left[\cos \omega \left(t - \frac{x_1}{v}\right) + \cos \omega \left(t - \frac{x_2}{v}\right) \right]. \quad (29)$$

Wykorzystując tożsamość trygonometryczną: $\cos \alpha + \cos \beta = 2 \cos \frac{\alpha + \beta}{2} \cos \frac{\alpha - \beta}{2}$

uzyskamy:

$$u = 2A \cos \frac{\omega \left(2t - \frac{x_1 + x_2}{v}\right)}{2} \cos \frac{\omega \frac{x_2 - x_1}{v}}{2} = 2A \cos \pi \frac{x_2 - x_1}{\lambda} \cos 2\pi \left(\frac{t}{T} - \frac{x_1 + x_2}{2\lambda}\right), \quad (30)$$

co w skrócie, po oznaczeniu $2A \cos \pi \frac{x_2 - x_1}{\lambda} = B$ oraz $\frac{x_1 + x_2}{2} = x$, można zapisać w postaci:

$$u = B \cos 2\pi \left(\frac{t}{T} - \frac{x}{\lambda} \right). \quad (31)$$

Amplituda drgania wypadkowego uzyskuje maximum dla $B = 2A$ czyli wtedy gdy:

$$x_2 - x_1 = n\lambda \quad n = 0, 1, 2, \dots, \quad (32)$$

czyli dla różnicy dróg będącej wielokrotnością długości fali (fale w fazie).

Amplituda wynosi zero gdy:

$$x_2 - x_1 = (2n + 1) \frac{\lambda}{2}, \quad (33)$$

czyli dla różnicy dróg będącej nieparzystą wielokrotnością połowy długości fali (fale w przeciwfazie).

Fale stojące

Ciekawy rezultat otrzymuje się przy nałożeniu dwóch fal o jednakowych amplitudach i częstotliwościach propagujących się w przeciwnym kierunku. Sytuacja taka ma często miejsce gdy fala zostaje w całości odbita od przeszkody i interferuje ze sobą.

$$u = A \left[\cos \omega \left(t - \frac{x}{v} \right) - \cos \omega \left(t + \frac{x}{v} \right) \right]. \quad (34)$$

Wykorzystując tożsamość trygonometryczną $\cos \alpha - \cos \beta = -2 \sin \frac{\alpha + \beta}{2} \sin \frac{\alpha - \beta}{2}$ otrzymamy:

$$u = 2A \sin \frac{2\pi x}{\lambda} \sin \omega t. \quad (35)$$

Wyrażenie to opisuje drgania harmoniczne ośrodka o amplitudzie zależnej od położenia i pulsacji ω . Fala stojąca nie przenosi zatem energii, a więc nie jest falą.

Maksymalną amplitudę fali stojącej zwaną strzałką otrzymamy dla $x = (2n + 1) \frac{\lambda}{4}$ zaś zerową

zwaną węzłem dla $x = \frac{n\lambda}{2}$. Odległość między sąsiednimi strzałkami wynosi $\lambda/2$ podobnie jak odległość między sąsiednimi węzłami. Położenie węzłów jest niezmiennie, a zmianie podlega tylko amplituda.

Fale stojące powstają w różnego typu ośrodkach ograniczonych, przez które przechodzą fale sprężyste.

Rozważmy falę odbijającą się całkowicie (zupełnie) pomiędzy dwoma równoległymi płaszczyznami. Przyjmijmy dla wygody, że oś x jest prostopadła do obydwu płaszczyzn. Niech pierwsza płaszczyzna znajduje się w $x=0$, a druga w $x=L$. W punktach tych muszą znaleźć się węzły fali, co oznacza spełnienie warunku:

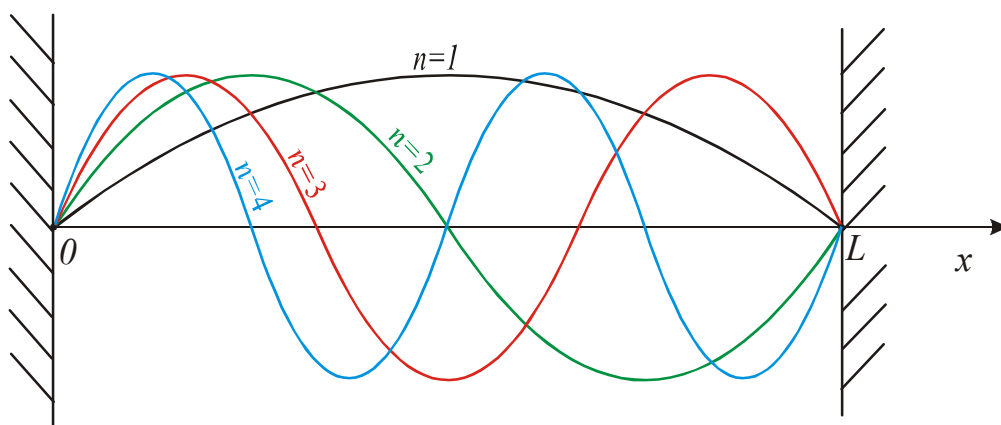
$$2A \sin \frac{2\pi L}{\lambda} = 0 \quad (36)$$

i prowadzi do zależności:

$$\lambda = \frac{2L}{n} \text{ lub też } f = n \frac{v}{2L}. \quad (37)$$

Między dwoma płaszczyznami odległymi o L mogą więc powstawać tylko fale stojące o długościach będących podwielokrotnością $2L$ czyli o częstotliwościach będących wielokrotnością $\frac{v}{2L}$.

Własność ta stanowi podstawę działania rezonatorów. Pierwsze trzy fale stojące (mody rezonatora) przedstawia rys. 9.



Rys 9. Mody rezonatora - zupełne fale stojące

Jak widać mody nieparzyste są symetryczne, a mody parzyste antysymetryczne względem płaszczyzny dzielącej rezonator na połowy.

Zupełne odbicie fali od przeszkody jest sytuacją wyidealizowaną. Przypadkiem realistycznym jest odbicie częściowe (niezupełne). Załóżmy, że na przeszkodę pada fala o amplitudzie A_{PAD} , odbija się częściowo od tej przeszkody i propaguje w przeciwnym kierunku z amplitudą A_{ODB} . Po nałożeniu tych fal otrzymuje się:

$$\begin{aligned} u &= A_{PAD} \cos \omega \left(t - \frac{x}{v} \right) - A_{ODB} \cos \omega \left(t + \frac{x}{v} \right) = \\ &= \underbrace{(A_{PAD} - A_{ODB}) \cos \omega \left(t - \frac{x}{v} \right)}_{\text{fala biegnąca}} + \underbrace{A_{ODB} \left[\cos \omega \left(t - \frac{x}{v} \right) - \cos \omega \left(t + \frac{x}{v} \right) \right]}_{\text{fala stojąca}} \end{aligned} \quad (38)$$

Jak widać rezultat nałożenia fali na siebie po jej częściowym odbiciu jest złożeniem fali biegnącej z falą stojącą. Wyrażenie (38) można też na mocy (35) zapisać w postaci:

$$u = (A_{PAD} - A_{ODB}) \cos \omega \left(t - \frac{x}{v} \right) + 2A_{ODB} \sin \frac{2\pi x}{\lambda} \sin \omega t \quad (39)$$

Miarą zupełności (czy też niezupełności) odbicia jest tzw. współczynnik fali stojącej określany z zależności:

$$WFS \stackrel{\text{def}}{=} \frac{A_{PAD} + A_{ODB}}{A_{PAD} - A_{ODB}}. \quad (40)$$

Dla odbicia zupełnego $WFS = \infty$ zaś w sytuacji gdy fala przechodzi przez przeszkodę nie ulegając odbiciu $WFS = 1$. W przypadkach odbić niezupełnych WFS przyjmuje wartości z przedziału $(1, \infty)$.

Inną miarą zupełności odbicia jest współczynnik odbicia określany przez relację:

$$r \stackrel{def}{=} \frac{A_{ODB}}{A_{PAD}}. \quad (41)$$

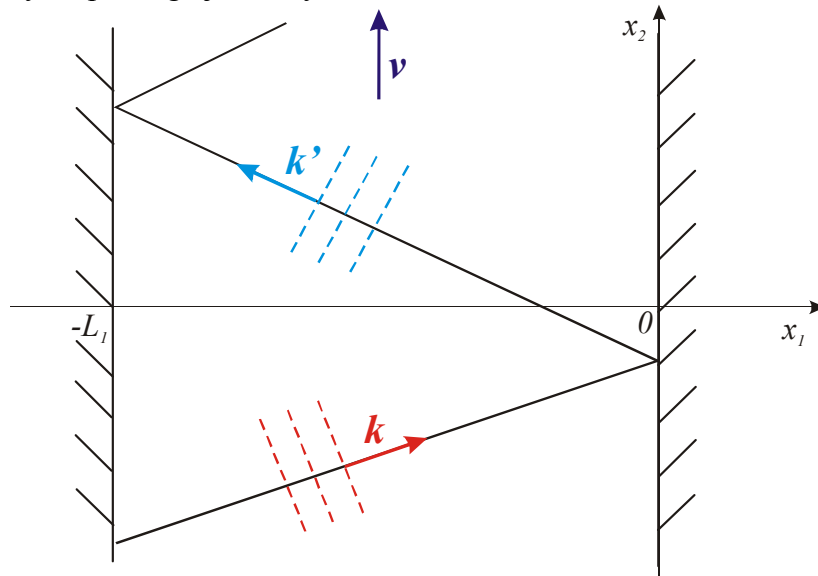
Współczynnik ten przyjmuje wartości od 0 w sytuacji gdy fala przechodzi przez przeszkodę nie ulegając odbiciu do 1 dla odbicia zupełnego.

Współczynniki WFS i r związane są ze sobą następująco:

$$WFS = \frac{1+r}{1-r} \quad (42)$$

Podobne współczynniki stosuje się także dla fal elektromagnetycznych.

W omówionych wyżej przypadkach fale padały na przeszkodę pod kątem prostym. Taki sposób padania jest jednak szczególny. Rozważmy teraz ogólniejszą sytuację kiedy kąt odbicia jest różny od prostego jak na rys. 10.



Rys. 10. Fala propagująca się w kierunku osi x_2 dzięki wielokrotnym odbiciom od równoległych ścian. Przerzywanymi liniami oznaczone są czoła fali

W układzie współrzędnych widocznym na rys. 10 fale padająca i odbita mają postać:

$$u_{PAD} = A \cos(\omega t - k_1 x_1 - k_2 x_2), \quad (43)$$

$$u_{ODB} = -A \cos(\omega t - k'_1 x_1 - k'_2 x_2). \quad (44)$$

Złożenie tych fal ma postać fali stojącej:

$$u = 2A \sin\left(\omega t - \frac{k_1 + k'_1}{2} x_1 - \frac{k_2 + k'_2}{2} x_2\right) \sin\left(\frac{k_1 - k'_1}{2} x_1 + \frac{k_2 - k'_2}{2} x_2\right). \quad (45)$$

Aby jej węzły znajdowały się w miejscach odbić od prawej ściany $x_1 = 0$ składowe wektorów falowych k_2 i k'_2 muszą być równe, a z kolei z zasady zachowania liczby falowej $k = k'$ wynika równość:

$$k'_1 = \pm k_1. \quad (46)$$

Nietrywialne rozwiązanie (45) uzyskać można tylko dla $k'_1 = -k_1$. Ma ono postać:

$$u = 2A \sin(\omega t - k_2 x_2) \sin k_1 x_1. \quad (47)$$

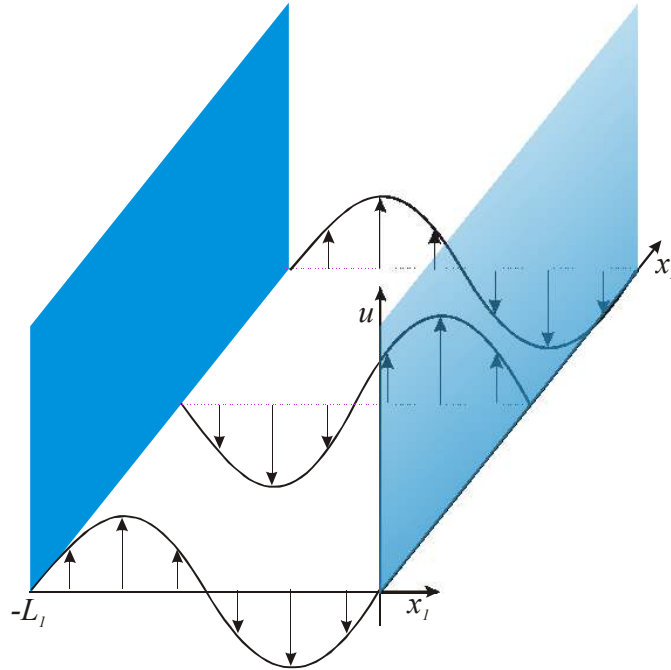
Aby węzły powstawały także w miejscach odbić od lewej ściany $x_1 = -L_1$ musi spełniony być warunek:

$$k_1 = n_1 \frac{\pi}{L_1}, \quad (48)$$

który po uwzględnieniu w (47) prowadzi do zależności:

$$u = 2A \sin(\omega t - k_2 x_2) \sin \frac{n_1 \pi}{L_1} x_1 \quad (49)$$

opisującej falę propagującą się w kierunku x_2 pomiędzy ścianami w $x_1=0$ i $x_1=-L_1$ czyli mod falowodowy. Przykład takiego modu pokazany jest na rys. 11.



Rys. 11. Drugi mod $n_l=2$, pierwszy antysymetryczny (por. rys 9)

Prędkość falowa takiej fali jest większa od prędkości w przestrzeni nieograniczonej. Wektory falowe w przestrzeni nieograniczonej \mathbf{k} i w falowodzie \mathbf{k}_G związane są zależnością:

$$\mathbf{k}_G^2 = \mathbf{k}^2 - \left(\frac{n_1 \pi}{L_1} \right)^2. \quad (50)$$

Najmniejsza możliwa długość \mathbf{k}_G ozn. k_c wynosi π/L_1 , co odpowiada maksymalnej długości fali:

$$\lambda_c = 2L_1. \quad (51)$$

Wielkość ta nazywana jest długością krytyczną fali lub też długością odcięcia fali. W falowodzie mogą się zatem propagować fale o długościach mniejszych od $2L_1$.

Zbliżanie się wartości długości do długości fali krytycznej odpowiada fizycznie zbliżaniu się kąta padania fali do wartości 90° , przy której fala zachowuje się jak w rezonatorze.

Znając prędkość fazową w falowodzie można też określić częstotliwość odcięcia:

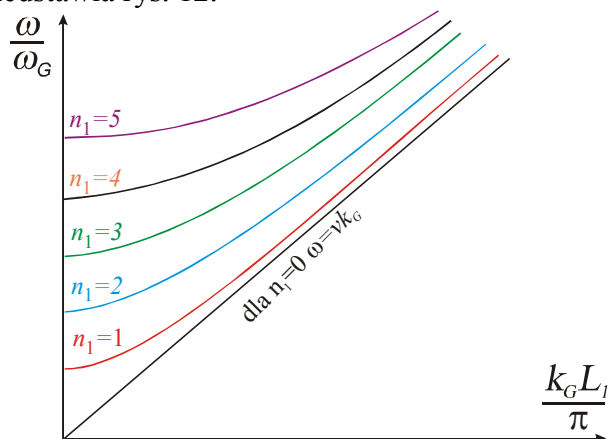
$$f_c = \frac{v}{2L_1}. \quad (52)$$

W falowodzie mogą propagować się fale o częstotliwościach wyższych od f_c .

Falowód jest przykładem struktury dyspersyjnej. Dyspersja czyli zależność parametrów fali od częstotliwości wprowadzona jest w tym przypadku poprzez geometrię falowodu. Z zależności (50) wynika związek:

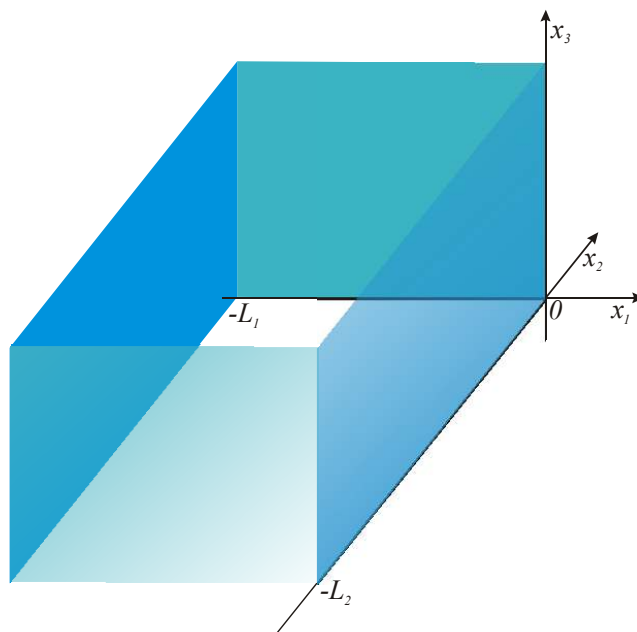
$$\omega = vk = v \sqrt{k_G^2 + \left(\frac{n_1 \pi}{L_1} \right)^2}. \quad (53)$$

Oczywiście przy nieograniczonym wzroście szerokości falowodu L_1 wyrażenie (53) przechodzi w relację dla przestrzeni nieograniczonej. Zależności dyspersyjne dla kilku pierwszych modów przedstawia rys. 12.



Rys. 12. Zależności dyspersyjne. Asymptotę $\omega = vk_G$ uzyskuje się dla $n_1=0$

Dołóżmy do układu ścian z rys. 11 kolejne dwie ściany w miejscach $x_2=0$ oraz $x_2=-L_2$ jak na rys. 13.



Rys. 13. Układ czterech ścian odbijających fale

W takim układzie otrzymamy falę padającą w postaci:

$$u_{PAD} = 2A \sin(\omega t - k_2 x_2) \sin\left(\frac{n_1 \pi}{L_1} x_1\right) \quad (54)$$

i falę odbitą

$$u_{ODB} = -2A \sin(\omega t - k_2 x_2) \sin\left(\frac{n_1 \pi}{L_1} x_1\right), \quad (55)$$

których superpozycja wynosi:

$$u = u_{PAD} + u_{ODB} = -4A \sin k_2 x_2 \sin\left(\frac{n_1 \pi}{L_1} x_1\right) \cos \omega t \quad (56)$$

Z faktu istnienia węzła fali w $x_2 = -L_2$ wynika $k_2 = \frac{n_2\pi}{L_2}$, $n_2 = 1, 2, 3, \dots$, co po uwzględnieniu w (56) daje:

$$u = -4A \sin\left(\frac{n_1\pi}{L_1}x_1\right) \sin\left(\frac{n_2\pi}{L_2}x_2\right) \cos\omega t. \quad (57)$$

Wiedząc, że $\left(\frac{\omega}{v}\right)^2 = \left(\frac{2\pi f}{v}\right)^2 = \mathbf{k}^2 = k_1^2 + k_2^2$ otrzymamy:

$$f = \frac{v}{2} \sqrt{\left(\frac{n_1}{L_1}\right)^2 + \left(\frac{n_2}{L_2}\right)^2}. \quad (58)$$

W falowodzie takim fala propaguje się w kierunku x_3 z wektorem falowym k_3 i ma postać:

$$u_{PAD} = 4A \sin\left(\frac{n_1\pi}{L_1}x_1\right) \sin\left(\frac{n_2\pi}{L_2}x_2\right) \cos(\omega t - k_3x_3). \quad (59)$$

Jeśli obszar propagacji fali ponownie zawężymy wstawiając ścianę w $x_3=0$ to fala odbita będzie wyrażona przez:

$$u_{ODB} = -4A \sin\left(\frac{n_1\pi}{L_1}x_1\right) \sin\left(\frac{n_2\pi}{L_2}x_2\right) \cos(\omega t + k_3x_3). \quad (60)$$

Superpozycja tych fal wyrazi się poprzez:

$$u = 8A \sin\left(\frac{n_1\pi}{L_1}x_1\right) \sin\left(\frac{n_2\pi}{L_2}x_2\right) \sin k_3x_3 \cos\omega t. \quad (61)$$

Jeśli postawimy ścianę w $x_3 = -L_3$ to uzyskamy zamkniętą przestrzeń w postaci prostopadłościanu – czyli rezonator.

Uwzględniając, podobnie jak poprzednio, węzeł w $x_3 = -L_3$ otrzymuje się $k_3 = \frac{n_3\pi}{L_3}$,

$n_3 = 1, 2, 3, \dots$, co daje:

$$u = 8A \sin\left(\frac{n_1\pi}{L_1}x_1\right) \sin\left(\frac{n_2\pi}{L_2}x_2\right) \sin\left(\frac{n_3\pi}{L_3}x_3\right) \cos\omega t, \quad (62)$$

oraz analogicznie do (58):

$$f = \frac{v}{2} \sqrt{\left(\frac{n_1}{L_1}\right)^2 + \left(\frac{n_2}{L_2}\right)^2 + \left(\frac{n_3}{L_3}\right)^2}. \quad (63)$$

Uzyskane proste wyrażenia dla falowodu i rezonatora obowiązują w przypadkach prostych odbić fali, bez towarzyszących im zjawisk transformacji modów falowych, które to zjawiska powszechnie występuje w ciałach stałych – po odbiciu fala przekształca się (transformuje) w szereg różnych fal o własnych polaryzacjach i prędkościach.

1.1. Fale w ciałach stałych

Jeśli odkształcony ośrodek dąży do przywrócenia stanu w jakim znajdował się przed odkształceniem mówi się, że wykazuje on cechy sprężyste. Sprężystość ośrodka może zależeć od tego, który jego fragment jest rozpatrywany oraz od tego w jaki sposób, bądź w jakim kierunku ciało jest odkształcane.

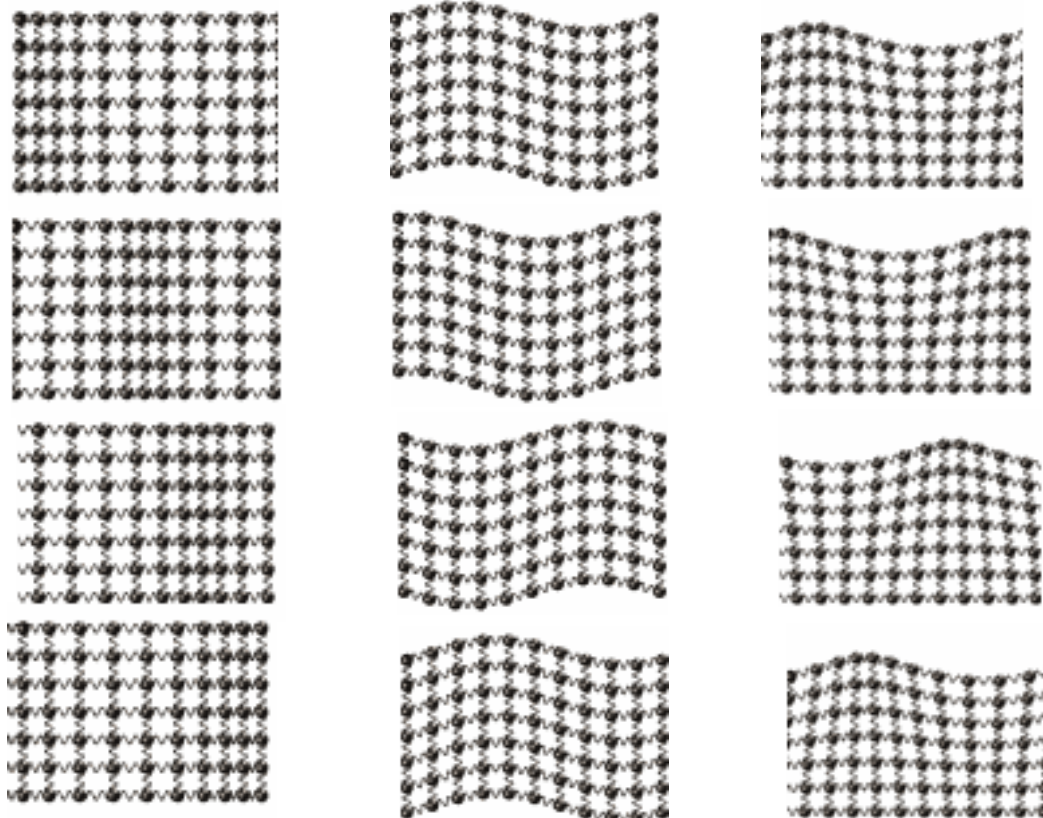
W ośrodkach sprężystych propagować się mogą dwa rodzaje fal: fale podłużne i poprzeczne.

Fale podłużne tworzone są poprzez cykliczne zagęszczanie i rozrzedzanie ośrodka i mogą występować zarówno w gazach, cieczech, jak i ciałach stałych. Wektor odkształceń dla tych fal jest zgodny z kierunkiem ich propagacji.

Fale poprzeczne występują tylko w ciałach stałych, które posiadają sprężystość postaci tzn. dążą do przywrócenia swojego kształtu (wyjątek stanowi powierzchnia cieczy posiadająca sprężystość postaci dzięki zjawisku napięcia powierzchniowego). Szczególną cechą fal poprzecznych jest prostopadłość wektora odkształceń do kierunku rozchodzenia się fali.

Fale powierzchniowe propagują się po powierzchni ciał stałych i są wynikiem faktu, że przy powierzchni fale poprzeczne i podłużne nie mogą propagować się w sposób niezależny, co powoduje, że obydwa te typy fal propagują się łącznie pod postacią fali powierzchniowej.

Przykłady fal sprężystych prezentuje rys. 9.



Rys. 1. Fale podłużne

Fale poprzeczne

Fale na powierzchni cieczy

Prędkość rozchodzenia się fal sprężystych określona jest przez wzór Newtona:

$$v = \sqrt{\frac{E}{\rho}}, \quad (1)$$

gdzie E jest wyznaczanym empirycznie modułem sprężystości. Ściślej rzecz ujmując dla fal podłużnych rolę modułu sprężystości pełni moduł ściśliwości K , a dla fali poprzecznej moduł sztywności G . Moduł ściśliwości dla gazów wynosi $p \frac{c_p}{c_v}$ tzn. jest iloczynem ciśnienia i stosunku ciepła właściwego przy stałym ciśnieniu do ciepła właściwego przy stałej objętości.

Moduł sprężystości równa się z def.

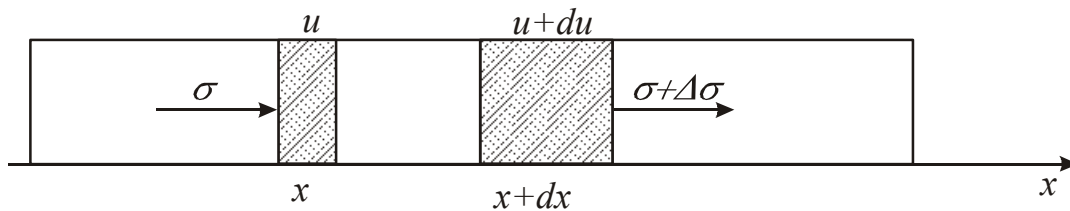
$$E = \frac{F}{S} : \frac{\Delta l}{l}, \quad (2)$$

czyli

$$\frac{\Delta l}{l} = \frac{1}{E} \cdot \frac{F}{S}, \quad (3)$$

gdzie F jest siłą, S powierzchnią przekroju, l długością ośrodka, zaś Δl spowodowanym przez falę jego wydłużeniem. Stosunek F/S nazywa się naprężeniem i ozn. symbolem σ , zaś stosunek $\Delta l/l$ nazywany jest odkształceniem i ozn. przez ε , E jest współczynnikiem empirycznym naz. modułem Younga. Równanie (3) można więc przedstawić jako proporcjonalność odkształceń do naprężeń. Twierdzenie o takiej proporcjonalności nazywa się prawem Hooke'a. **Prawo to obowiązuje tylko dla małych odkształceń.**

Aby przyrzeć się dokładniej falam sprężystym, rozważmy propagację fali podłużnej w pręcie o przekroju S i długości l jak na rys. 2.



Rys. 2. Fala podłużna w pręcie

W pewnej chwili t wychylenie w punkcie x wynosi u zaś w punkcie $x+dx$ $u+du$. Wychylenie w punkcie x spowodowane jest naprężeniem σ , a w punkcie $x+dx$ $\sigma+\Delta\sigma$.

Pręt na dystansie dx wydłużył się o du . Zatem:

$$\frac{\Delta l}{l} = \frac{du}{dx} = \frac{\partial u}{\partial x}. \quad (4)$$

stąd też

$$du = \frac{\partial u}{\partial x} dx. \quad (5)$$

Z prawa Hooke'a wiadomo, że względny przyrost długości zachodzący pod wpływem naprężenia wynosi:

$$\frac{\partial u}{\partial x} = \frac{1}{E} \sigma \quad (6)$$

Rozpatrzmy ruch elementu masy pręta zawartej na odcinku pomiędzy x a $x+dx$. Element ten ma postać:

$$dm = S \rho dx. \quad (7)$$

Po podziałaniu na obydwie strony (7) operatorem przyspieszenia otrzymamy:

$$dm \frac{\partial^2 u}{\partial t^2} = S \rho dx \frac{\partial^2 u}{\partial t^2} \quad (8)$$

Z II zasady dynamiki Newtona wiadomo, że lewa strona (8) jest siłą działającą na element masy dm . Pamiętając, że naprężenie to stosunek działającej siły do przekroju, siłą działającą na element dm można również przedstawić jako różnicę iloczynów naprężeń i przekrojów na dystansie dx :

$$F = S(\sigma + d\sigma) - S\sigma = S \left(\sigma + \frac{\partial \sigma}{\partial x} dx \right) - S\sigma = S \frac{\partial \sigma}{\partial x} dx. \quad (9)$$

Porównując siły wyrażone w (8) i (9) otrzymamy:

$$\rho \frac{\partial^2 u}{\partial t^2} = \frac{\partial \sigma}{\partial x} \quad (10)$$

Posługując się równaniem (6) można z równania (10) wyeliminować naprężenia. Różniczkując obie strony równania (6) po dx otrzymamy:

$$\frac{\partial^2 u}{\partial x^2} = \frac{1}{E} \frac{\partial \sigma}{\partial x} \quad (11)$$

Z porównania (10) i (11) wynika:

$$\rho \frac{\partial^2 u}{\partial t^2} = E \frac{\partial^2 u}{\partial x^2}. \quad (12)$$

Równanie (12) korzystając z (1) można przekształcić do postaci:

$$\frac{\partial^2 u}{\partial t^2} = v^2 \frac{\partial^2 u}{\partial x^2} \quad (13)$$

Równanie (13) nazywa się równaniem falowym. Równanie jest jednowymiarowe, ale można je łatwo uogólnić na przypadek trójwymiarowy:

$$\frac{\partial^2 u}{\partial t^2} = v^2 \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} \right) = v^2 \nabla^2 u \quad (14)$$

Równanie falowe ma zasadnicze znaczenie dla analizy zagadnień technicznych dot. ruchu falowego.

Rozpatrzmy przy użyciu tego równania najprostsze zagadnienie, tzn. drgania obustronnie zamocowanej cienkiej struny. Prędkość fali w strunie analogicznie do (1) określimy jako:

$$v = \sqrt{\frac{T}{\rho}}$$

gdzie T jest napięciem struny, zaś ρ gęstością materiału, z którego struna jest wykonana.

Rozwiązanie jednowymiarowego równania falowego ma ogólną postać:

$$u(t, x) = (A \sin kx + B \cos kx) \sin(\omega t + \varphi).$$

W celu jej uszczegółowienia rozważmy warunki brzegowe ujmujące fakt, że w miejscach zamocowania struny jej wychylenia są zerowe:

$$u|_{x=0} = 0 \text{ oraz } u|_{x=l} = 0.$$

Warunki te po podstawieniu do rozwiązania jednowymiarowego równania falowego są źródłem następującego układu dwóch równań:

$$u|_{x=0} = 0 \Rightarrow A \cdot 0 + B \cdot 1 = 0$$

$$u|_{x=l} = 0 \Rightarrow A \sin kl + B \cos kl = 0$$

Wyznacznik tego układu ma postać:

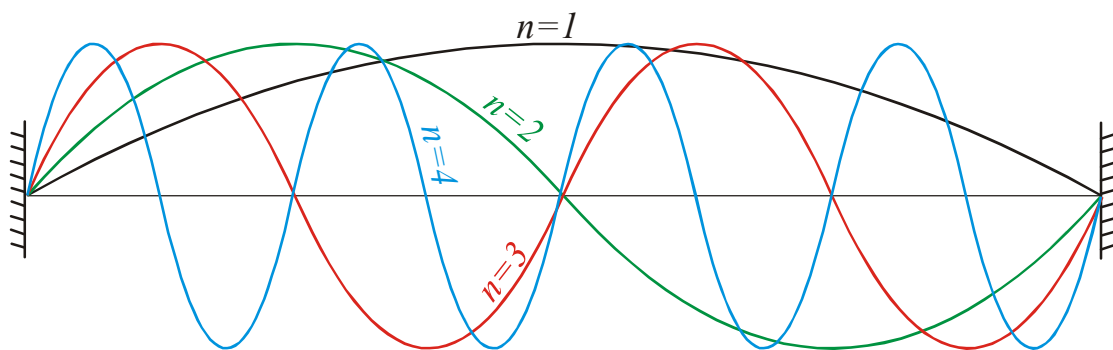
$$\begin{vmatrix} 0 & 1 \\ \sin kl & \cos kl \end{vmatrix}$$

Przyrównując ten wyznacznik do zera otrzyma się wyrażenie na tzw. wartości własne zagadnienia czyli związek pozwalający wyznaczyć wszystkie możliwe drgania układu. Ma ono postać:

$$\sin kl = 0.$$

Stąd:
$$k_n = \frac{n\pi}{l} \Rightarrow \lambda_n = \frac{2l}{n} \quad n=1, 2, 3 \dots,$$

Pierwsze 4 rodzaje drgań prezentuje poniższy rysunek.



1.2. Sprężyste własności materiałów anizotropowych

Sprężystość wynika z elektrycznych oddziaływań (sił) między atomami ciała, które, po ich wzajemnym przemieszczeniu (tj. odkształceniu tegoż ciała) dążą do odzyskania poprzedniego, zwykle energetycznie korzystniejszego, ustawienia. Ze względu na olbrzymią liczbę atomów nawet w ciele o niewielkich rozmiarach geometrycznych zjawiska związane z elektrycznym ich oddziaływaniem opisuje się zwykle wielkościami makroskopowymi. Takimi wielkościami są choćby zdefiniowane już wcześniej naprężenie σ i odkształcenie ε . W przytoczonych wcześniej definicjach mieliśmy do czynienia z działaniem sił powodujących odkształcenia w jednym tylko kierunku. W anizotropowych ciałach stałych, tzn. ciałach, których własności zależne są od rozpatrywanego kierunku, trzeba zazwyczaj uwzględnić w obliczeniach wszystkie kierunki, co powoduje pewną komplikację przedstawionych wyżej zagadnień – naprężenia i odkształcenia takiego ciała są wielkościami tensorowymi.

Terminu tensor użył pierwszy raz Wolfgang Voigt ok. 1910 r. właśnie w celu umożliwienia opisu naprężeń mechanicznych w ciałach anizotropowych. Łatwiej będzie zrozumieć czym jest tensor rozważając następujący przykład.

Pole elektryczne \mathbf{E} wywołuje w przewodniku przepływ prądu o określonej gęstości \mathbf{j} . Jeśli przewodnik jest izotropowy to obydwa wektory są do siebie równoległe, a wielkość gęstości prądu jest proporcjonalna do natężenia pola elektrycznego, co zapisuje się w postaci związku:

$$\mathbf{j} = a\mathbf{E}, \quad (\text{a})$$

lub po rozłożeniu wektorów na składowe:

$$j_1 = aE_1, \quad j_2 = aE_2, \quad j_3 = aE_3.$$

Jeśli przewodnik nie jest anizotropowy to wektory \mathbf{E} oraz \mathbf{j} nie są do siebie równoległe, a ich związek ma postać układu równań:

$$\begin{cases} j_1 = a_{11}E_1 + a_{12}E_2 + a_{13}E_3 \\ j_2 = a_{21}E_1 + a_{22}E_2 + a_{23}E_3, \\ j_3 = a_{31}E_1 + a_{32}E_2 + a_{33}E_3 \end{cases} \quad (\text{b})$$

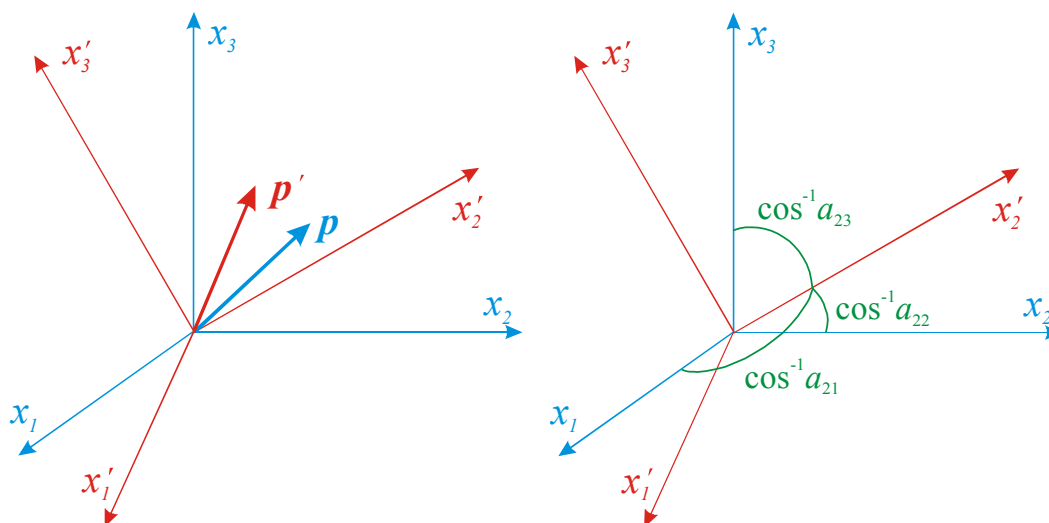
lub w skrócie:

$$j_i = a_{ij}E_j.$$

Każda składowa \mathbf{j} jest tu liniową kombinacją trzech składowych \mathbf{E} . Musi tak być z powodu zależności związku pomiędzy tymi wektorami od kierunku. Współczynniki a_{ij} tworzą tensor drugiego rzędu (rzęd tensora jest równy liczbie wskaźników). Każda składowa tensora ma swoją ścisłą fizyczną interpretację. W celu pełnego określenia przewodnictwa ośrodka anizotropowego należy więc podać dziewięć współczynników. Zwykle robi się to w postaci tabeli.

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \quad (\text{c})$$

Słuszność wzoru (b) łatwiej zauważyć rozpatrując związek między współrzędnymi tego samego wektora w dwóch różnych układach kartezjańskich o wspólnym początku.



Oznaczmy przez \mathbf{p} wektor w układzie pierwszym, zaś przez \mathbf{p}' ten sam wektor w układzie

drugim (przetransformowanym). Jednostki miary w obydwu układach są identyczne. Związki kątowe pomiędzy osiami oryginalną a przetransformowaną wyznaczają kosinusy kierunkowe. Jak widać z rysunku każda z trzech osi układu przetransformowanego tworzy z trzema osiami układu oryginalnego trzy kąty (na rysunku pokazane są tylko kąty jakie tworzy oś x'_2 układu przetransformowanego z trzema osiami układu oryginalnego). Jednoznaczne określenia związku pomiędzy tymi dwoma układami wymaga zatem podania dziewięciu kątów albo dziewięciu kosinusów kierunkowych. Jak widać z rysunku kosinusami kierunkowymi osi x'_2 względem x_1, x_2, x_3 są a_{21}, a_{22}, a_{23} przy czym pierwszy wskaźnik odnosi się do osi przetransformowanej, a drugi do oryginalnej. Kosinusy kierunkowe można zebrać w tabeli:

		współrzędne oryginalne		
		x_1	x_2	x_3
współrzędne przetransformowane	x'_1	a_{11}	a_{12}	a_{13}
	x'_2	a_{21}	a_{22}	a_{23}
	x'_3	a_{31}	a_{32}	a_{33}

Tabela zawiera 9 współczynników (niektóre są od siebie zależne i ich liczba może być mniejsza). Dowolną składową wektora p'_i otrzymuje się przez rzutowanie wszystkich składowych wektora p (p_1, p_2, p_3 potraktowanych jako wektory odpowiednio skierowane wzdłuż osi oryginalnych) na kierunek x'_i , zatem:

$$\begin{aligned} p'_1 &= p_1 \cos(x_1, x'_1) + p_2 \cos(x_2, x'_1) + p_3 \cos(x_3, x'_1), \\ p'_2 &= p_1 \cos(x_1, x'_2) + p_2 \cos(x_2, x'_2) + p_3 \cos(x_3, x'_2), \\ p'_3 &= p_1 \cos(x_1, x'_3) + p_2 \cos(x_2, x'_3) + p_3 \cos(x_3, x'_3) \end{aligned}$$

lub w skrócie za pomocą tabeli:

$$\begin{aligned} p'_1 &= a_{11}p_1 + a_{12}p_2 + a_{13}p_3, \\ p'_2 &= a_{21}p_1 + a_{22}p_2 + a_{23}p_3, \\ p'_3 &= a_{31}p_1 + a_{32}p_2 + a_{33}p_3 \end{aligned}$$

albo też:

$$p'_i = a_{ij}p_j. \tag{d}$$

Współczynniki a_{ij} tworzą tensor drugiego rzędu.

Powtarzając rozumowanie w przeciwnym kierunku tzn. dla sytuacji kiedy osie przetransformowane są oryginalnymi i vice versa otrzymamy:

$$p_i = a_{ji}p'_j. \tag{e}$$

W wyrażeniach tych wskaźnik (indeks) j powtarzający się po jednej stronie znaku równości, nazywany jest wskaźnikiem niemy, zaś wskaźnik i występujący po obu stronach równości wskaźnikiem wolnym. Jak widać we wzorach transformacja wyrażająca przetransformowane współczynniki za pomocą oryginalnych daje zależność (d), w której wskaźniki nieme sąsiadują ze sobą. Przy transformacji odwrotnej (e) są one rozdzielone.

Niech pewna cząstka materialna (punkt materialny) znajduje się w miejscu $\mathbf{x} = (x_1, x_2, x_3)$. Jeśli cząstkę ta wytrącić z położenia równowagi to przemieści się ona o $\mathbf{u} = (u_1, u_2, u_3)$, którego składowe będą zależne od położenia początkowego (w przeciwnym wypadku uważa się, że żadne siły nie działają na cząstkę, a przemieszcza się ona razem z ciałem). Aby odróżnić przemieszczenia cząstki spowodowane deformacją ciała od przemieszczeń spowodowanych ruchem ciała deformację (odkształcenie) określa się następująco:

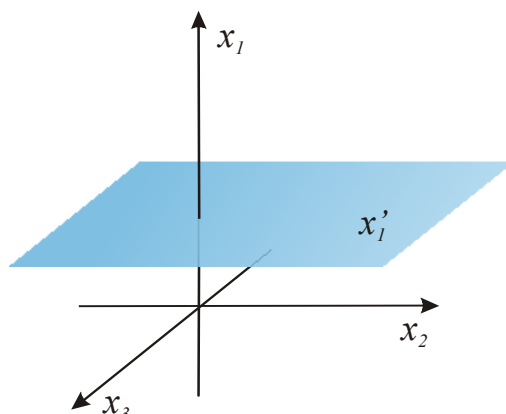
$$\varepsilon_{ij}(x_1, x_2, x_3) = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right), \quad i, j=1, 2, 3 \quad (1)$$

Zależność ta obowiązuje w każdym punkcie ciała. Można pokazać, że deformacja (odkształcenie) jest tensorem drugiego rzędu, który jak większość tensorów w mechanice jest symetryczny tzn. spełnia równość:

$$\varepsilon_{ij} = \varepsilon_{ji}, \quad (2)$$

co oznacza niezależność 6 pośród 9 składowych.

Siły wewnętrzne działające w ciele anizotropowym opisuje się tensorem naprężeń σ_{ij} . (W literaturze anglojęzycznej tensor deformacji (tensor odkształceń) oznaczany jest zwykle literą S , zaś tensor naprężeń literą T .) W celu określenia tensora naprężeń wprowadźmy w kartezjańskim układzie współrzędnych płaszczyznę x'_1 jak na rys. 3.



Rys. 3. Pomocniczy rysunek do definicji tensora sprężystości

Założmy, że płaszczyzna ta leży wewnątrz nieskończenie rozciągniętego ciała stałego. Założmy, że pod działaniem deformacji w dowolnym obszarze sąsiadującym z płaszczyzną materiał z jednej strony płaszczyzny będzie działać na materiał po drugiej stronie płaszczyzny (zachowana jest ciągłość materiału). Oczywiście siły mogą działać w dowolnych kierunkach i zmieniać się. Naprężenie definiuje się jako siłę działającą na jednostkę powierzchni płaszczyzny x'_1 w kierunku x_1 . Dla rys. 3 jeśli $x_1 < x'_1$ będzie to $\sigma_{i1}(x'_1, x_2, x_3)$, a jeśli $x_1 > x'_1$ to $-\sigma_{i1}(x'_1, x_2, x_3)$, $i=2,3$. Definicja ta obowiązuje dla każdej płaszczyzny równoległej do x'_1 zatem można zapisać: $\sigma_{i1}(x_1, x_2, x_3)$. Podobnie też można rozważać płaszczyzny przecinające pod kątem prostym pozostałe osie tj. x_2 oraz x_3 i tym samym określić pozostałe składowe tensora naprężeń. Ostatecznie, we wszystkich kierunkach będziemy mieli:

$$\sigma_{ij}(x_1, x_2, x_3).$$

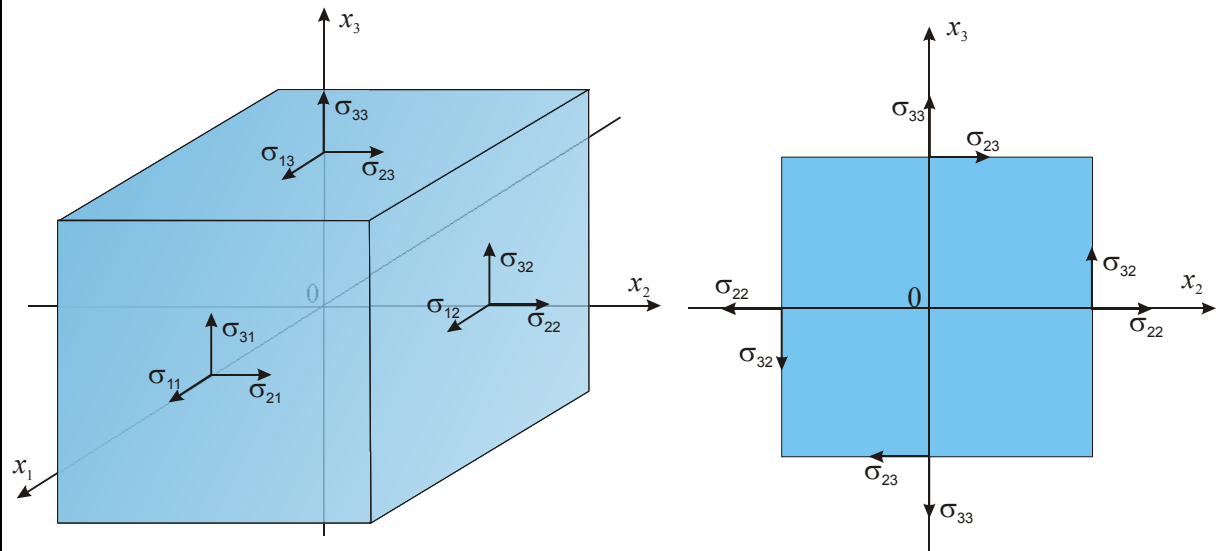
Przytoczona tu konstrukcja, określa tensor naprężeń, ale fakt, że wielkość ta jest tensorem (i podlega odpowiednim prawom transformacji przy przechodzeniu z dowolnego układu

kartezjańskiego do innego) wymaga oddzielnego dowodu, który wykracza poza ramy prowadzonych tu rozważań.

Podobnie jak tensor odkształceń także tensor naprężeń jest symetryczny:

$$\sigma_{ij} = \sigma_{ji}.$$

Znaczenie naprężeń ilustruje poniższy rysunek:

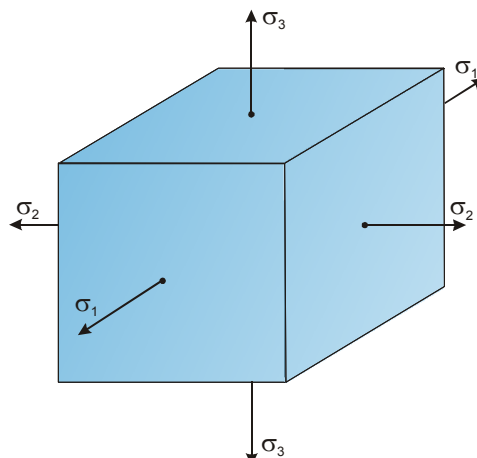


Po lewej: siły działające na ścianki wyciętego z przestrzeni sześcianu jednostkowego dostatecznie małego aby występujące w nim naprężenia były jednorodne. Po prawej przekrój sześcianu płaszczyzną prostopadłą do osi x_1 i działające nań siły. Naprężenia σ_{ii} są typu ściskającego natomiast σ_{ij} ścinającego.

Z symetrii tensora naprężeń wynika możliwość sprowadzenia go do osi głównych tj. takich, dla których znikają naprężenia typu ścinającego:

$$\begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{bmatrix} \rightarrow \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{bmatrix}.$$

Zabieg ten sprowadza się do takiego wyboru układu współrzędnych, dla którego krawędzie sześcianu jednostkowego są równoległe do trzech głównych kierunków naprężeń.



Także dla ciał anizotropowych deformacja związana jest z naprężeniem poprzez prawo Hook'a, które także i w tym przypadku obowiązuje dla małych odkształceń. Postać tego prawa musi być dostosowana do wielkości, które w nim występują. W przypadku jednowymiarowym współczynnikiem proporcjonalności była odwrotność modułu Younga. Dla ciał anizotropowych składowe tensora naprężeń będą wyrażane poprzez liniowe kombinacje składowych tensora deformacji i współczynnikami proporcjonalności będą odpowiednie składowe tensora sztywności (in. tensora stałych sprężystych) C :

$$\sigma_{ij} = \sum_k \sum_l C_{ijkl} \varepsilon_{kl}, \quad i, j, k, l = 1, 2, 3. \quad (3)$$

lub wykorzystując, wygodną i oszczędną w zapisie konwencję Einsteina automatycznego sumowania po powtarzających się indeksach:

$$\sigma_{ij} = C_{ijkl} \varepsilon_{kl}. \quad (4)$$

Po prawej stronie równania (4) powtarzają się indeksy k oraz l , co przy takiej notacji oznacza konieczność sumowania po nich i znaki sum można pominąć. Konwencja ta stosowana będzie w zapisie następnych wyrażań.

Tensor sztywności jest jak widać czwartego rzędu i posiada $3 \cdot 3 \cdot 3 \cdot 3 = 81$ składowych. Wartości tych składowych są specyficzne dla każdego ciała i w pełni charakteryzują jego własności sprężyste. Symetria tensora deformacji i tensora naprężeń wymaga, aby składowe tensora sztywności były inwariantne względem zamiany indeksów i oraz j lub k oraz l tj.:

$$C_{ijkl} = C_{jikl},$$

$$C_{ijkl} = C_{ijlk}.$$

Dzięki temu tylko 36 elementów tensora sztywności jest niezależnych. Jeśli dalej skorzystać z praw termodynamiki to okazuje się, że indeksy można zamieniać parami:

$$C_{ijkl} = C_{klij},$$

co powoduje kolejną redukcję składowych tensora sztywności do 21. Często liczbę niezależnych składowych udaje się jeszcze zredukować biorąc pod uwagę symetrię materiału. Dla perowskitów czyli kryształów o strukturze kubicznej liczba ta redukuje się do 3.

Szczególnym przypadkiem są materiały izotropowe (o własnościach niezależnych od kierunku), dla których tensor sztywności redukuje się do dwóch stałych Lamé λ i μ :

$$C_{ijkl} = \lambda \delta_{ij} \delta_{kl} + \mu (\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}), \quad (5)$$

gdzie: $\delta_{ij} = \begin{cases} 1 & \text{dla } i = j \\ 0 & \text{dla } i \neq j \end{cases}$ jest deltą Kroneckera.

Równanie ruchu

Przytoczone wyżej prawo Hook'a jest prawem statycznym (niezależnym od czasu). Jeśli naprężenia i deformacje zależą od czasu i współrzędnych przestrzennych to prawo Hook'a powinno być uzupełnione o część dynamiczną. Uzupełnienia tego można dokonać wykorzystując II zasadę dynamiki Newtona:

$$\mathbf{F} = m\mathbf{a}. \quad (6)$$

Siła działająca na jednostkową objętość ciała, w którym istnieje naprężenie σ_{ij} ma postać wynikającą z definicji tensora naprężeń:

$$F_i = \frac{\partial \sigma_{ij}}{\partial x_j}. \quad (7)$$

Zgodnie z (6) sile tej odpowiada przyspieszenie w kierunku i , zaś masie, w przypadku objętości jednostkowej, odpowiada gęstość (ponieważ $m=\rho V$). Porównując siły (6) i (7) otrzymamy:

$$\rho \frac{\partial^2 u_i}{\partial t^2} = \frac{\partial \sigma_{ij}}{\partial x_j}. \quad (8)$$

Wstawiając do prawa Hook'a (4) wyrażenie na tensor deformacji (1) można prawo to można wyrazić następująco:

$$\sigma_{ij} = C_{ijkl} \frac{\partial u_j}{\partial x_k}. \quad (9)$$

Wstawiając następnie (9) do (8) uzyskamy:

$$\rho \frac{\partial^2 u_i}{\partial t^2} = C_{ijkl} \frac{\partial^2 u_l}{\partial x_j \partial x_k}. \quad (10)$$

Równanie to ma postać ogólnego równania falowego:

$$\frac{\partial^2 \Psi}{\partial t^2} = v^2 \nabla^2 \Psi, \quad (11)$$

dla funkcji Ψ , która ma w ogólności postać fali płaskiej. Każde odkształcenie zmienne w czasie jest więc w ciele stałym źródłem fali.

Jeśli ośrodek jest nieskończenie rozciągliwy we wszystkich kierunkach (tzn. jest przestrzenią sprężystą) to fale płaskie rozwiązujące równanie (10) opisane są częścią rzeczywistą wyrażenia:

$$\mathbf{u} = \mathbf{A} e^{i\mathbf{k}(l_i x_i - vt)}, \quad (12)$$

gdzie $\mathbf{A} = \mathbf{j} \alpha_j$, \mathbf{j} jest wersorem osi x_i , α_j składową wektora własnego przemieszczeń, zaś v prędkością fazową fali mierzoną wzdłuż wektora falowego \mathbf{k} o kosinusach kierunkowych l_i .

Jednorodne fale płaskie tego typu nazywane są falami objętościowymi.

Prędkości fazowe i odpowiedni wektor własny przemieszczenia o składowych α_j w przypadku przestrzeni sprężystej dają się wyrazić w jawnej postaci poprzez podstawienie rozwiązania (12) do równania (10). Podstawienie to, po wykonaniu odpowiednich działań prowadzi do równania Greena-Christoffela:

$$\left(\Gamma_{jk} - \delta_{jk} \rho v^2 \right) \alpha_j = 0, \quad (13)$$

w którym $\Gamma_{jk} = l_i l_l C_{ijkl}$ jest pierwszym tensorem Christoffela, zaś δ_{jk} deltą Kroneckera. Równanie (13) jest *de facto* układem równań jednorodnych. Z teorii wyznaczników wiadomo, że układ taki posiada rozwiązanie wtedy i tylko wtedy gdy jego wyznacznik jest równy zero, stąd:

$$|\Gamma_{jk} - \delta_{jk} \rho v^2| = 0. \quad (14)$$

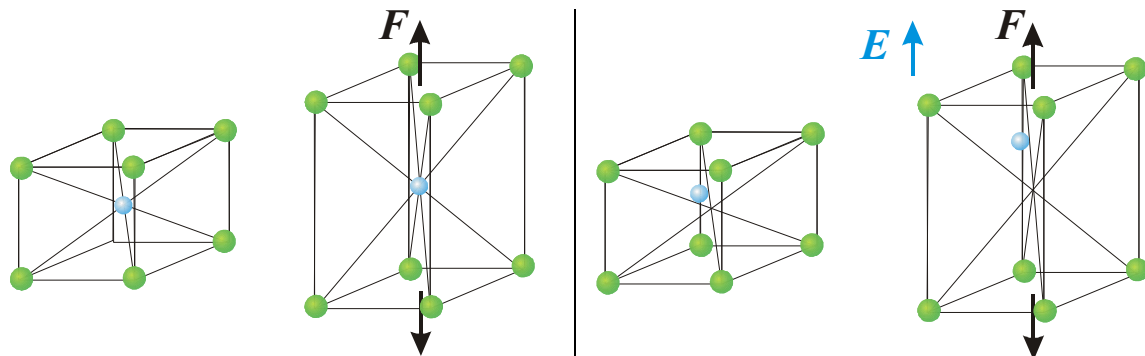
Trzy rozwiązania tego równania są kwadratami prędkości dwóch fal objętościowych, zaś wektory własne odpowiadające tym rozwiązaniom są wzajemnie prostopadłe. Jeden z nich należy do fali podłużnej o prędkości $v_l = \sqrt{\frac{C_{11}}{\rho}}$, a dwa pozostałe do fal poprzecznych o prędkościach $v_t = \sqrt{\frac{C_{44}}{\rho}}$. Dla większości ciał stałych $v_t = 0,63v_l$.

Uwaga.

Ponieważ słowa poprzeczny i podłużny zaczynają się w języku polskim na tą samą literę, zwykle w literaturze polskojęzycznej stosuje się przy indeksowaniu nazewnictwo łacińskie lub angielskie. Stąd też prędkości fal poprzecznych i podłużnych oznacza się odpowiednio v_t transversus (ang. transversal) – poprzeczny oraz v_l longitudo (ang. longitudinal) – podłużny.

1.3. Materiały piezosprężyste

Piezosprężystość jest zjawiskiem polegającym na wzajemnym sprzężeniu deformacji z natężeniem pola elektrycznego lub wektorem indukcji elektrycznej. Zjawisko to występuje w materiałach anizotropowych o strukturze pozbawionej centrów symetrii (rys. 1).



Rys. 1. Schemat zjawiska piezoelektrycznego. Po lewej: deformacja komórki kryształu o symetrii punktowej nie powoduje naruszenia równowagi elektrycznej (symetria zostaje zachowana). Po prawej: brak symetrii punktowej jest przyczyną zachwiania równowagi elektrycznej po deformacji (symetria ulega zmianie).

Sprężystość jest w istocie makroskopowym efektem mikroskopowych oddziaływań elektrycznych. W kryształach o wysokiej symetrii żadne odkształcenie nie jest jednak w stanie zaburzyć istniejącej równowagi elektrycznej. Jeśli kryształ pozbawiony jest centrum symetrii to przy odpowiednio skierowanym naprężeniu można tą równowagę zaburzyć. Wtedy atomy, których oddziaływania elektryczne nie równoważą się tworzą dipol elektryczny. Pojawienie się na skutek naprężeń takich dipoli makroskopowo objawia się istnieniem wypadkowego wektora natężenia pola elektrycznego. Jest to tzw. proste zjawisko piezoelektryczne (in. prosty efekt piezoelektryczny). Istnieje też odwrotny efekt piezoelektryczny - przyłożenie pola elektrycznego indukuje dipole elektryczne, co jest równoznaczne z przemieszczaniem się odpowiednich atomów i deformacją ośrodka.

Często efekt ten jest słaby i jego wpływ na sprężystość nie ma wielkiego znaczenia, jednak dla materiałów o wysokim współczynniku sprzężenia deformacji z natężeniem pola elektrycznego (tzw. silnych piezoelektryków) musi ono zostać uwzględnione w równaniach.

Prawo Hook'a (4) dla takiego ośrodka należy więc uzupełnić, biorąc pod uwagę siły wynikające z piezoeфекtu. Przy założeniu małości odkształceń związek pomiędzy odkształceniami, naprężeniami i polem elektrycznym można zapisać w postaci:

$$\sigma_{ij} = C_{ijkl}^E \varepsilon_{kl} - e_{ijk} E_k, \quad (15)$$

gdzie C_{ijkl}^E jest tensorem sztywności (stałych sprężystych) określonym przy stałym natężeniu pola elektrycznego \mathbf{E} , zaś e_{ijk} jest tensorem piezoelektrycznym wiążącym pole elektryczne z polem odkształceń).

Wektor indukcji elektrycznej \mathbf{D} , zwykle zależny tylko od wektora natężenia pola elektrycznego \mathbf{E} i tensora przenikalności elektrycznej ε_{ij} w materiałach piezoelektrycznych zależy również od tensora deformacji ε_{ij} :

$$D_i = \varepsilon_{ij}^e E_j + e_{ijk} \varepsilon_{jk}, \quad (16)$$

gdzie ε_{ij}^e jest tensorem przenikalności elektrycznej przy stałej deformacji.

Wektor indukcji elektrycznej \mathbf{D} analogicznie można też wyrazić poprzez tensor naprężeń:

$$D_i = \sum_j \varepsilon_{ij}^\sigma E_j + \sum_{j,k} d_{ijk} \sigma_{jk} \quad (17)$$

gdzie ε_{ij}^σ jest tensorem przenikalności elektrycznej przy stałym naprężeniu, a d_{ijk} tensorem podatności piezoelektrycznej.

Wyprowadzone wcześniej równanie ruchu (10) jest spełnione także w ośrodkach piezoelektrycznych jednakże z uwzględnieniem towarzyszącego deformacjom wektora natężenia pola elektrycznego. Wektor ten, ze względu na fakt, że fale sprężyste propagują się znacznie wolniej niż elektromagnetyczne ($\sim 10^5$ razy) można w przybliżeniu uznać za niezmienny, co upoważnia do wyrażenia go poprzez gradient potencjału:

$$E_i = -\frac{\partial \varphi}{\partial x_i}. \quad (18)$$

Po wstawieniu (18) i (15) do równania ruchu (10) otrzymuje się równanie ruchu dla ośrodka piezoelektrycznego w postaci:

$$\rho \frac{\partial^2 u_i}{\partial t^2} - C_{ijkl}^E \frac{\partial^2 u_k}{\partial x_j \partial x_l} - e_{ijk} \frac{\partial^2 \varphi}{\partial x_j \partial x_k} = 0. \quad (19)$$

Ponieważ piezoelektryk jest równocześnie dielektrykiem (używa się też określenia piezodielektryk) czyli ośrodkiem nie zawierającym ładunków swobodnych to:

$$\operatorname{div} \mathbf{D} = 0.$$

Biorąc powyższe pod uwagę i wykorzystując związek (16) uzyskamy:

$$\varepsilon_{ij}^e \frac{\partial^2 \varphi}{\partial x_i \partial x_j} - e_{ijk} \frac{\partial^2 u_j}{\partial x_i \partial x_k} = 0. \quad (20)$$

Równania (19) i (20) stanowią układ 4 równań wiążących trzy przemieszczenia i potencjał, który w przestrzeni swobodnej spełnia także równanie Laplace'a $\nabla^2 \varphi = 0$.

Rozwiązania układu (19) (20) poszukuje się w postaci:

$$u_j = \alpha_j e^{ik[l_1 x_1 - vt]}, \quad (21)$$

$$\varphi = \alpha_4 e^{ik[l_1 x_1 - vt]}, \quad (22)$$

gdzie l_1 i l_2 wskazują kierunek propagacji (kosinusy kierunkowe), a l_3 , v i α_j są wielkościami szukanymi. Podstawiając przewidywane rozwiązania do układu równań (19) i (20) uzyskuje się układ 4 równań jednorodnych względem α :

$$(\Gamma_{rs} - \delta'_{rs} v^2) \alpha_r = 0, \quad (23)$$

gdzie $\Gamma_{ij} = C_{ijkl} l_i l_j$, $\Gamma_{j4} = \Gamma_{4j} = e_{ij} l_i l_k$, $\Gamma_{44} = -\epsilon_{ik} l_i l_k$, $r, s=1, 2, 3, 4$, zaś $i, j=1, 2, 3$.

Postępując jak poprzednio, tzn. przyrównując wyznacznik układu (23) do zera otrzymamy równanie ósmego stopnia (o współczynnikach rzeczywistych) względem l_3 .

Równanie to ma osiem pierwiastków odpowiadających przemieszczeniom i kwadratowi prędkości fal poprzecznych i podłużnych propagujących się w objętości piezoelektryka.

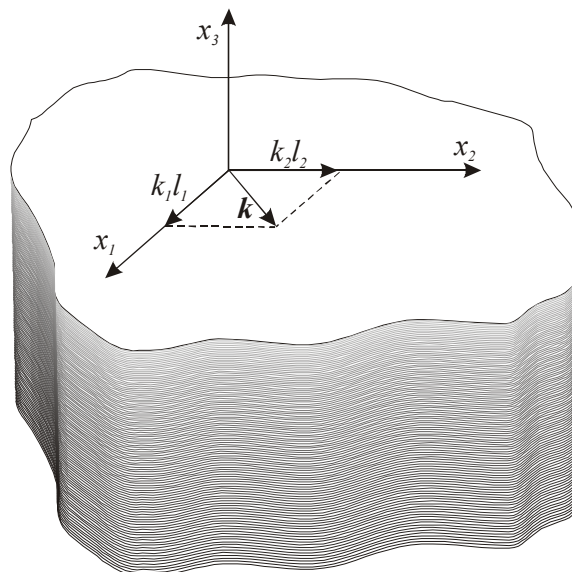
2. Akustyczne fale powierzchniowe

Jak już wiadomo z poprzednich rozdziałów w jednorodnej przestrzeni sprężystej mogą propagować się dwa rodzaje fal tj. fale podłużne i fale poprzeczne. Wiadomo też, że ograniczenie tej przestrzeni dwiema lub czterema ścianami powoduje powstawanie modów falowodowych, a ograniczenie tejże przestrzeni do pewnej objętości (poprzez trzecią parę ścian) powoduje, że w rozwiązaniach równań mają postać modów rezonatorowych.

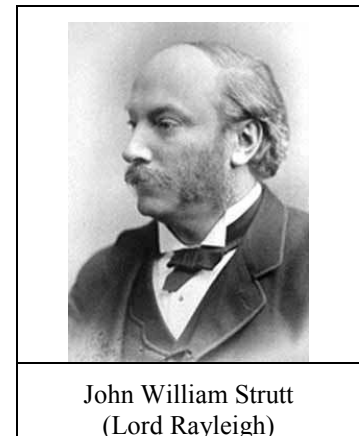
Interesujące rozwiązania, szczególnie z punktu widzenia akustoelektroniki (a także sejsmologii), uzyskuje się po ograniczeniu przestrzeni sprężystej płaszczyzną dzielącą tę przestrzeń na dwie półprzestrzenie sprężyste. Uzyskana po takim ograniczeniu powierzchnia staje się nośnikiem odrębnego typu fal. Propagujące się wewnątrz kryształu fale podłużne i poprzeczne ulegają złożeniu i propagują się w pobliżu powierzchni jako jedna fala. Tego typu falę nazwa się falą powierzchniową typu Rayleigha [wym. *Rejleja*]. Półprzestrzeń sprężysta jest więc nośnikiem nie tylko fal podłużnych i poprzecznych, które daleko od jej powierzchni propagują się tak jak w przestrzeni nieograniczonej, ale i fali powierzchniowej. Ogólną własnością takich fal jest szybkie zanikanie wraz ze wzrostem głębokości wywołanych przez nie przemieszczeń ośrodka - w ruchu falowym uczestniczy tylko leżąca pod powierzchnią warstwa ośrodka o grubości rzędu kilku długości fali.

Dla jasności trzeba tutaj stwierdzić, że nośnikiem tylko jednego typu fal powierzchniowych, wspomnianych fal Rayleigha, jest powierzchnia idealnie gładka i swobodna (bez naprężeń). Każde odstępstwo od tej idealizacji powoduje powstawanie fal powierzchniowych różnorodnych typów. Także odstępstwo od jednorodności wewnątrz kryształu (np. w postaci defektu liniowego) powoduje powstawanie różnych typów fal objętościowych.

Rozważmy pokazany na rys. 1 wycinek brzegu półprzestrzeni sprężystej wybierając kartezjański układ współrzędnych z osią x_3 prostopadłą do powierzchni. Będziemy poszukiwać fali o wektorze falowym \mathbf{k} mającej tę własność, że amplituda wszystkich składowych wywołanych przez nią przemieszczeń znika w nieskończonej odległości od powierzchni.



Rys. 1. Wycinek brzegu półprzestrzeni sprężystej z naniesionym układem współrzędnych



John William Strutt
(Lord Rayleigh)

Założmy, że rozwiązaniem równania falowego jest liniowa kombinacja członów w postaci:

$$u = \alpha_j e^{ikl_3 x_3} e^{ik(l_1 x_1 + l_2 x_2 - vt)}, \quad (24)$$

l_i są kosinusami kierunkowymi.

Rozwiązanie to będzie odpowiadało fali powierzchniowej typu Rayleigha jeśli spełniony będzie warunek:

$$u \Big|_{x_3 \rightarrow -\infty} = 0. \quad (25)$$

W takim ujęciu składnik $\alpha_j e^{ikl_3 x_3}$ wyrażenia (24) można uważać za amplitudę fali zaś składnik $e^{ik(l_1 x_1 + l_2 x_2 - vt)}$ za wielkość opisującą propagację fali. Oznacza to, że wektor falowy jest zawsze równoległy do powierzchni, front fali do niej prostopadły, a amplituda przemieszczeń zależna od x_3 .

Po wstawieniu próbnego rozwiązania do równania falowego otrzymuje się układ równań identyczny jak w przypadku płaskich fal objętościowych w ośrodku nieograniczonym. Układ ten jak już wiadomo prowadzi do układu równań Christoffela-Greena i równania charakterystycznego:

$$|\Gamma_{ij} - \delta_{ij} \rho v^2| = 0. \quad (26)$$

Równanie (26) można rozpatrywać jako równanie trzeciego stopnia względem v_2 z parametrem l_3 lub, w przypadku rozważania fal powierzchniowych, jako równanie szóstego stopnia względem l_3 z parametrem v . Wtedy dla każdej prędkości v dowolne rozwiązanie układu (26) względem l_3 daje równanie o postaci (24). Rozwiązanie takie spełnia równanie falowe dla ośrodka anizotropowego i opisuje falę rozprzestrzeniającą się w kierunku $n_1 l_1 + n_2 l_2$ (n_i jest wersorem x_i) z prędkością fazową v .

W przypadku ogólnym, dla każdej wartości v równanie charakterystyczne (26) ma trzy pary pierwiastków zespolonych sprzężonych.

W celu uzyskania rozwiązań jednoznacznych przekształćmy liniową kombinację (24) w ten sposób aby składała się ona z trzech członów opisujących fale o jednakowych prędkościach fazowych ale różnych wartościach l_3 równych pierwiastkom równania charakterystycznego. Otrzymamy wtedy:

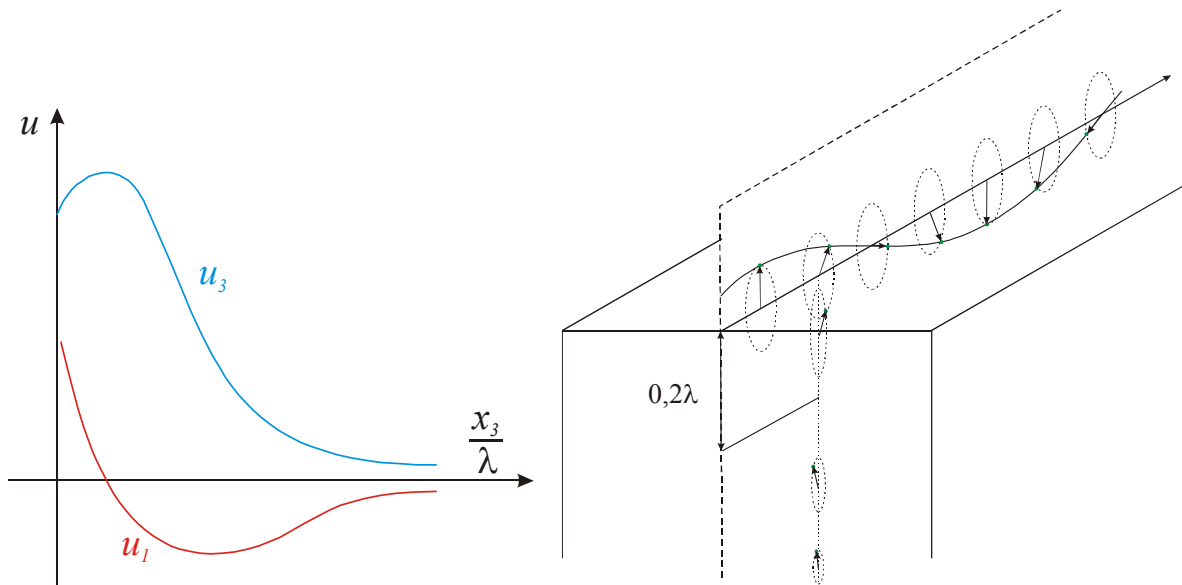
$$u_j = \sum_{n=1}^3 C_n \alpha_j^{(n)} e^{[ik(l_1 x_1 + l_2 x_2 + l_3^{(n)} x_3 - vt)]}. \quad (27)$$

Równanie to będzie spełniać równanie falowe jeśli tylko $\alpha_j^{(n)}$ będzie składową wektora własnego równania Christoffela-Greena odpowiadającą pierwiastkowi $l_3^{(n)}$.

Współczynniki C_n określić można przyjmując warunek brzegowy opisujący znikanie naprężeń na powierzchni:

$$\sigma_{3j} \Big|_{x_3=0} = 0. \quad (28)$$

Wstawiając (27) do warunku (28) otrzymuje się układ trzech równań jednorodnych. Przyrównanie do zera wyznacznika tego układu daje równanie zawierające w niejawny sposób prędkość fali Rayleigha i prędkości fal objętościowych. Znalezienie prędkości pozwala też na określenie współczynników C_n oraz przemieszczeń fali Rayleigha. W przypadku izotropowym fala ta ma tylko dwie składowe przemieszczeń u_1 i u_3 ($u_2=0$), co pokazano na rys. 2.

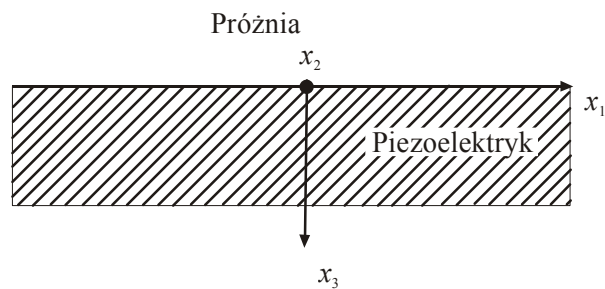


Rys. 2. Składowe przemieszczeń fali Rayleigha. Ich amplituda szybko maleje i w odległości od powierzchni rzędu długości fali można przyjąć, że praktycznie jest ona równa zero

Rysunek ten dotyczy idealnej powierzchni sprężystej. W ogólności na powierzchniach z różnego typu defektami może się pojawić składowa poprzeczna u_2 fali Rayleigha

2.1. Fala Rayleigha w piezoelektrykach

Zagadnienie rozpatruje się zakładając, że akustyczna fala powierzchniowa rozchodzi się bez tłumienia wzdłuż swobodnej płaszczyzny ($x_3=0$) jednorodnego kryształu, w kierunku osi x_1 (rys. 1).



Rys. 1. Analizowany układ piezoelektryk – próżnia

Swobodna płaszczyzna kryształu graniczy z próżnią i może być pokryta nieskończenie cienką warstwą idealnego przewodnika. Osie falowego układu współrzędnych (x_1, x_2, x_3) są obrócone względem osi układu współrzędnych krystalograficznych (X, Y, Z) o dowolny kąt.

Równania różniczkowe cząstkowe dla składowych przemieszczeń mechanicznych u_j i potencjału φ w obróconym układzie współrzędnych będą miały postać następującą:

$$c'_{ijkl} u_{k,li} + e'_{kij} \varphi_{,ki} = \rho \ddot{u}_j, \quad (1a)$$

$$e'_{ikl} u_{k,li} - \epsilon'_{ik} \varphi_{,ki} = 0, \quad (1b)$$

gdzie c'_{ijkl} , e'_{kij} i ϵ'_{ik} są tensorami w obroconym układzie współrzędnych.

Zakłada się, że u_j i φ nie zależą od współrzędnej x_2 , a fala powierzchniowa ma harmoniczną zależność od czasu ($e^{j\omega t}$). Powyższe równania przyjmą wtedy postać:

$$\begin{aligned} L_{11}u_1 + L_{12}u_2 + L_{13}u_3 + L_{14}\varphi &= 0, \\ L_{21}u_1 + L_{22}u_2 + L_{23}u_3 + L_{24}\varphi &= 0, \end{aligned} \quad (2)$$

$$L_{31}u_1 + L_{32}u_2 + L_{33}u_3 + L_{34}\varphi = 0,$$

$$L_{41}u_1 + L_{42}u_2 + L_{43}u_3 + L_{44}\varphi = 0,$$

gdzie L_{ij} są równaniami różniczkowymi.

Jeśli wszystkie elementy macierzy operatorowej L_{ij} , utworzonej ze współczynników występujących w równaniach (2), są różne od zera, to akustyczna fala powierzchniowa będzie miała trzy składowe przemieszczeń mechanicznych u_1 , u_2 , u_3 i potencjał elektryczny φ . W celu uzyskania jawnych wyrażeń na elementy macierzy L_{ij} , zakłada się, że fala powierzchniowa rozchodzi się wzdłuż osi x_1 i zanika wykładniczo wzdłuż osi x_3 . Poszukuje się więc rozwiązań na u_j i φ w postaci:

$$u_j = \beta_j e^{-\alpha k x_3} e^{i(\omega t - k x_1)},$$

$$\varphi = \beta_4 e^{-\alpha k x_3} e^{i(\omega t - k x_1)}, \quad j = 1, 2, 3. \quad (3)$$

Po obliczeniu elementów macierzy L_{ij} i powierzchni wstawieniu wyrażeń (3) do układu równań (2) otrzymuje się:

$$\sum_{j=1}^4 A_{ij} \beta_j = 0, \quad (4)$$

gdzie:

$$\begin{aligned} A_{11} &= c'_{55} \alpha^2 + 2c'_{15} i\alpha - c'_{11} + \rho V^2, \\ A_{12} = A_{21} &= c'_{45} \alpha^2 + (c'_{14} + c'_{56}) i\alpha - c'_{16}, \\ A_{13} = A_{31} &= c'_{35} \alpha^2 + (c'_{13} + c'_{55}) i\alpha - c'_{15}, \\ A_{14} = A_{41} &= e'_{35} \alpha^2 + (e'_{15} + e'_{31}) i\alpha - e'_{11}, \\ A_{22} &= c'_{44} \alpha^2 + 2c'_{46} i\alpha - c'_{66} + \rho V^2, \\ A_{23} = A_{32} &= c'_{34} \alpha^2 + (c'_{36} + c'_{45}) i\alpha - c'_{56}, \\ A_{24} = A_{42} &= e'_{34} \alpha^2 + (e'_{14} + e'_{36}) i\alpha - e'_{16}, \\ A_{33} &= c'_{33} \alpha^2 + 2c'_{35} i\alpha - c'_{55} + \rho V^2, \\ A_{34} = A_{43} &= e'_{33} \alpha^2 + (e'_{13} + e'_{35}) i\alpha - e'_{15}, \\ A_{44} &= \epsilon'_{33} \alpha^2 + 2\epsilon'_{13} i\alpha - \epsilon'_{11}, \end{aligned} \quad (5)$$

V – jest prędkością fali powierzchniowej.

Rozwiązanie układu równań (4) istnieje wtedy, gdy wyznacznik A_{ij} stanie się równy zeru:

$$|A_{ij}| = 0.$$

Po rozwinięciu tego wyznacznika, otrzymuje się:

$$A_8\alpha^8 + iA_7\alpha^7 + A_6\alpha^6 + iA_5\alpha^5 + A_4\alpha^4 + iA_3\alpha^3 + A_2\alpha^2 + iA_1\alpha + A_0 = 0,$$

gdzie współczynniki A_n dla $n = 0, 1, 2, \dots, 8$ są liczbami rzeczywistymi.

Jeżeli przyjąć za niewiadomą $i\alpha$ zamiast α , to współczynniki wielomianu będą również liczbami rzeczywistymi. Wówczas pierwiastki $i\alpha$ będą liczbami rzeczywistymi lub będą występować w postaci par zespolonych sprzężonych:

$$j\alpha_1 = a + ib,$$

$$j\alpha_2 = a - ib,$$

gdzie:

$$\alpha_1 = b - ia,$$

$$\alpha_2 = -b - ia.$$

Pierwiastki α występują parami i mają dodatnie i ujemne części rzeczywiste lub mogą być liczbami czysto urojonymi.

Wybierając cztery pierwiastki z dodatnimi częściami rzeczywistymi (warunek zanikania fali wzdłuż osi x_3), można obliczyć stałe β_j (z dokładnością do stałej) dla każdej wartości α . Rozwiązania ogólne na u_j i φ w obszarze $x_3 \geq 0$ można przedstawić w formie liniowej kombinacji wyrażeń (3) przy dozwolonych czterech wartościach współczynnika $\alpha^{(l)}$

$$u_j = \sum_{l=1}^4 C_l \beta_j^{(l)} e^{-\alpha^{(l)} k x_3} e^{i(\omega t - k x_1)}, \quad (6)$$

$$\varphi_j = \sum_{l=1}^4 C_l \beta_4^{(l)} e^{-\alpha^{(l)} k x_3} e^{i(\omega t - k x_1)}. \quad (7)$$

Każda składowa przemieszczeń mechanicznych u_j i potencjał φ fali powierzchniowej są więc sumą czterech tzw. fal cząstkowych.

W próżni potencjał elektryczny φ musi spełniać równanie Laplace'a:

$$\nabla^2 \varphi = 0. \quad (8)$$

Rozwiązanie tego równania, w przyjętym układzie współrzędnych będzie miało postać:

$$\varphi = \varphi_0 e^{k x_3} e^{i(\omega t - k x_1)}. \quad (9)$$

Powyższe rozwiązanie spełnia warunek zanikania potencjału φ , gdy $x_3 \rightarrow -\infty$.

2.1.1. Warunki brzegowe

Dla przypadku powierzchni swobodnej nie występują naprężenia mechaniczne skierowane wzdłuż osi x_3 dla $x_3 = 0^+$, czyli $\sigma_{31} = \sigma_{32} = \sigma_{33} = 0$ lub ogólnie:

$$\sigma_{3j} \Big|_{x_3=0^+} = 0.$$

Po wstawieniu odpowiedniego wyrażenia na σ_{3j} z równań konstytutywnych otrzymuje się:

$$\left[c'_{3jkl} u_{k,l} + e'_{k3j} \varphi_{,k} \right]_{x_3=0^+} = 0. \quad (10)$$

Podobnie, z warunku ciągłości składowej normalnej indukcji elektrycznej:

$$D_3 \Big|_{x_3=0^+} = D_3 \Big|_{x_3=0^-},$$

uzyskuje się:

$$\left[e'_{3kl} u_{k,l} + \epsilon'_{3k} \varphi_{,k} \right]_{x_3=0^+} = \left[-\epsilon \varphi_{,3} \right]_{x_3=0}. \quad (11)$$

Ostatni warunek dotyczy ciągłości potencjału dla $x_3 = 0$ i ma postać:

$$\varphi \Big|_{x_3=0^+} = \varphi \Big|_{x_3=0^-}. \quad (12)$$

W przypadku powierzchni metalizowanej powinien być spełniony warunek (10) oraz

$$\varphi \Big|_{x_3=0} = 0. \quad (13)$$

2.1.2. Rozwiązanie ogólne

Z warunku (12) oraz uwzględniając wyrażenia (7) i (9) otrzymuje się następujące wyrażenie na potencjał elektryczny φ w próżni:

$$\varphi = \sum_{l=1}^4 C_l \beta_4^{(l)} e^{kx_3} e^{i(\omega t - kx_1)}. \quad (14)$$

Wstawiając wyrażenia (6), (7) i (14) do warunków brzegowych (10) i (11), otrzymuje się następujący układ równań liniowych dla współczynników C_l przy powierzchni swobodnej:

$$\sum_{l=1}^4 \left[\beta_1^{(l)} (ic'_{15} + \alpha^{(l)} c'_{55}) + \beta_2^{(l)} (ic'_{56} + \alpha^{(l)} c'_{45}) + \beta_3^{(l)} (ic'_{55} + \alpha^{(l)} c'_{35}) + \right. \\ \left. + \beta_4^{(l)} (ie'_{15} + \alpha^{(l)} e'_{35}) \right] C_l = 0, \quad (15)$$

$$\sum_{l=1}^4 \left[\beta_1^{(l)} (ic'_{14} + \alpha^{(l)} c'_{45}) + \beta_2^{(l)} (ic'_{46} + \alpha^{(l)} c'_{44}) + \beta_3^{(l)} (ic'_{45} + \alpha^{(l)} c'_{34}) + \right. \\ \left. + \beta_4^{(l)} (ie'_{14} + \alpha^{(l)} e'_{34}) \right] C_l = 0, \quad (16)$$

$$\sum_{l=1}^4 \left[\beta_1^{(l)} (ic'_{13} + \alpha^{(l)} c'_{35}) + \beta_2^{(l)} (ic'_{36} + \alpha^{(l)} c'_{34}) + \beta_3^{(l)} (ic'_{35} + \alpha^{(l)} c'_{33}) + \right. \\ \left. + \beta_4^{(l)} (ie'_{13} + \alpha^{(l)} e'_{33}) \right] C_l = 0, \quad (17)$$

$$\sum_{l=1}^4 \left[\beta_1^{(l)} (ie'_{31} + \alpha^{(l)} e'_{35}) + \beta_2^{(l)} (ie'_{36} + \alpha^{(l)} e'_{34}) + \beta_3^{(l)} (ie'_{35} + \alpha^{(l)} e'_{33}) + \right. \\ \left. + \beta_4^{(l)} (i \epsilon'_{13} + \alpha^{(l)} \epsilon'_{33} + \epsilon_0) \right] C_l = 0. \quad (18)$$

W przypadku powierzchni metalizowanej muszą być spełnione warunki (10) i (13). Tak więc równania (15) ÷ (17) pozostaną bez zmiany, natomiast zamiast równania (18) otrzymuje się:

$$\sum_{l=1}^4 \beta_4^{(l)} C_l = 0. \quad (19)$$

Rozwiązanie problemu istnieje wówczas, gdy wyznacznik powyższego układu równań stanie się równy zero. Z warunku zerowania się wyznacznika układu równań (15) ÷ (18) można obliczyć prędkość fazową fali powierzchniowej przy powierzchni swobodnej V_f , a z warunku zerowania się wyznacznika układu równań (15) ÷ (17) i (19) można wyznaczyć prędkość fali powierzchniowej przy powierzchni metalizowanej V_m . Obliczenia przeprowadza się w ten sposób, że oblicza się wartość wyznacznika w funkcji zadawanej prędkości V . Po zlokalizowaniu zakresu V , dla którego wyznacznik przypuszczalnie zeruje się, obliczenia przeprowadza się metodą kolejnych przybliżeń tak długo, aż wartość wyznacznika stanie się równa zero (w granicach dokładności).

Po określeniu prędkości V_f i V_m , można obliczyć współczynniki C_l (z dokładnością do stałej), a następnie amplitudy wszystkich interesujących wielkości fizycznych, takich jak potencjał elektryczny φ , przemieszczenia mechaniczne u_j , natężenie pola elektrycznego E_j oraz składowe tensorów odkształceń ϵ_{ij} i naprężeń σ_{ij} .

2.1.3. Współczynnik sprzężenia piezoelektrycznego

Zmiana prędkości akustycznych fal powierzchniowych przy przejściu od powierzchni swobodnej do metalizowanej jest miarą sprzężenia piezoelektrycznego. Względna zmiana prędkości wynosi:

$$\frac{\Delta V}{V_f} = \frac{V_f - V_m}{V_f}. \quad (20)$$

Parametr ten pozwala na bezpośrednie obliczenie współczynnika sprzężenia piezoelektrycznego k_p^2 dla fal powierzchniowych:

$$k_p^2 = 2(1 + \epsilon_0 / \epsilon_p^T) \frac{\Delta V}{V_f} \left(1 - \frac{\Delta V}{V_f} \right), \quad (21)$$

gdzie $\epsilon_p^T = \sqrt{\epsilon'_{11} \epsilon'_{33} - \epsilon'^2_{13}}$, a ϵ'_{11} , ϵ'_{33} i ϵ'_{13} są składowymi tensora stałych dielektrycznych w układzie współrzędnych falowych. Ponieważ zwykle $\epsilon_0 / \epsilon_p^T \ll 1$ i $\frac{\Delta V}{V_f} \ll 1$, to:

$$k_p^2 \approx 2 \frac{\Delta V}{V_f}. \quad (22)$$

2.1.4. Temperaturowy współczynnik czasu opóźnienia

Czas opóźnienia τ fali między punktami położonymi na osi x_1 w odległości l wynosi:

$$\tau = \frac{l}{V}, \quad (23)$$

czyli:

$$\frac{1}{\tau} \frac{\partial \tau}{\partial T} = \frac{1}{l} \frac{\partial l}{\partial T} - \frac{1}{V} \frac{\partial V}{\partial T}, \quad (24)$$

lub

$$T_\tau = T_l - T_V, \quad (25)$$

gdzie: T_τ – temperaturowy współczynnik czasu opóźnienia,

T_l – współczynnik rozszerzalności termicznej w kierunku x_1 ,

T_V – temperaturowy współczynnik zmian prędkości.

Współczynnik T_V można obliczyć znając współczynniki temperaturowe wszystkich parametrów fizycznych wchodzących do równań. Tak więc, znając współczynnik rozszerzalności termicznej w kierunku propagacji fali T_l oraz T_V możliwe jest obliczenie temperaturowego współczynnika czasu opóźnienia T_τ .

2.1.5. Kąt przepływu mocy

Ważnym parametrem charakteryzującym własności fal powierzchniowych jest kąt między kierunkiem propagacji fali a kierunkiem przepływu mocy lub między kierunkiem prędkości fazowej i grupowej. Składowe uśrednionego w czasie wektora Poyntinga dla ośrodka piezoelektrycznego można zapisać w następującej postaci:

$$P_i = -\frac{1}{2} \operatorname{Re}[\sigma_{ij} \dot{u}_j^* - \phi \dot{D}_i^*]. \quad (26)$$

Ponieważ gęstość mocy fali powierzchniowej jest rozłożona niejednorodnie w głąb kryształu, to całkując powyższą zależność wzdłuż kierunku x_3 , możemy obliczyć moc przypadającą na jednostkę długości wzdłuż osi x_2 , czyli:

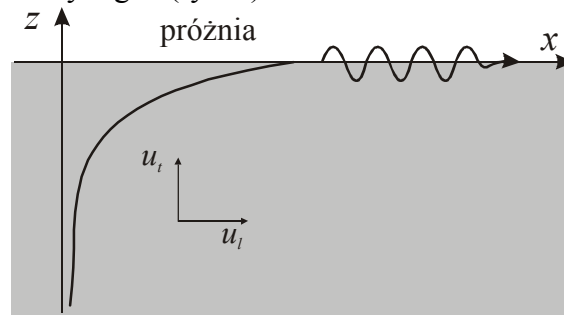
$$P'_i = -\frac{1}{2} \int_0^\infty \operatorname{Re}[\sigma_{ij} \dot{u}_j^* - \phi \dot{D}_i^*] dx_3. \quad (27)$$

W przypadku gdy nie ma odpromieniowania mocy w głąb kryształu, $P'_3 = 0$, a wyrażenie na kąt przepływu mocy Φ (w stosunku do osi x_1) ma postać:

$$\Phi = \operatorname{arctg} \frac{P'_2}{P'_1}. \quad (28)$$

2.2. Wybrane rodzaje akustycznych fal powierzchniowych

Jak już stwierdzono na powierzchni swobodnej (idealnie gładkiej i bez wstępnych naprężeń) mogą propagować się fale Rayleigha (rys. 1).

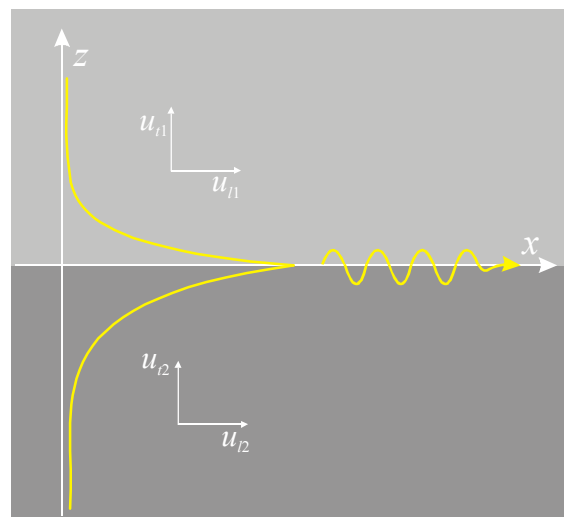


Rys. 1. Schemat propagacji fali Rayleigha

Jakakolwiek odchyłka od tego idealnego przypadku stwarza warunki propagacji innych rodzajów fal powierzchniowych. Ogólnie rzecz ujmując mianem fal powierzchniowych określa się fale, które propagują się na powierzchni ciała stałego graniczącego z próżnią, gazem, cieczą lub innym ciałem stałym i zanikają wraz z oddalaniem się od tej granicy. Generalnie wyróżnić można dwie grupy takich fal:

fale o polaryzacji pionowej (wertykalnej) zwane też falami typu relejowskiego, których wektor przemieszczeń leży na płaszczyźnie prostopadłej do płaszczyzny rozdziału, fale o polaryzacji poziomej (horyzontalnej), których wektor przemieszczeń leży na płaszczyźnie równoległej do płaszczyzny rozdziału.

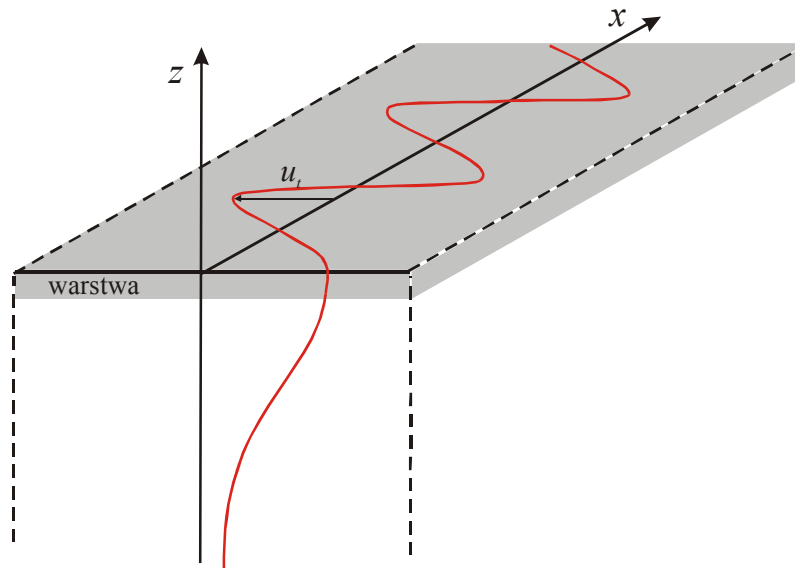
Typowym przykładem fali typu relejowskiego jest fala Stonleya (rys. 2).



Rys. 2. Schemat propagacji fali Stonleya

Ten typ fali propaguje się na idealnej granicy dwóch ciał stałych o zbliżonych własnościach sprężystych, a jej amplituda zanika w jednym i drugim ciele na odległości podobnej jak w przypadku fali Rayleigha. Z tego powodu falę Stonleya można uważać za złożenie dwóch fal Rayleigha. Prędkość fazowa fali Stonleya jest mniejsza od prędkości fal poprzecznych i podłużnych w obydwu ośrodkach.

Reprezentatywnym przykładem fal powierzchniowych o polaryzacji horyzontalnej jest fala Love'a (rys. 3).



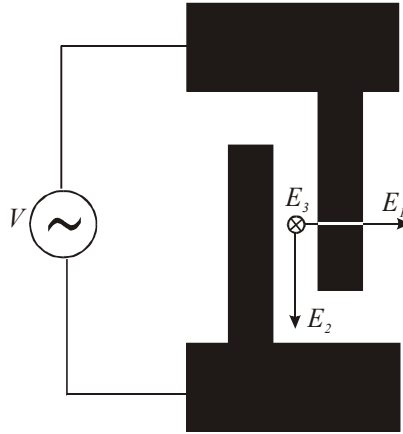
Rys. 3. Schemat propagacji fali Love'a. wektor przemieszczeń leży na płaszczyźnie równoległej do powierzchni podłoża i posiada tylko poprzeczną składową prostopadłą do kierunku propagacji

Jak widać na rys. 3. ten typ fal może propagować się przy powierzchni ciała stałego, na którą nałożona jest cienka warstwa. Wektor przemieszczeń posiada tylko jedną, poprzeczną składową, która jest prostopadła do kierunku propagacji. Przesunięcia w warstwie mają stałą wartość, natomiast zanikają w głąb podłoża. Szybkość zanikania zależy od grubości i sprężystych własności warstwy (lub warstw- może być ich wiele) oraz częstotliwości fali. Prędkość fazowa fali Love'a jest większa od prędkości fali poprzecznej w warstwie, a mniejsza od prędkości objętościowej fali poprzecznej w podłożu. Falam tym towarzyszy dyspersja. Dla dostatecznie cienkich warstw tzn. dla $\omega h/v_{tp} \rightarrow 0$, gdzie h – grubość warstwy, v_{tp} prędkość objętościowej fali poprzecznej w podłożu, prędkość fazowa fali Love'a zbliża się do prędkości objętościowej fali poprzecznej w podłożu. Jeśli natomiast $\omega h/v_{tp} \gg 1$ w warstwie może propagować się kilka odrębnych fal Love'a o nieznacznie różnych parametrach.

Istnienie konkretnej warstwy staje się zbędne przy bardzo wysokich częstotliwościach. W takim przypadku rolę warstwy pełnić mogą niejednorodności geometryczne lub materiałowe powierzchni. Niejednorodności takie można też wprowadzać celowo (np. za pomocą implantacji jonów) dla odpowiedniego wyprofilowania współczynników zanikania fal. Fale Love'a są ważne z technicznego punktu widzenia bowiem po pierwsze można je profilować poprzez dobór warstwy (lub odpowiednią modyfikację podłoża), a po drugie ich wektor przemieszczeń nie posiada składowej normalnej do powierzchni. Fala taka może więc np. propagować się na granicy z cieczą praktycznie bez oddawania do niej energii (przeciwnie zachowuje się fala Rayleigha).

Jak już wiadomo, każda modyfikacja parametrów ośrodków pociąga za sobą także modyfikację fali, a nawet konwersję jej rodzaju. Istnieją np. ośrodki piezoelektryczne, w których w pewnych kierunkach na powierzchni swobodnej (bez warstwy) mogą propagować się fale czysto poprzeczne (typu Love'a) nazywane falami Bleusteina-Gulajewa o prędkości zbliżonej do prędkości poprzecznych fal objętościowych (w praktyce różnica ta jest trudna do wychwycenia).

Opisane typu fal powierzchniowych wykorzystywane są w akustoelektronice i zazwyczaj generowane oraz detekowane przez tzw. przetwornik międzypalczasty, czyli leżący na podłożu piezoelektrycznym grzebieniowy układ metalowych elektrod. Jedną parę takich elektrod przedstawia rys. 4.



Rys. 4. Jedna para elektrod przetwornika międzypalczastego (widok z góry w kierunku prostopadłym do powierzchni). Przetwornik może zawierać nawet kilka tysięcy takich par

Sygnal napięciowy przyłożony do pary elektrod przetwornika, wytwarza pole elektryczne o składowych:

$$E_i = -\frac{\partial \varphi}{\partial x_i}, \quad i = 1 \dots 3$$

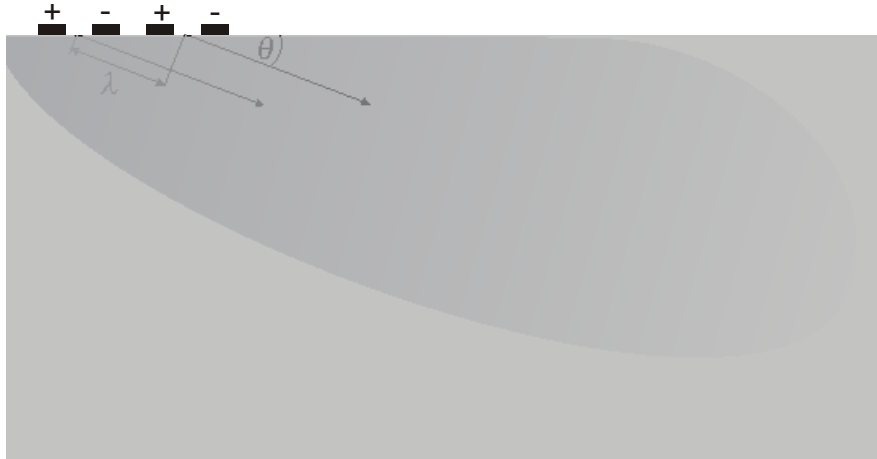
gdzie φ jest potencjałem elektrycznym.

Na skutek odwrotnego zjawiska piezoelektrycznego pole elektryczne wywołuje odkształcenie mechaniczne powierzchni. Jeśli sygnał jest harmoniczny to odkształcenia wywołane np. przez składową pola E_3 pochodzącą od jednej pary elektrod będą rozchodzić się po powierzchni piezoelektryka w postaci AFP (dla powierzchni swobodnej będzie to fala typu Rayleigha).

Cały zaś przetwornik emituje sygnał w postaci fali mechanicznej będący superpozycją zaburzeń generowanych przez wszystkie pary elektrod.

Jeśli np. odległość między sąsiednimi parami elektrod jest równa połowie długości fali, faza odkształcenia przy przejściu drogi o długości $\lambda/2$ zmieni się o π . Jednocześnie o tę samą wartość zmieni się faza sygnału wejściowego, co spowoduje sumowanie się odkształceń generowanych przez kolejne pary elektrod i liniowe narastanie amplitudy AFP.

Dzięki prostemu zjawisku piezoelektrycznemu, propagującym się odkształceniom mechanicznym towarzyszy ładunek elektryczny, który zbierany jest przez elektrody połączonego z obciążeniem przetwornika odbiorczego, zbudowanego podobnie jak przetwornik nadawczy. Teoria przetworników międzypalczastych zostanie przedstawiona w kolejnych rozdziałach. Zamieszczone tu skrótowe informacje są jednak konieczne do zrozumienia zasadniczych cech fal przypowierzchniowych typu SSBW (ang. surface skimming bulk waves – muskające powierzchnię fale objętościowe). Są to spolaryzowane horyzontalnie równoległe do powierzchni fale poprzeczne, zbliżone do fal Love'a, jednakże istniejące bez obecności na powierzchni warstwy i zanikające w swoisty sposób. Istnienie tych fal jest konsekwencją pracy przetwornika międzypalczastego, który można rozpatrywać jako ekwiwalent n elementowego szyku antenowego o elementach rozmieszczonych co pół fali i zasilanych przeciwfazowo (n – ilość elektrod przetwornika). Szyk taki emituje energię nie tylko w kierunku głównym (listku głównym), ale i w kierunkach (listkach) bocznych pod kątem $\theta \sim 1/\sqrt{n}$ (rys. 5).



Rys. 5. Generacja fali SSBW. Zaznaczono listek boczny emitowany przy powierzchni do objętości podłoża

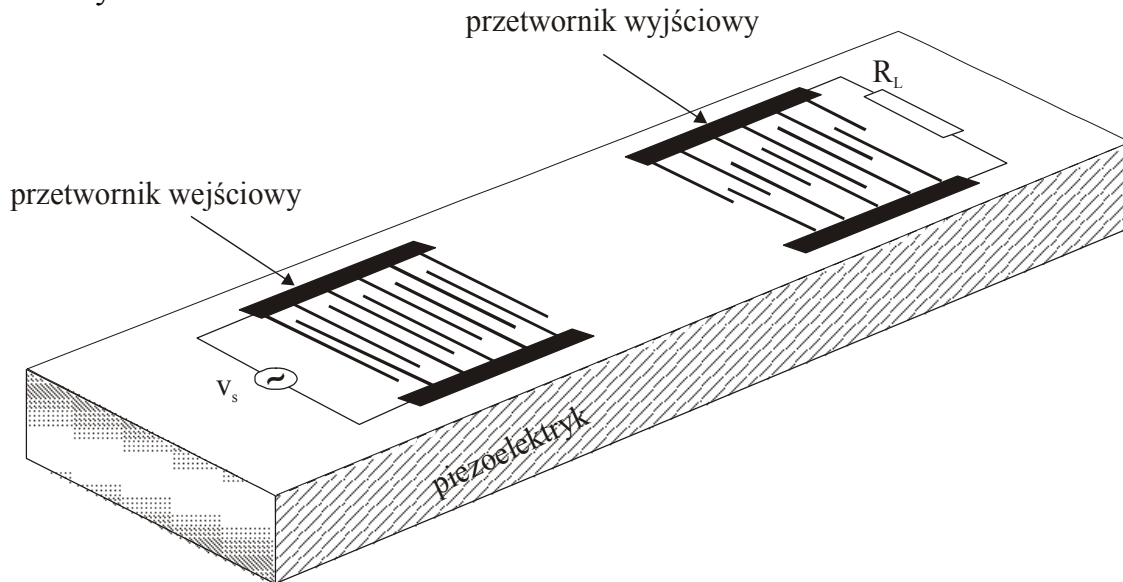
Fale te nabierają coraz większego znaczenia w akustoelektronice ze względu na swe unikalne cechy łączące zalety fal powierzchniowych z zaletami fal objętościowych. Charakteryzują się one dużo większą prędkością niż fale powierzchniowe, podlegają też znacznie mniejszemu tłumieniu a co za tym idzie mogą przenosić większe moce na większych częstotliwościach. SSBW są bardziej temperaturowo niż AFP i podobnie jak fale Love'a mało czułe na oddziaływania zewnętrzne.

Trzeba tu zaznaczyć, że rozdział ten nie wyczerpał wszystkich rodzajów fal związanych z powierzchnią ciała stałego, nakreślił jednakże istniejący obecnie, główny obszar zainteresowań akustoelektroników.

3. Synteza filtrów z AFP

3.1. Model funkcji δ

Klasyczny filtr z AFP składa się z układu dwóch przetworników międzypalczastych jak na rys. 1.



Rys 1. Układ przetworników międzypalczastych na powierzchni piezoelektryka. Przetwornik międzypalczasty jest układem elektrod usytuowanym na powierzchni piezoelektryka. Układ taki wykonuje się najczęściej techniką fotolitografii

Doprowadzenie do elektrod przetwornika zewnętrznego potencjału V powoduje powstanie na powierzchni piezoelektryka pola elektrycznego E_{\perp} . Na skutek istnienia odwrotnego efektu piezoelektrycznego z polem tym związane są naprężenia adekwatne do układu elektrod. Jeśli naprężenia te ustąpią, lub generowane są cyklicznie to rozchodzą się po powierzchni piezoelektryka w obie strony od przetwornika w postaci fali Rayleigha.

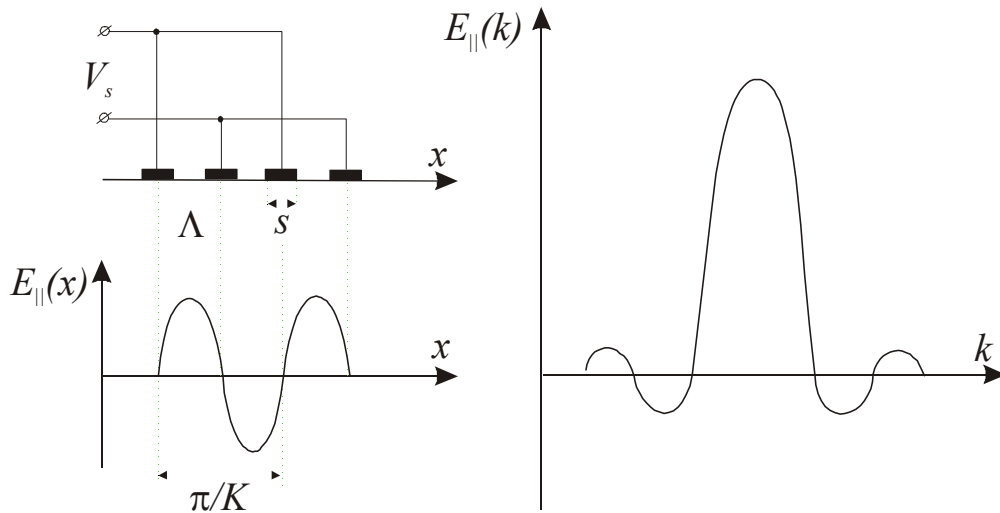
Propagująca się fala na skutek prostego efektu piezoelektrycznego stowarzyszona jest z ładunkiem elektrycznym. Generuje on, w trakcie przechodzenia fali pod drugim przetwornikiem, ładunki na jego elektrodach, które następnie zbierane są przez szyny zbiorcze i odprowadzane w postaci prądu do obciążenia.

Jeśli założymy się, że zjawiska charakterystyczne dla propagacji fali jak np. dyfrakcja czy dyspersja oraz same elektrody nie wpływają w istotny sposób na pracę urządzenia, to wtedy wektor falowy fali Rayleigha będzie wtedy określony: prędkość

$$k_R = \omega/v \quad (1)$$

gdzie: ω jest pulsacją napięcia przyłożonego na przetwornik, zaś v prędkością fali Rayleigha.

Pole elektryczne wymuszone przez potencjał elektrod można w przybliżeniu przedstawić jako odcinek sinusoidy, którego widmo ma postać przedstawioną na rys. 2.



Rys. 2. Dwie pary elektrod, generowany przez nie odcinek sinusoidy i fragment jego widma

Składowa widma dla $r = k_w$ odpowiada amplitudzie generowanej fali powierzchniowej. Ponieważ w funkcji częstotliwości sygnału widmo pola nie zmienia się, a zmienia się k_w , to amplituda fali powierzchniowej odwzorowuje to pole:

$$A(\omega) \approx E(\omega/v) \quad (2)$$

Przyłożenie do przetwornika impulsu napięciowego $V(\omega)=1$ spowoduje powstanie fali powierzchniowej w postaci:

$$A(x,t) \approx \int_{-\infty}^{\infty} E\left(\frac{\omega}{v}\right) e^{j\omega(t-x/v)} d\omega = E(x-vt) \quad (3)$$

Akustyczna odpowiedź przetwornika na impuls elektryczny jest odtworzeniem kształtu przetwornika.

Lepsze przedstawienie pracy przetwornika można uzyskać przykładając w obszarach położonych między elektrodami impulsy w postaci funkcji δ -Diraca, o amplitudach proporcjonalnych do różnicy napięć między sąsiadującymi elektrodami. Wtedy:

$$A(\omega) \approx \sum E_n e^{-j\omega x_n/v} \quad (4)$$

E_n są amplitudami impulsów przyłożonych między elektrodami w miejscach x_n . Przedstawione podejście nazywane jest modelem funkcji δ . Opisuje on zasadnicze własności przetwornika czyli tzw. efekty pierwszego rzędu.

3.2. Teoria widmowa

Teoria widmowa umożliwia analizę niektórych własności energetycznych przetwornika, o ile generowana fala Rayleigha jest płaska.

Generowana przez przetwornik fala niesie pewną energię:

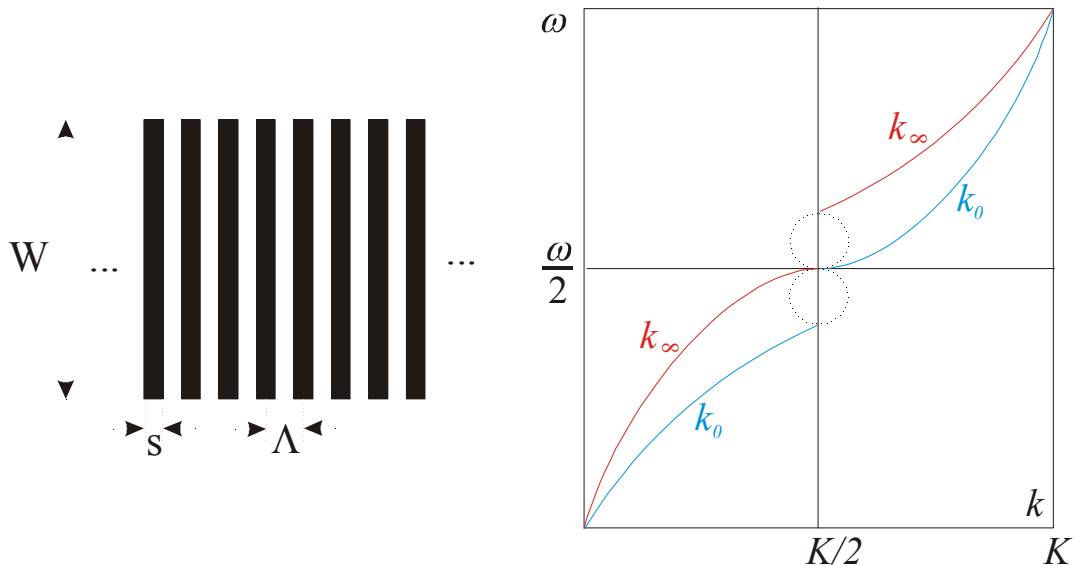
$$G(\omega) \approx |H(\omega)|^2, \quad (1)$$

gdzie G jest konduktancją promieniowania przetwornika, zaś

$$H(\omega) = A(\omega)/V(\omega) \quad (2)$$

jego charakterystyką amplitudową.

Istotne informacje o pracy przetwornika daje analiza zależności wektora falowego propagującej się pod elektrodami fali od częstotliwości.



Rys. 1. Periodyczny układ elektrod oraz przebieg części urojonej zależności wektora struktury $K=2\pi/\Lambda$ od częstotliwości. Rozróżniono tu dwa przypadki warunków elektrycznych tj. zwarcie (ozn. indeksem 0) i rozwarcie elektrod (ozn. indeksami odpowiednio 0 i ∞). Zmiana tych warunków prowadzi do zmiany prędkości fali

Dla przedstawionego na rys. 1 periodycznego układu elektrod:

$$K = \frac{2\pi}{\Lambda} > k_w = \frac{\omega}{v} \quad (3)$$

gdzie $k_w = \frac{k_0 + \varepsilon_r k_\infty}{1 + \varepsilon_r}$, a ε_r jest efektywną względną stałą dielektryczną.

Widoczna nieciągłość przebiegu części urojonej $K(\omega)$ (część rzeczywista jest stała i wynosi $K/2$) jest typowa dla propagacji fali w układzie periodycznie zaburzonym.

Fala powierzchniowa o wektorze falowym k , propagując się pod periodycznym układem elektrod, wymusza pole elektryczne. Z polem tym związana jest moc akustyczna, określona amplitudami fali postępującej A_0 oraz fali wstecznej A_{-1} . Zależności między tymi amplitudami dla $s = \Lambda/2$ ma postać:

$$A_{-1} = -\frac{k - \frac{k_0 + k_w}{2}}{K - k - \frac{k_0 + k_w}{2}} A_0 \quad (4)$$

Fala A_{-1} jest związana z określoną energią mechaniczną co wpływa na prędkość fali postępującej A_0 najsilniej w pobliżu $k - K \approx -k_0$. Nie wymusza ona fali wstecznej, jeżeli $k_0 - K \neq -k_0$, ponieważ wtedy wektor falowy pola wymuszającego nie pokrywa się z wektorem falowym fali wstecznej.

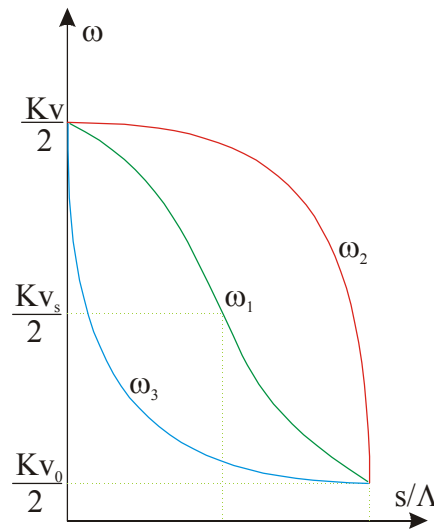
W przypadku synchronizmu fali A_0 z układem elektrod, tzn. dla $Re(k_0) = K/2$ wektor falowy pola A_{-1} jest równy $k_0 - K = -r - r_0^*$ jest wektorem fali wstecznej.

Część urojona k_0 jest odpowiedzialna za zanikanie fali postępującej kosztem generacji fali wstecznej. Podobna sytuacja ma miejsce przy rozwarciu elektrod.

W pobliżu synchronizmu relacje dyspersyjne mają postać:

$$\begin{aligned} k_0 &= \frac{K}{2} + \frac{1}{v} \sqrt{(\omega - \omega_1)(\omega - \omega_3)} \\ k_\infty &= \frac{K}{2} + \frac{1}{v} \sqrt{(\omega - \omega_1)(\omega - \omega_2)} \end{aligned} \quad (5)$$

Częstotliwości charakterystyczne $\omega_1, \omega_2, \omega_3$ w funkcji parametru s/Λ przedstawione są na rys. 2.



Rys. 2. Zależność częstotliwości charakterystycznych od parametru s/Λ

Różnica pomiędzy k_0 i k_∞ jest niewielka i wynosi:

$$k_0 - k_\infty = \frac{\omega}{v} \frac{\Delta v}{v} \frac{\sin \frac{\pi k}{K}}{\pi \frac{k}{K} P_{-k/K} \left(\cos \frac{\pi s}{A} \right) P_{-k/K} \left(-\cos \frac{\pi s}{A} \right)}, \quad (6)$$

gdzie: $\frac{\Delta v}{v} = \frac{1 - k_w}{k_0} \ll 1$, zaś $P_{-k/K} \left(\cos \frac{\pi s}{A} \right)$ jest funkcją Legendre'a o postaci:

$$P_{-k/K} \left(\cos \frac{\pi S}{A} \right) = \frac{1}{\pi} \int_0^\pi \left(\cos \frac{\pi S}{A} \pm \cos \varphi \sqrt{\left(\cos \frac{\pi S}{A} \right)^2 - 1} \right)^{-k/K} d\varphi.$$

Poza synchronizmem można przyjąć z dobrym przybliżeniem równość tych wielkości:

$$k_0 \approx k_\infty \approx \frac{\omega}{v} = k \quad (7)$$

Wektory falowe k_0 i k_∞ mogą być zespolone nie tylko w warunkach synchronizmu. Jeżeli $|k_0 - K| < k_\infty$ wzbudzone są fale objętościowe zasilane energią fali postępującej. Wynikające stąd tłumienie generuje część urojoną wektora falowego.

Przedstawione zależności mają zastosowanie w skrajnych sytuacjach: zwarcia i rozwarca elektrycznego elektrod. Oczywiście wektor falowy ulega również zmianie w przypadku obciążenia każdej elektrody skończoną admitancją Y_k . Wtedy zależność dyspersyjną uzyskuje się przyjmując równość admitancji, a co za tym idzie także prądów i potencjałów elektrod:

$$\frac{J}{V} = Y_k = j2\omega\varepsilon W \frac{k - k_\infty}{k - k_0} \frac{k + k_\infty - K}{k + k_0 - K} \frac{P_{-k/K} \left(\cos \frac{\pi S}{\Lambda} \right)}{P_{-k/K} \left(-\cos \frac{\pi S}{\Lambda} \right)} \sin \frac{\pi k}{K} \quad (8)$$

Transmitancję przetwornika można wyznaczyć przy założeniu $K > k_0$ z przybliżonej zależności dla pary elektrod m -tej i n -tej:

$$Y_{mn} = \frac{J_n}{E_m} = y_{mn}^c + y_{mn}^R, \quad (9)$$

gdzie: $E_m = V_m - V_{m-1}$, J_n gęstością prądu dopływającego do n -tej elektrody, V_k potencjałem m -tej elektrody przy uziemieniu pozostałych.

Składowe transmitancji y_{mn}^c oraz y_{mn}^R wynikające odpowiednio z pojemności wzajemnych i faktu promieniowania fal powierzchniowych można obliczyć z przybliżonych zależności:

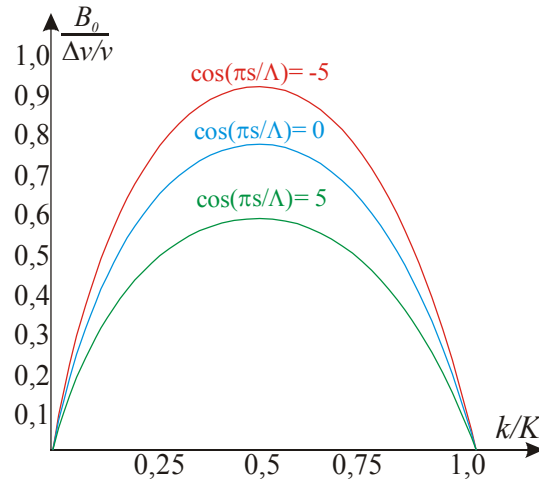
$$y_{mn}^c \approx \omega\varepsilon \int_0^K \frac{P_{-k/K} \left(\cos \frac{\pi S}{\Lambda} \right)}{P_{-k/K} \left(-\cos \frac{\pi S}{\Lambda} \right)} e^{-j \left(n - m + \frac{1}{2} \right) k \Lambda} dk / K, \quad (10)$$

$$y_{lk}^R = 2\pi\omega\varepsilon B_0 \sqrt{\frac{\omega - \omega_1}{\omega - \omega_3}} e^{-j \left| n - m + \frac{1}{2} \right| k_0 \Lambda} \text{sign} \left(n - m + \frac{1}{2} \right), \quad (11)$$

gdzie:

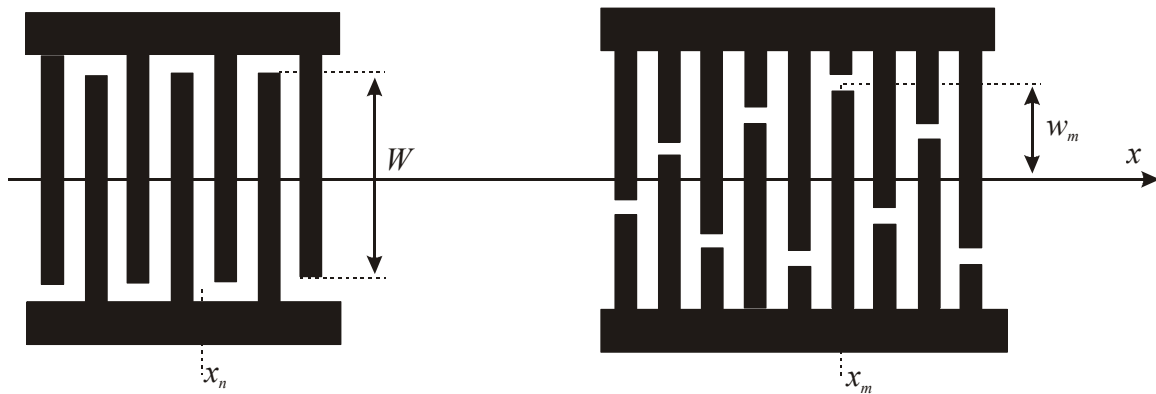
$$B_0 = \frac{\Delta v}{v\pi} \frac{\sin \pi k_0 / K}{\left[P_{-r_0/K} \left(-\cos \frac{\pi S}{\Lambda} \right) \right]^2} \quad (12)$$

nazywa się skutecznością promieniowania. W powyższych zależnościach zakłada się $\text{Im}(k_0) < 0$. Znormalizowaną zależność skuteczności promieniowania od geometrii układu elektrod przedstawia rys. 3.



Rys. 3. Znormalizowana zależność skuteczności promieniowania od geometrii struktury

Opisane relacje obowiązują w nieskończonym periodycznym układzie elektrod. Mogą być one jednak zastosowane dla konfiguracji dwóch oddzielnych przetworników, o ile tylko liczba elektrod jest wystarczająco duża i pracują one daleko od synchronizmu. Przykład takiego układu prezentuje rys. 4.



Rys. 4. Układ dwóch przetworników międzypalczastych

Przetwornik nadawczy posiada stałą aperturę i generuje falę płaską. Przetwornik odbiorczy może być ważony i ewentualnie dla wyrównania frontu fazowego uzupełniony przez tzw. elektrody ślepe, wychodzące z przeciwnej szyny zbiorczej.

Fala, generowana przez przetwornik nadawczy, indukuje w każdej elektrodzie przetwornika odbiorczego ładunek zbierany do szyny górnej lub dolnej. Przy zwarciu szyn tego przetwornika prąd od każdej elektrody będzie dawał prąd zwarcia proporcjonalny do amplitudy fali oraz różnicy długości elektrody i odpowiadającej jej elektrody ślepej. Definiując charakterystykę częstotliwościową H jako stosunek amplitudy prądu zwarcia przetwornika odbiorczego do amplitudy napięcia podawanego na przetwornik nadawczy otrzymuje się zależność:

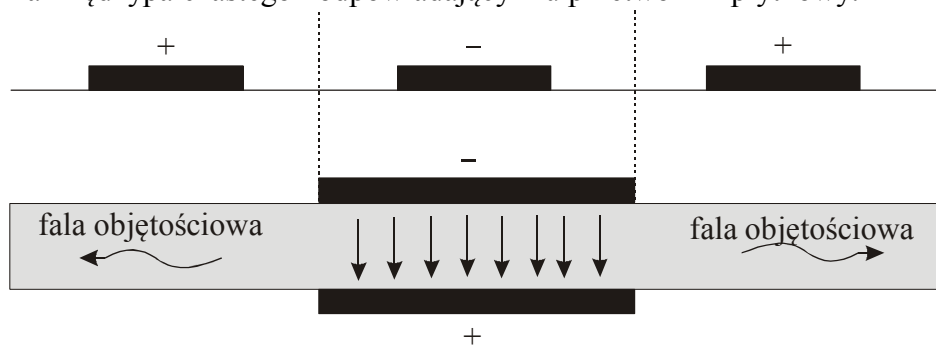
$$H(\omega) \sim \sum_{m,n} E_n e^{-j\omega \frac{x_n}{v}} w_m e^{j\omega \frac{x_n}{v}}, \quad (13)$$

E_n jest potencjałem n -tej elektrody.

Relacja (13) opisuje własności filtracyjne podzespołów z akustyczną falą powierzchniową.

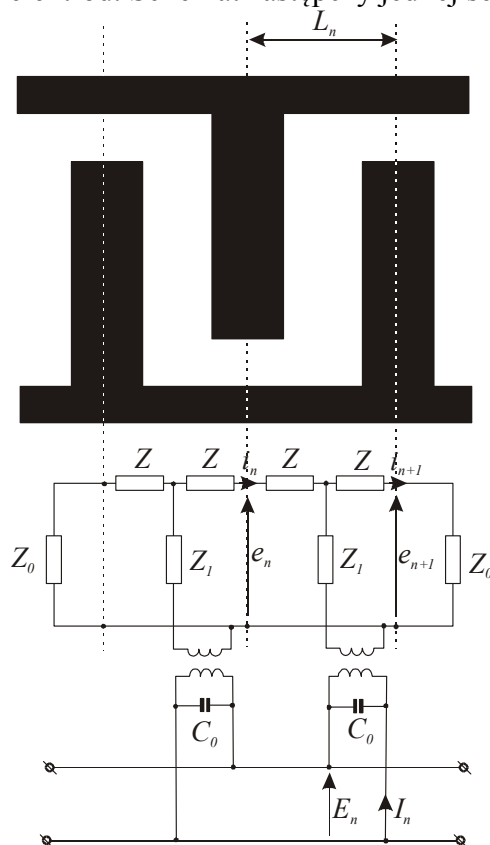
3.3. Model schematu zastępczego

Model schematu zastępczego wprowadzony został przez analogię do schematu Mezona stosowanego do obliczeń przetworników wzbudzających fale objętościowe w ośrodku monofalowym. Fala akustyczna generowana jest przez przetwornik na swej drodze propagacji analogicznie jak przy pobudzeniu fali objętościowej w pakiecie przetworników płytkowych. Analogia ta stała się podstawą zastąpienia pary elektrod jednym przetwornikiem płytkowym i jego elektrycznym schematem zastępczym. Rys. 7 przedstawia fragment przetwornika międzypalczastego i odpowiadający mu przetwornik płytkowy.



Rys. 7. Przetwornik międzypalczasty i analogiczny przetwornik płytkowy

W takim ujęciu analiza pracy przetwornika międzypalczastego sprowadza się do rozpatrzenia równoważnego schematu elektrycznego. Przetwornik dzieli się na sekcje, położone między środkami dwóch sąsiednich elektrod. Schemat zastępczy jednej sekcji przedstawia rys. 8.



Rys. 8. Jedna sekcja przetwornika i jej schemat zastępczy

Parametry elektryczne określone są następującymi zależnościami:

$$Z = jZ_0 \operatorname{tg} \frac{\alpha_n}{2}, \quad \alpha_n = \frac{\pi f}{2f_n}, \quad f_n = \frac{v}{2L_n}, \quad Z_1 = \frac{jZ_0}{\sin \alpha_n}, \quad D_n = jZ_0 i_n, \quad \alpha_n = \frac{\pi \omega}{\omega_0},$$

$$I_n = j\omega C_o E_n + \frac{j(D_n - D_{n+1})}{Z_o}, \quad Z_0 = \frac{2\pi}{\omega_0 C_o K_s^2}$$

jest impedancją falową ośrodka, zaś K_s^2 współczynnikiem sprzężenia elektromechanicznego oraz równaniem:

$$\begin{bmatrix} D_{n+1} \\ e_{n+1} \end{bmatrix} = \begin{bmatrix} \cos \alpha_n & \sin \alpha_n \\ -\sin \alpha_n & \cos \alpha_n \end{bmatrix} \begin{bmatrix} D_n \\ e_n \end{bmatrix} + E_n \begin{bmatrix} -\sin \alpha_n \\ 1 - \cos \alpha_n \end{bmatrix}.$$

Zaciski akustyczne poszczególnych sekcji połączone są kaskadowo, a elektryczne równolegle, z uwzględnieniem polaryzacji poszczególnych elektrod. Na zewnątrz przetwornika układ obciążony jest impedancją falową ośrodka Z_0 .

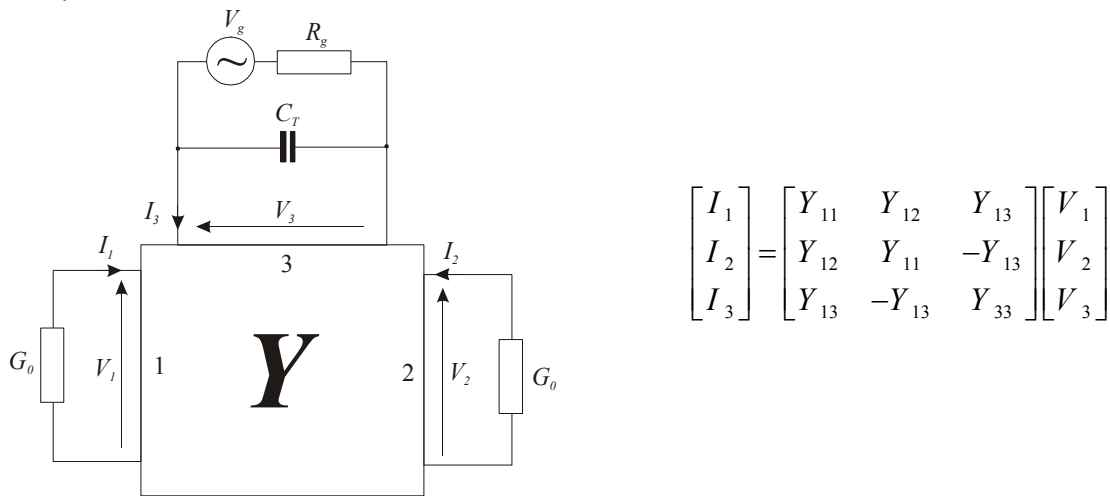
Model schematu zastępczego umożliwia obliczenie charakterystyki częstotliwościowej przetwornika oraz jego impedancji.

Możliwa jest jednak modyfikacja przedstawionego schematu zastępczego pozwalająca na uwzględnienie efektów dodatkowych występujących w pracy przetwornika tzw. efektów drugiego rzędu, do których należy zaliczyć:

- odbicia międzyelektrodowe,
- wpływ dyfrakcji fali na charakterystyki,
- generację fal objętościowych,
- wpływ rezystancji elektrod,
- zależność skuteczności promieniowania fali powierzchniowej od częstotliwości i geometrii przetwornika.

Na podstawie schematu zastępczego można także przeprowadzić analizę macierzową.

dwóch współpracujących ze sobą przetworników międzypalczastych połączonych ze źródłem sygnału o rezystancji wewnętrznej R_g i obciążeniem R_0 i tworzących filtr z AFP.



Rys. 9. Macierz admitancyjna dwóch współpracujących przetworników

Charakterystyka częstotliwościowa oraz admitancje wejścia i wyjścia jednoznacznie określają jego parametry układu dwóch współpracujących przetworników.

Zgodnie z teorią obwodów elektrycznych równania opisujące układ przedstawiony na rys. 9 mają postać:

$$\begin{bmatrix} I_N \\ I_0 \end{bmatrix} = [Y] \begin{bmatrix} U_N \\ U_0 \end{bmatrix}$$

$$\begin{aligned} E_g &= I_N R_g + U_N \\ U_0 &= -I_0 R_0 \end{aligned} \quad (18)$$

Niezależne składowe macierzy $[Y]$:

$$[Y] = \begin{bmatrix} y_{nn} & y_{no} \\ y_{on} & y_{oo} \end{bmatrix}$$

zdefiniowane są przez relacje:

$$\begin{aligned} y_{nn} &= \frac{I_N}{U_N} \Big|_{U_0=0} & y_{oo} &= \frac{I_0}{U_0} \Big|_{U_N=0} \\ y_{no} &= \frac{I_N}{U_0} \Big|_{U_N=0} & y_{on} &= \frac{I_0}{U_N} \Big|_{U_0=0} \\ & & y_{no} &= y_{on} \end{aligned}$$

Z rozwiązania układu równań (18) otrzymuje się charakterystykę częstotliwościową układu w postaci:

$$H(\omega) = \frac{U_0}{E_g} = \frac{y_{on} R_0}{(1 + y_{nn} R_g)(1 + y_{oo} R_0) - y_{on}^2 R_0 R_g} \quad (19)$$

Wprowadzona definicja admitancji wzajemnej y_{on} odpowiada charakterystyce układu dwóch współpracujących przetworników międzypalczastych określonej w modelu δ funkcji jako iloczyn widma struktury przetwornika nadawczego i odbiorczego.

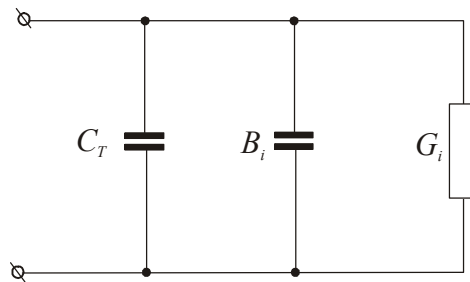
Charakterystyka częstotliwościowa układu bez uwzględnienia składnika $y_{on}^2 R_0 R_g$ opisującego tzw. sygnał trzeciego echa ma postać

$$H(\omega) = \frac{H_\delta(\omega) R_0}{(1 + y_{nn} R_g)(1 + y_{oo} R_0)} \quad (20)$$

Admitancja wejściowa i wyjściowa zależą od częstotliwości i są opisane za pomocą konduktancji i susceptancji:

$$y_{ii}(j\omega) = G_i(\omega) + jB_i(\omega) + j\omega C_T,$$

które można przedstawić na schemacie zastępczym:



Rys. 10. Schemat zastępczy do obliczania admitancji układu

C_T jest pojemnością statyczną przetwornika z elektrodami o długości W i jest obliczana jako suma ładunku elektrostatycznego Q na wszystkich elektrodach przetwornika międzypalczastego:

$$C_T = W \sum_{n=1}^N \sum_{m=1}^N b_n b_m Q_{m-n} \quad (21)$$

Susceptancja B jest związana z konduktancją G poprzez transformację Hilberta:

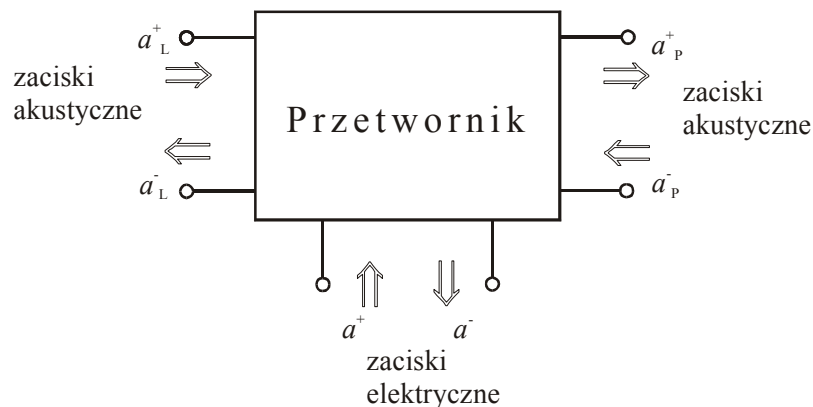
$$B_i(\omega) = -\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{G_i(\omega')}{\omega - \omega'} d\omega' \quad (22)$$

Straty wtrąceniowe filtra definiowane są jako iloraz mocy dostarczonej do obciążenia oraz mocy dysponowanej przez generator:

$$ST = -10 \log \frac{P_o}{P_g} = -10 \log \left(\frac{4R_g |H(\omega)|^2}{R_o} \right) \quad (23)$$

3.3.1. Zależności energetyczne

Informacji o energetycznych własnościach przetwornika międzypalczastego dostarcza jego macierz rozproszenia S . Wprowadza się ją rozważając przetwornik w postaci trójwrotnika jak na rys. 11.



Rys. 11. Rysunek pomocniczy do analizy własności energetycznych przetwornika za pomocą macierzy S

Układ posiada dwa wyjścia na akustyczne linie transmisyjne oraz jedno na elektryczną linię transmisyjną o impedancji falowej Z . Na rysunku zaznaczono amplitudy fal padających i odbitych, które dotyczą fal rozchodzących się w liniach transmisyjnych.

Z zależności transadmitancyjnych wynikają następujące relacje:

$$\begin{bmatrix} a_P^+ \\ a_L^- \\ a^- \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{12} & S_{11} & S_{13} \\ S_{13} & S_{13} & S_{33} \end{bmatrix} \begin{bmatrix} a_L^+ \\ a_P^- \\ a^+ \end{bmatrix} \quad (24)$$

gdzie:

$$S_{11} = 1 - \frac{ZG}{1 + ZY_T}, \quad S_{12} = \frac{ZG}{1 + ZY_T}, \quad S_{13} = 1 - \frac{\sqrt{2ZG}}{1 + ZY_T}, \quad S_{33} = 1 - \frac{1 - ZY_T}{1 + ZY_T},$$

Y_T jest admitancją przetwornika, zaś G jego konduktancją promieniowania.

Z analizy macierzy rozproszenia S_{ik} , wynika że przetwornik dwukierunkowy przekształca padającą moc elektryczną na falę elektryczną odbitą od przetwornika oraz dwie fale akustyczne o równych amplitudach, propagujące się w dwóch kierunkach. Maksymalna moc elektryczna jaka może być odebrana od przetwornika odbiorczego wynosi $P/4$, gdzie P oznacza moc wypromieniowaną przez przetwornik nadawczy. Stąd wynika, że minimalne straty wnoszone przez linię opóźniającą z dwoma dwukierunkowymi przetwornikami wynoszą co najmniej 6dB, o ile spełniony jest warunek elektrycznego dopasowania przetworników do linii elektrycznej. Moc wypromieniowana do elektrycznej linii transmisyjnej nawet przy dopasowaniu do tej linii jest nie większa niż $P/2$. Reszta mocy rozpraszana jest na falę przechodzącą pod przetwornikiem i na falę odbitą od niego.

Minimalne straty dla przetworników jednokierunkowych wynoszą 0 dB, ponieważ cała moc dostarczona do przetwornika jest przekształcana w falę powierzchniową przemieszczającą się w jednym kierunku. O przetwornikach jednokierunkowych będzie mowa w dalszej części tej pracy.

3.3.2. Dopasowanie energetyczne

Z zależności na S_{33} wynika, że znaczna część strat wnoszonych przez układ dwóch przetworników może pochodzić z niedopasowania przetworników do źródła i obciążenia o impedancji Z , ponieważ najczęściej $Z = 50 \Omega$, zaś Y_T ma dużą składową pojemnościową oraz najczęściej małą część rzeczywistą. Najprostszym sposobem dopasowania byłoby włączenie indukcyjności o wartości dobranej do rezonansu, z odpowiednimi pojemnościami statycznymi równoległe do przetworników. Wówczas źródło i obciążenie „widziałoby” tylko

konduktancję promieniowania. Najczęściej jest ona bardzo mała ($\frac{\Delta v}{v} N^2$ jest zwykle małe),

rezystancja byłaby więc dużo większa od stosowanych rezystancji źródeł i obciążeń.

Innym rodzajem dopasowania, zmniejszającym rezystancję wejściową przetworników jest dopasowanie polegające na zastosowaniu indukcyjności szeregowej L . Wynikowa rezystancja w rezonansie (tj. dla częstotliwości środkowej) dla małych konduktancji G w stosunku do susceptancji (ωC) przetworników jest wtedy równa:

$$R = \frac{L}{C} G \quad (25)$$

Najczęściej rezystancja R powinna wynosić 50Ω , należy więc tak dobrać wymiary przetwornika, aby zależność (25) spełniała ten warunek. Obwód rezonansowy utworzony z indukcyjności dopasowującej L i pojemności statycznej C przetwornika ma pewną dobroć „elektryczną” wynoszącą:

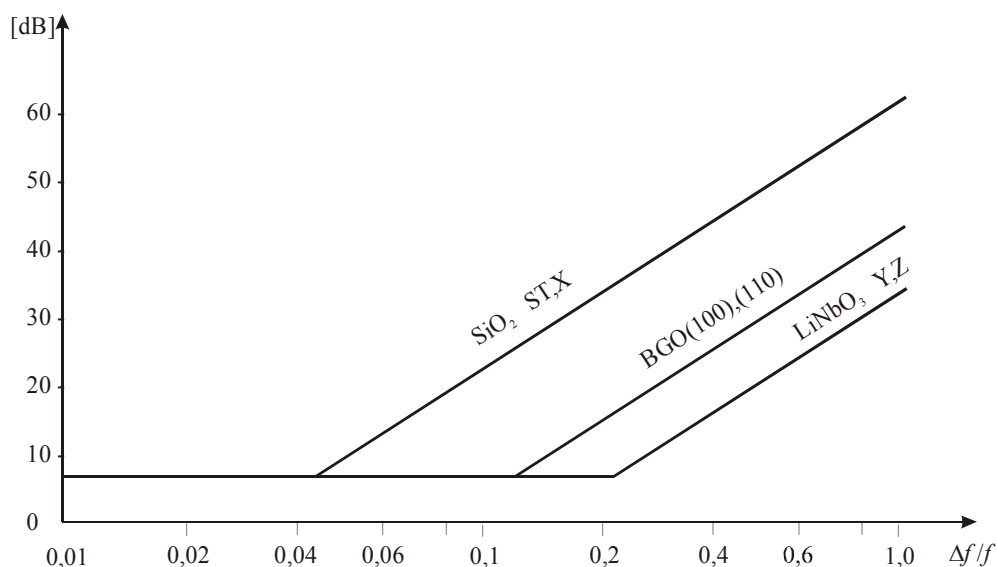
$$Q_e = \omega_o C / G. \quad (26)$$

Dobroć ta jest całkowicie określona (podobnie jak C i L) przez wymiary przetwornika.

Jako warunek dopuszczalnego ograniczenia pasma układu (obwód dopasowujący – przetwornik) przyjmuje się równość:

$$Q_e = Q_a = f_o / \Delta f \quad (27)$$

gdzie Δf jest pasmem „akustycznym” wynikającym z kształtu przetwornika. Rys. 12 przedstawia minimalne straty wnoszone w zależności od pasma względnego przetworników dopasowanych szeregową indukcyjnością do rezystancji 50Ω dla trzech podłoży.



Rys. 12. Minimalne straty wnoszone w zależności od pasma względnego przetworników dopasowanych szeregową indukcyjnością do rezystancji 50Ω

Dla przetwornika prostokątnego dobroć akustyczna równa jest liczbie par elektrod N .

W przypadku małych N , czyli dla przetworników szerokopasmowych, może okazać się niemożliwe spełnienie warunku (27) przy dopasowaniu do 50Ω . W takim niekorzystnym

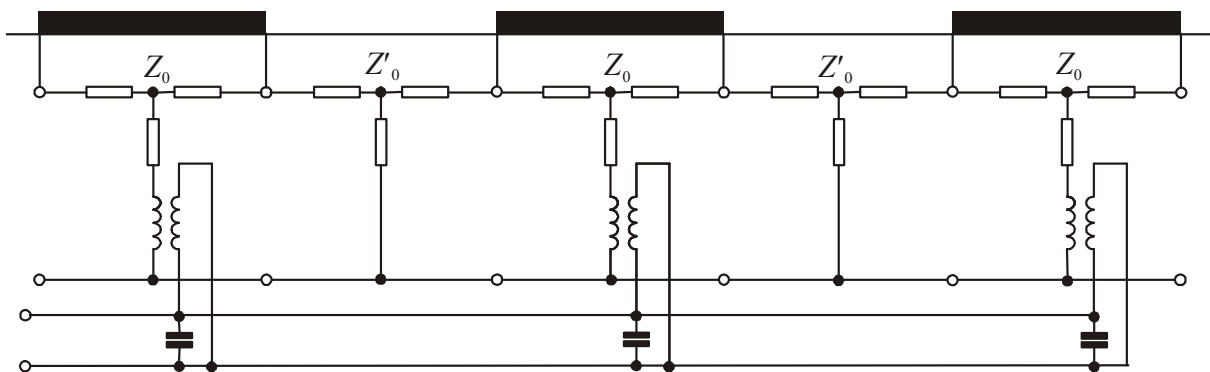
przypadku trzeba obniżyć dobroć elektryczną poprzez włączenie dodatkowej rezystancji, jednakże kosztem wzrostu strat. Możliwe są dopasowania szerokopasmowe przy zastosowaniu bardziej złożonych obwodów LC , zwłaszcza że wprowadzone przez nie zniekształcenia amplitudowo-fazowe można uwzględnić w charakterystyce projektowanych przetworników. Należy zwrócić uwagę, że wraz z maleniem wnoszonych strat rośnie sygnał potrójnego echa. W przypadku pracy układu w zakresie wyższych mikrofal dopasowania za pomocą elementów RLC zwykle nie udaje się zrealizować i trzeba w takim przypadku posługiwać się układami mikrofalowymi takimi jak transformatory ćwierćfalowe czy też stroiki.

3.4. Efekty drugiego rzędu

Współpracujące przetworniki międzypalczaste, tworzące filtr z AFP, nanoszone są na powierzchnię materiałów piezoelektrycznych, o ściśle określonych własnościach. Także same przetworniki posiadają szereg cech nie branych pod uwagę w procesie projektowania. Okoliczności te powoduje, że w praktyce charakterystyka częstotliwościowa wykonanego urządzenia jest zniekształcona wskutek pojawiania się sygnałów niepożądanych.

3.4.1. Odbicia międzyelektrodowe

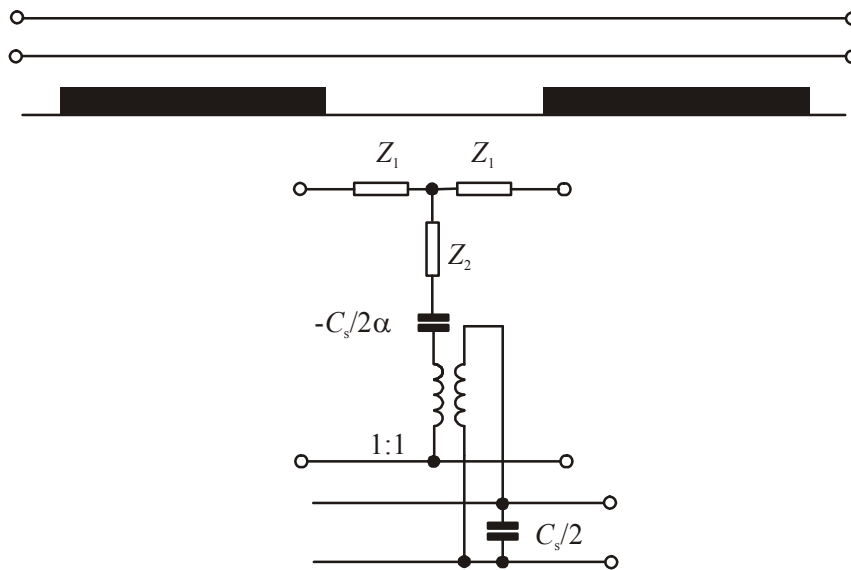
Odbicia międzyelektrodowe powstają w wyniku zaburzeń związanych z bezwładnością (masą) elektrod oraz ich własnościami sprężystymi. Zaburzenia te prowadzą do odbić fali powierzchniowej od elektrod przetworników. Za pomocą modelu układu zastępczego można dokonać przybliżonej analizy tych efektów poprzez wprowadzenie do schematu zastępczego segmentów o odmiennej impedancji falowej dla odcinków propagacji pod elektrodami i między elektrodami jak na rys. 13.



Rys. 13. Modyfikacja schematu zastępczego

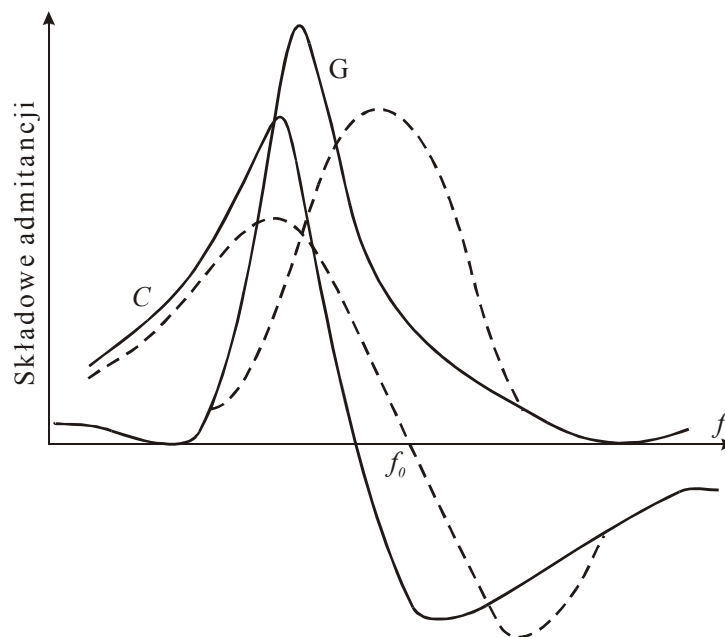
Względna różnica impedancji falowej zwykle dobierana jest eksperymentalnie. W przypadku niejednorodności tylko typu piezoelektrycznego (powodowane przez elektrody odcinkowe zwarcia pola elektrycznego) modyfikację schematu zastępczego można wprowadzić na

podstawie zależności dyspersyjnych. Taki schemat postępowania jest słuszny tylko w przypadku przetworników o elektrodach niedzielonych (o dzieleniu elektrod będzie mowa w dalszej części). Analiza zjawisk dyspersyjnych prowadzi do następującego schematu:



Rys. 14. Modyfikacja schematu zastępczego na podstawie zależności dyspersyjnych

Poniżej pokazano przykładowe zniekształcenia charakterystyk przetworników wprowadzane przez odbicia międzyelektrodowe.



Rys. 14. Spowodowane odbiciami międzyelektrodowymi zniekształcenie admittancej przetwornika prostokątnego o pojedynczych elektrodach

Eliminacji tego szkodliwego efektu można dokonać stosując dzielone elektrody. W przypadku pojedynczych elektrod fale odbite od każdej z nich interferują konstruktywnie, zaś dla dzielonych destruktywnie. Zanika wówczas pojemność określona przez parametr α i

otrzymuje się model tzw. „crossed field”, nie zawierający efektu odbić międzyelektrodowych.

Eliminacja wpływu odbić międzyelektrodowych na charakterystyki przetworników polega zatem na odejściu od warunków synchronizmu w paśmie pracy przetwornika

3.4.2. Efekt trzeciego echa (potrójnego przejścia)

Sygnał trzeciego echa (ang. TTS - triple transit signal), spowodowany jest odbiciem akustycznej fali powierzchniowej od przetwornika odbiorczego. W wyniku tego właściwy sygnał detekowany przez przetwornik odbiorczy zawiera echa, z których największą amplitudę posiada echo trzecie. Kolejne echa nieparzyste 5, 7 ... na skutek wydłużania się toru propagacji są coraz silniej tłumione. Z tego powodu zasadniczo tylko trzecie echo powoduje znaczące zniekształcenie okresowe charakterystyk amplitudowych i fazowych. Sygnał ten jest łatwo rozróżnialny, ponieważ częstotliwość periodycznych zniekształceń charakterystyki częstotliwościowej jest związana z czasem opóźnienia τ między przetwornikami międzypalczastymi przez zależność: $f_{TTS}=1/2\tau$.

Jedyną praktyczną metodą eliminacji ech wielokrotnych na charakterystykę filtru jest odpowiedni dobór jego konstrukcji. Ponieważ amplituda kolejnych ech jest coraz mniejsza, zatem problem ich eliminacji sprowadza się do ograniczenia wielkości trzeciego echa.

Jeżeli pasmo pracy przetwornika położone jest poniżej częstotliwości środkowej jego struktury trzecie echo powstaje jedynie w wyniku elektrycznego odbicia fali i można je wyeliminować przez wprowadzenie dodatkowego sygnału o tej samej amplitudzie lecz przeciwnej fazie. Sygnał ten docierając do przetwornika wzbudzającego falę powierzchniową jednocześnie z sygnałem trzeciego echa będzie wtedy interferował z nim destruktywnie. Ta specjalna konstrukcja polega na zastosowaniu dwóch przetworników odbiorczych położonych w jednakowej odległości od przetwornika nadawczego. Jeden z nich połączony jest z obciążeniem i pełni rolę typowego przetwornika odbiorczego. Zadaniem drugiego przetwornika odbiorczego jest odbicie fali o amplitudzie równej fali odbitej od właściwego przetwornika odbiorczego, lecz o przeciwnej fazie. Spełnienie tego warunku wymaga obciążenia go odpowiednio dobraną impedancją.

Możliwość wykorzystania tej metody ogranicza się do przetworników prostych, ponieważ przy konstrukcjach złożonych trudne jest uzyskanie odpowiednich relacji pomiędzy falami odbitymi w całym paśmie pracy filtru.

3.4.3. Inne efekty drugiego rzędu

Efektów drugiego rzędu jest bardzo wiele. Obok najważniejszych, opisanych już zjawisk występują też:

- sygnały pochodzące od odbić AFP od elektrod spowodowane istnieniem nieciągłości impedancji na powierzchni podłoża,
- sygnały związane z falami objętościowymi wzbudzonymi poprzez przetwornik międzypalczasty. Ze względu na większą prędkość fali objętościowej w porównaniu z prędkością fali powierzchniowej, znaczne zniekształcenia charakterystyki częstotliwościowej występują dla zakresu częstotliwości powyżej częstotliwości środkowej filtru,
- wpływ impedancji źródła, a przede wszystkim obciążenia na charakterystyki filtru,
- dyfrakcja AFP związana ze skończoną długością elektrod przetwornika,
- wzbudzanie przez przetwornik sygnałów na częstotliwościach harmonicznych (czasem wykorzystywanych w różnych konstrukcjach), które związane jest z geometrią elektrod i zależy od współczynnika metalizacji powierzchni.

Istnieją możliwości ograniczenia wpływu na pracę filtru przynajmniej dwóch zjawisk pasożytniczych tzn. generacji sygnałów odbitych oraz fal objętościowych.

Znajomość przebiegu efektywnej przenikalności powierzchniowej podłoża przy określonym cięciu i kierunku propagacji AFP pozwala na wybór parametrów przetworników tak, aby ich pasma pracy były położone poza obszarami silnego wzbudzenia fal objętościowych. Możliwe jest też odpowiednie przygotowanie spodu podłoża w taki sposób aby fale objętościowe ulegały silnemu rozpraszaniu i docierały do powierzchni silnie stłumione.

W przypadku technologii produkcji mikrofalowych układów z AFP wymienione zjawiska stanowią główną barierę ograniczającą możliwość poprawy ich parametrów.

3.5. Metody kształtowania przetworników

Jak już wiadomo charakterystyka częstotliwościowa przetwornika prostego odtwarza funkcję $\text{sinc}x$. Realizacja charakterystyki o określonym przebiegu wymaga zatem odpowiedniego kształtowania przetwornika (tzw. ważenia) tzn. odpowiedniego zaprojektowania jego geometrii. Istnieje cały szereg metod kształtowania przetworników. Do zasadniczych zalicza się zwykle:

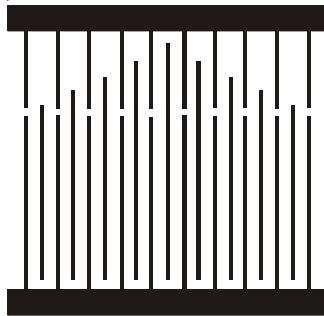
- zmianę długości elektrod (tzw. apodyzacja), ta metoda była już stosowana.
- zmianę szerokości elektrod i odstępów między nimi.

Największe możliwości kształtowania daje zmiana długości elektrod i odstępów między nimi. Szerokość elektrod można zmieniać w niewielkich granicach, zależnych od odległości między nimi. Metoda ta ponadto wymaga znacznego zwiększenia dokładności wykonywania procesów fotolitograficznych. W przypadku urządzeń mikrofalowych z AFP jest zatem skrajnie trudna do zrealizowania.

Przykłady innych metod ważenia

Metody pojemnościowe

Pierwsza z nich polega ona na zaprojektowaniu pojemnościowych dzielników napięcia dla każdej elektrody oddzielnie (rys. 1).

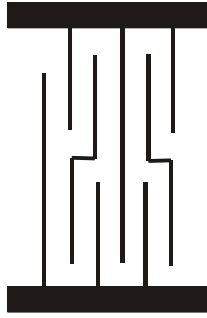


Rys. 1. Przetwornik ważony pojemnościowo

Jedną z pojemności dzielnika może być pojemność statyczna danej elektrody względem elektrod sąsiednich. W pierwszym przybliżeniu można uwzględnić wyłącznie podział napięcia na pojemnościach statycznych. W przypadku występowania dużych pojemności rzeczywiste napięcia mogą jednak znacznie odbiegać od przybliżenia.

Problem ten można rozwiązać zwiększając sztucznie pojemność wzajemną elektrod przez zrównoleglenie jej pojemnością. Charakterystyka przetwornika zależy wtedy tylko od potencjału elektrod. Wartości potencjałów elektrod można znaleźć stosując np. przedstawiony wcześniej model schematu zastępczego.

Druga metoda polega na wykonaniu dzielników pojemnościowych wewnątrz przetwornika (rys. 2).



Rys. 2. Elektrody schodkowe

Jeśli elektroda nie dołączona do żadnej szyny zbiorczej, ma „schodek” w połowie, to jej potencjał będzie równy średniej potencjałów elektrod sąsiednich. Napięcie między elektrodą schodkową, a elektrodami sąsiednimi będzie więc o połowę mniejsze niż w przetworniku prostym, przy zachowaniu kierunku spadku napięcia.

Pochylenie elektrod

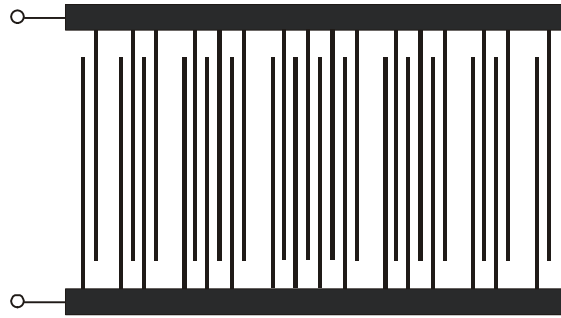
Interesującą metodą kształtowania przetworników jest pochylanie elektrod. Płaska fala powierzchniowa, propagując się ukośnie względem układu elektrod, generuje w każdym punkcie elektrody pewien ładunek, którego faza odpowiada fazie fali w tymże punkcie. Całkowity ładunek wypływający z danej elektrody jest całą ładunku po całej długości elektrody. Z tego względu jeśli np. końce elektrod są przesunięte względem frontu fali o całkowitą wielokrotność długości fali, to całkowity prąd zwarciaowy pochodzący od elektrod będzie zerowy. Zmiana pochylenia elektrody daje więc podobny efekt do apodyzacji (rys. 3).



Rys. 3. Przetwornik kątowy

Wycinanie elektrod

Prostą w realizacji metodą kształtowania przetwornika jest tzw. „wycinanie” elektrod tj projektowanie przetwornika z połączonymi odczepami (rys. 4). Wewnątrz każdego z odczepów można stosować indywidualne metody kształtowania (n.p. takie jak ww). Tego typu metoda może być stosowana np. do obniżania admitancji wejściowych długich przetworników, bez wyraźnej zmiany ich charakterystyk częstotliwościowych.



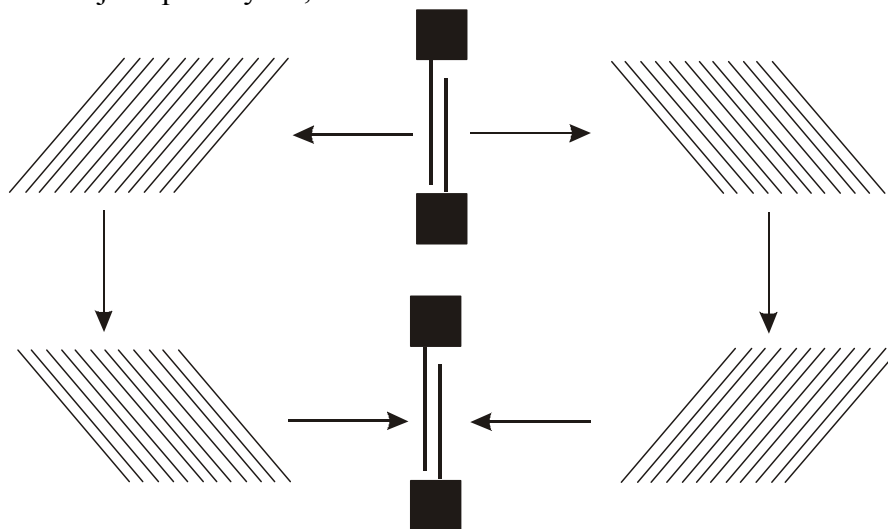
Rys. 4. Ważenie poprzez wycinanie partii elektrod

Podstawową metodą ważenia pozostaje apodyzacja. Przedstawione przykłady są raczej rzadko wykorzystywane. Często producenci zadowolają się też parametrami układów bez ważenia.

3.6. Przetworniki jednokierunkowe

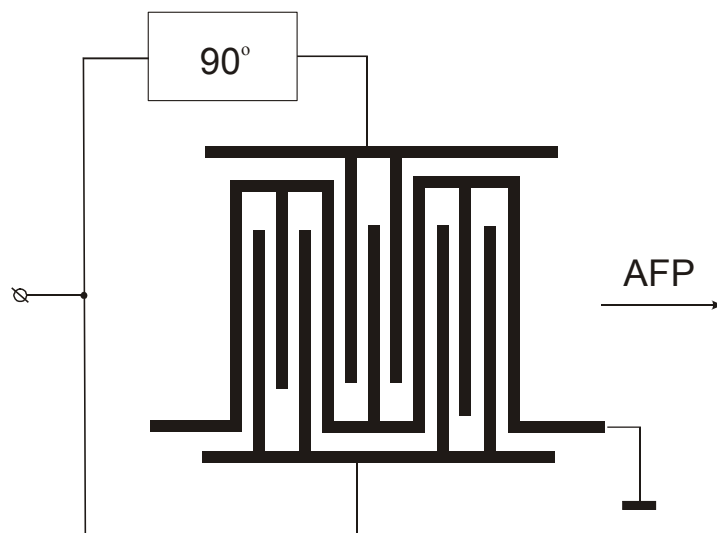
Jedną z ważniejszych wad klasycznych przetworników międzypalczastych jest ich dwukierunkowość. Niedogodność ta jest na tyle istotna, że od początku istnienia tych układów poszukiwano metod jej likwidacji bądź obejścia, co umożliwiłoby konstruowanie przyrządów z AFP o znacznie mniejszych stratach wnoszonych. Można wyodrębnić przynajmniej cztery nurty tych poszukiwań:

- urządzenia jednofazowe z symetrycznie ulokowanymi reflektorami bądź przetwornikami jak np. na rys. 1,



Rys. 1. Filtr rezonansowy z symetrycznymi strukturami odbijającymi

- urządzenia wielofazowe, w których jednokierunkowość uzyskuje się poprzez odpowiednio przesunięte w fazie zasilanie przetwornika podzielonego na sekcje – przykład na rys. 2,



Rys. 2. Dwufazowy przetwornik międzypalczasty

- układy uzyskujące jednokierunkowość poprzez odpowiedni układ wstępnych naprężeń podłoża,
- przetworniki typu jednofazowe (SPUDT - single phase unidirectional transducer).

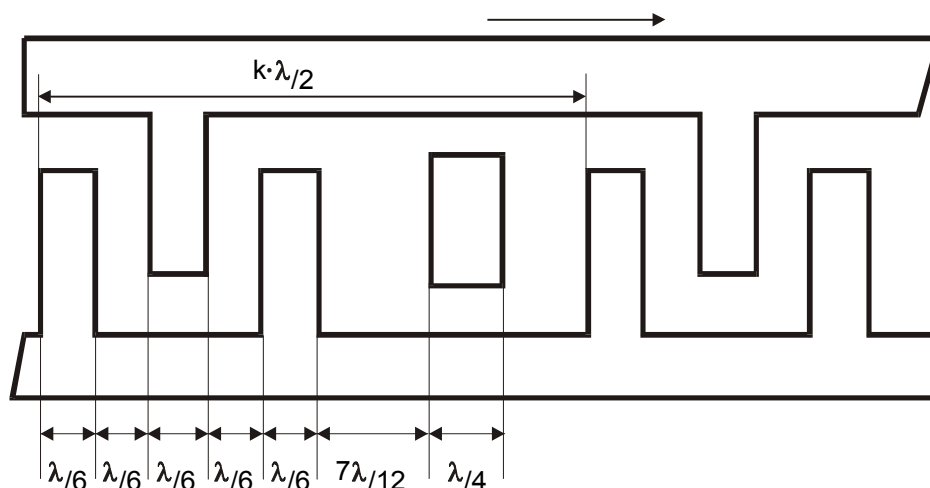
Metodę generacji AFP wykorzystującą przetworniki typu SPUDT opatentowano w 1982 r. Przetworniki tego typu zasadniczo są odmianą przetworników typu rezonansowego (współpracujących z reflektorami akustycznymi). Podstawową jednak różnicę stanowi fakt, że w przetwornikach typu SPUDT reflektory przeniesiono do wnętrza przetwornika, na skutek czego cały układ stał się zwarty. Cecha zwartości jest niezwykle korzystna ponieważ umożliwia łatwe ważenie przetwornika a także redukcję efektów drugiego rzędu. Koncepcja tego typu filtra okazała się bardzo trafna i do dziś powstało wiele odmian tej konstrukcji. Dwa przykłady pokazane są na rys. 3.



Rys. 3. Przykładowe rozwiązania przetworników SPUDT. Po lewej z elektrodami prostymi, po prawej z elektrodami pochylonymi

Najszerzej rozpowszechnione są struktury typu DART (Distributed Acoustic Reflection Transducers) oraz EWC (Electrode Width Controlled). Obie struktury mają jedną lub dwie

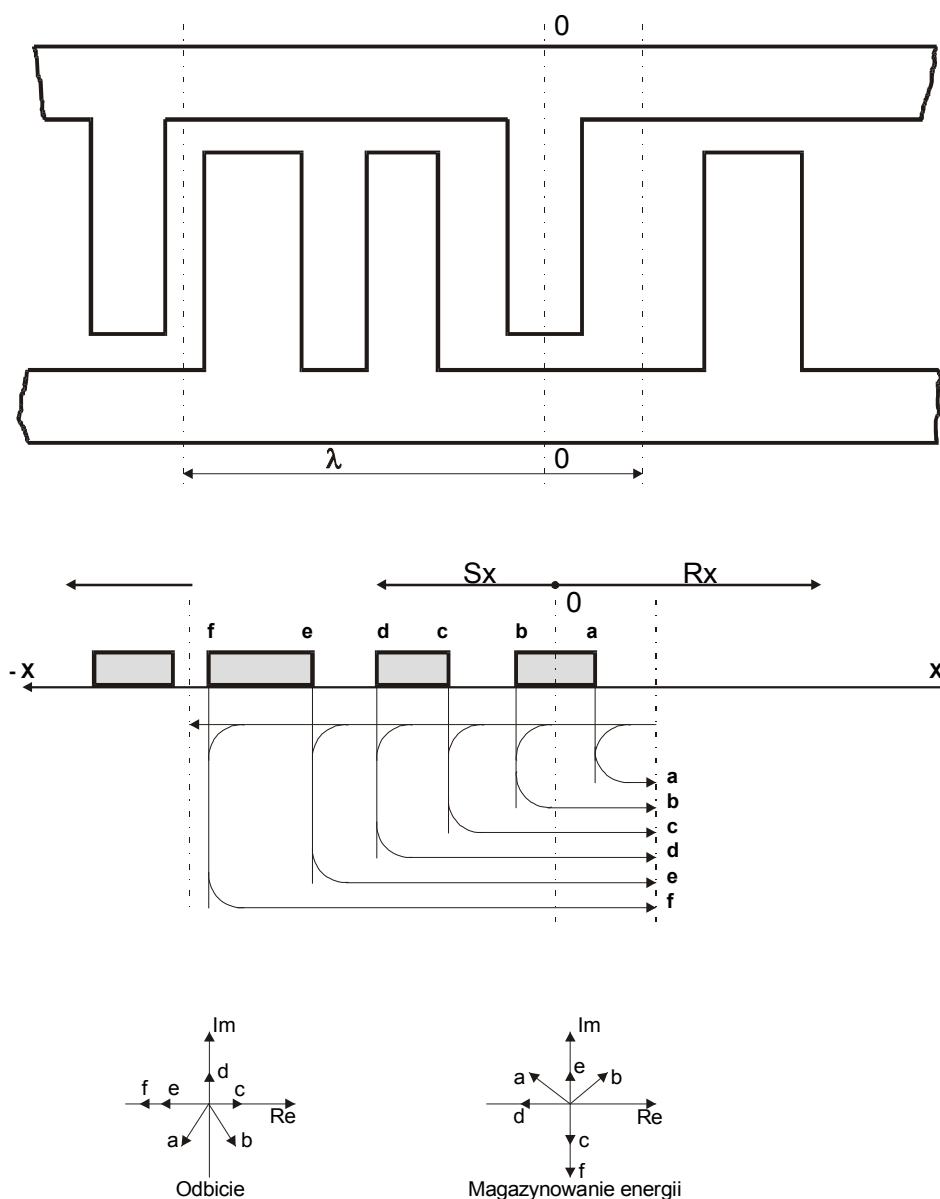
elektrody generujące o szerokości $\lambda_0/8$ (λ_0 - jest długością fali częstotliwości podstawowej) z przerwą $\lambda_0/4$ oraz reflektory akustyczne o szerokości $\lambda_0/4$ (EWC) lub o szerokości $3\lambda_0/8$ (DART). Oprócz niewątpliwych zalet, struktury te mają pewne wady. Przede wszystkim mała szerokość elektrod $\lambda_0/8$ sprawia trudności technologiczne i ogranicza górny zakres częstotliwości konstruowanych filtrów. Po drugie odstęp między źródłami AFP w DART i EWC wynoszący $\lambda_0/4$ nie odpowiada optymalnemu odstępowi $\lambda_0/2$ w dwukierunkowych przetwornikach IDT. W efekcie sprawność generacji AFP w wymienionych przetwornikach SPUDT wynosi jedynie 0,448 – 0,528 sprawności dwukierunkowego IDT. Dlatego straty wtrąceniowe dla średniopasmowych filtrów SPUDT są raczej duże. Po trzecie, jeżeli szerokość pasma filtru odpowiada przypadkowi synchronizacji obserwuje się zazwyczaj zniekształcenie szerokości pasma, spowodowane częścią urojoną impedancji. Efekt ten pogarsza odpowiedź częstotliwościową filtrów pasmowych, szczególnie z wysokim współczynnikiem kształtu $SF = 1,5 \div 1,7$. Sprawność SPUDT może być jednak zwiększona dwoma sposobami. Pierwszym z nich jest optymalne ustawienie źródeł AFP w ramach okresu przetwornika, drugim zaś użycie materiałów z wysokim współczynnikiem sprzężenia elektromechanicznego. Opracowano jednak struktury SPUDT częściowo wolne od wad typowych dla EWC i DART. Struktury te oparte są na IDT z okresem źródeł AFP $n\lambda_0/3$. W takich strukturach kompensacja niepożądanych odbić od elektrod następuje w ramach powtarzających się grup złożonych z trzech elektrod o szerokości $\lambda_0/3$ lub $\lambda_0/6$. W pierwszej strukturze TES SPUDT (Triple Electrode Sections SPUDT), grupy trzech źródeł oraz wewnętrznych reflektorów zapewniających jednokierunkowość działania są rozdzielone, zaś w drugiej reflektory są skojarzone ze źródłami AFP. W celu optymalizacji położenia, szerokości źródeł i reflektorów wewnętrznych przetwornika niezbędne są odpowiednie modele analityczne przede wszystkim ujmujące efekty drugiego rzędu, własności podłoża oraz grubości warstwy metalu. Struktury te optymalizuje się najczęściej wykorzystując model schematu zastępczego. Dla uzyskania jednokierunkowości stosuje się dwie różne struktury SPUDT - pierwsza jest zwykle strukturą typu grupowego. Zastosowane reflektory sytuowane są między konwencjonalnymi TES z elektrodami $\lambda/6$ (rys. 4).



Rys. 4. Reflektory usytuowane między konwencjonalnymi TES z elektrodami $\lambda/6$

Jeżeli minimalne odstępów przerw są ograniczone, okres pomiędzy TES wynosi: $L_p = n\lambda/2$, gdzie: $n = 4, 5, 6 \dots$. Maksymalna liczba reflektorów w każdej przerwie wynosi $(n-3)$. Odpowiednio sprawność generacji i odbicia AFP przypadająca na długość fali jest nieco mniejsza niż w DART czy EWC ze względu na podział na sekcje. Jednak przetworniki tego typu mogą być efektywnie wykorzystane w filtrach wąskopasmowych i wysokoczęstotliwościowych z nadmiarową liczbą źródeł i reflektorów.

Drugim sposobem zapewnienia jednokierunkowości jest wykorzystanie TES jednocześnie do generacji i odbijania AFP podobnie jak to ma miejsce w przypadku konwencjonalnych struktur SPUDT. Taki TES SPUDT został schematycznie przedstawiony na rys. 5.



Rys. 5. Budowa i mechanizm działania przetwornika typu TES SPUDT

Długość elementarnej sekcji przetwornika TES SPUDT wynosi λ . Składa się ona z trzech elektrod o różnych szerokościach i niesymetrycznym usytuowaniu wewnątrz sekcji. Z powodu asymetrii zmiana fazy centrum wzbudzenia w odniesieniu do środkowej elektrody wynosi około 5° . Sprawność generacji AFP w tej strukturze jest wyższa o około 30% niż w strukturze EWC. Mechanizm jednokierunkowości wyjaśniony został na rys. 5 przedstawiającym odbicia od brzegów elektrod oraz wykresy wektorowe fal składowych. Jak wynika z tego wykresu amplituda fali w kierunku $+x$ wynosi:

$$R_x^r = E[(1-r) - j 0,73 r]. \quad (1)$$

Dla fal rozchodzących się w kierunku $-x$ amplituda wynosi:

$$S_x^r = E[(1+r)-j 0,73 r], \quad (2)$$

gdzie E jest amplitudą AFP, zaś r współczynnikiem odbicia od krawędzi elektrody. Jednokierunkowe działanie TES SPUDT spowodowane jest głównie przez efekt magazynowania energii AFP. Dla cienkich elektrod metalowych ($h/\lambda=0,1-0,5\%$, h – grubość elektrody) na podłożu kwarcu, składowa rzeczywista znormalizowanego współczynnika odbicia na długość fali wynosi mniej niż 40% w porównaniu ze strukturą DART i około dwa razy mniej niż w strukturze EWC. Przy użyciu grubych elektrod metalowych ($h/\lambda = 1 - 4\%$) współczynnik odbicia zwiększa się znacznie dzięki efektowi magazynowania energii. Można wywnioskować, że współczynnik odbicia wynosi:

$$S_x^B = E[1- B/2] \quad (3)$$

gdzie B jest współczynnikiem magazynowania energii. Zsumowanie równań (2) i (3) daje zależność:

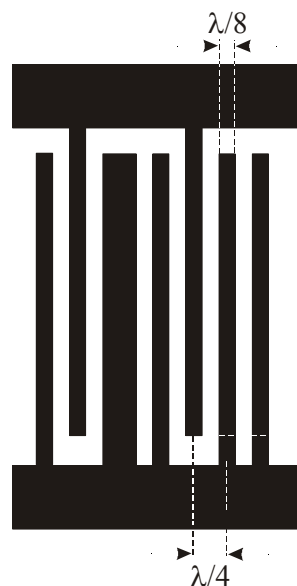
$$S_x = E[1-(r+B/2)-j 0,73 r], \quad (4)$$

z której wynika, że część rzeczywista współczynnika odbicia jest równa odpowiedniej składowej w strukturze EWC jeżeli grubość elektrody metalowej $h/\lambda \approx 1,8\%$.

Opisany TES SPUDT typu grupowego może być używany w filtrach wąskopasmowych i poziomem listków bocznych powyżej 60dB. Struktura TES, z asymetrycznym rozmieszczeniem elektrod zapewnia lepszą sprawność generacji AFP niż DART lub EWC. Rozwiązanie to pozwala także zmniejszyć straty wtrąceniowe o 1,3 - 1,5 dB. Tego typu przetworniki mogą być używane w filtrach na podłożu kwarcowym z szerokością pasma przenoszenia 1- 3% lub na podłożu LiTaO₃ z szerokością pasma 2 - 5%. Zastosowanie TES SPUDT pozwala na zmniejszenie zniekształceń pasma przenoszenia w filtrach z wymaganą wysoką selektywnością ($SF = 1,5 - 1,8$).

3.6.1. Elementy analizy przetworników SPUDT

Podstawowa sekcja takiego przetwornika składa się z uziemionych palców o szerokości $\lambda_0/8$, zasilanych palców $\lambda/8$ oraz reflektorów o szerokości $3\lambda_0/8$ jak na rys. 6.



Rys. 6. Podstawowa sekcja przetwornika SPUDT

Najprostszy algorytm projektowania wykorzystuje model funkcji δ do uzyskania próbek pożądanej odpowiedzi czasowej oraz do obliczania odpowiedzi impulsowych poszczególnych sekcji przetwornika. Podobnie jak w przetwornikach konwencjonalnych amplituda fali generowanej przez sąsiednie elektrody przetwornika jest proporcjonalna do ich apertury. Odpowiedź częstotliwościowa przetwornika z palcami o współrzędnych X_k , Y_k , i szerokości W_k wynosi:

$$A(f) = \sum_{k=1}^{N-1} (Y_{k+1} - Y_k) e^{-j2\pi f T_k}, \quad (5)$$

gdzie:

$$T_k = (X_k + X_{k+1} + 0,5(W_k - W_{k+1})) / (2v_s),$$

$$X_k = kv_s / (4f_c), \quad k = 1 \dots N,$$

$$W_k = v_s / (8f_c), \quad k = 1 \dots N,$$

$$Y_1 = 0, \quad Y_k = Y_{k-1} + WT_k, \quad k = 2 \dots N,$$

T_k – efektywna współrzędna czasowa,

v_s – prędkość fali na powierzchni półmetalizowanej,

N – liczba palców.

Algorytm projektowania rozpoczyna się on od obliczenia odpowiedzi impulsowej dla czterech próbek na długości fali przy użyciu odwrotnej transformacji Fouriera.

$$AT_k = \text{Re} \left(\sum_{i=1}^{N_f} D(f_i) e^{(2\pi f_i t_k + \phi)j} \right), \quad (6)$$

gdzie: $t_k = t_0 + k / (4f_0)$, $k = 1 \dots N-1$.

Przesunięcie czasowe t_0 wybiera się tak aby uzyskać poprawny zakres odpowiedzi czasowej, natomiast przesunięcie fazowe ϕ tak aby wartości AT_k znajdowały się w pobliżu -1, -1, +1, +1 itd.

Wartości AT_k poddaje się następnie ważeniu którego wynikiem są wartości WT_k wybrane tak aby wartość funkcji błędu ET_k była minimalna:

$$ET_k = \left| \sum_{m=1}^{k+4} AT_m z^m - 2 \sum_{m=1}^{k+4} WT_m z^m \right|, \quad (7)$$

gdzie: z - jest zmianą fazy na próbkę, $z = e^{-j\frac{\pi}{2}}$

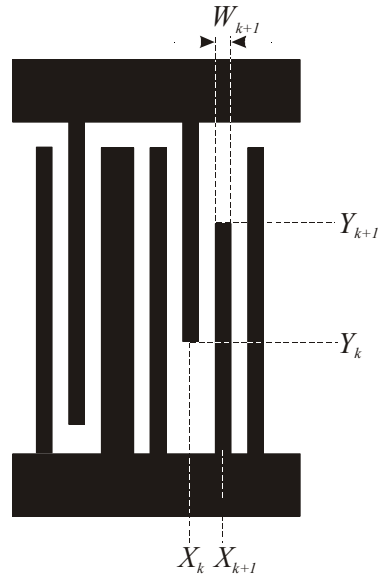
Reasumując, procedura projektowania sprowadza się do:

- syntezy ograniczonej w czasie odpowiedzi przetwornika,
- wykorzystania odwrotnej DFT do obliczenia geometrii przetwornika,
- ustalenie odpowiedzi częstotliwościowej zaprojektowanego przetwornika,
- porównanie odpowiedzi z zakładaną,
- dokonanie odpowiedniej korekcji (można z niej zrezygnować jeśli wymagana charakterystyka jest zadowolająca).

Alternatywną metodą może być algorytm zakładający generowanie amplitudy fali nie jak poprzednio proporcjonalne do apertury sąsiednich elektrod ale do ich długości. W takim przypadku odpowiedź częstotliwościowa przetwornika z palcami w punktach X_k , Y_k i szerokością W_k , dana jest wyrażeniem:

$$A(f) = \sum_{k=1}^N (Y_k - Y_0) e^{-j2\pi f X_k / v_s}, \quad (8)$$

Znaczenia parametrów widoczne są na rys. 7.



Rys. 7. Ważona sekcja SPUDT

Podobnie jak poprzednio algorytm korzysta z zaprojektowanej, ograniczonej w czasie odpowiedzi impulsowej. W tym jednak przypadku oblicza się odpowiedź impulsową przy jednej próbie na długość fali używając także odwrotnej dyskretnej transformacji Fouriera w formie:

$$TX_n = \sum_{i=1}^{N_f} D(f_i) e^{2\pi f_i (nv_s / f_c - x_0) / v_s + \theta} j \quad (9)$$

gdzie N_c jest liczbą sekcji, $n = 1 \dots N_c$.

Przesunięcie x_0 jest dobierane tak aby uzyskać poprawny zakres czasowy. Przesunięcie fazowe θ wybiera się tak, by część urojona TX_n przyjmowała minimum w pobliżu maksymalnych odpowiedzi przetwornika.

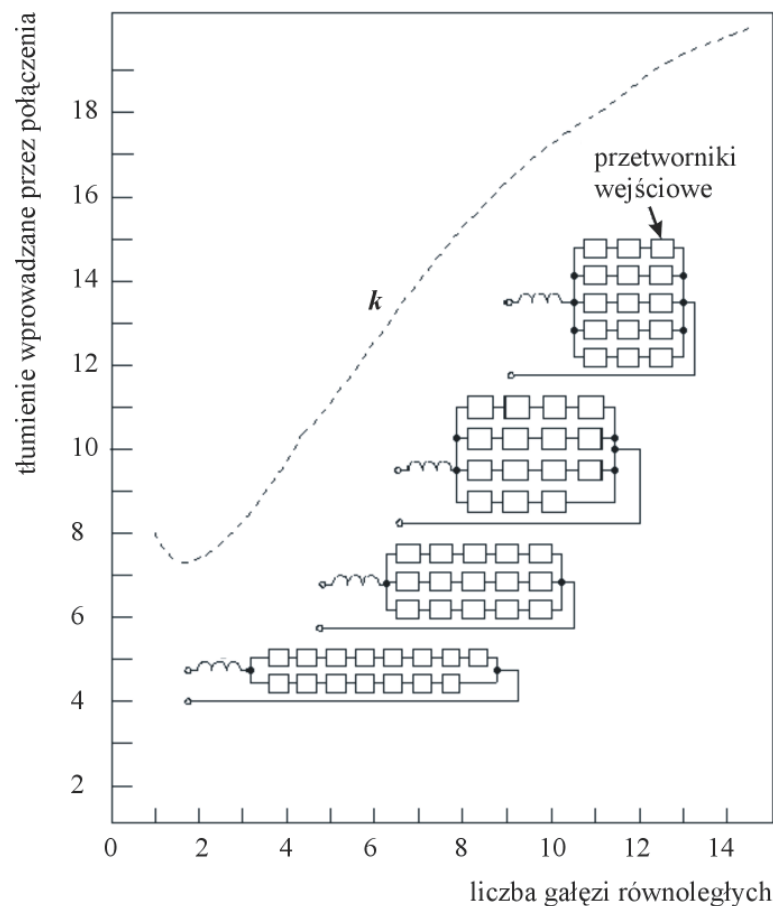
4. Urządzenia z AFP

4.1. Banki filtrów

Banki filtrów są urządzeniami skupiającymi w ramach jednej struktury szereg filtrów, z których każdy kolejny obsługuje inne, zwykle sąsiednie pasmo.

Charakterystyki przenoszenia poszczególnych kanałów stanowią strukturę ortogonalną tzn., taką, że środki pasm sąsiednich filtrów są umieszczone w jednakowych odstępach.

Przykładowym rozwiązaniem może być metoda w której źródło sygnału połączone jest z m kanałowym bankiem. Połączenie pojedynczego, 50Ω źródła z każdym kanałem akustycznym (lub połączenie akustycznych kanałów z pojedynczym 50Ω obciążeniem) jest bardzo ważnym problemem w projektowaniu banku filtrów tego typu (szczególnie w zakresie częstotliwości mikrofalowych). Głównym kryterium projektowania jest konieczność zapewnienia małej tłumienności wtrąceniowej i stałego poziomu sygnału na wyjściu. Są to zazwyczaj dwa przeciwstawne kryteria. Przykłady różnych konfiguracji 15-kanałowego szeregowo-równoległego banku filtrów pokazano na rys. 1.

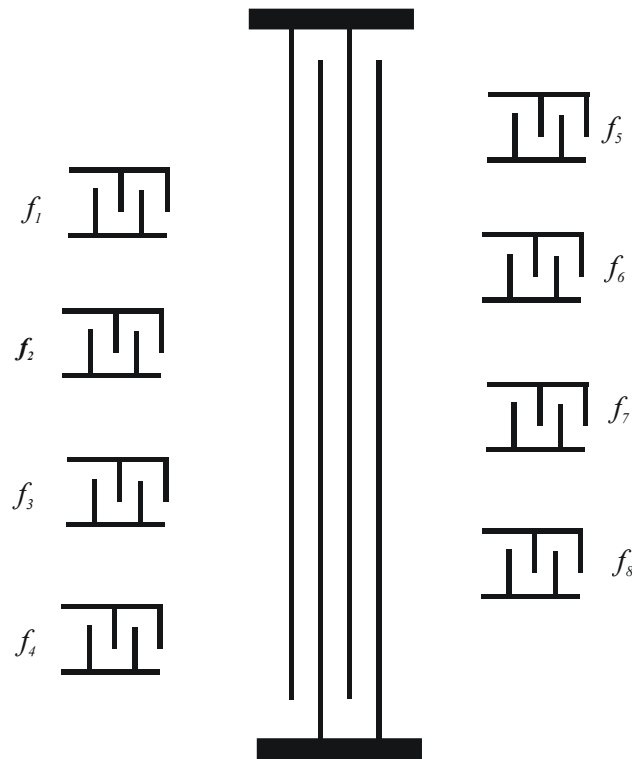


Rys. 1. Tłumienie w funkcji różnych połączeń wejściowych przetworników banku filtrów. k jest obwiednią poziomu tłumienia w funkcji liczby gałęzi równoległych

Zewnętrzna indukcyjność jest wybierana pod kątem optymalizacji środkowego kanału. Minimum tłumienności wtrąceniowej otrzymuje się dla dwóch gałęzi równoległych. W

przedstawionym układzie różnica poziomów sygnałów poszczególnych kanałów nie przekracza 2 dB.

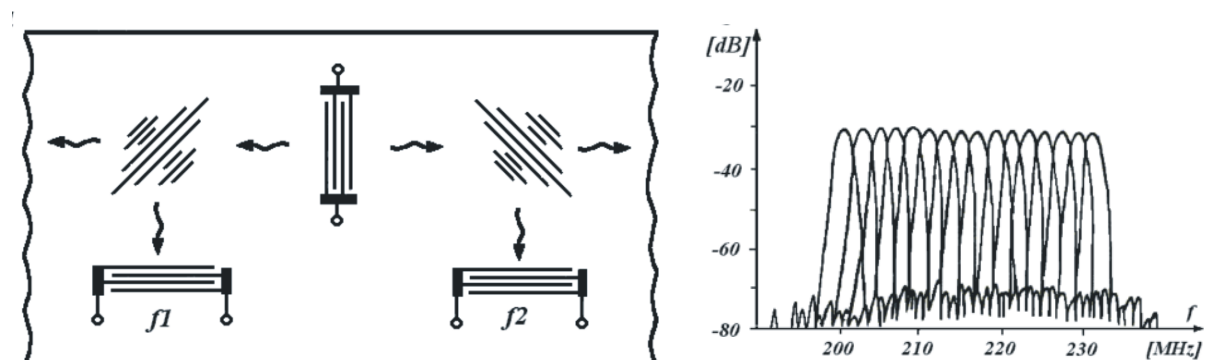
Rys. 2 prezentuje konstrukcję 8-kanałowego banku filtrów z jednym przetwornikiem nadawczym. Wykorzystano tu dwukierunkowość przetwornika.



Rys. 2. Konstrukcja banku filtrów oparta na jednym przetworniku wejściowym

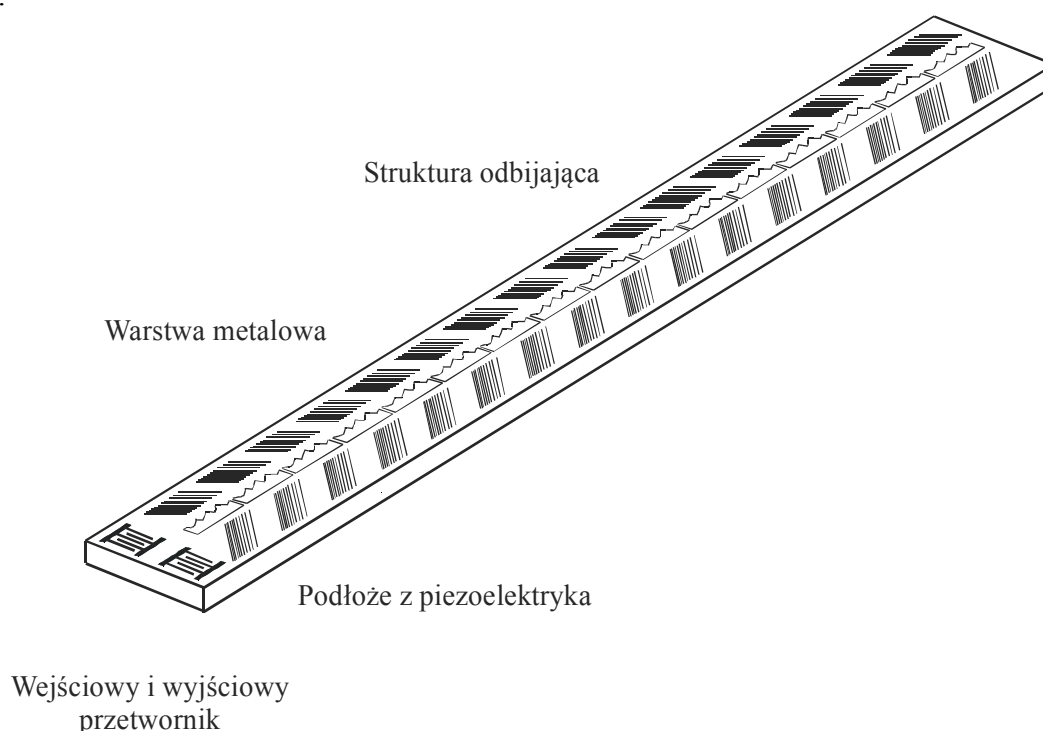
Przetwornik wejściowy powinien pokrywać odpowiednie do przetworników wyjściowych pasmo częstotliwości. Przetworniki wyjściowe są wąskopasmowe i zaprojektowane na konkretną częstotliwość mieszczącą się w granicach pasma przenoszenia przetwornika wejściowego. Zastosowanie takiej konstrukcji ogranicza się do urządzeń, które dysponują dużym poziomem mocy. Istotną zaletą tej konfiguracji jest wykorzystanie całkowitej energii wysyłanej przez przetwornik nadawczy.

Interesującym przykładem konstrukcji jest bank filtrów wykorzystujący struktury odbijające. Schemat takiego układu przedstawiono na rys. 3.



Rys. 3. Po lewej fragment banku ze strukturami odbijającymi, po prawej jego charakterystyka częstotliwościowa

W układzie tym fala wzbudzana przez przetwornik nadawczy pada kolejno na struktury odbijające sygnały o różnych częstotliwościach dzięki różnym odstępom między rowkami. Zwykle stosowane są struktury z małym współczynnikiem odbicia (płytke rowki). Warunkiem poprawności działania jest jednak, aby amplituda fali pierwotnej nie uległa wyraźnemu zmniejszeniu wskutek generacji fali odbitej. Wtedy można przyjąć, że amplituda fali odbitej w danym miejscu struktury zależy tylko od parametrów rowka odbijającego, a nie od jego położenia. W układach z ukośnymi strukturami odbijającymi występuje wiele efektów szkodliwych takich jak wielokrotne odbicia, nierówności kąta padania i odbicia fali wskutek anizotropii podłoża oraz zmiany prędkości fali w obszarze pokrytym rowkami. Z tego powodu projektowanie takiego banku filtrów jest dość skomplikowane. Główny problem wiąże się koniecznością kształtowania struktur. Zabieg ten polega najczęściej na zmianie długości bądź głębokości rowków lub stosowaniu zamiast rowków struktur kropkowych. Gęstość tych struktur wzdłuż danej linii jest ekwiwalentna zmianie głębokości rowków. Inne zastosowanie struktur odbijających do budowy banku filtrów z AFP przedstawiono na rys. 4.



Rys. 4. Struktura banku filtrów składająca się z 16 struktur odbijających

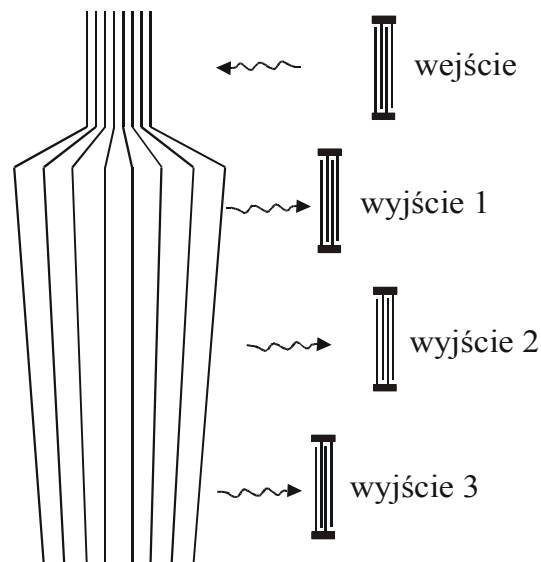
Układ zawiera szesnaście par struktur odbijających usytuowanych na wspólnym podłożu. Każdy kanał składa się z szerokopasmowego przetwornika wejściowego i wyjściowego oraz z pary struktur odbijających. Najczęściej rolę tych struktur pełnią rowki o zmiennym odstępem. Fala odbita od odpowiedniej struktury pada na strukturę przeciwną, gdzie odbija się powtórnie. W wyniku tego zjawiska fala jest odbierana przez przetwornik leżący obok przetwornika nadawczego. W ten sposób długość podłoża jest wykorzystywana podwójnie. W zależności od częstotliwości sygnału odbicie fali w obydwu strukturach zachodzi w różnych miejscach struktury i tylko rowki o odpowiednim odstępem efektywnie odbijają falę o danej częstotliwości. Każda fala odbijana jest głównie w zlokalizowanym miejscu wzdłuż struktury odbijającej, w obszarze, gdzie spełniony jest warunek synchronizmu długości fali i rowków. Umożliwia to korekcję błędów fazowych przez zmianę opóźnienia fali propagującej się między reflektorami. Zmianę opóźnienia realizuje się przez napylenie warstwy metalu w miejscu odpowiadającym fali odbitej dla danej częstotliwości. Tego typu bank filtrów stosuje

się w urządzeniach radiolokacyjnych do określania częstotliwości Dopplera. Dzięki zastosowaniu struktur odbijających z liniową modulacją częstotliwości uzyskuje się większy poziom sygnału na wyjściu banku filtrów, proporcjonalny do pierwiastka współczynnika kompresji oraz większą rozróżnialność w odległości i częstotliwości.

Do konstrukcji banku filtrów można też wykorzystać sprzęgacze kierunkowe z AFP. Sprzęgacz kierunkowy stanowi periodyczny układ oddzielnych elektrod na powierzchni piezoelektryka. Istotną, wykorzystywaną cechą sprzęgacza, jest zdolność przenoszenia AFP do sąsiednich kanałów, nawet jeśli w kanale pobudzonym fala nie ma jednorodnej amplitudy. Podstawowe zastosowania sprzęgaczy kierunkowych w filtrach pasmowych z AFP polegają na wykorzystaniu ich zalet, do których należy zaliczyć:

- odsunięcie przetwornika odbiorczego od obszaru, gdzie propagują się fale objętościowe, wzbudzone przez przetwornik nadawczy; obniża to poziom sygnałów fałszywych,
- „iloczynowa” współpraca przetworników kształtowanych przez zmianę długości ich elektrod; umożliwia zwiększenie tłumienia poza pasmem filtru.

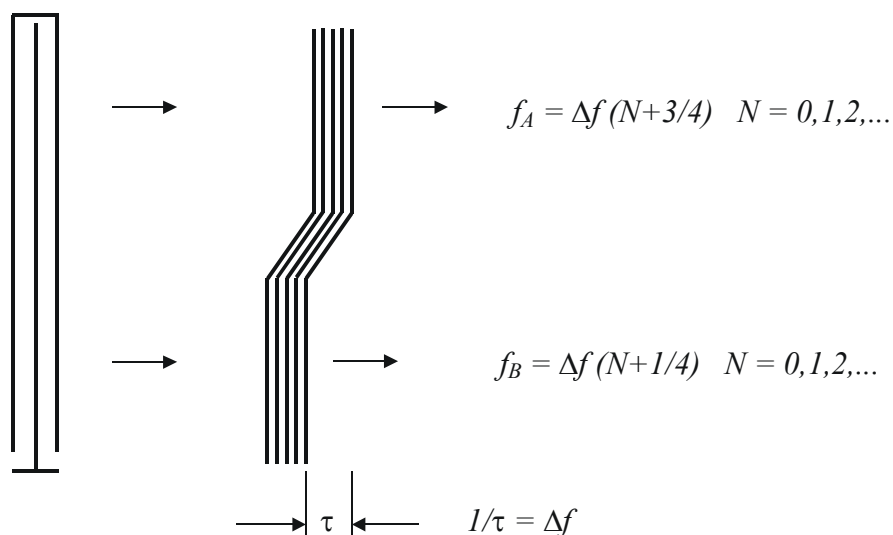
Przykład banku ze sprzęgaczem pokazano na rys. 5.



Rys. 5. Wykorzystanie sprzęgacza kierunkowego do konstrukcji banku filtrów

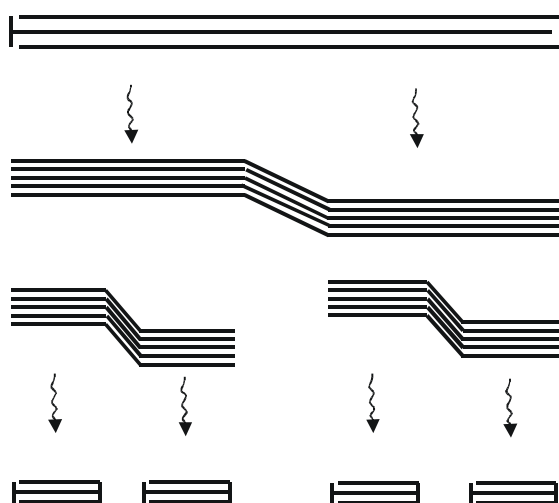
Zastosowano tu ukośny układ elektrod sprzęgacza o zmiennym odstępnie w obszarach kanałów wyjściowych, co pozwoliło na rozdzielenie sygnału wejściowego na szereg trzy podpasma.

W urządzeniach gdzie jest ważny niski poziom strat i małe wymiary banku filtrów używane są sprzęgacze 3 dB. Szerokość pasma przenoszenia zdeterminowana jest przez szerokopasmowy przetwornik nadawczy, natomiast częstotliwości wyjściowe przez wąskopasmowe przetworniki odbiorcze. Charakterystyka przenoszenia sprzęgacza jest typu $\sin x/x$ i może być kształtowana poprzez dobór geometrii elektrod. Rozdzielenie sygnału, ze względu na zajmowane pasmo uzyskuje się przez zastosowanie sprzęgaczy 3dB, wzbudzanych w obu kanałach przez fale przesunięte w fazie. Konfigurację taką przedstawiono na rys. 6.



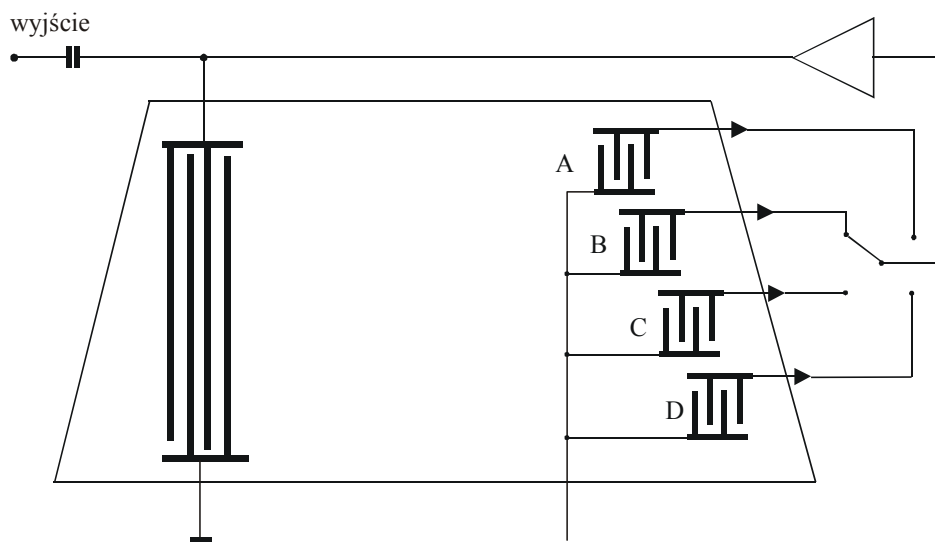
Rys. 6. Rozdział sygnału wejściowego na dwa podpasma z wykorzystaniem sprzęgacza 3 dB

Wyjściowy sygnał, będzie pojawiać się w kanałach z odstępem równym połowie pasma $\Delta f = 1/\tau$, gdzie τ jest opóźnieniem między jednym, a drugim kanałem sprzęgacza realizowanym poprzez „załamanie” sprzęgacza. W podobny sposób zbudować można bank filtrów z większą liczbą sprzęgaczy 3 dB (rys. 7).



Rys. 7. Schemat 4-kanałowego banku filtrów z wykorzystaniem trzech sprzęgaczy 3dB

Przykładem zastosowania banku filtrów może być prosty syntezer częstotliwości z AFP. Jego budowę przedstawiono na rys. 8.

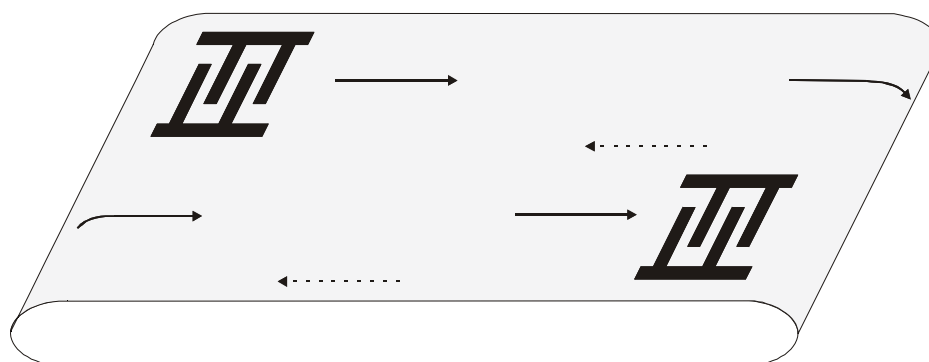


Rys. 8. Syntezer częstotliwości

Układ jest zbudowany z szerokopasmowego przetwornika nadawczego oraz z czterech przetworników odbiorczych o określonym paśmie przepustowym. Poprzez przełącznik obrotowy i wzmacniacz, który zasila układ, sygnał zostaje podany na wejściowy przetwornik nadawczy zamykając tym samym pętlę dodatniego sprzężenia zwrotnego. Jednakowe przetworniki odbiorcze są przesunięte względem siebie o stałą wartość, a różnice częstotliwości wyjściowych uzyskuje się różnicami długości dróg propagacji dla poszczególnych kanałów.

4.2. Linie opóźniające

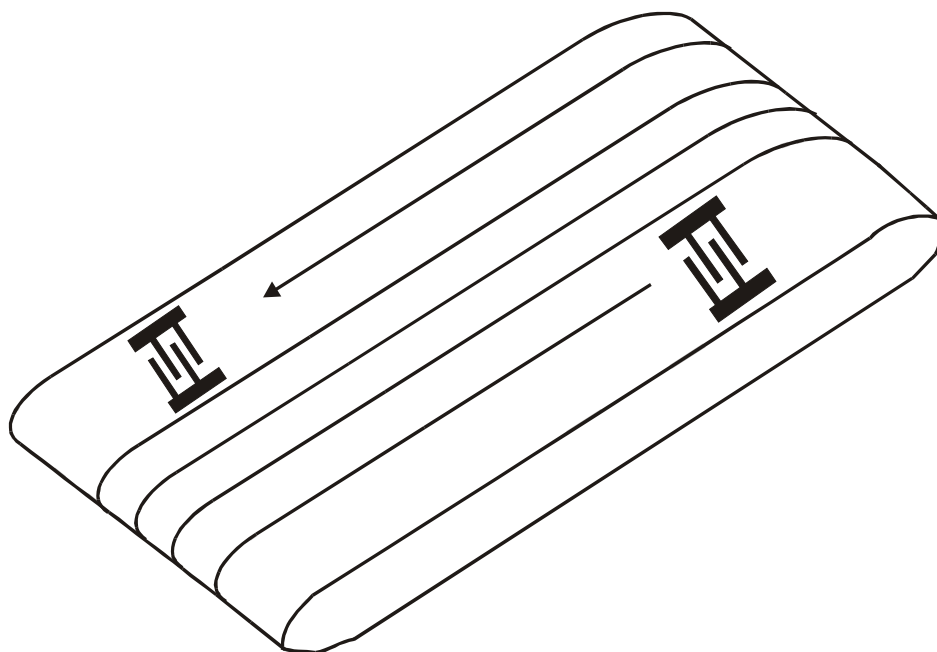
Zapotrzebowanie na linie opóźniające o dużych czasach opóźnień doprowadziło do szeregu interesujących rozwiązań konstrukcyjnych tych urządzeń. Wykonanie prostej linii opóźniającej o czasie opóźnienia powyżej 100 μs wymaga kryształów o bardzo dużych rozmiarach. Fizycznie dostępne kryształy nie są jednak tak duże. Jednym z rozwiązań umożliwiających uzyskanie dużych czasów opóźnienia jest zaokrąglanie krawędzi podłoża jak na rys. 1.



Rys. 1. Linia opóźniająca z zaokrąglonymi krawędziami

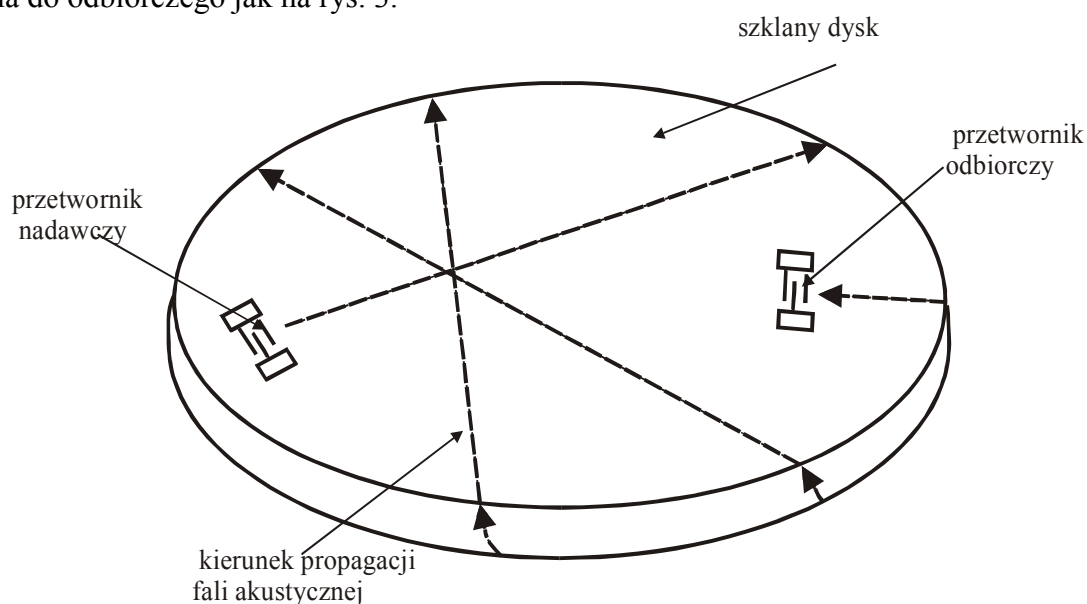
Odpowiednie oszlifowanie krawędzi bocznych podłoża powoduje, że AFP wzbudzona przez przetwornik nadawczy propaguje się po powierzchni piezoelektryka najpierw po stronie wierzchniej a później także po stronie spodniej. Na spodniej stronie umieszczony jest zwykle sprzęgacz kierunkowy (nie widoczny na rysunku) pozwalający skierować falę powierzchniową do przetwornika odbiorczego.

Inne rozwiązanie linii opóźniającej z zaokrąglonymi krawędziami przedstawiono na rys. 2.



Rys. 2. Inny wariant linii opóźniającej z zaokrąglonymi krawędziami

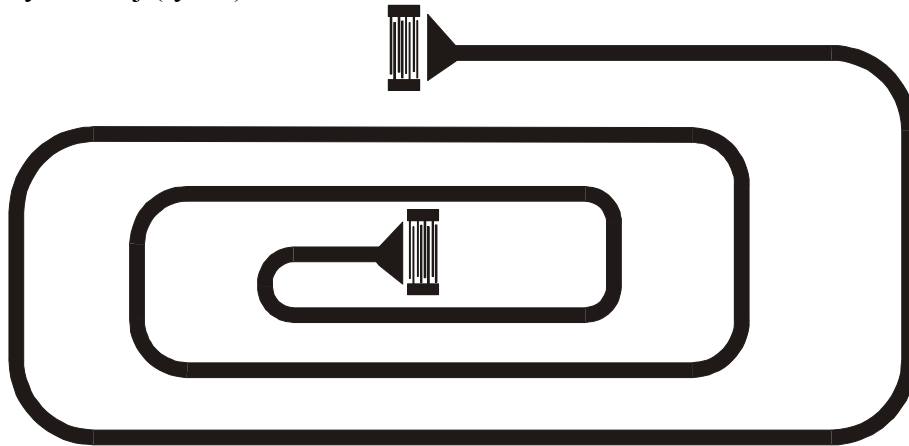
To rozwiązanie nie posiada żadnych struktur odbijających czy sprzęgaczy. Fala powierzchniowa generowana tu jest pod pewnym kątem, dobranym tak, by po pokonaniu drogi kilka razy wzdłuż piezoelektryka po obu stronach, trafiła do przetwornika odbiorczego. Opóźnienie jak widać na rysunku jest proporcjonalne do długości kryształu i liczby obiegów. Z uwagi na efekty niepożądane, szczególnie tłumienie, nie może to być jednak droga zbyt długa, ponieważ w skrajnym przypadku sygnał nie dotrze do przetwornika odbiorczego. Istnieją także konstrukcje przypominające kształtem dysk, którego krawędzie są zaokrąglone, a przetworniki umieszczone na jednej z powierzchni. Przetworniki ustawione są pod odpowiednim kątem tak, by fala generowana przez przetwornik nadawczy po wielu obiegach trafiła do odbiorczego jak na rys. 3.



Rys. 3. Linia opóźniająca w postaci dysku z zaokrąglonymi krawędziami

Istotną wadą takiego rozwiązania może okazać się jednak rozszerzalność temperaturowa podłoża. Przy dużych współczynnikach rozszerzalności zmieniają się wartości kątów przy „zawijaniu” się fali na zaokrąglonej krawędzi, co może doprowadzić do dużych strat sygnału, a nawet jego utraty.

Spotyka się też linie opóźniające współpracujące z przewodnicami falowymi. Są to rozwiązania, które wydłużenie czasu opóźnienia uzyskują dzięki wydłużeniu drogi propagacji w przewodnicy falowej (rys. 4).



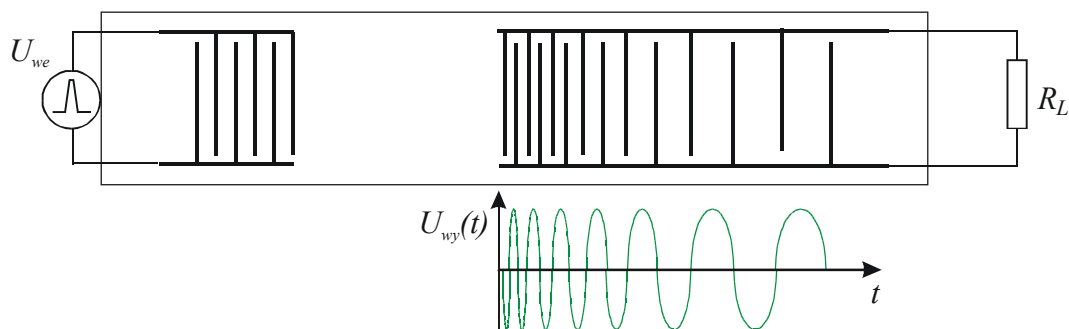
Rys. 4. Linia opóźniająca z przewodnicą falową

Jak widać generacja fali powierzchniowej odbywa tu się w analogiczny sposób do poprzednich typów linii czyli za pomocą przetworników międzypalczastych. Istotną różnicą jest zwinięta metalowa przewodnica, która pozwala na efektywne wykorzystanie powierzchni piezoelektryka. Fala rozchodzi się wzdłuż falowodu i trafia do przetwornika odbiorczego, gdzie jest zamieniana na sygnał elektryczny. Opóźnienie proporcjonalne jest więc do długości przewodnicy. Przy projektowaniu tego typu linii należy szczególną uwagę zwrócić na odpowiedni dobór podłoża i kierunku propagacji, ponieważ fala jest prowadzona w dwóch prostopadłych do siebie kierunkach. Należy nadmienić, że przewodnica falowa może być wykonana także w postaci rowka o odpowiedniej głębokości.

4.3. Linie dyspersyjne

Jeżeli przyjąć, że elementarny rezonator tworzy para elektrod, to przetwornik można porównać do zbioru połączonych układów drgających, z których każdy może mieć nieco inną częstotliwość rezonansową. Jeżeli w przetworniku odbiorczym odstęp między sąsiadującymi elektrodami będą rosły, to wówczas na końcu tego przetwornika największą sprawność przetwarzania, a więc i największą amplitudę uzyskają harmoniczne niskoczęstotliwościowe. Ponieważ droga jaką przebędą te składowe będzie dłuższa niż ta, którą pokonają harmoniczne wysokoczęstotliwościowe różne harmoniczne wystąpią w różnych momentach odpowiedzi impulsowej przetwornika, a tym samym wystąpi zjawisko dyspersji.

Dyspersyjna linia opóźniająca składa się więc z szerokopasmowego przetwornika wejściowego oraz dyspersyjnego przetwornika wyjściowego (rys. 1).



Rys. 1. Poglądowy schemat linii dyspersyjnej.

Istotny jest sposób pobudzania przetwornika wejściowego. Możliwe są tu w zasadzie dwa warianty:

- pobudzanie za pomocą krótkich impulsów wizyjnych,
- pobudzanie za pomocą krótkich impulsów radiowych o częstotliwości środkowej równej częstotliwości środkowej przetwornika.

Impuls pobudzający musi zapewniać dostatecznie dużą wartość amplitudy harmonicznych, które znajdują się w paśmie przetwornika dyspersyjnego, pracującego na częstotliwości środkowej f_0 , przy czym amplitudy sąsiednich harmonicznych powinny nieznacznie się różnić. Warunki te będą tym lepiej spełnione im impuls będzie krótszy przy czym amplituda impulsu wizyjnego musi być rzędu kilkudziesięciu V, co stwarza niebezpieczeństwo wystąpienia przebiegów międzyelektrodowych szczególnie przy pracy na wysokich częstotliwościach (bardzo krótkie odstępki). W przypadku pobudzania przetwornika impulsem radiowym jego amplituda może być rzędu pojedynczych V, ponieważ maksimum jego widma przesunięte jest na częstotliwość środkową przetwornika.

Sygnaly typu „chirp” są to długie paczki falowe z liniową modulacją częstotliwości o czasie trwania T sięgającym kilkudziesięciu mikrosekund. Sygnaly te mają prawie prostokątne widmo o szerokości B , położone wokół częstotliwości f_0 . Sygnal typu „chirp” może być wygenerowany za pomocą podzespołu z akustyczną falą powierzchniową, jako jego odpowiedź impulsowa. Zadanie to realizuje tzw. filtr dyspersyjny (dyspersyjna linia opóźniająca), który składa się zwykle z szerokopasmowego przetwornika nadawczego oraz dyspersyjnego przetwornika odbiorczego ustalającego charakterystykę częstotliwościową całej linii. Istnieją dwa warianty konstrukcji takich przetworników odbiorczych:

- elektrody położone są w punktach wzdłuż drogi propagacji akustycznych fal powierzchniowych odpowiadających kolejnym maksimum sygnalu „chirp”
- elektrody są położone w stałych odstępach, a ich długości dobierane są tak, aby uzyskać założoną odpowiedź impulsową. (zmiana długości elektrod pozwala na odpowiednie „ważenie” próbek, o różnych opóźnieniach, odpowiadających położeniu elektrod).

Filtry dyspersyjne stosowane są przede wszystkim w radiolokacji, gdyż pozwalają na rozwiązanie sprzeczności polegającej na tym iż dłuższy sygnał daje lepszą wykrywalność, a krótszy lepszą rozróżnialność. Sygnal „chirp” jest w stanie pogodzić te przeciwstawne wymagania. Realizując filtrację optymalną pozwala on na takie skompromowanie sygnalu o czasie trwania T i paśmie B , że po filtracji ma on postać impulsu o czasie trwania $1/B$, znacznie krótszym od T . Uzyskuje się to dzięki odpowiedniej budowie nadajnika radiolokacyjnego, gdzie sygnal „chirp” wysyłany jest w przestrzeń z modulacją „w górę”, a więc wyższe częstotliwości są wysyłane później niż niższe. Po odbiciu od obiektu sygnal radiolokacyjny jest kierowany do dyspersyjnej linii opóźniającej z modulacją „w dół” tj. wykazującej odwrotną zależność opóźniania od częstotliwości. Na wyjściu filtru o

parametrach B i T , takich samych jak w sygnale „chirp”, wszystkie składowe częstotliwościowe sygnału nakładają się, co w przybliżeniu daje wąski impuls o czasie trwania $1/B$. W praktyce stosuje się dodatkowe „ważenie” charakterystyki amplitudowej filtru dyspersyjnego, dzięki czemu zamiast sygnału typu $\text{sinc}x$ (o niekorzystnych listkach bocznych) na wyjściu otrzymuje się sygnał zbliżony do trójkąta, o szerokości nieco większej od $1/B$, ale o bardzo słabych listkach bocznych. Poziom tych listków świadczy o jakości filtrów dyspersyjnych.

Filtry dyspersyjne stosowane są także jako elektronicznie przestrajalne linie opóźniające, różnego rodzaju filtry pasmowe oraz w układach widmowej analizy sygnałów, rezonatorach (o bardzo wysokiej dobroci) czy też układach dopasowanej filtracji sygnałów.

4.4. Rezonatory

4.4.1. Struktury odbijające

Budowa klasycznych rezonatorów opiera się na wykorzystaniu pełnego odbicia fali objętościowej od płaszczyzn ograniczających płytkę rezonansową. W przypadku fal powierzchniowych analogiczne odbicie od krawędzi podłoża prowadzi do znacznych strat wynikających z rozpraszania AFP do fal objętościowych.

Niemal pełne odbicie można uzyskać przez wykorzystanie efektu sprzężenia fali postępującej i wstecznej, propagujących się pod okresowo nałożonymi elektrodami. Zjawisko to omawiano wcześniej jako szkodliwe w przypadku przetworników międzypalczastych. Analogiczne odbicie fali otrzymuje się też przy propagacji AFP pod okresowo zaburzoną powierzchnią.

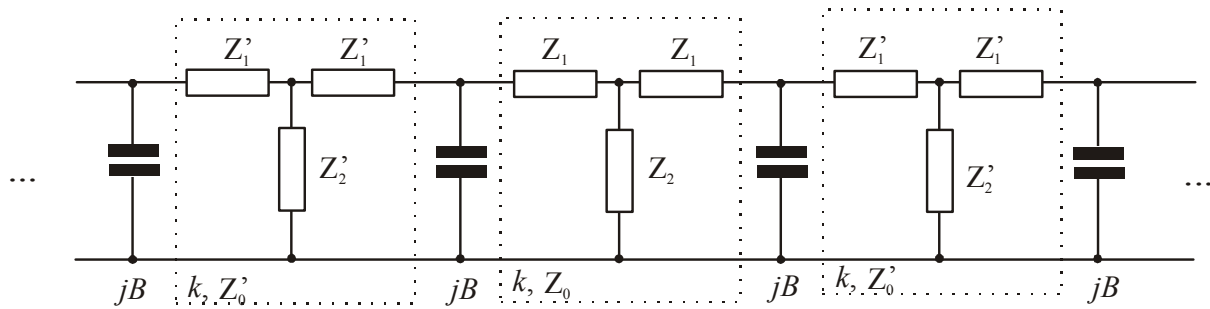
Istnieje kilka metod realizacji okresowych zaburzeń powierzchni:

- wykorzystaniu niejednorodności elektrycznej wprowadzanej przez elektrody,
- zastosowaniu okresowych rowków na powierzchni podłoża,
- naruszeniu struktury podłoża przez dyfuzję lub implantację jonów,
- nałożeniu na powierzchnię „ciężkich” elektrod.

Najczęściej stosowane są struktury rowkowe, wykonywane przez trawienie (np. jonowe). Stosuje się rowki o głębokościach w granicach $0,01 \div 0,03$ długości odbijanej fali powierzchniowej. Pasma pracy struktury odbijającej leży wokół częstotliwości środkowej, gdzie spełniony jest warunek synchronizmu fali i układu elektrod.

Od głębokości rowków zależy wielkość współczynnika odbicia fali powierzchniowej od każdego z nich i od całej struktury zawierającej N rowków. W porównaniu ze strukturami wykorzystującymi niejednorodności elektryczne stanowi to zaletę struktur rowkowych, ponieważ praktycznie niezależnie współczynnik odbicia od zmiany prędkości fali na powierzchni.

Struktury odbijające służą do konstrukcji rezonatorów typu Fabry-Perota. Najczęściej stosowanymi podłożami dla rezonatorów są dużym współczynnikiem sprzężenia piezoelektrycznego (np. LiNbO_3) oraz materiałami stabilnymi temperaturowo (SiO_2 ST). Podstawowe własności struktur odbijających można analizować wykorzystując metody układu zastępczego. Niezbędne parametry określone są zwykle empirycznie. W przypadku struktur rowkowych, gdzie zarówno w obszarze rowka, jak i poza nim prędkość fali powierzchniowej jest jednakowa, efekt odbicia fali wprowadza się przez zastąpienie tych obszarów odcinkami linii transmisyjnych o różnych impedancjach falowych (Z_0 i Z'_0 , rys. 1).



Rys. 1. Schemat zastępczy struktury rowkowej

Obserwowany w strukturach rowkowych efekt przesunięcia częstotliwości pracy struktury odbijającej w stronę niższych częstotliwości wraz ze wzrostem głębokości rowków oraz efekt odbijania fali na drugiej harmonicznej uwzględnia się przez włączenie susceptancji B między ogniwa układu zastępczego. Zmiana częstotliwości pracy struktury odbijającej nie ma jednak większego znaczenia daje się ją bowiem łatwo skorygować poprzez odpowiedni wybór periodu struktury. Przesunięcie środkowej częstotliwości pracy struktury odbijającej, uwzględniane przez susceptancję B ma wartość :

$$\frac{\Delta f}{f_0} = \frac{1}{\pi} \frac{B}{Y_0}$$

Empioryczne parametry układu z rys. 1 można zapisać w postaci zależności :

$$\frac{Y'_0}{Y_0} = 1 - 0,67 \frac{h}{\lambda} \quad \text{dla LiNbO}_3(\text{YZ}),$$

$$\frac{Y'_0}{Y_0} = 1 - 0,54 \frac{h}{\lambda} \quad \text{dla SiO}_2(\text{ST, X})$$

oraz:

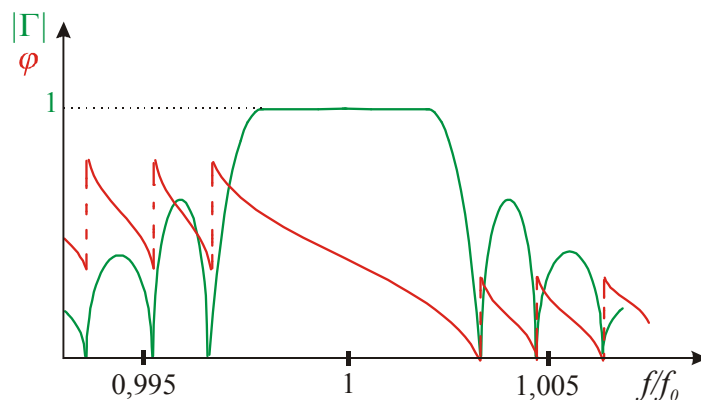
$$\frac{B}{Y_0} = -42 \left(\frac{h}{\lambda} \right)^2 \quad \text{dla LiNbO}_3(\text{YZ}),$$

$$\frac{B}{Y_0} = -35 \left(\frac{h}{\lambda} \right)^2 \quad \text{dla SiO}_2(\text{ST, X}).$$

Zależności te obowiązują dla $0,01 < h/\lambda < 0,03$ oraz periodu $\lambda/2$.

Dla dostatecznie długich struktur odbijających można przyjąć, że podobnie jak w przypadku nieskończonego układu elektrod, w obszarze struktury propagują się fale postępujące i wsteczne, poza obszarem struktury propagują się zaś fale padające i odbite.

Biorąc pod uwagę amplitudy tych fal oraz nakładając warunek zachowania energii sumy fal propagujących się w poszczególnych kierunkach otrzymuje się współczynnik odbicia dla całej struktury odbijającej (rys. 2).



Rys. 2. Przebieg współczynnika odbicia i fazy odbicia fali od struktury

Przebieg współczynnika odbicia $|\Gamma|$ jest symetryczny względem częstotliwości środkowej. Dla dostatecznie długich struktur wynosi on 1 przy czym ze względu na zjawiska szkodliwe nie stosuje się struktur dłuższych niż kilkanaście λ^2/h .

Jak widać na rys. 2 faza odbicia ma liniowy przebieg w pobliżu częstotliwości środkowej, co pozwala na wprowadzenie pojęcia ekwiwalentnego lustra (w analogii do rezonatora Fabry-Perota), w pewnej odległości od początku struktury. Odległość ta wynika z warunku na zmianę fazy fali odbitej od „lustra” i wynosi:

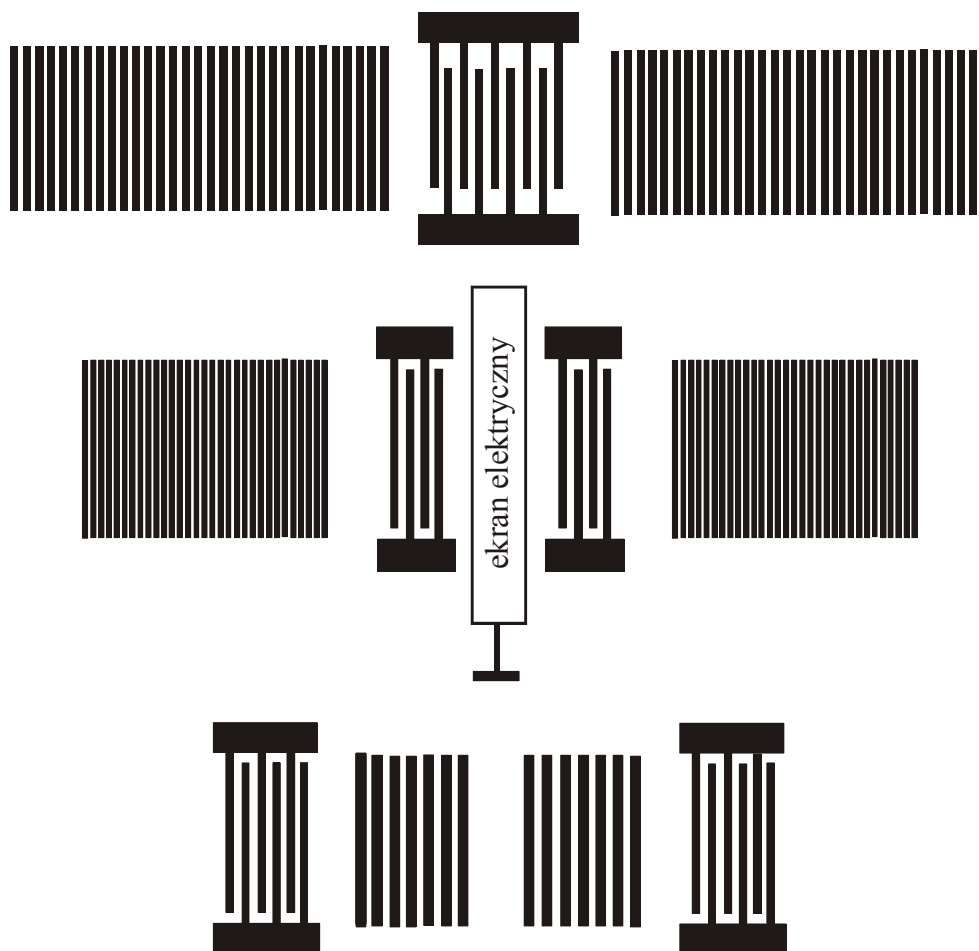
$$L_p = -\frac{1}{4\pi} \frac{\partial \varphi}{\partial (f/f_0)} \lambda.$$

4.4.2. Podstawowe konfiguracje rezonatorów

Zależnie od usytuowania lusterek i przetworników można zasadniczo otrzymać trzy typy rezonatorów:

- rezonatory z jednym przetwornikiem między strukturami; efekt rezonansowy polega tu na zmianie impedancji przetwornika,
- rezonator z dwoma przetwornikami między strukturami; efekt rezonansowy polega tu na zmianie współczynnika transmisji między przetwornikami,
- rezonator z dwoma przetwornikami na zewnątrz lusterek, które położone są w pewnym odstępie zwanym obszarem rezonansu; efekt rezonansowy polega tu również na zmianie transmisji między przetwornikami.

Konfiguracje te kolejno przedstawia rys. 3



Rys. 3. Kolejno od góry: rezonator z jednym przetwornikiem, rezonator z dwoma przetwornikami, rezonator z dwoma przetwornikami na zewnątrz struktury odbijającej

Dwie pierwsze konfiguracje stanowią tzw. rezonator obciążony, zaś trzecia – nieobciążony, gdyż fala stojąca wewnątrz struktur nie jest odprowadzana do elektrycznego obciążenia.

Własności rezonatorów w istotny sposób zależą od odległości struktur odbijających. Powinny one być tak dobrane aby zapewnić ułożenie się całkowitej liczby połówek fal między strukturami tj. powstanie fali stojącej dla częstotliwości, dla których współczynnik odbicia jest maksymalny. Ponieważ struktury odbijają falę w pewnym paśmie częstotliwości, to warunek stojącej fali (warunek fazy) może być spełniony wielokrotnie.

Dobroć rezonatora nieobciążonego jego wynosi:

$$Q = \frac{2\pi|\Gamma|L_{ef}}{\lambda(1-|\Gamma|^2)},$$

gdzie L_{ef} jest efektywną odległością luster równą mniej więcej odległości od ich środków.

Czynnikami zmniejszającej dobroć rezonatorów są:

- zły wybór odległości struktur powodujący, że warunek fazy jest spełniany dla częstotliwości, dla której Γ nie jest maksymalne,
- tłumienie fali powierzchniowej,
- zjawisko dyfrakcji powodujące zbaczanie fali z kierunku osi układu,
- niepełne odbicie – transmisja fali poza struktury odbijające.

Dobroć rezonatorów nieobciążonych osiąga wartość 60 000. Dla rezonatorów obciążonych jest ona znacznie mniejsza i dochodzi do 20 000.

najczęściej stosowane są rezonatory z dwoma przetwornikami pomiędzy strukturami odbijającymi ponieważ wprowadzają najniższe straty w rezonansie (kilka dB), co z kolei umożliwia szeregowe łączenie rezonatorów w celu zwiększenia ich dobroci.

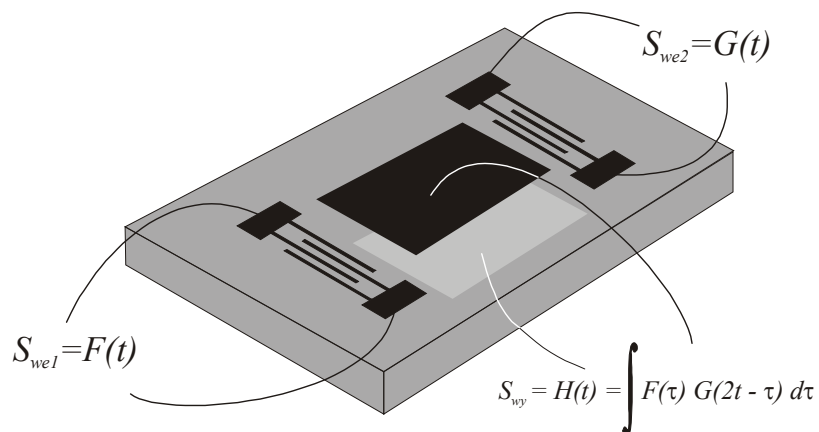
4.5. Konwoluty

Konwoluty są podzespołami realizującymi analogowe obliczanie całki splotowej. Na wejście takiego układu podawane są dwa sygnały, które w postaci fali przemieszczają się przeciwbieżne i mieszają. Ich mnożenie zachodzi na słabych nieliniowościach podłoża, a całkowanie na długości, na której sygnały się pokrywają.

Można wyróżnić dwie płaszczyzny zastosowań konwolutorów:

- ultra szybkie przetwarzanie informacji - szerokopasmowość oznacza w tym przypadku dużą szybkość, a konwolutor pełni rolę analogowego komputera,
- detekcja i dekodowanie sygnałów w systemach szerokopasmowych. Szerokie pasma są wykorzystywane do poprawy poziomu ochrony informacji oraz zwiększenia ogólnego zysku przetwarzania. Konwolutor pełni w takim przypadku rolę programowalnego korelatora o mikrosekundowych czasach odpowiedzi.

Najprostszym i jednocześnie najstarszym rozwiązaniem konwolutora jest układ składający się z dwóch przetworników międzypalczastych i leżącej pomiędzy nimi metalowej elektrody (rys. 1).



Rys. 1. Najprostszy konwolutor

Przyrząd ten wykonuje tzw. splot zdegenerowany, (w zwykłej całce splotowej drugi czynnik podcałkowy ma postać $G(t - \tau)$) degeneracja jest rezultatem względnego podwojenia prędkości fal propagujących się w przeciwnych kierunkach. Możliwa jest jednak, po dodaniu do układu dodatkowych struktur opóźniających, konstrukcja konwolutora wykonującego zwykły splot.

Aby zrozumieć działanie układu z rys. 1 zapiszmy falę biegnącą od pierwszego i od drugiego przetwornika w postaci odpowiednio:

$$A_1 \left(t - \frac{x}{v} \right) \exp \left[i \omega_1 \left(t - \frac{x}{v} \right) \right], \quad (1)$$

$$A_2 \left(t + \frac{x-L}{v} \right) \exp \left[i \omega_2 \left(t - \frac{x-L}{v} \right) \right], \quad (2)$$

L jest długością elektrody splatającej.

Wytworzone pole elektryczne od wzajemnie oddziaływających fal wyniesie:

$$E = cA_1 \left(t - \frac{x}{v} \right) A_2 \left(t - \frac{x-L}{v} \right) \exp \left\{ i \left(\omega_1 + \omega_2 \right) t - \left(\omega_1 - \omega_2 \right) \frac{x}{v} \right\}, \quad (3)$$

co odpowiada wyjściowemu napięciu:

$$U(t) = c \exp(2i\omega t) \int_0^L A_1 \left(t - \frac{x}{v} \right) A_2 \left(t + \frac{x-L}{v} \right) dx, \quad (4)$$

stała c charakteryzuje oddziaływania nieliniowe w podłożu.

Stosując podstawienie $\tau = t - x/v$ i przechodząc z granicami całkowania do nieskończoności uzyskuje się postać zdegenerowanej całki splotowej:

$$V_0(2\omega) = M \exp(2i\omega t) \int_{-\infty}^{\infty} A_1(\tau) A_2(2t - \tau) d\tau \quad (5)$$

Wykonując całkowanie (5) można uzyskać zależność napięcia na wyjściu konwolutora od gęstości mocy fal akusytcznych:

$$V_0(2\omega) = \frac{M}{W} \sqrt{P_1 P_2} \quad (6)$$

Wyjściowa wartość skuteczna napięcia w stanie rozwarcia zdegenerowanego konwolutora jest więc proporcjonalna do iloczynu dwóch gęstości mocy wejściowych sygnałów akustycznych. gdzie M jest współczynnikiem określającym siłę oddziaływań nieliniowych materiału, a W szerokością elektrody splatającej.

W celu określenia warunków optymalnej pracy układu, należy wziąć pod uwagę maksymalną moc wydzielaną na wyjściu. W tym celu należy obliczyć wewnętrzną impedancję urządzenia, definiowaną jako stosunek V_0 do natężenia prądu na wyjściu układu w stanie jego zwarcia I_0 . Impedancję tą można wyrazić poprzez rozwiązanie problemu liniowego i traktowanie zjawisk nieliniowych jako dodatkowego członu generującego ładunki. Część rzeczywista R_a impedancji wewnętrznej daje się wtedy wyznaczyć jako wartość proporcjonalna do stosunku współczynnika sprzężenia piezoelektrycznego podłoża do wielkości powierzchni regionu oddziaływania wzajemnego fal $W \cdot L$ zaś część urojona związana jest z pojemnością konwolutora C_a , która z kolei jest proporcjonalna do WL . Należy także wziąć pod uwagę pojemność związaną z grubością podłoża C_0 , która generalnie jest mniejsza niż C_a . Całkowita impedancja wewnętrzna wynosi wtedy:

$$Z_i = (r_a + 1/jc_a 2\omega)WL + 1/jC_0 2\omega \quad (7)$$

gdzie $r_a = WLR_a$, zaś $c_a = C_a/WL$. Zależność na maksymalną moc osiągalną na wyjściu jest wtedy następująca:

$$P_{\max} = V_0^2 / 4R_a = V_0^2 WL / 4r_a, \quad (8)$$

a przy wykorzystaniu zależności (6):

$$P_{\max} = CP_{1a} P_{2a}, \quad (9)$$

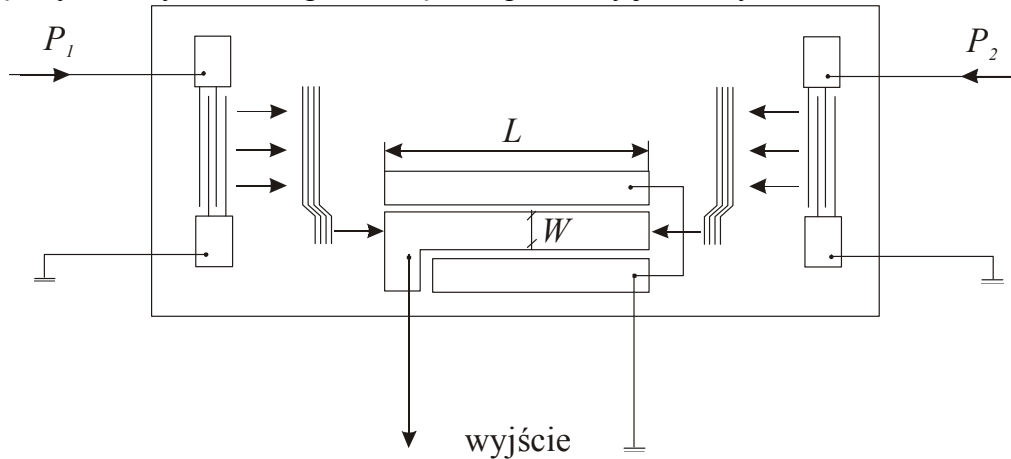
gdzie $C = M^2(L/W)/4r_a$ jest tzw. współczynnikiem biliniowości struktury.

Doświadczalnie stwierdzono, że im węższe jest pasmo, tym współczynnik C jest większy. Współczynnik proporcjonalności M dla kryształów piezoelektrycznych jest raczej mały. Zmierzone wartości dla YZ LiNbO₃ są rzędu 10^{-5} Vm/W, co powoduje, że moce wyjściowe konwolutorów klasycznych są niewielkie.

Badania konwolutorów klasycznych doprowadziły do wniosku, że w celu zwiększenia mocy wyjściowej urządzenia konieczna jest minimalizacja szerokości regionu wzajemnego oddziaływania fal. Dolną granicę szerokości tego regionu wyznacza zjawisko dyspersji falowodowej. Z analizy zależności dyspersyjnych wynika, że jeżeli szerokość elektrody

splatającej jest przynajmniej dwa razy większa od długości fali, zjawisko dyspersji można zaniedbać.

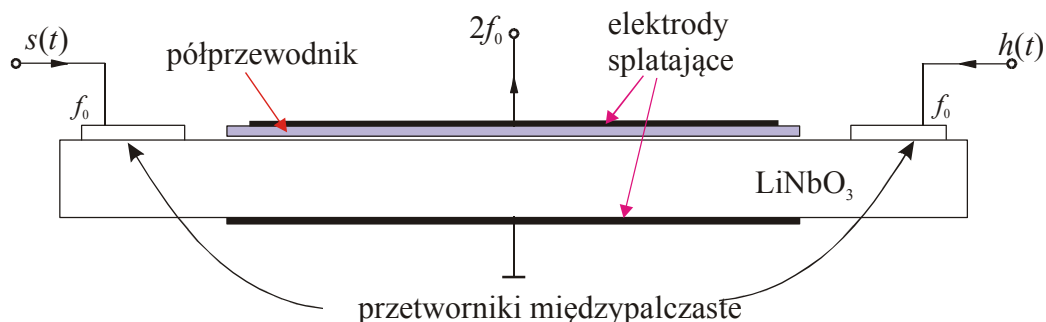
Zakres dynamiczny konwolutorów można rozszerzać poprzez kompresję szerokości pasma za pomocą wielopaskowych kompresorów pasma (BWC - ang. bandwidth wide compressor), wymagających niewielkich powierzchni i pozwalających na redukcję sygnałów niepożądanych. Przykład takiego rozwiązania pokazany jest na rys. 2.



Rys. 2. Konwolutor z wielopaskowym kompresorem pasma

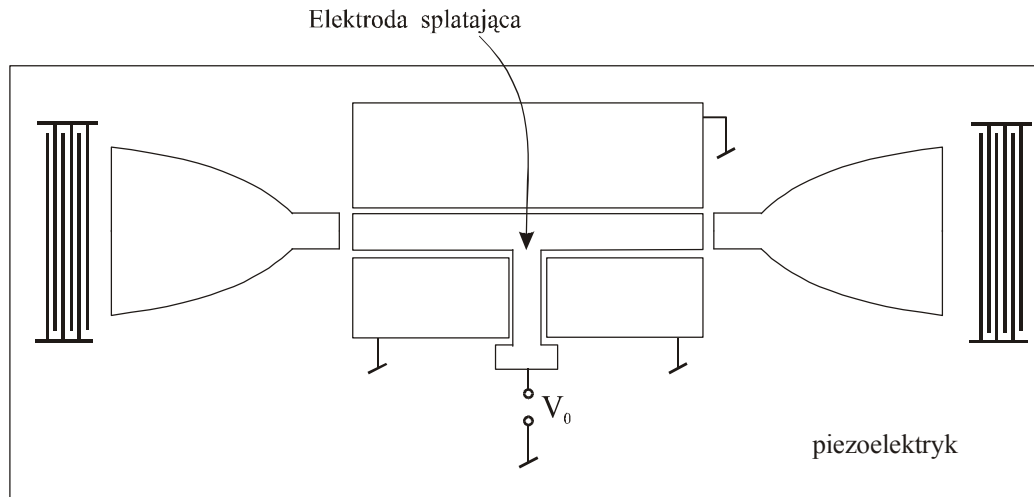
Przedstawione urządzenie może pracować z sygnałami o $BT = 600$ (50 MHz pasmo i 12 μ s opóźnienie), $W = 5\lambda_0 = 0,11$ mm, $f_0 = 156$ MHz. Niesymetryczne sprzęgacze wielopaskowe są składają się z 235 pasków i pozwalają na uzyskanie kompresji 15:1 dla pięciopalcowego przetwornika o aperturze 1.65 mm.

Zwiększenie nieliniowości, a tym samym poprawę sprawności konwolutorów uzyskać można w układzie konwolutora półprzewodnikowego. Fale powierzchniowe propagujące się w przeciwnych kierunkach (sygnały $s(t)$ i $h(t)$) wchodzą do obszaru gdzie nałożona jest warstwa półprzewodnika (najczęściej warstwa Si). Pole elektryczne, związane z tymi falami, moduluje gęstość nośników prądu w półprzewodniku i wywołuje generację prądu, którego mała, nieliniowa część, jest proporcjonalna do gęstości nośników i pola elektrycznego. W wyniku powstaje składowa prądu wzdłuż grubości warstwy półprzewodnika, proporcjonalna do iloczynu sygnałów $s(t)$ i $h(t)$. Prąd nieliniowy będący rezultatem tych oddziaływań, jest zbierany z pewnej długości wzdłuż drogi propagacji fali przez elektrodę nałożoną na powierzchnię warstwy półprzewodnika. Druga elektroda wyjściowa jest nałożona na dolną powierzchnię piezoelektryka (rys. 3).



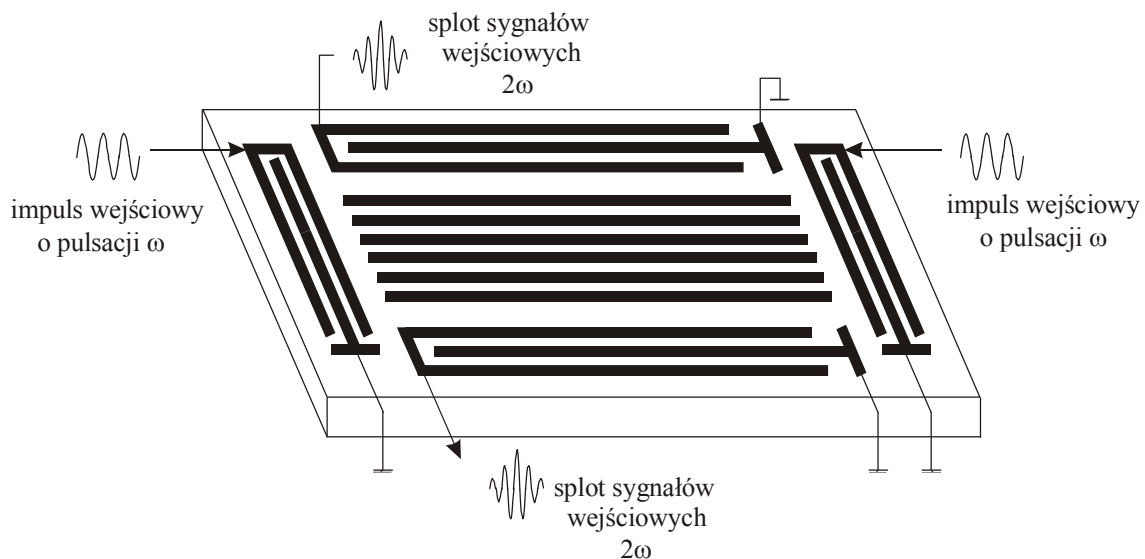
Rys. 3. Konwolutor półprzewodnikowy

Inną metodą wywołania silniejszych oddziaływań nieliniowych w konwolutorze jest zwiększenie gęstości mocy fal poprzez stosowanie przewodnic typu horn (ang. róg). Charakterystyczne dla tego rozwiązania jest występowanie „lejków” pomiędzy przetwornikami wejściowymi a elektrodą splatającą (rys. 4).



Rys. 4. Konwolutor z przewodnicą typu horn

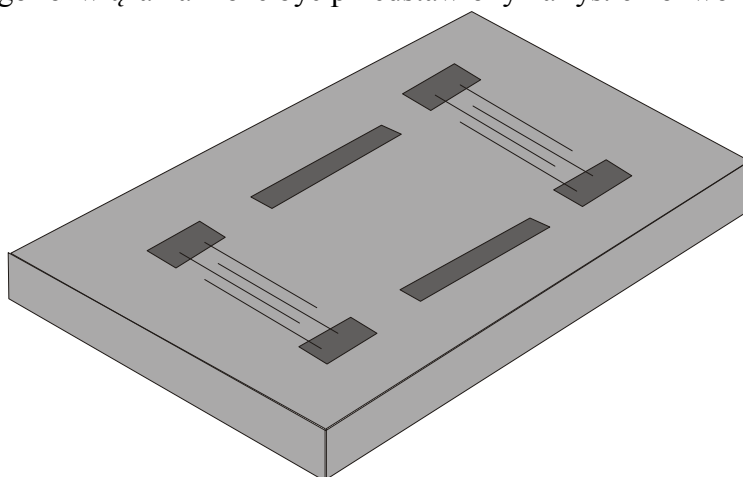
Gęstość mocy na szczytach stożków może być tak duża, że pojawiają się pęknięcia w strukturze krystalicznej piezoelektryka. Badania tych konwolutorów pokazały, że pęknięcia te mają podłoże elektryczne i powstają na skutek nierównych ładunków po obu stronach połączenia stożka z falowodem już podczas procesu produkcji. Aby wyeliminować to zjawisko stosuje się m.in. przerwy elektryczne pomiędzy „lejkami” a przewodnicą falową. Interesującym typem konwolutora jest układ z falowodem wielokanałowym (rys. 5). Falowód taki składa się z kilku linii paskowych. Sygnał splotu jest otrzymywany na IDT, który jest pobudzany przez AFP propagując się w kierunku normalnym do falowodu.



Rys. 5. Schemat konwolutora pracującego z falowodem wielokanałowym

Mechanizm generacji splotu oparty jest na zjawisku parametrycznego mieszania fal, które prowadzi do powstania nieliniowej składowej o podwójnej częstotliwości 2ω propagującej się w kierunku normalnym do falowodu.

Konwencjonalne konwolatory z AFP wykorzystują głównie poprzeczne pole polaryzacyjne, indukowane przez nieliniowości sieci krystalicznej w obecności dwóch, propagujących się przeciwbieżnie sygnałów wejściowych. Możliwa jest też konstrukcja konwolatorów wykrywających zarówno składowe poprzeczne jak i wzdłużne pola polaryzacyjnego wytworzonego przez współdziałanie AFP i SSBW (ang.- surface skimming bulk waves) na LiNbO_3 . Najlepsze, opisane w literaturze akustyczne konwolatory tego typu osiągają wartość iloczynu czas - pasmo, rzędu 1500 i efektywność splotu na poziomie – 65 dBm. Przykładem takiego rozwiązania może być przedstawiony na rys. 6 konwolator THC.



Rys. 6. Przykład realizacji konwolatora THC

W układzie THC biliniowe pole powstałe dzięki AFP nie propaguje się zgodnie z normalną do powierzchni sieci krystalicznej podłoża. Jego kierunek propagacji można natomiast opisać trzema ortogonalnymi składowymi, pionową (TV), poprzeczną (TH) i wzdłużną (L). Składowa poprzeczna leży w płaszczyźnie sieci krystalicznej i jest normalną do wektora falowego, składowa wzdłużna jest równoległa do wektora falowego. Detekcja nowych składowych wymaga zastosowania specjalnych elektrod (rys. 6).

Konwolator THC ma tą zaletę, że nie posiada metalizacji na drodze propagacji AFP i dlatego można w nim obniżyć udział efektów niepożądanych powstających przez odbicia fal.

Początkowo prace związane ze splotami sygnałów technologią akustyczną, wykorzystywały raczej fale objętościowe zamiast powierzchniowych, jednakże okazało się, że większe gęstości mocy można osiągnąć, stosując te drugie, dlatego obecnie stosowane układy wykorzystują wyłącznie fale powierzchniowe lub przypowierzchniowe (SSBW) charakteryzujące się gęstościami mocy tego samego rzędu, co AFP.

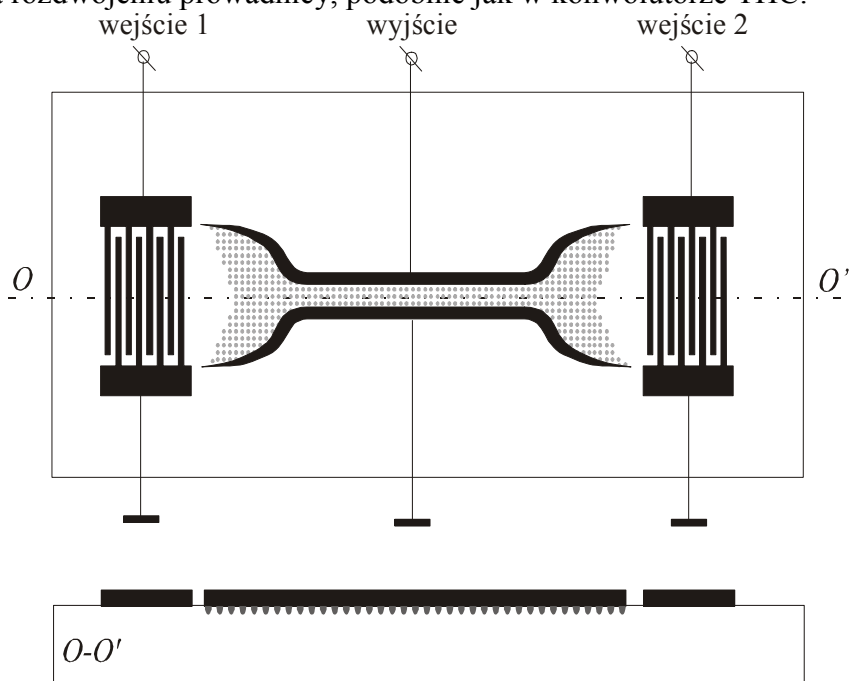
Sprawność konwolatora THC charakteryzowana jest za pomocą wielkości M , definiowanej jako:

$$M = \frac{V \cdot a}{P}$$

gdzie V jest napięciem wyjściowym układu w stanie rozwarcia, a szerokością obszaru oddziaływania wzajemnego, zaś P średnią geometryczną mocy akustycznej dwóch fal wejściowych.

Efektywność działania konwolutora jest, jak już wiadomo, przede wszystkim uzależniona od stopnia nieliniowości podłoża, ten zaś ujawnia się tym silniej im większa jest energia oddziaływujących ze sobą fal. Stąd dążenia do kumulacji tejże energii. Kumulację taką uzyskano np. stosując pokazaną wcześniej strukturę typu horn.

Konwolutor pokazany na rys. 7 zawiera podobną strukturę z dodatkową modyfikacją polegającą na rozdwojeniu przewodnicy, podobnie jak w konwolutorze THC.



Rys. 7. Konwolutor z przestrzenną kumulacją energii; widok od góry i przekrój

Wewnątrz przewodnicy, dzięki wytworzeniu odpowiedniej makrostruktury podłoża w jej obrębie propagują się głównie fale poprzeczne i SSBW. Wytworzona makrostruktura posiada także zdolność przechwytywania modów przypowierzchniowych i transformowania ich do modów powierzchniowych typu poprzecznego. Tym samym pozwala ona na kumulację energii fal mechanicznych w płaszczyźnie prostopadłej do powierzchni podłoża. Taka kumulacja energii w trzech wymiarach umożliwia znaczne zwiększenie efektywności oddziaływań nieliniowych w podłożu.

Makrostruktura tego typu może być uzyskana na drodze dyfuzji bądź za pomocą implantacji jonów. Szczególnie silny wpływ na własności sprężyste podłoży używanych w akustoelektronice przejawiają jony ziem rzadkich oraz jony grupy żelaza.

4.6. Procesory Fouriera

Są to pracujące w czasie rzeczywistym układy obliczające w sposób analogowy transformaty Fouriera sygnałów szerokopasmowych (najczęściej o specjalnym przeznaczeniu). Oferują one liczbę punktów transformacji większą niż 2^{10} . W porównaniu z cyfrowymi procesorami *FFT*, procesory analogowe z AFP pracują w czasie rzeczywistym, z dużo szerszym pasmem, z niższym poborem mocy i z dużo większą niezawodnością. Charakteryzują się też małymi (wręcz mikroskopowymi) rozmiarami. Wadą procesorów Fouriera z AFP jest nieco niższa dokładność ($\sim 1\%$) oraz ograniczony szumem zakres dynamiczny (60 – 70 dB).

Jak wiadomo z teorii filtrów z AFP odpowiedź impulsowa filtru w dziedzinie czasu odpowiada jego charakterystyce częstotliwościowej:

$$h(t) \Leftrightarrow H(\omega).$$

Jeżeli odpowiedź impulsowa filtru z AFP przyjmie postać:

$$h(t) = e^{j\mu t^2}, \quad (1)$$

gdzie $\mu = B/T$ jest nachyleniem charakterystyki fazowej filtru z liniową modulacją częstotliwości (B - pasmo, T - czas trwania odpowiedzi impulsowej), to z twierdzenia o splocie otrzyma się wyrażenie na postać sygnału wyjściowego w postaci:

$$s_{wy}(t) = \int_{-\infty}^{\infty} s_{we}(\tau) e^{j\mu(t-\tau)^2} d\tau = e^{j\mu t^2} \int_{-\infty}^{\infty} s_{we}(\tau) e^{j\mu\tau^2 - j2\mu t\tau} d\tau \quad (2)$$

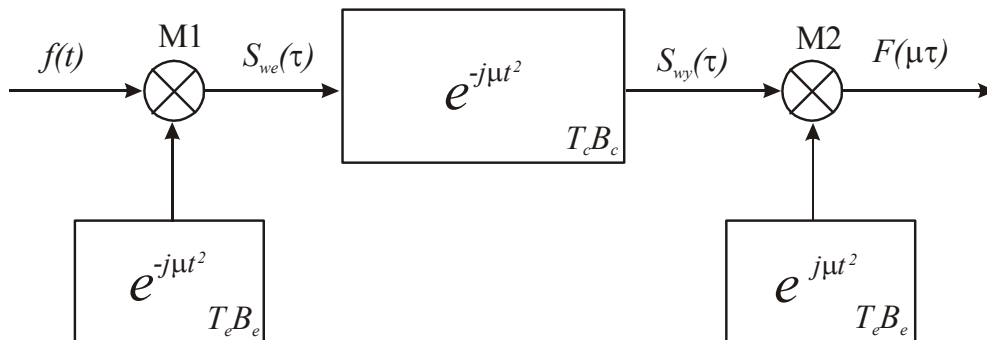
Podobieństwo między transformacją (2) a transformacją Fouriera sygnału $f(t)$ w postaci:

$$F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-j\omega t} dt \quad (3)$$

można pokazać przez dokonanie podstawienia $\omega = 2\mu\tau$ oraz rozpisanie $-2\omega t$ na postać $(t - \omega)^2 - t^2 - \omega^2$. Przy założeniu liniowości relacji czasowo częstotliwościowych można wtedy otrzymać:

$$F(\omega) = F(\mu t) = e^{-j\mu t^2} \int_{-\infty}^{\infty} f(\tau) e^{-j\mu\tau^2} e^{j\mu\{t-\tau\}^2} d\tau \quad (4)$$

co sugeruje, że wymnożenie wstępnie sygnału wejściowego $f(t)$ z przebiegiem posiadającym LMC, następnie splot w filtrze z AFP oraz kolejne końcowe mnożenie z przebiegiem z LMC, da w wyniku transformatę Fouriera o postaci (3). Urządzenie realizujące te operacje pokazano na rys. 1.

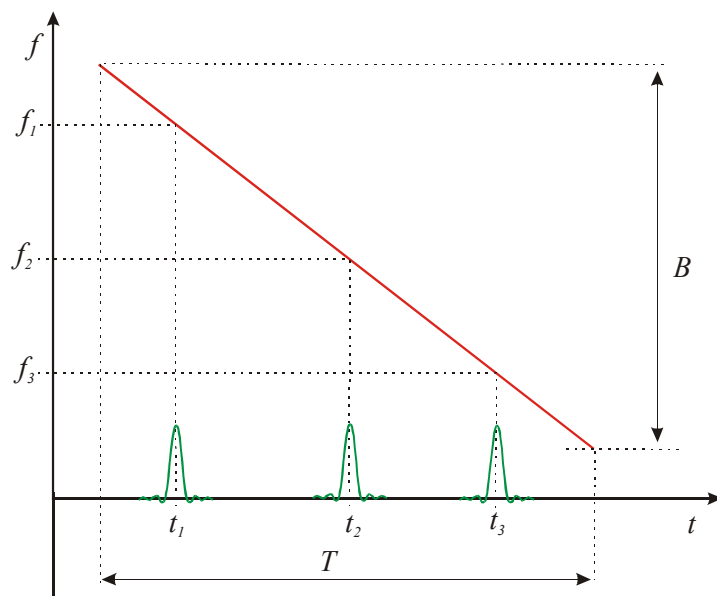


Rys. 1. Schemat prostego procesora Fouriera

Jak widać na rys. 1 procesor Fouriera powinien zawierać przynajmniej trzy podzespoły z AFP kształtujące nieliniową charakterystykę fazową.

Wejściowy sygnał $f(t)$ jest mieszany w mieszaczu M1 z sygnałem z LMC o nachyleniu charakterystyki fazowej $-\mu$. Sygnał $S_{wy}(\tau)$ jest komprimowany w filtrze dyspersyjnym o nachyleniu charakterystyki fazowej μ . Jeżeli sygnał analizowany $f(t)$ zawiera składowe częstotliwości f_1, f_2, f_3, \dots to sygnał $S_{wy}(\tau)$ będzie zawierał serię sygnałów skomprimowanych typu $\text{sinc}(x)$, których amplituda będzie proporcjonalna do transformaty Fouriera sygnału $f(t)$. Sygnały skomprimowane, odpowiadające poszczególnym składowym częstotliwości, będą przesunięte względem siebie w czasie. Sygnał $S_{wy}(\tau)$ można wtedy zaobserwować na oscyloskopie, przy czym skala czasu będzie odpowiadała skali częstotliwości ($\omega = 2\mu t$). Współczynnikiem proporcjonalności jest nachylenie charakterystyki fazowej μ filtru

dyspersyjnego. Przesunięcie czasowe sygnałów skompresowanych w filtrze kompresyjnym dla dwóch przykładowych składowych częstotliwości badanego sygnału przedstawiono na rys. 2.



Rys. 2. Czasowe rozdzielanie sygnałów skompresowanych dla składowych f_1, f_2 i f_3 sygnału wejściowego

W wielu zastosowaniach wymagana jest jedynie amplitudowa charakterystyka sygnału. W takich przypadkach informację tą będzie już zawierał sygnał $S_{wy}(\tau)$, co oznacza możliwość rezygnacji z mieszacza M2 oraz związanego z nim drugiego filtra dyspersyjnego i znacznego uproszczenia układu.

Stosowane linie dyspersyjne mają skończony czas trwania odpowiedzi impulsowej T oraz skończoną szerokość pasma B . Z tego powodu rozdzielczość częstotliwościowa takiego procesora będzie miała postać:

$$\Delta f_s = \frac{1}{\min(T_e, T_c)} \quad (5)$$

Optymalnym pasmem pracy układu jest zakres, w którym spełniony jest warunek:

$$T_e = \frac{T_c}{2}, \quad (6)$$

zatem:

$$T_e B_e = \frac{T_c B_c}{4}. \quad (7)$$

Szerokość pasma procesora wynosi:

$$BW_s = \mu(T_c - T_e), \quad (8)$$

i dla warunków optymalnych uzyskuje się:

$$BW_s = \frac{B_c}{2}. \quad (9)$$

Maksymalna liczba punktów transformacyjnych jest związana z iloczynem czas-pasma filtru splatającego i jest tym większa im większą wartość ma ten iloczyn.

Przedstawiony na rys. 1 schemat nie jest jedynym możliwym rozwiązaniem układowym. Architektury procesorów Fouriera mogą mieć inne konfiguracje, które na ogół złożone są z podukładów mnożących splatających i generujących sygnały z LMC. W zależności od wzajemnego połączenia tych bloków otrzymuje się następujące dwie architektury podstawowe:

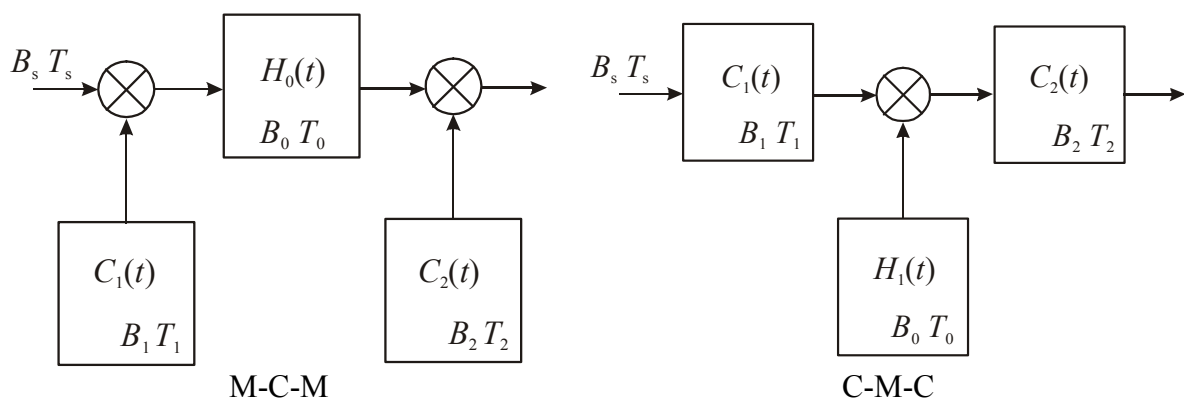
- M-C-M - mnożący – splatający – mnożący (układ z rys. 1),
- C-M-C - splatający – mnożący – splatający.

Architektura M-C-M posiada dwie konfiguracje. W pierwszej z nich okres czasu sygnału z LMC wykorzystywanego do mnożenia wstępnego jest krótki w porównaniu z impulsowymi sygnałami wyjściowymi filtra splatającego (architektura $M(s)-C(l)-M$), w drugiej natomiast okres trwania przebiegu z LMC wykorzystywanego do mnożenia wstępnego jest długi (układ $M(l)-C(s)-M$). Notacja l oznacza tu przebieg z LMC o długim okresie (ang. long), a s oznacza przebieg o krótkim okresie (ang. short).

Stwierdzono, że w schemacie $M(l)-C(s)-M$, mnożenie końcowe nie pozwala uzyskać pożądanych składowych transformaty Fouriera w związku z czym taka architektura może być wykorzystana tylko do analizy widma mocy sygnałów, których składowe pozostają stałe przez okres przebiegu wykorzystywanego do mnożenia wstępnego.

Podstawowe równania dla architektury C-M-C można wyznaczyć z równań obowiązujących dla architektury M-C-M poprzez zamianę operacji mnożenia na operacje splotu i odwrotnie.

Rys. 3 wraz z tabelą ilustrują dualizm tych dwóch architektur.



	M-C-M	C-M-C
Maksymalny czas trwania sygnału	T_1	$T_2=T_0-T_1$
Maksymalna szerokość pasma sygnału	$B_2=B_0-B_1$	B_1
Obciążenie pasma transformaty wyjściowej dla	$B_s>B_0-B_1$	$B_s>B_1$
Obciążenie czasu trwania sygnału wejściowego dla	$T_s>T_1$	$T_s>T_0-T_1$

Rys. 3. Porównanie architektur M-C-M i C-M-C

Konfiguracja C-M-C jest wyborem optymalnym ze względu na rozdzielczość częstotliwościową procesora i jego pasmo robocze. Wymagania stawiane wartościom iloczynu czasowo-częstotliwościowego układów splatających w tej architekturze są mniej ostre niż w architekturze M-C-M przy jednakowych stratach, tym samym stosunku sygnału do szumu oraz zakresie dynamicznym obu konfiguracji procesora. Wadą architektury C-M-C jest potrzeba stosowania trzech układów AFP, nawet do analizy widma, podczas gdy w przypadku struktury M-C-M potrzebne są tylko 2 układy.

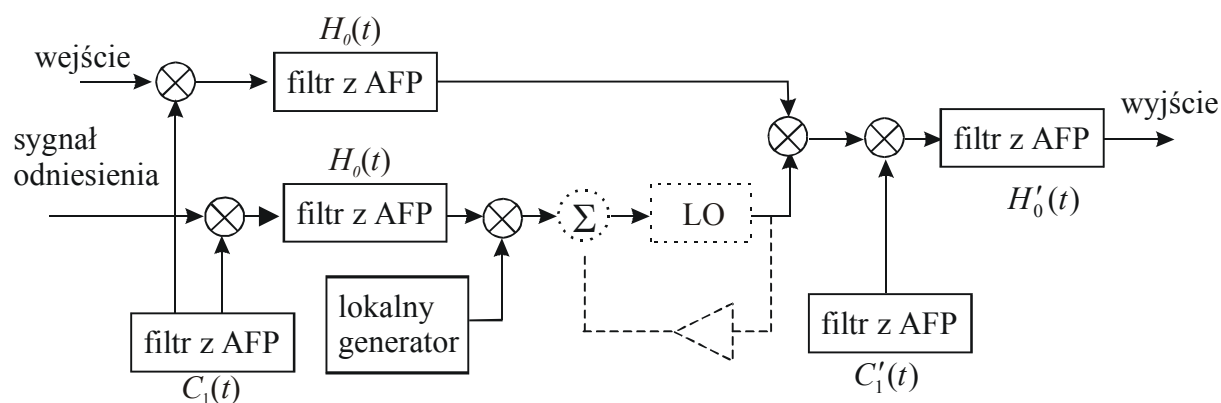
Za pomocą układów z AFP można też realizować inne operacje takie jak transformacja Hadamarda, transformacja Fresnela czy też transformacja falkowa.

4.7. Korelatory

Korelacja w dziedzinie czasu może zostać obliczona poprzez mnożenie transformat dwóch sygnałów w dziedzinie częstotliwości, a następnie obliczenie odwrotnej transformaty wyniku tego mnożenia. Układ realizujący takie operacje (korelator) daje możliwość filtracji dopasowanej do wielu typów sygnałów.

Przebieg odniesienia wykorzystywany w korelatorze może zawierać wszelkie potrzebne kompensacje. Dodatkową zaletę stanowi fakt dostępu do widma sygnałów pozwalającego na ich korekcję i kompensację w celu tłumienia listków bocznych i usuwania interferencji.

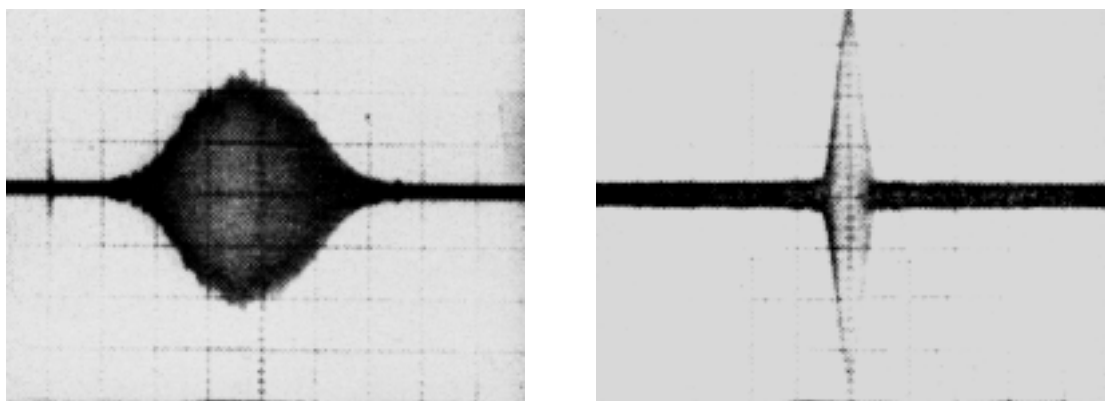
Na rys. 1. pokazano schemat blokowy prostego programowalnego korelatora, wykorzystującego procesory Fouriera z AFP.



Rys. 1. Programowalny korelator

W urządzeniu tym dwa identyczne procesory generują transformaty sygnału odbieranego i sygnału odniesienia. Generator lokalny w torze sygnału odniesienia zmienia częstotliwość sygnału i poprzez dobór odstępu częstotliwościowego, tworzy jego zespoloną wartość sprzężoną.

Rys. 2 przedstawia sygnały zaobserwowane za pomocą oscyloskopu w rzeczywistym korelatorze wykorzystującego transformaty sygnału z LMC.



Rys. 2. Po lewej: sygnał wejściowy korelatora o czasie trwania $T = 10$ ms zajmujący pasmo $B = 20$ MHz, skala $2 \mu\text{s/dz}$, po prawej sygnał wyjściowy reprezentujący autokorelację sygnału wejściowego, skala 200 ns/dz

Ważną, z punktu widzenia systemów radarowych, zaletą przedstawionego urządzenia jest możliwość wykorzystania jako przebiegu odniesienia sygnału z LMC o postaci identycznej z

tą, jaka jest emitowana przez system radarowy. Wtedy wszelkie błędy powstałe w wyniku pracy systemu będą automatycznie kompensowane. Kompensowane będą ponadto błędy powstałe w wyniku pracy układu mnożącego. Błędy te są równo podzielone pomiędzy transformatę Fouriera sygnału i transformatę sygnału interferującego oraz wszystkie dopasowane filtry splatające co w rezultacie powoduje, że błędy wprowadzone do wyjściowego sygnału korelacji są zminimalizowane (za wyjątkiem błędów powstałych przy obliczaniu odwrotnej transformaty Fouriera). Wykazano jednak, że błędne wyniki pracy korelatora pracującego z sygnałem odniesienia z LMC są mniej szkodliwe niż błędy pracy pętli kompresji impulsu prostego.

Dodanie pamięci do układu programowalnego korelatora pozwala na zapisanie informacji o sygnale odniesienia do przeprowadzenia korelacji z opóźnionymi kopiami powracającego sygnału. Taką pamięć można zrealizować stosując recyrkulacyjną linię opóźniającą AFP, co daje możliwość dalszej redukcji błędów. Próbkę sygnału z anteny systemu radarowego poddana transformacji i zachowana w pamięci umożliwia korelacyjną redukcję błędów i zniekształceń wprowadzonych przez stopnie przemiany i wzmocnienia systemu.

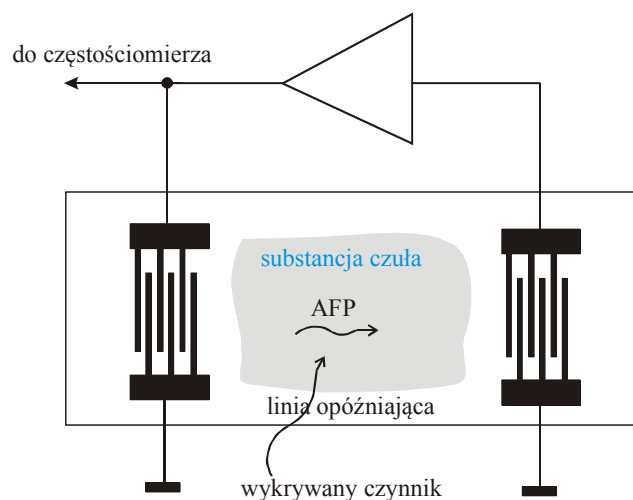
4.8. Czujniki

Interesującym zastosowaniem układów z AFP są czujniki wielkości elektrycznych, a szczególnie nieelektrycznych. Pierwsze konstrukcje czujników pracowały z falami objętościowymi, jednak dążenie do zwiększenia ich częstotliwości pracy powodowało stosowanie coraz cieńszych kryształów, co w konsekwencji doprowadziło do przejścia na struktury z AFP, które mogą pracować nawet na częstotliwościach z zakresu dolnych mikrofal i których podstawową zaletą jest większa czułość w porównaniu do czujników z falami objętościowymi.

Za pomocą układów z falą powierzchniową daje się wykrywać takie wielkości jak:

- naprężenia mechaniczne,
- przyspieszenie,
- temperaturę,
- wilgotność,
- stężenie różnorodnych substancji chemicznych (np. NO₂),
- stężenie różnorodnych substancji biologicznych,
- inne.

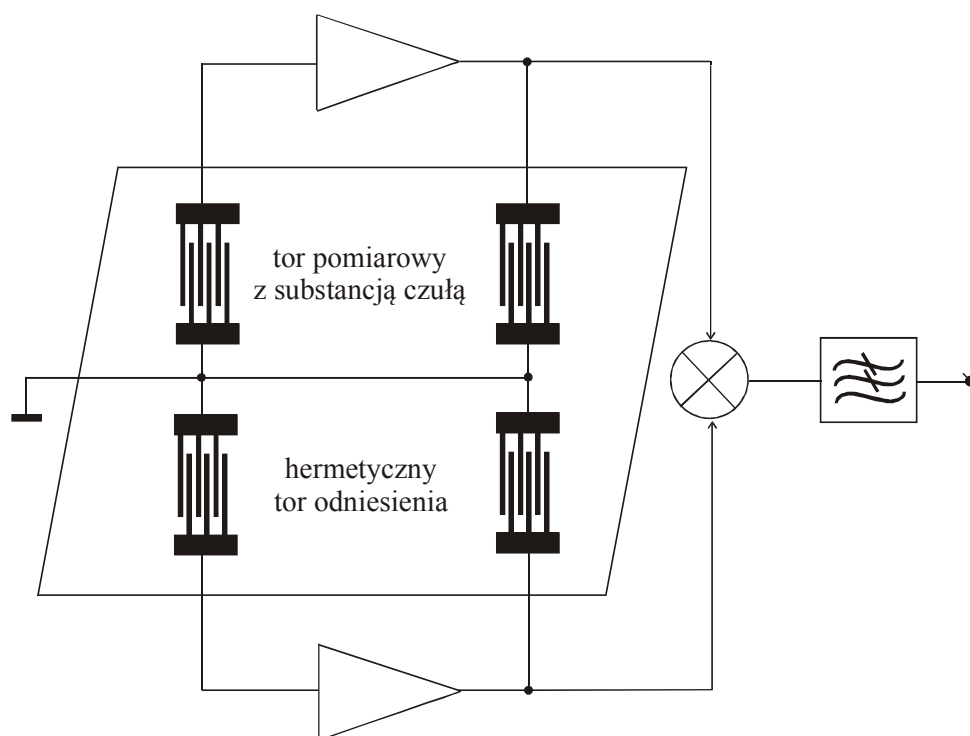
Najprostszym rozwiązaniem czujnika jest układ przedstawiony na rys. 1.



Rys. 1. Układ czujnika z AFP

W układzie tym linia opóźniająca wraz ze wzmacniaczem tworzą układ generacyjny. Na drodze propagacji AFP znajduje się odpowiedni czynniki czuły na mierzony parametr. W przypadku czujników substancji chemicznych jest to warstwa chemoczuła, zaś w przypadku czujników przyspieszeń może to być masa bezwładna. Badany czynnik oddziałując na linię opóźniającą (zwykle poprzez zmianę obciążenia masowego, naprężeń lub warunków elektrycznych na powierzchni) zmienia częstotliwość drgań oscylatora, która mierzona jest w mierniku częstotliwości.

Do celów profesjonalnych budowane są czujniki skompensowane pracujące z liniami odniesienia i generujące sygnał różnicowy. Jeden z wariantów takiego rozwiązania jest pokazany na rys. 2.



Rys. 2. Skompensowany czujnik z AFP

Przedstawiony na rys. 2 układ czujnika składa się z dwóch identycznych linii opóźniających. Jedną z nich, nie pokrytą substancją czułą stanowi tor odniesienia, natomiast druga, pokryta odpowiednią substancją, tor pomiarowy. Sygnał wyjściowy jest sygnałem różnicowym. Powodowana czynnikiem mierzonym zmiana prędkości AFP powoduje zmianę czasu opóźnienia w linii, co przekłada się na zmianę częstotliwości w obwodzie, w którym linia jest objęta pętlą sprzężenia zwrotnego. Uzyskany na wyjściu układu sygnał różnicowy będzie zatem proporcjonalny do zmiany prędkości fali w torze pomiarowym, a tym samym do wielkości czynnika wykrywanego przez czujnik.

Prędkość fali powierzchniowej jest zależna od następujących parametrów:

$$\frac{\Delta v}{v_0} = \frac{1}{v_0} \left[\left(\frac{\delta v}{\delta m} \right) \Delta m + \left(\frac{\delta v}{\delta C} \right) \Delta C + \left(\frac{\delta v}{\delta \sigma} \right) \Delta \sigma + \left(\frac{\delta v}{\delta \varepsilon} \right) \Delta \varepsilon + \left(\frac{\delta v}{\delta T} \right) \Delta T + \left(\frac{\delta v}{\delta p} \right) \Delta p \right], \quad (1)$$

gdzie v_0 jest prędkością fali powierzchniowej na powierzchni swobodnej, m masą obciążającą powierzchnię, C odpowiednią stałą sprężystości, σ przewodnictwem elektrycznym, ε przenikalnością elektryczną, T temperaturą, zaś p ciśnieniem na powierzchni.

Zmiana prędkości fali powierzchniowej związana jest przede wszystkim z obciążeniem masowym i elektrycznym powierzchni kryształu. Dla obciążenia masowego:

$$\frac{\Delta v}{v_0} = -nC_m m d, \quad (2)$$

gdzie C_m jest współczynnikiem mechanicznego sprzężenia podłoża z warstwą, n liczbą molekuł zakumulowanych w warstwie, m masą molową molekuł zakumulowanych, d grubością warstwy adsorpcyjnej.

Potencjał elektryczny związany z AFP może oddziaływać z ładunkami w warstwie czulej. W wyniku tego oddziaływania następuje zmiana prędkości AFP (obciążenie elektryczne):

$$\frac{\Delta v}{v_0} = -\frac{k^2}{2} \frac{\sigma^2}{\sigma^2 + v_0^2 C_s^2} \quad (3)$$

gdzie: k^2 - współczynnik sprzężenia elektromechanicznego, σ - przewodnictwo elektryczne, C_s - pojemność elektryczna jednostkę długości.

Zmianę częstotliwości drgań generatora z rys. 1. opisuje zależność:

$$\frac{\Delta f_0}{f_0} = \frac{\Delta v}{v_0} \frac{\tau}{\tau_0} \quad (4)$$

gdzie: τ - czas propagacji w obszarze oddziaływania z czynnikiem mierzonym, τ_0 - całkowity czas opóźnienia AFP. Dokładność pomiaru zmian prędkości w układzie różnicowym może być rzędu pojedynczych procentów, dla zmian częstotliwości rzędu 10^{-6} .

W czujnikach chemicznych z AFP wykorzystuje się zjawisko zmiany prędkości rozchodzenia się fali na podłożu piezoelektrycznym pokrytym warstwą chemiczną na skutek adsorpcji przez tę warstwę cząstek określonego gazu. Zmiana masy lub przewodnictwa elektronowego jest proporcjonalna do zmiany stężenia określonej substancji w środowisku pracy czujnika. Substancje wykorzystywane jako receptory czujników chemicznych powinny charakteryzować się szeregiem cech, takich jak: czułość, selektywność, trwałość, krótki czas odpowiedzi, odwracalność reakcji itd. Często jednak nie wszystkie wymagania mogą być spełnione jednocześnie.

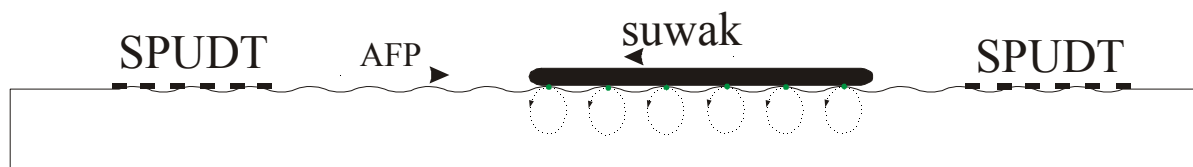
W poniższej tabeli zebrano kilka przykładów z licznej grupy substancji nadających się do konstrukcji czujników gazów z AFP.

substancja czuła	wykrywany czynnik
SnO ₂ , ZnO, TiO ₂ , Ta ₂ O ₅ , IrO _x , Ga ₂ O ₃ , WO ₃ , NiO	O ₂ , H ₂ O, CO, NO _x , CH _x , H ₂ S, CN, HN ₃ , (CH ₃) ₃ N, R-OH
ZrO ₂ +CaO, MgO, Y ₂ O ₃ , CeO ₂ +CaO, MgO, BiO ₂ -MoO ₃ , PbSnF ₄ , AgCl + AgS ₂	O ₂ , CO, H ₂ , H ₂ O, Cl ₂ , I ₂ w gazach i cieczach H ⁺ , K ⁺ , Na ⁺ , Cu ²⁺ , Pb ²⁺ , La ³⁺ , Ag ⁺ , F ⁻ , Cl ⁻ , Br ⁻ , J ⁻ w elektrolitach ciekłych
LiCl, Ba _{1-x} Sr _x TiO ₃ +MgCr ₂ O ₄ , MgCr ₂ O ₄ +TiO ₂	wilgoć

Zastosowanie enzymów pozwala też wykrywać złożone substancje organiczne, a nawet przeciwciała stąd możliwe jest wykorzystanie układów z AFP także w diagnostyce medycznej.

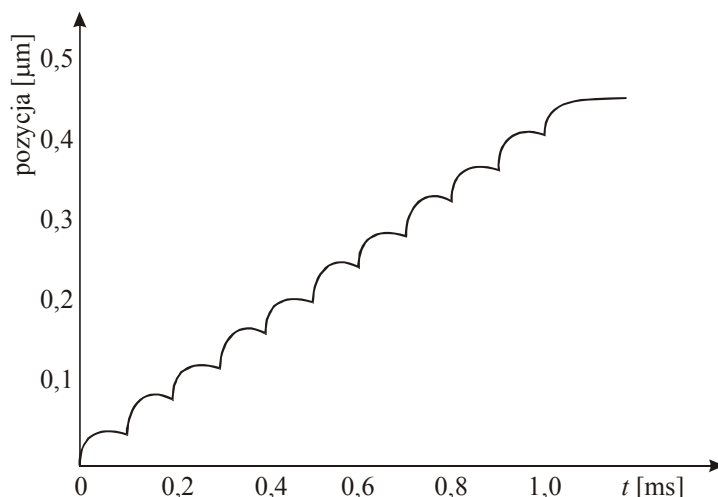
4.9. Silniki

Powszechna miniaturyzacja dosięga też układów mechanicznych takich jak silniki. Istnieje bardzo duże zapotrzebowanie na mikrosilniki, które znajdują zastosowania w takich urządzeniach jak kieszonkowe aparaty fotograficzne, samoogniskujące minikamery, sztuczne organy do przeszczepów, minisamoloty i in. Zmniejszanie silników indukcyjnych ma jednak swój kres, co związane jest z ich znaczną komplikacją i koniecznością wykonywania uzwojeń. Z tego też powodu w wielu laboratoriach prowadzone są badania nad konstrukcją mikrosilników piezoelektrycznych. Okazało się, że do napędu można wykorzystać fale Rayleigha. Rys. 1. przedstawia schemat takiego silnika.



Rys. 1. Schemat działania silnika liniowego z AFP. Ze względu na specyfikę fali Rayleigha suwak porusza się w kierunku przeciwnym do fali

Jak wiadomo punkt materialny na powierzchni, po której propaguje się fala Rayleigha porusza się po elipsie w płaszczyźnie prostopadłej do powierzchni i równoległej do kierunku propagacji. Ruch ten daje się wykorzystać do kontrolowanego przesuwania (na skutek tarcia) płaskich obiektów położonych na powierzchni. Silnik wykonany na $Y 128^\circ \text{LiNbO}_3$, pracujący na częstotliwości 10 MHz jest w stanie poruszać suwakiem z prędkością 1 m/s, wytwarzając siłę rzędu kilku N. Osiągalna dokładność pozycjonowania suwaka jest rzędu nm. Suwak przemieszczany jest krokowo, co ilustruje rys. 2.

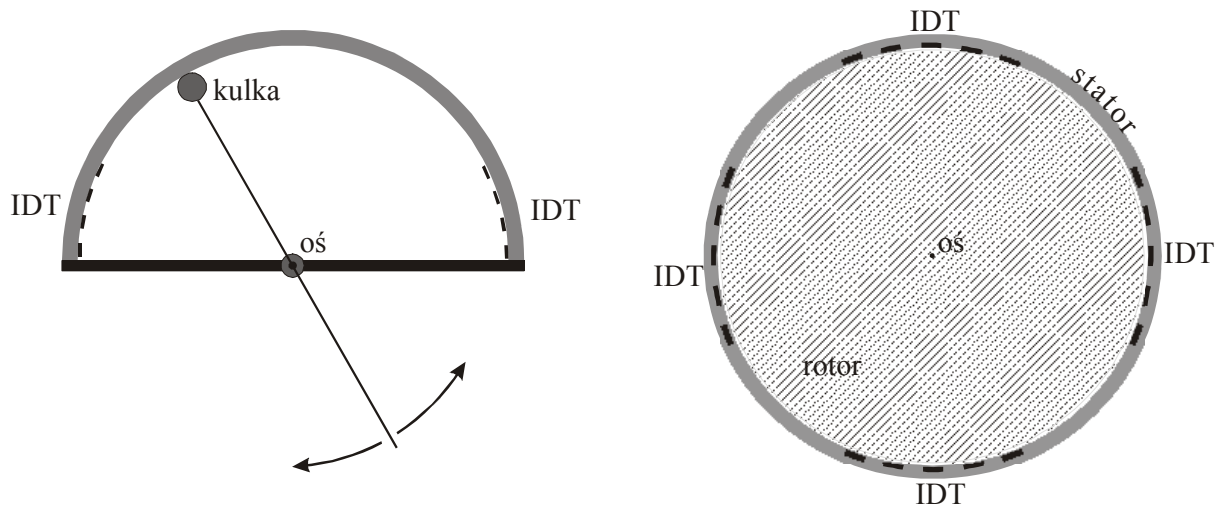


Rys. 2. Przemieszczanie się suwaka

Prędkość przemieszczania się suwaka jest regulowana poprzez zmianę napięcia na przetworniku IDT. Przy napięciu międzyszczytowym wynoszącym 170 V suwak porusza się z prędkością 1,1 m/s, zaś maksymalne przyśpieszenie suwaka dochodzi do 200 m/s^2 .

Szybkość poruszania się suwaka zależna jest od jego masy. W opisywanym silniku wynosiła ona 30 N.

Obok silników liniowych konstruuje się także silniki rotujące – przykładowe schematy uwidacznia rys. 3.



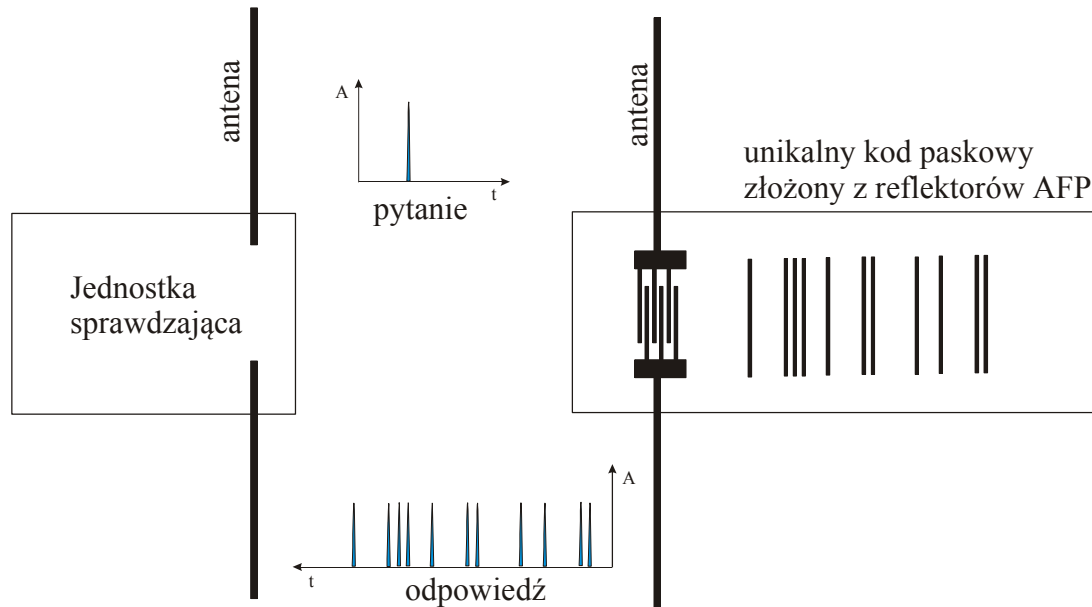
Rys. 3. Po lewej: silnik z obrotowym ramieniem. Po prawej silnik pracujący z przetwornikami „wkopanymi”. W obydwu przypadkach podłożem AFP jest kształtka wykonana z ceramiki piezoelektrycznej

Do głównych zalet silników ultrasonicznych z AFP należy zaliczyć:

- brak pola magnetycznego,
- mały pobór mocy (i tylko w trakcie ruchu),
- praktyczna bezawaryjność,
- możliwość pracy w próżni,
- wysoka czystość,
- możliwość pracy w ekstremalnie niskich temperaturach (nawet bliskich 0 K),
- generacja dużych sił/momentów sił (nawet do kilkunastu KN),
- bardzo wysoka precyzja ruchu,
- małe wymiary i niewielka masa,
- ultraniski poziom hałasu.

4.10. Układy zdalnej identyfikacji

Do interesujących zastosowań AFP należą też układy zdalnej identyfikacji. Z racji swej trwałej konstrukcji wykazują one w niektórych przypadkach cechy daleko korzystniejsze niż systemy magnetyczne czy optyczne, zaś nielegalna duplikacja takiego układu bez dostępu do mikrotechnologii jest praktycznie wykluczona. Schemat przykładowego systemu zdalnej identyfikacji przedstawia rys. 1.



Rys. 1. System zdalnej identyfikacji z AFP

Działanie systemu jest bardzo proste. Jednostka sprawdzająca wysyła impuls sondujący (grupę impulsów), który po odebraniu przez antenę dołączoną do przetwornika zamieniany jest na AFP. Po wielokrotnym odbiciu fali akustycznej od zestawu reflektorów stanowiącego unikalny kod, sygnał ponownie zamieniany jest na postać impulsów elektryczny i tą samą anteną odsyłany do jednostki sprawdzającej.

Jednostka taka może też współpracować z wieloma czujnikami wykonanymi w technice AFP, z których każdy wysyła razem z informacją o czynniku mierzonym swój indywidualny kod. System taki pozwala zdalnie monitorować duże obszary.

4.11. Urządzenia na podłożach modyfikowanych

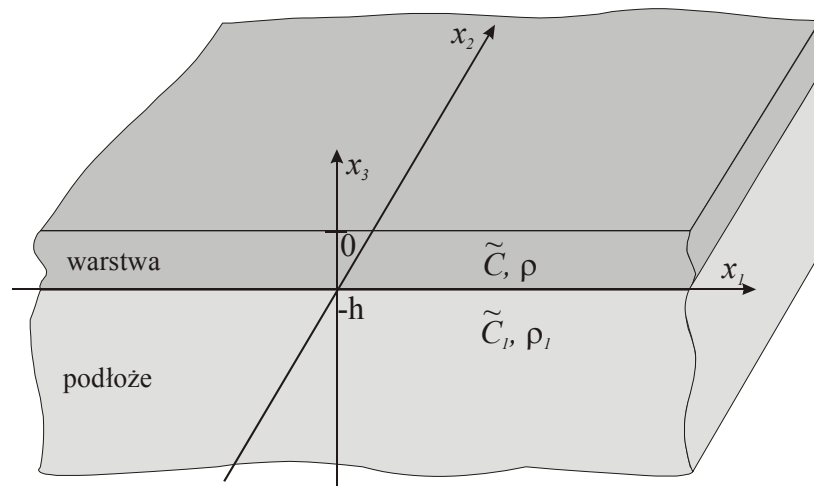
Kompozyty, kojarzone najczęściej z materiałami o podwyższonej wytrzymałości mechanicznej, stosowanymi w różnych gałęziach techniki, znajdują obecnie coraz szersze zastosowania we współczesnej elektronice. Łączenie różnych materiałów powoduje bowiem nie tylko zmianę własności mechanicznych powstającej w ten sposób całości, lecz także jej własności elektrycznych. Fakt ten stanowi jedną z głównych przesłanek intensywnego rozwoju całego szeregu dyscyplin zajmujących się teorią, technologią oraz zastosowaniem kompozytów (materiałów zaawansowanych) w różnych gałęziach techniki w tym również w akustoelektronice (piezokompozyty). Ta dziedzina współczesnej elektroniki bazuje bowiem zarówno na własnościach elektrycznych jak i mechanicznych materiałów, od których to własności zależne są parametry wykorzystywanych w niej fal akustycznych. Duże znaczenie kompozytów w tym zakresie jest też związane z dość silnym uzależnieniem akustoelektroniki od niewielkiej grupy materiałów piezoelektrycznych nadających się do zastosowań technicznych. Grupa ta jest jeszcze zawężona na skutek konieczności spełnienia różnorodnych dodatkowych wymagań (np. niskiej higroskopijności, wysokiej stabilności temperaturowej itp). Z tego powodu cały asortyment materiałów nadających się na podłoża akustoelektroniczne sprowadza się w zasadzie tylko do kilku kryształów. Wytwarzanie podłoży piezoceramicznych niewiele poprawia sytuację w tym zakresie ponieważ zwykle nie nadają się one do konstrukcji akustoelektronicznych układów pracujących na wyższych częstotliwościach. Piezokompozyty wytwarzane na bazie monokryształów pozwalają na wyjście poza to ograniczenie. Mają też one cały szereg innych bardzo cennych zalet. Przede wszystkim zaliczyć do nich należy, wynikającą z istoty budowy, możliwość kontrolowania ich własności elektromechanicznych, a nawet sterowania nimi, co daje dużo większą, w porównaniu z klasycznymi metodami, swobodę przy konstruowaniu działających na ich bazie urządzeń. Bardzo ważny jest także fakt, że niektóre typy piezokompozytów mogą oddziaływać z różnorodnymi rodzajami pól zewnętrznymi dając tym samym, bardzo cenną w elektronice, możliwość sterowania elektromechanicznymi własnościami podłoża, a tym samym, wykonanego w oparciu o nie całego urządzenia akustoelektronicznego.

Urządzenia, których zasada działania oparta jest o własności fal akustycznych (szczególnie w ich wariacie powierzchniowym - AFP) znalazły szereg ważnych zastosowań. Do najbardziej znanych należą, wykorzystywane w radiolokacji, układy filtracji dopasowanej takie jak linie dyspersyjne, konwolatory, korelatory oraz różnego typu filtry pasmowe, linie opóźniające i czujniki zarówno wielkości elektrycznych jak i nieelektrycznych.

Bardzo dobre parametry użytkowe tych urządzeń, jak np. małe wymiary, adaptacyjność, wysoka niezawodność i duża szybkość obróbki sygnałów oraz specyfika pełnionych funkcji w układach elektronicznych sprawia, że są one niezwykle użyteczne, a w wielu aplikacjach wręcz bezkonkurencyjne. Wykorzystanie do ich konstrukcji podłoży kompozytowych niewątpliwie znacznie podniesie ich użyteczność, a także rozszerzy zakres zastosowań. Do najbardziej interesujących wśród nich zaliczyć należy urządzenia lokalizacji i zobrazowania (także trójwymiarowego) obiektów znajdujących się w ośrodkach optycznie nieprzeźroczystych (tzw. all-weather radars) oraz spokrewnione z nimi układy przeznaczone do badań nieniszczących i diagnostyki medycznej.

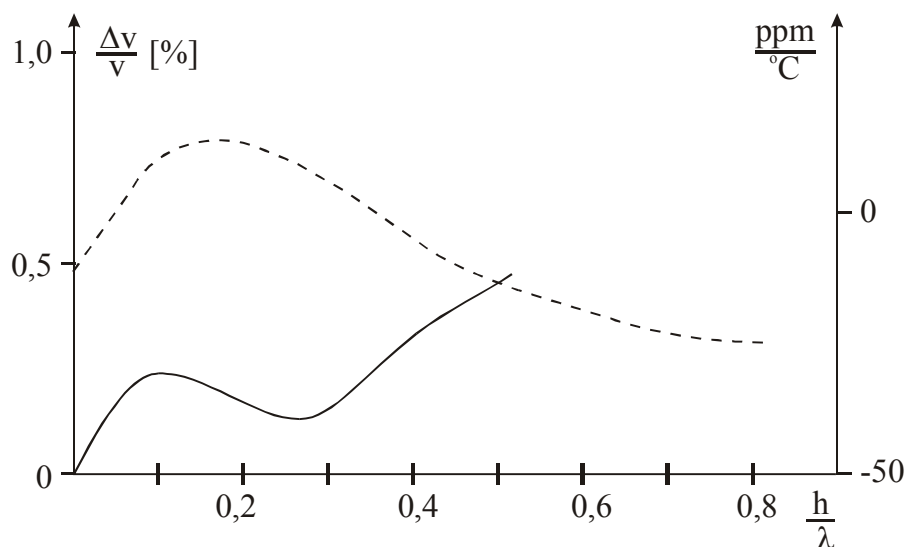
Kompozyty warstwowe

Do najlepiej poznanych i najstarszych rodzajów kompozytów należą ich odmiany warstwowe (laminaty). Materiały takie, w ogólności złożone są z n jednorodnych warstw o dobranej grubości nałożonych na podłoże bazowe. W najprostszym, przedstawionym na rys. 1, przypadku $n = 1$.



Rys. 1. Kompozyt jednowarstwowo

W strukturze takiej, w zależności od konfiguracji, struktury warstw mogą propagować się, obok modów Rayleigha także mody Love'a i Stonleya. Występowanie tychże modów jest silnie uzależnione od stosunku długości propagujących się fal do grubości warstwy (lub warstw). Tego typu kompozyty, obok zastosowań związanych z wykorzystaniem własności konkretnych rodzajów fal, stosuje się też w celu uzyskania pożądanych parametrów elektromechanicznych, a także temperaturowych podłoża. Rys. 2 przedstawia zmianę prędkości fazowej fali mechanicznej oraz temperaturowego współczynnika jej opóźnienia w zależności od stosunku grubości warstwy ZnO na powierzchni kwarcu do długości fali.

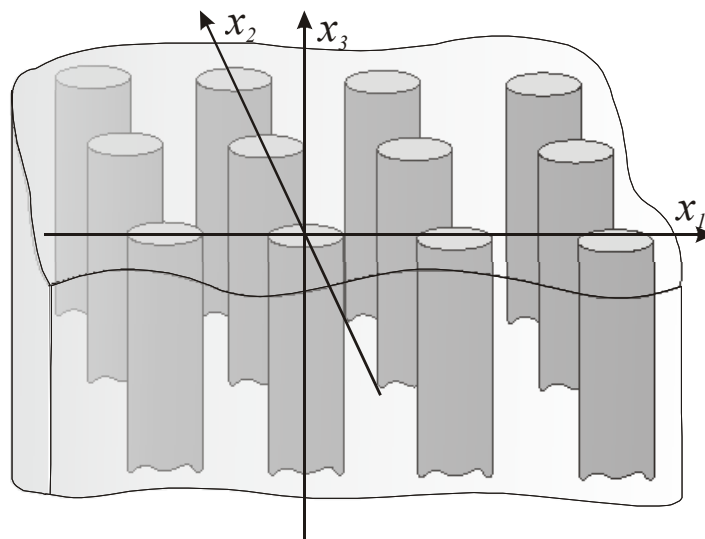


Rys. 2. Zmiana prędkości fazowej (linia ciągła) oraz temperaturowego współczynnika opóźnienia (linia przerywana) w zależności od stosunku grubości warstwy ZnO na powierzchni [111] Si do długości fali Rayleigha propagującej się w kierunku $(1\bar{1}1)$

Metody obliczania własności takich materiałów można znaleźć w wielu pozycjach literaturowych. Pozwalają one na określenie własności fal akustycznych propagujących się wzdłuż powierzchni kompozytu (najczęściej jednowarstwowego) oraz wyznaczyć odpowiednie relacje dyspersyjne. Istnieją także metody analizy elektromechanicznych własności kompozytów n – warstwowych. Jak dotąd nie zdołano jednak otrzymać ścisłych rozwiązań analitycznych, a uzyskane formuły mogą służyć jedynie jako podstawa do obliczeń numerycznych.

Kompozyty zbrojone

Interesującą wersję kompozytów stanowią struktury uzyskiwane poprzez wprowadzenie do podłoża układu prętów (najczęściej równoległych i równoodległych). Przykład takiego materiału przedstawia rys. 3.

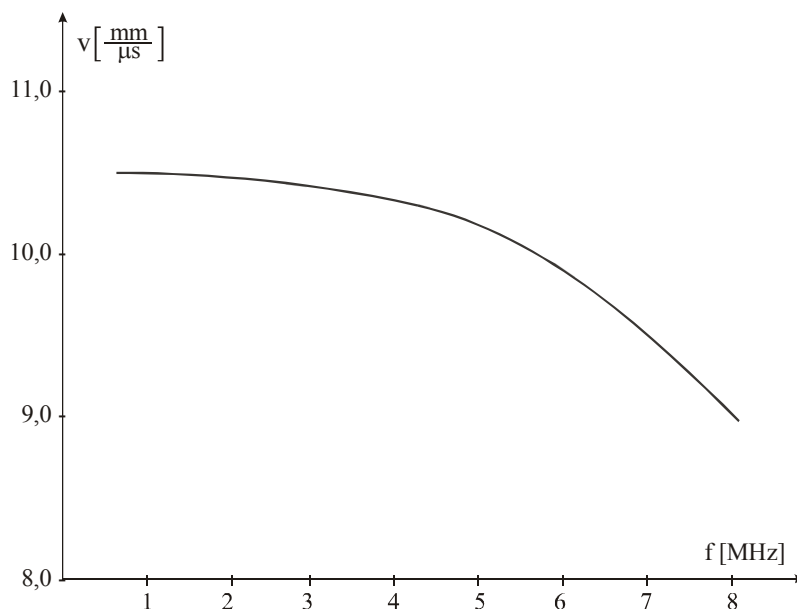


Rys. 3. Przykład kompozytu zbrojonego prętami prostopadłymi do jego powierzchni

O ile w kompozytach warstwowych pożądane parametry elektromechaniczne podłoża można uzyskiwać poprzez odpowiedni dobór rodzaju i grubości poszczególnych warstw, to w przypadku kompozytów zbrojonych dobierać można parametry geometryczne i materiałowe prętów oraz ich rozkład w płaszczyźnie podłoża, co daje projektującym większą, w porównaniu z laminatami, liczbę stopni swobody. Ta zaleta jest jednak okupiona, obok problemów natury analitycznej, dużymi trudnościami związanymi z technologią wytwarzania takich materiałów. Mimo to, kompozyty tego typu są wykonywane ze względu na ich unikalne własności elektromechaniczne. Możliwe jest także uzyskiwanie piezokompozytów na drodze wszczepiania prętów z materiału piezoelektrycznego do elektrycznie obojętnego podłoża. Zabieg taki powoduje nadanie całemu materiałowi własności piezoelektrycznych przy czym takie parametry jak współczynnik sprzężenia elektromechanicznego czy temperaturowy współczynnik opóźnienia zależą od rozkładu wszczepionych prętów oraz własności materiału, z którego są one wykonane. Od tych parametrów zależy także prędkość fal akustycznych zarówno objętościowych jak i powierzchniowych oraz impedancja akustyczna. Wielkości te mogą być zatem zmieniane np. pod kątem realizacji dopasowania energetycznego do otoczenia, z którym graniczy kompozyt.

Technika ta otwiera też możliwości tworzenia całkiem nowych materiałów o własnościach elektrostrykcyjnych czy też ferroelektrycznych.

Na uwagę zasługuje też typowa dla tego typu materiałów silna dyspersja, która może także znaleźć wiele zastosowań technicznych. Rys 3. przedstawia krzywą dyspersyjną dla modów wzdłużnych propagujących się po powierzchni kompozytu złożonego z prętów boru zatopionych w neutralnym polimerze.



Rys. 3. Krzywa dyspersyjna dla modów wzdłużnych propagujących się po powierzchni kompozytu złożonego z prętów boru zatopionych w polimerze. Pręty o średnicy $100 \pm 3 \mu\text{m}$ zajmują ok. 45 % objętości całego kompozytu

Struktury tego typu posiadają jeszcze jedną istotną właściwość polegającą na zdolności przechwytywania modów przypowierzchniowych, a nawet objętościowych i transformowania ich do postaci poprzecznych modów powierzchniowych. Właściwość ta jest bardzo cenna z punktu widzenia urządzeń z AFP. Wykorzystanie bowiem do ich konstrukcji podłoży kompozytowych daje możliwość częściowego odzyskiwania energii promieniowanej przez przetworniki do objętości podłoża i prowadzenia jej w obrębie cienkiej warstwy przypowierzchniowej. Cecha ta w zestawieniu z możliwością kontrolowania własności tej warstwy stwarza nowe możliwości zastosowań kompozytów zbrojonych do konstrukcji wysokoparametrowych układów obróbki sygnałów radiolokacyjnych.

Kwazikompozyty

Obok niewątpliwych zalet, kompozyty zbrojone posiadają kilka zasadniczych wad. Jak już wspomniano, możliwości ich wytwarzania silnie ograniczone są przez trudności natury technologicznej zawężające ilość dostępnych podłoży bazowych w zasadzie tylko do materiałów polimerowych, którymi zalewa się odpowiednio ustawiony rząd prętów. Niedogodność ta powoduje też silne, dolne ograniczenia wymiarów tychże prętów.

Wad tych, przy jednoczesnym zachowaniu większości zalet, pozbawione są tzw. kwazikompozyty uzyskiwane poprzez termiczną bądź kinetyczną dyfuzję specjalnie dobranych jonów do odpowiedniego podłoża. Szczególnie przydatna w tym zakresie jest

dyfuzja kinetyczna (implantacja jonów) ze względu na swobodę wyboru zarówno rodzaju jonów jak i podłoża.

Istnieje kilka mechanizmów umożliwiających tą drogą generację pożądaných zmian w materiale. Należą do nich m.in. przypowierzchniowe przejścia fazowe zarówno typu ferrosprężystego jak i ferroelektrycznego. Mogą być one indukowane, zależnie od rodzaju materiału podłoża, poprzez zewnętrzne pola elektryczne, pola naprężeń mechanicznych, a nawet zjawiska kapilarne. Wykazano, że efekt kapilarny może indukować niestabilności powierzchni, które mają miejsce przy określonej liczbie falowej modu powierzchniowego i często są stowarzyszone z formowaniem się „zamrożonej” mikrostruktury przypowierzchniowej.

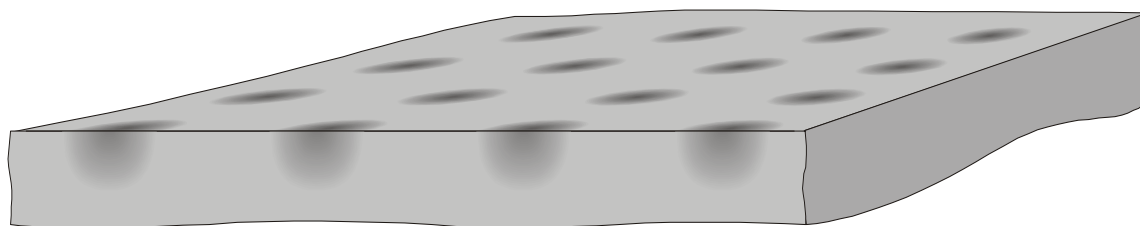
Do generacji wymienionych tu efektów szczególnie nadają się tzw. kryształy jahn-tellerowskie i spokrewnione z nimi roztwory stałe o własnościach ferroelektrycznych.

Niemniej jednak, co jest bardzo ważne w omawianym tutaj aspekcie, istnieje możliwość stymulowania określonych zmian właściwości elektromechanicznych konwencjonalnych podłoży AFP (jak np. LiNbO_3). Można je uzyskać poprzez wprowadzenie do przypowierzchniowej struktury sieci krystalicznej jonów pierwiastków grup przejściowych, które stanowią szczególnie efektywne centra dystorsji. Bardzo użyteczny w zastosowaniach technicznych może być w tym zakresie fakt, że ich własności, w odpowiednich warunkach, mogą być sterowane polowo. Możliwość ta stwarza olbrzymi potencjał wytwarzania nowych, np. napięciowo sterowanych urządzeń z AFP.

Kwazikompozyty, w zależności od sposobu implantacji mogą stanowić szczególny przypadek zarówno laminatów jak i kompozytów zbrojonych. Analogiem ostatnich z wymienionych są, uzyskiwane w procesie implantacji selektywnej obszary o zmienionych własnościach (inkluzyje) o określonym kształcie i rozkładzie.

Możliwości wykorzystania takich struktur są podobne do przedstawionych w poprzednim paragrafie kompozytów zbrojonych przy czym, w przypadku kwazikompozytów projektujący ma dodatkowo wpływ na kształt implantacyjnie indukowanych klasterów poprzez dobór energii, kąta implantacji oraz w przypadku implantacji multienergetycznej odpowiednich parametrów całego procesu.

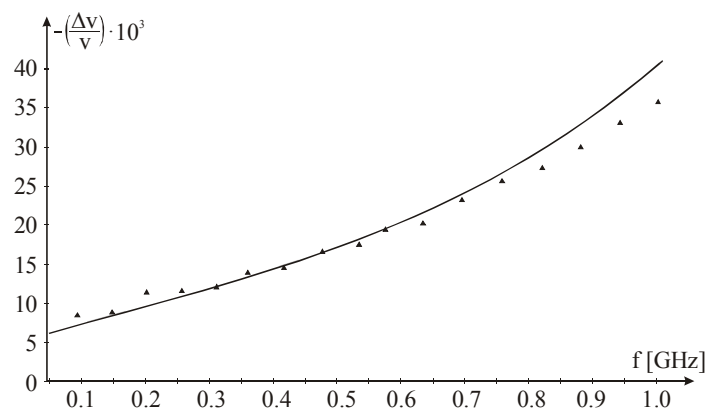
Przykład takiego kompozytu przedstawia rys. 4.



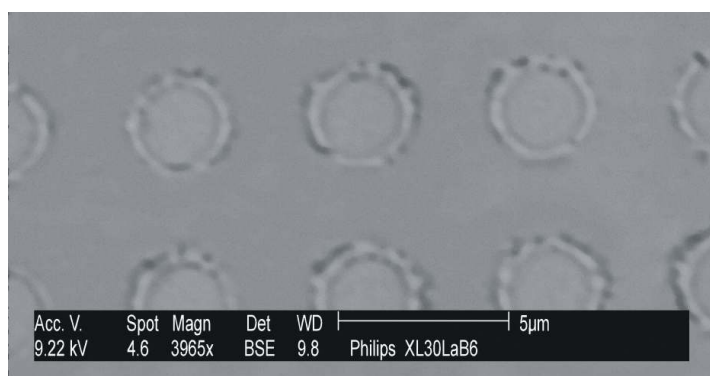
Rys. 4. Fragment przekroju kwazikompozytu. Obszary ciemniejsze symbolizują zmianę mikrostruktury podłoża w miejscach wprowadzenia domieszek

Jedną z ważniejszych wad kwazikompozytów przy zastosowaniach akustoelektronicznych stanowi ograniczenie zakresu pracy wykonanych na nich urządzeń do częstotliwości mikrofalowych. Fakt ten jest spowodowany niewielką osiągalną głębokością implantacji pociągającą za sobą stosunkowo płytkie zmiany parametrów mechanicznych podłoża, które mogą być „odczuwane” tylko przez fale dostatecznie krótkie. Inny ważny niedostatek tej technologii wiąże się z ograniczeniem dawki i energii implantowanych jonów. Jest ono

szczególnie silne w przypadku piezoelektryków. Nawet przy niewielkich energiach i dawkach jonów modyfikujących, na skutek generowania silnych naprężeń, dochodzi do ich pęknięcia. Na podstawie analizy najważniejszych własności kwazikompozytów wyznaczono, widoczną na rys. 5 krzywą dyspersyjną.



Rys. 5. Krzywa dyspersyjna dla modyfikowanego LiNbO₃. Linią ciągłą oznaczono przebieg teoretyczny, trójkątami zaś wartości zmierzone dla kompozytu przedstawionego na fot. 1



Fot. 1. Mikroskopowa fotografia fragmentu powierzchni badanego kompozytu. Średnica inkluzji wynosi 2,4 μm, period sieci inkluzji 4,8 μm. Kompozyt uzyskano drogą selektywnej implantacji jonów Nd⁺ o energii 100 keV, dawką 10¹⁶/cm² pod kątem 90° do płaszczyzny powierzchni podłoża. Materiał wykonano we współpracy z Instytutem Technologii Materiałów Elektronicznych

Perspektywy zastosowań kompozytów w radio i sonolokacji

Układy bazujące na propagacji akustycznych fal powierzchniowych znalazły już szereg zastosowań w urządzeniach radiolokacyjnych. Do najczęściej wykorzystywanych w tym zakresie należą różnego typu dyspersyjne linie opóźniające, konwolatory i korelatory pracujące w układach filtracji dopasowanej.

Zastosowanie podłoży kompozytowych do konstrukcji tych urządzeń może znacznie poprawić ich parametry techniczne, a także rozszerzyć zakres zastosowań.

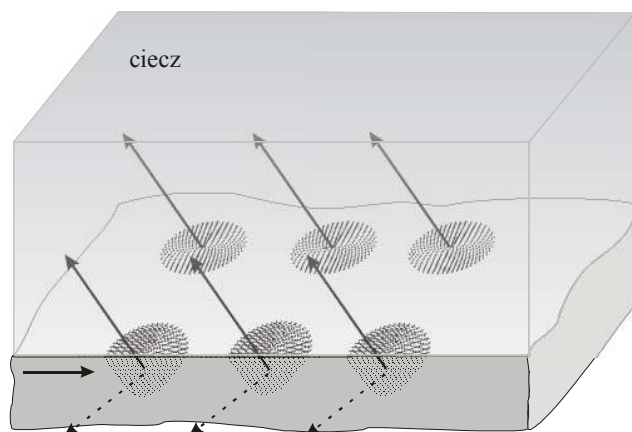
Podłoże kwazikompozytowe zastosowano też do konstrukcji konwolatora pracującego z poprzeczną falą powierzchniową. W porównaniu z klasycznymi rozwiązaniami wykorzystującymi fale Rayleigha charakteryzuje się on dużo lepszą sprawnością. Parametr ten ulega poprawie w związku z występowaniem silniejszych, w porównaniu z klasycznymi materiałami, oddziaływań nieliniowych w kompozycie. Ważną rolę w tym zakresie odgrywa

też efekt falowodowy w cienkiej, przypowierzchniowej warstwie kompozytu pozwalający uzyskanie wymaganej koncentracji energii. Współczynniki zanikania, propagującej się w takiej strukturze fali, w głąb podłoża zależą wyłącznie od parametrów wspomnianej warstwy. Wykorzystana może zostać także silna dyspersja podłoża kompozytowych. Konstrukcja konwencjonalnych układów kompresji/dekompresji sygnałów radiolokacyjnych bazuje na technice odpowiedniego kształtowania geometrii przetworników AFP. Z tego powodu długość klasycznych linii dyspersyjnej ograniczona jest rozmiarami dostępnych podłoży. W przypadku kompozytów ograniczenie to jest znacznie słabsze, a ponieważ dyspersja stanowi cechę samego tylko podłoża, możliwa staje się konstrukcja znacznie dłuższych, w porównaniu z konwencjonalnymi, linii dyspersyjnych. Wykorzystanie oddziaływania pól zewnętrznych ze strukturą przypowierzchniową (np. przy zastosowaniu kryształów jahn-tellerowskich) może też pozwolić na sterowanie ich parametrami częstotliwościowo - czasowymi.

Kompozyty są także wykorzystywane w urządzeniach sonolokacji tj. układach akustoelektronicznych pracujących na granicy z badanym ośrodkiem. Najczęściej jest to ośrodek ciekły m.in. ze względu na względną łatwość kontroli współczynnika odbicia fali na granicy cieczy z powierzchnią kompozytu i tym samym realizacji dopasowania energetycznego.

Również w tym przypadku ważne zastosowanie mogą znaleźć kompozyty zbrojne oraz kwazikompozyty. W takich bowiem strukturach każdy z wszczepionych prętów (lub każda inkluzja) może pełnić rolę elementu promieniującego, a ich odpowiedni rozkład będzie wtedy pełnił funkcję sfazowanego szyku antenowego. Zależności fazowe w takim układzie mogą być kształtowane poprzez dobór rozkładu prętów, a także sposobu ich zasilania (za pośrednictwem odpowiednio umieszczonego źródła lub źródeł drgań mechanicznych) przy czym rolę przesuwników fazy będzie pełniła sama struktura podłoża.

Podobnie jak poprzednio zastosowanie modyfikowanych kryształów jahn-tellerowskich może pozwolić na konstrukcję szyków antenowych o sterowanych połowo charakterystykach promieniowania. Przykład takiej struktury przedstawia rys. 6.



Rys. 6. Kwazikompozyt pracujący jako szyk antenowy na granicy z cieczą

Tego rodzaju kompozyty, obok strictly radiolokacyjnych, znajdują także zastosowania w defektoskopii oraz diagnostyce medycznej.

Podłoża kompozytowe mogą też zostać wykorzystane do konstrukcji miniaturowych linii opóźniających o połowo zmiennym opóźnieniu, które mogą stanowić podstawę konstrukcji szyków klasycznych anten elektromagnetycznych umożliwiając tym samym elektroniczną syntezę ich apertury.

4.12. Przykłady zastosowań

Przyrządy AFP można z punktu widzenia zastosowań zakwalifikować do czterech podstawowych grup:

- rezonatory i filtry;
- liniowe przyrządy z przetwornikami jednokierunkowymi;
- liniowe przyrządy z przetwornikami dwukierunkowymi;
- przyrządy nieliniowe.

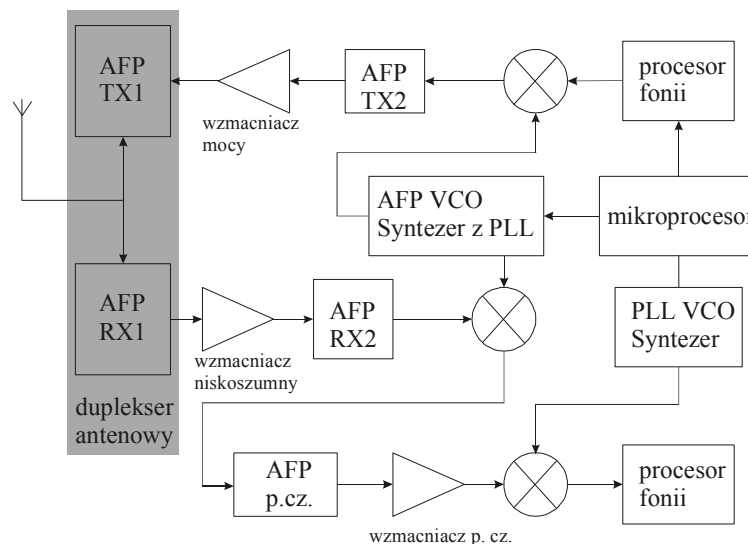
W grupie pierwszej mieszczą się takie układy jak dupleksery antenowe o mocy 2 – 4 W, filtry środkowoprzepustowe, filtry rezonansowe, układy generacji częstotliwości ze stałą bądź zmienną częstotliwością pracy.

W grupie drugiej można wymienić niskostratne filtry pośredniej częstotliwości (p.cz.), wielomodowe oscylatory (stosowane w komunikacji szerokopasmowej), niskostratne linie opóźniające przeznaczone dla odbiorników z podziałem czasowym.

Do grupy trzeciej zalicza się filtry Nyquista do radiowych odbiorników cyfrowych, oscylatory sterowane napięciem VCO pracujące w obwodach mieszaczy, przenośne urządzenia nadawczo – odbiorcze, linie opóźniające generujące i detekujące kody Barkera, regeneratory akustooptyczne pracujące w traktach światłowodowych.

W grupie czwartej lokują się synchroniczne i asynchroniczne konwoluty, korelatory, procesory Fouriera, procesory falkowe i wiele innych.

Najwięcej zastosowań i jednocześnie największą produkcję przyrządów z AFP notuje się w takich dziedzinach techniki jak: telewizja kolorowa, telekomunikacja przewodowa (filtry w centralach telefonicznych – gros produkcji Instytutu Technologii Materiałów Elektronicznych wykorzystywane jest właśnie w centralach telefonicznych), telefonia komórkowa. Na rys.1 przedstawiony jest przykładowy schemat telefonu komórkowego typu AMPS (Advanced Mobile Phone Service) używanego w USA z zaznaczonymi miejscami zastosowań przyrządów z AFP.

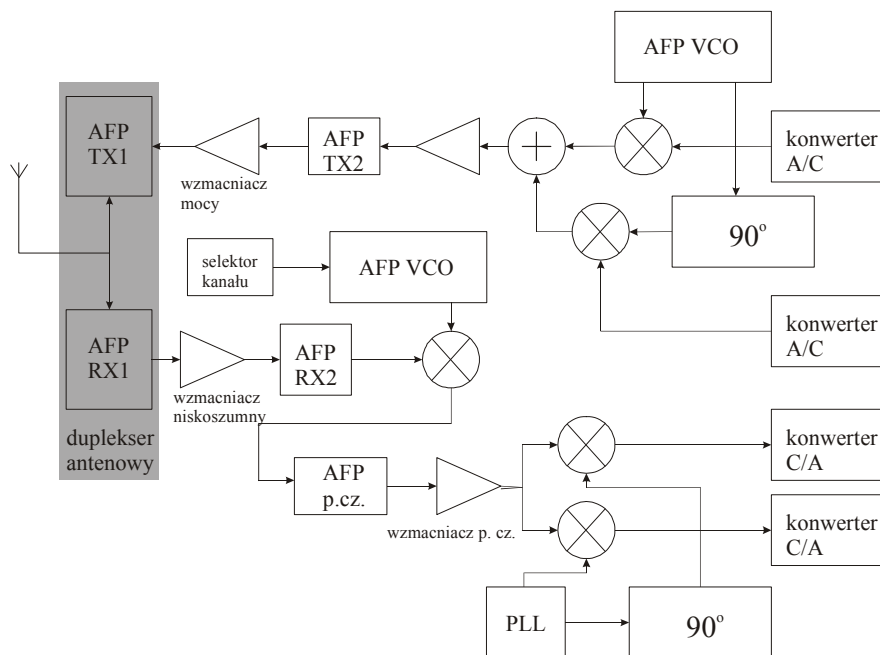


Rys. 1. Schemat telefonu komórkowego typu AMPS

Przedstawiony układ pracuje z modulacją wąskopasmową z podziałem częstotliwościowym. Pasma nadawcze i odbiorcze wynoszą odpowiednio 824 859 i 869 - 894 MHz. Mieszczą one 832 kanały rozmieszczone w odstępach 30 kHz. Duplexer antenowy widoczny na schemacie jest zrealizowany za pomocą struktur typu SPUDT pracujących z falą Rayleigha i częściowo z falami przypowierzchniowymi (typu SSBW), co zapewnia możliwość przenoszenia dużych

mocy. Filtr odbiorczy RX1 wprowadza niskie straty, mniejsze niż 3dB i posiada wysoką selektywność zabezpieczając przed przeciążeniem następujący po nim wzmacniacz niskoszumowy. Następujący po wzmacniaczu filtr RX2 jest filtrem rezonansowym którego zadaniem jest tłumienie niepożądanych harmonicznnych oraz szumów wzmacniacza. Dzięki takiemu rozwiązaniu możliwe staje się uzyskanie dynamiki rzędu 120 dB.

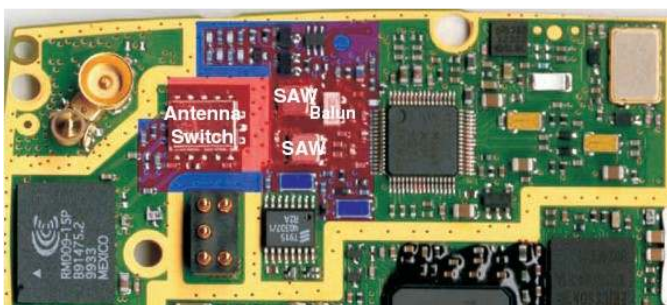
Filtr TX1 przenosi moc rzędu 30 dBm i wyższą. Filtr TX2 jest zbudowany podobnie jak RX2. Ma za zadanie tłumienie szumów. Układ z AFP współpracujący z VCO, w pierwszym mieszaczu najczęściej jest dwumodowym filtrem rezonansowym lub szerokopasmową linią opóźniającą. Ponieważ odstęp kanałów wynosi tylko 30 kHz, filtr p. cz. również wykorzystujący AFP musi być wysokoselektywny i stabilny temperaturowo. Typowo jest to filtr rezonansowy wykonany na podłożu kwarcu ST. W Europie najbardziej rozpowszechniony jest system GSM (Global System for Mobile Communications). Rys. 2 przedstawia schemat telefonu komórkowego z zaznaczonymi elementami typu AFP.



Rys. 2. Schemat telefonu komórkowego z zaznaczonymi elementami z AFP

System ten nadaje w paśmie 890 - 915 MHz a odbiera w paśmie 925 - 960 MHz. Posiada on 124 kanały, przy czym na jeden kanał przypada ośmiu użytkowników. Odstęp międzykanałowy wynosi 1250 kHz. Przyrządy z AFP mają podobną konstrukcję do wcześniej omawianych ale wymagania na ich parametry są mniej ostre.

Rys. 3. prezentuje fotografię takiego urządzenia z zaznaczonymi zintegrowanymi układami z AFP oraz widok takiego układu na tle zapalki.



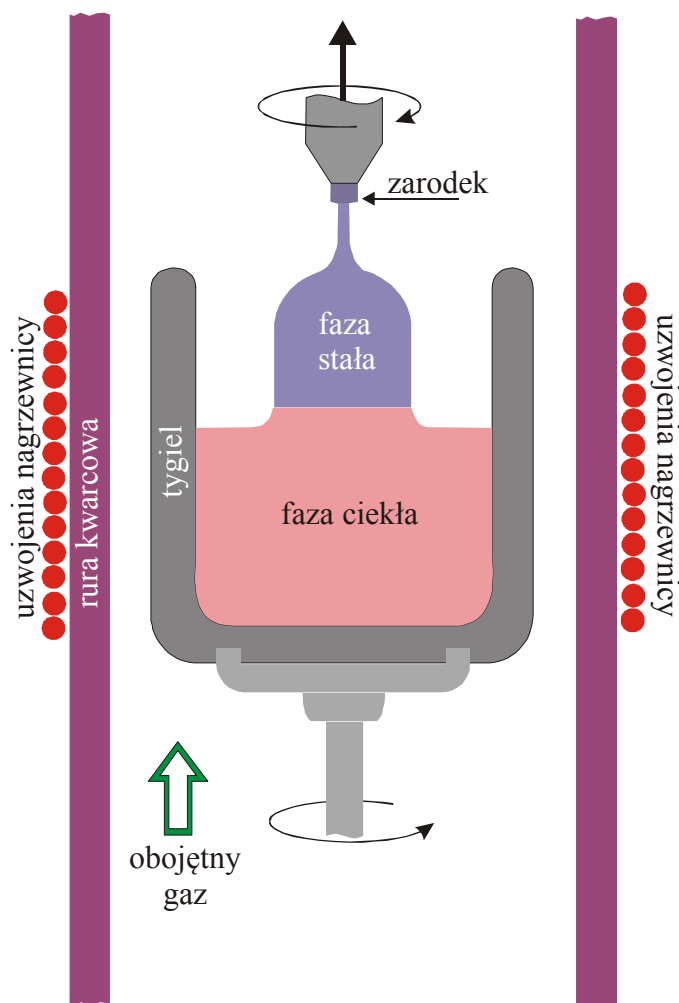
Rys. 3. Po prawej: płytką drukowaną telefonu komórkowego z zaznaczonymi układami z AFP, po lewej: układ z AFP na tle zapalki

Filtry z AFP są także wykorzystywane w szerokopasmowej komunikacji satelitarnej zazwyczaj jako filtry p.cz. z pasmem 50%. Pracują one często z przetwornikami nachylonymi (slanted- 7°) dzięki czemu uzyskuje się bardzo płaską charakterystykę amplitudową (z pofalowaniami mniejszymi od 0,6 dB) i niemal idealnie liniową fazę (nierówności mniejsze od 5%). Typowy odstęp listków bocznych od głównego wynosi 50 dB.

5. Technologia urządzeń z AFP

5.1. Wytwarzanie kryształów

Pomimo znacznego postępu w wielu technologiach do dnia dzisiejszego nie udaje się wynaleźć innej metody uzyskiwania dużych monokryształów niż metoda Czochralskiego. Rozwój elektroniki półprzewodników pociągający za sobą wzmożony popyt na monokryształy sprawił, że to właśnie polskie nazwisko jest najczęściej cytowane w technicznej literaturze naukowej. Metoda Czochralskiego została opracowana w 1916 r. i jest dość prosta (rys. 1).



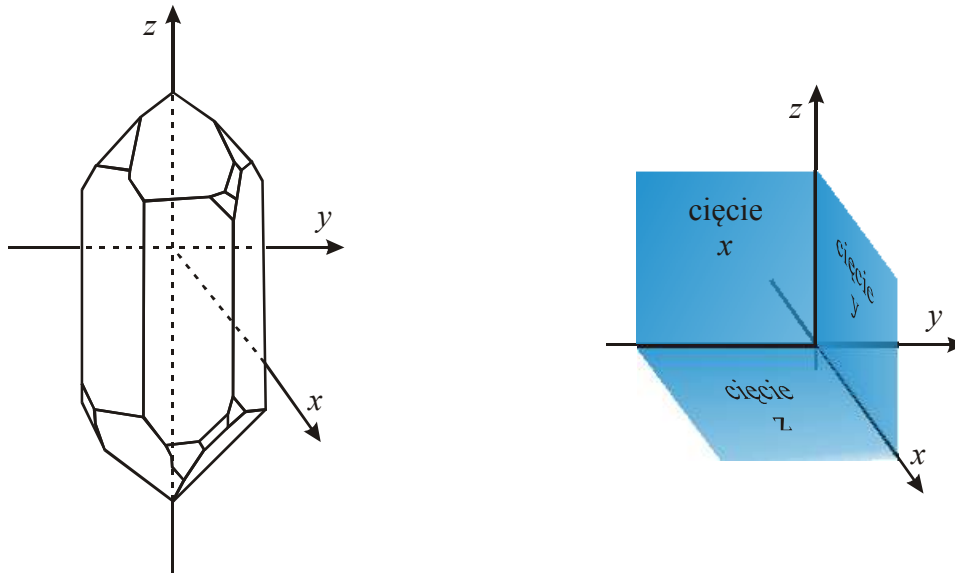
Rys. 1. „Wyciąganie monokryształu z fazy ciekłej metodą Czochralskiego

Prof. Jan Czochralski
1885-1953

Umieszczony w grafitowym lub kwarcowym tyglu związek polikrystaliczny ulega stopieniu, po czym zanurzany jest w nim zarodek kryształu (mały kryształ). W miejscu zanurzenia obniża się nieco temperatura i następuje krystalizacja, przy czym narastająca warstwa ciała stałego zachowuje ciągłość budowy krystalicznej zarodka. Przez powolne podnoszenie zarodka następuje wzrost kolejnych warstw monokryształu. Tak więc jest on niejako wyciągany z cieczy. Proces wzrostu monokryształu jest zależny od wielu parametrów technologicznych i z tego powodu hodowla kryształów wymaga wiedzy i staranności. Jak pokazuje doświadczenie kryształy tego samego związku uzyskiwane w podobnych

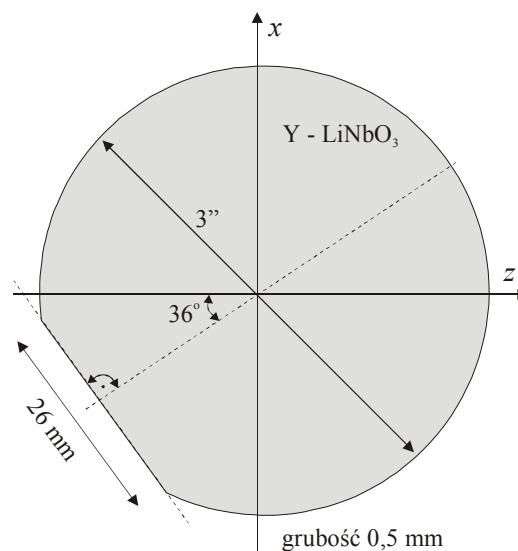
warunkach niejednokrotnie, z powodu różnego typu defektów i zanieczyszczeń, znacznie różnią się od siebie.

Uzyskany pręt krystaliczny jest odpowiednio orientowany (ważne ze względu na anizotropię) i szlifowany w celu oznaczenia płaszczyzn sieci krystalicznej. Rys. 2 przedstawia oszlifowany kryształ kwarcu z zaznaczonymi osiami krystalograficznymi oraz podstawowe rodzaje cięć.



Rys. 2. Po lewej: oszlifowany monokryształ kwarcu, po prawej: podstawowe płaszczyzny cięć

Po tych operacjach kryształ zostaje pocięty piłą diamentową na płytki, które poddawane są szlifowaniu i polerowaniu. Na gotowych płytkach wybrana orientacja zaznaczana jest poprzez wykonanie odpowiedniego ścięcia bazowego. Rys. 3 przedstawia płytkę LiNbO_3 ze ścieniem bazowym.



Rys. 3. Płytkę LiNbO_3 z wykonanym ścieniem bazowym

Ścięcie to jest ważne w dalszych procesach technologicznych i z tego względu powinno być wykonane precyzyjnie (podobnie jak orientacja kryształu przed pocięciem). Ważna jest nie

tylko jakość kryształu, ale i jakość powierzchni płytki. Do najważniejszych, z punktu widzenia akustoelektroniki, należą trzy parametry świadczące o jakości płytki:

- rozrzut grubości mierzony w czterech punktach i w środku płytki,
- promień krzywizny płytki,
- nierównomierności powierzchni.

Podane parametry powinny być oczywiście jak najmniejsze.

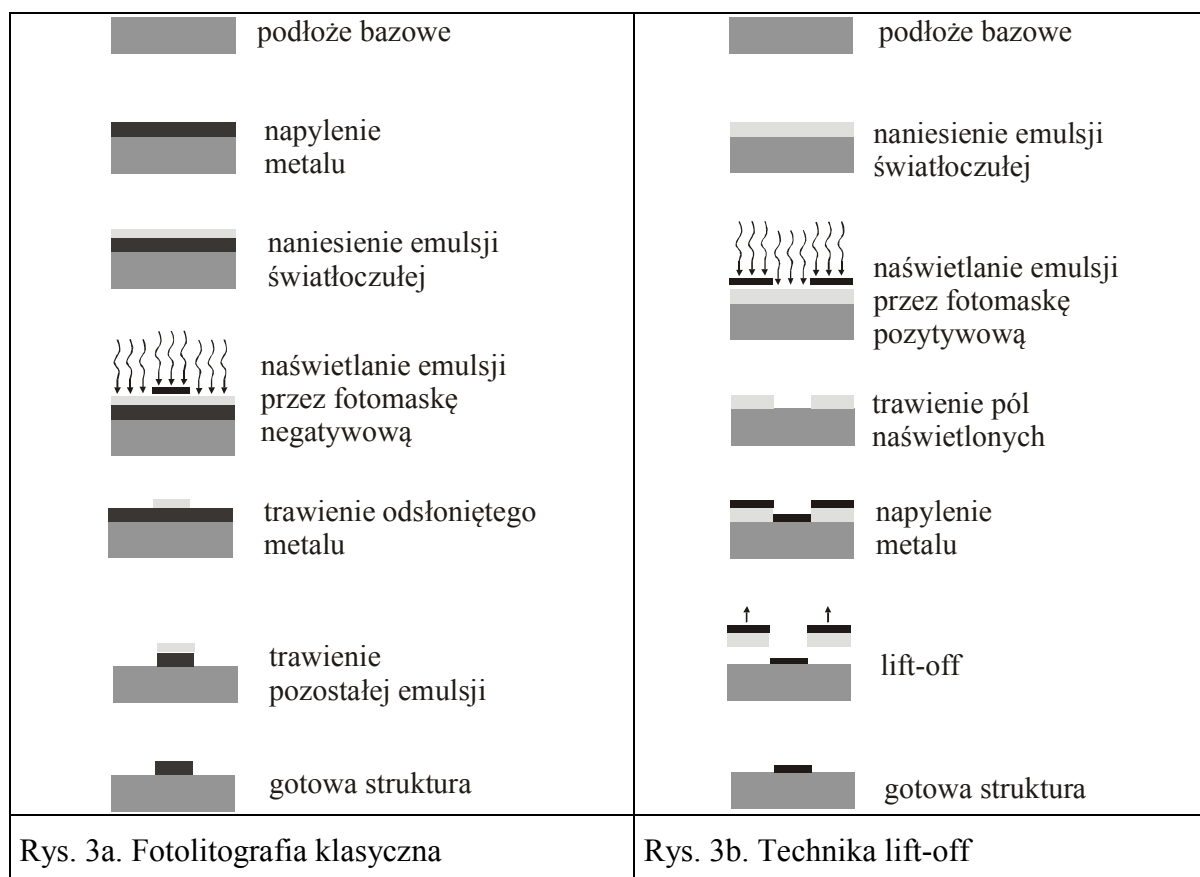
Poniższa tabela zawiera najczęściej spotykane kryształy piezoelektryczne, a także materiały kompozytowe i piezoceramikę (spieki polikrystaliczne).

materiał	cięcie	kierunek propagacji	prędkość fali Rayleigha [m/s]	współczynnik sprężenia elektromech. [%]	współczynnik temperaturowy γ [$10^{-6}/K$]	względna przenikalność elektryczna
kwarc	42,75° Y (ST)	X	3175	0,16	0	4,5
kwarc	- 75° Y	X	3960	0,11	9	4,5
LiNbO ₃	Y	Z	3488	4,82	-94	36,7
LiNbO ₃	128° Y	X	4000	5,56	-74	39,1
LiNbO ₃	64° Y	X	4742	11,3	-79	37,1
LiTaO ₃	X	112° Y	3295	0,64	-18	44,0
LiTaO ₃	36° Y	X	4178	4,8	-33	48,3
Li ₂ B ₄ O ₇	45° X	Z	3401	1,0	0	9,6
polikryst. ZnO na szkle	-	-	2576	1,4	-11	10,8
kryst. ZnO na szafirze	-	-	5500	3,4	-7	10
Pb(Sn _{1/2} Sb _{1/2})O ₃ + PbTiO ₃ + PbZrO ₃	-	-	2420	2,4	-38	270
0,1Pb(Mn _{1/3} Nb _{2/3})O ₃ + 0,9Pb(Zr _{0,74} Ti _{0,26})O ₃	-	-	2430	2,9	-17	460

5.2. Wytwarzanie struktur z AFP

Technika wytwarzania struktur na powierzchni kryształów czyli technologia planarna jest niewiele młodsza od metoda Czochralskiego. Opracowana w 1960 r. przetrwała do czasów współczesnych w niemal niezmienionej postaci. Bazują nawet na niej nowoczesne techniki trójwymiarowe.

Istnieją zasadniczo dwie odmiany tej technologii: klasyczna z wykorzystaniem integralnych fotomasek oraz tzw. lift-off. Schemat obydwu technik jest przedstawiony poniżej.

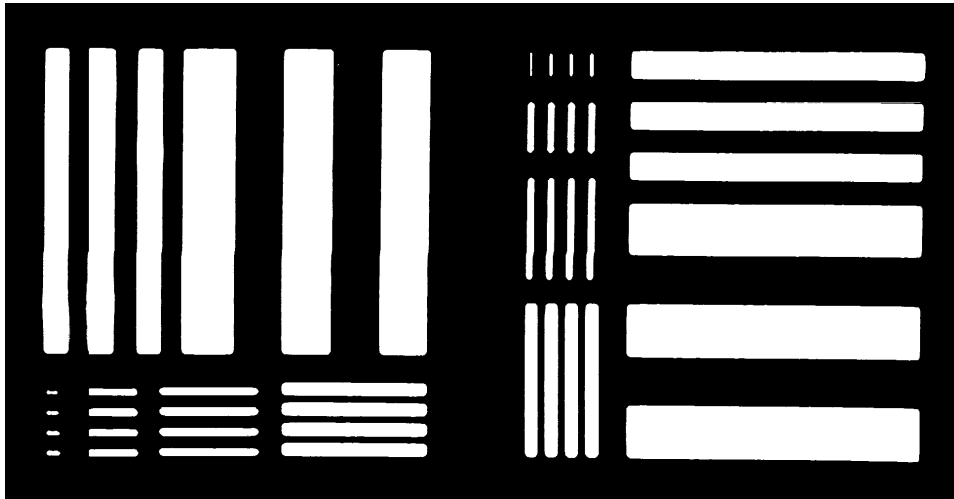


Jak widać zasadnicza różnica między tymi metodami tkwi w sposobie naświetlania. Klasyczna metoda wymaga zastosowania fotomaski negatywowej, zaś metoda lift-off posługuje się maską pozytywową. Niektóre urządzenia pracują z maskami wirtualnymi, tj. zapisanymi w pamięci komputera sterującego wąską wiązką naświetlającą, która pracuje jak stempel. Jest to technika prostsza ale i wymagająca dużo większej precyzji.

Z wiązką naświetlającą związana jest pewna długość fali z tego powodu w technologii o wysokiej precyzji wykorzystywany jest przede wszystkim ultrafiolet, promienie Roentgena (rentgenolitografia) lub elektrony (elektronolitografia). Uzyskiwana rozdzielczość zależy od zastosowanej wiązki (długości fali). Orientacyjne osiągi różnych systemów litografii przedstawione są w tabeli.

Rodzaj systemu	Uzyskiwana szerokość linii [μm]
litografia optyczna	3
litografia optyczna głęboki nadfiolet	2
rentgenolitografia	0,3
elektronolitografia	0,1

Rys. 4 prezentuje próbne naświetlenie maski za pomocą elektronolitografu ZBA 20.



Rys. 4. Naświetlenie testowe wykonane za pomocą elektronolitografu ZBA 20. Grubości pasków odpowiadają kolejno wymiarom 0,1; 0,2; 0,5; 1,0; 2,0 oraz 5,0 μm . Błędy w naświetlaniu pozwalają zauważyć, że odbywało się ono techniką prostokątnych „stępli”

Po uzyskaniu struktur, których może być na jednej płytce kilka tysięcy, płytka jest cięta, a wyselekcjonowane gotowe układy montowane są w obudowach, elektrycznie łączone z wyprowadzeniami i hermetycznie zamykane. Całość tych czynności wykonywana jest w pomieszczeniach utrzymywanych w najwyższej czystości. Przeciętny ludzki włos ma grubość ok. 60 μm , zaś uzyskiwane struktury nawet poniżej 0,1 μm .

Trzeba stanowczo podkreślić, że w rzeczywistości procesy technologiczne są znacznie bardziej skomplikowane niż wynikałoby to z powyższego krótkiego opisu. Szczegóły takiej technologii, zarówno jeżeli chodzi o uzyskiwanie i obróbkę krysztalów jak i o technikę planarną, stanowią często tajemnicę wytwórni.

Technologie planarne stosuje się także przy produkcji układów scalonych. W tym jednak wypadku etapy zilustrowane na rys. 3 są wielokrotnie powtarzane z różnymi maskami (w przypadku mikroprocesorów stosuje się nawet kilkaset masek) i przeplatane procesami domieszkania, implantacji i in.

Dodatek

Analiza harmoniczna sygnałów

Celem analizy harmonicznej jest przedstawienie dowolnej funkcji $f(t)$ w postaci sumy funkcji harmoniczych. Utwórzmy funkcję okresową $f_T(t)$, w której co pewien okres T sekund powtarza się dowolna funkcja $f(t)$. Jeśli $f(t)$ nie jest ograniczona w czasie tj. dla T dążącego do nieskończoności funkcje $f_T(t)$ i $f(t)$ są identyczne:

$$\lim_{T \rightarrow \infty} f_T(t) = f(t) \quad (1)$$

Wykładniczy szereg Fouriera, odpowiadający funkcji $f_T(t)$, można przedstawić w postaci:

$$f_T(t) = \sum_{n=-\infty}^{\infty} \hat{c}_n e^{jn\omega_0 t}, \quad (2)$$

$$\text{gdzie } \omega_0 = \frac{2\pi}{T}, \text{ zaś } \hat{c}_n = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} f_T(t) e^{-jn\omega_0 t} dt. \quad (3)$$

Współczynnik \hat{c}_n jest amplitudą zespoloną składowej harmoniczej o pulsacji $n\omega_0$.

W granicy, gdy $T \rightarrow \infty$, amplituda każdej składowej staje się nieskończenie mała, natomiast liczba składowych nieskończenie wielka. Składowe odpowiadają wszystkim pulsacjom tworząc widmo, które jest ciągłą funkcją pulsacji ω .

Oznaczmy:

$$n \cdot \omega_0 = \omega_n. \quad (4)$$

Współczynnik \hat{c}_n jest teraz funkcją zmiennej $n\omega_0$. Jeżeli podstawimy:

$$\hat{c}(\omega_n) = T \cdot \hat{c}_n(\omega_n), \quad (5)$$

to wzór (2) przyjmie formę:

$$f_T(t) = \frac{1}{T} \sum_{n=-\infty}^{\infty} \hat{c}(\omega_n) e^{j\omega_n t}, \quad (6)$$

gdzie teraz:

$$\hat{c}(\omega_n) = \int_{-\frac{T}{2}}^{\frac{T}{2}} f_T(t) \cdot e^{-j\omega_n t} dt. \quad (7)$$

Przyjmując, że okres wynosi $T = \frac{2\pi}{\omega_0}$, wyrażenie (6) można zapisać w postaci:

$$f_T(t) = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \hat{c}(\omega_n) \cdot e^{j\omega_n t} \cdot \omega_0 \quad (8)$$

Jeśli okres wydłużał się będzie do nieskończoności $T \rightarrow \infty$ to ω_0 będzie zmierzała do zera i można ją będzie zastąpić infinitezimalnym przyrostem $d\omega$. W takim przypadku wzór (8)

stanie się wyrażeniem całkowym. Dyskretne wartości ω_n przyjmą wartość zmiennej ω oraz, zgodnie z (1), funkcja $f_T(t)$ przejdzie w $f(t)$:

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{c}(\omega) e^{j\omega t} d\omega, \quad (9)$$

gdzie teraz

$$\hat{c}(\omega) = \int_{-\infty}^{\infty} f(t) e^{-j\omega t} dt. \quad (10)$$

Jak widać funkcję nieokresową $f(t)$ można przedstawić za pomocą funkcji wykładniczych, opisujących fale harmoniczne.

Równania (9) i (10) stanowią parę wzorów transformacyjnych Fouriera, przy czym (10) jest prostą transformacją Fouriera funkcji $f(t)$, zaś (9) odwrotną transformacją Fouriera funkcji $\hat{c}(\omega)$. Symbolicznie operacje te zapisuje się następująco:

$$\hat{c}(\omega) = F[f(t)], \quad (11a)$$

$$f(t) = F^{-1}[\hat{c}(\omega)] \quad (11b)$$

albo:

$$F[f(t)] = \int_{-\infty}^{\infty} f(t) e^{-j\omega t} dt \quad (12a)$$

$$F^{-1}[\hat{c}(\omega)] = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{c}(\omega) e^{j\omega t} d\omega \quad (12b)$$

Prosta transformata Fouriera (wynik transformacji Fouriera) funkcji $f(t)$, $\hat{c}(\omega) = F[f(t)]$ opisuje zespolone widmo częstotliwościowej funkcji $f(t)$ i nazywana jest funkcją widmową. Przedstawia ona zależność amplitud składowych harmonicznych przebiegu $f(t)$ od częstotliwości, a więc przedstawienie funkcji czasu $f(t)$ w dziedzinie częstotliwości. Przedstawienie funkcji w dziedzinie czasu określa jej wartość dla każdej chwili czasu, natomiast przedstawienie w dziedzinie częstotliwości określa zależność amplitud składowych harmonicznych danej funkcji od częstotliwości tych składowych. Obydwa te sposoby przedstawiania funkcji są jednoznaczne. Transformata Fouriera jest nazywana całką Fouriera o postaci zespolonej (wykładniczej),

Odwrotna transformacja Fouriera funkcji $\hat{c}(\omega)$, formalnie jest przeniesieniem funkcji z dziedziny częstotliwości w dziedzinę czasu.

Warunkiem istnienia prostej transformaty Fouriera jest skończona wartość (bezwzględna zbieżność) całki:

$$\int_{-\infty}^{\infty} f(t) e^{-j\omega t} dt \quad (13)$$

Ponieważ moduł $e^{-j\omega t}$ jest równy jedności, warunkiem wystarczającym istnienia prostej transformaty Fouriera jest skończona wartość całki:

$$\int_{-\infty}^{\infty} |f(t)| dt. \quad (14)$$

Warunek ten nie jest jednak warunkiem koniecznym. Istnieją bowiem funkcje, które nie są bezwzględnie całkowne, a posiadają transformaty.

Transformacja Fouriera przebiegu okresowego

Transformacja Fouriera została tutaj wyprowadzona jako przypadek graniczny szeregu Fouriera przy założeniu, że okres funkcji wydłuża się do nieskończoności. Metodę taką można też zastosować do wykazania, że szereg Fouriera jest granicznym przypadkiem transformaty Fouriera, co ujednotliło podejście do funkcji okresowych i nieokresowych. Choć jak wiadomo, dla dowolnej funkcji okresowej zachodzi związek:

$$\int_{-\infty}^{\infty} |f(t)| dt = \infty, \quad (15)$$

to jednak w granicy transformata istnieje.

Funkcję okresową $f(t)$ o okresie T można także przedstawić, jak już wiadomo, za pomocą szeregu Fouriera:

$$f(t) = \sum_{n=-\infty}^{\infty} \hat{c}_n e^{jn\omega_0 t}; \quad \omega_0 = \frac{2\pi}{T}. \quad (16)$$

Obliczmy transformatę Fouriera równania (16):

$$F[f(t)] = F\left[\sum_{n=-\infty}^{\infty} \hat{c}_n e^{jn\omega_0 t}\right] = \sum_{n=-\infty}^{\infty} \tilde{c}_n F[e^{jn\omega_0 t}]. \quad (17)$$

Transformatę Fouriera czynnika $e^{jn\omega_0 t}$ określić można korzystając ze wzoru Eulera:

$$F[e^{jn\omega_0 t}] = F[\cos n\omega_0 t + j \sin n\omega_0 t], \quad (18)$$

co daje:

$$F[f(t)] = 2\pi \sum_{n=-\infty}^{\infty} \hat{c}_n \delta(\omega - n\omega_0), \quad (19)$$

gdzie: $\delta(\omega - n\omega_0)$ jest deltą Diraca) w dziedzinie częstotliwości.

Otrzymany wzór świadczy o tym, że funkcja widmowa albo transformata Fouriera sygnału okresowego składa się z impulsów w dziedzinie częstotliwości, umieszczonych w punktach o częstotliwościach równych częstotliwościom harmonicznym sygnału, a amplituda tych impulsów jest 2 razy większa od wartości współczynników w szeregu wykładniczym Fouriera. Wynika z tego, że ciągła funkcja widmowa (prosta transformata Fouriera) przy przechodzeniu do funkcji okresowych, w granicy zmienia się w ciąg równoległych impulsów.

WŁAŚCIWOŚCI TRANSFORMACJI FOURIERA

Transformacja Fouriera jest podstawowym narzędziem analizy częstotliwościowej. Umożliwia przejście od funkcji w dziedzinie czasu do funkcji w dziedzinie częstotliwości. Z punktu widzenia analizy częstotliwościowej interesujący jest wpływ pewnych operacji w jednej dziedzinie na postać funkcji w drugiej i odwrotnie.

Równania określające obie dziedziny posiadają pewien rodzaj symetrii, co można zauważyć porównując wzory transformacyjne (12a) i (12b).

$$F[f(t)] = \int_{-\infty}^{\infty} f(t) e^{-j\omega t} dt \quad F^{-1}[\hat{c}(\omega)] = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{c}(\omega) e^{j\omega t} d\omega$$

Nietrudno zauważyć, że symetria ta będzie miała wpływ na własności transformat. Określony skutek w dziedzinie częstotliwości powstały w wyniku pewnej operacji w dziedzinie czasu będzie w jakiś sposób podobny do skutku, powstającego w dziedzinie czasu operacji w dziedzinie częstotliwości, co można symbolicznie zapisać:

$$f(t) \leftrightarrow \hat{c}(\omega). \quad (20)$$

Na tej podstawie można określić najważniejsze własności transformacji Fouriera.

Własność symetrii

Korzystając ze wzorów transformacyjnych można zapisać:

$$\omega(\pm t) = 2\pi f(\pm t) = \int_{-\infty}^{\infty} \hat{c}(\omega) \cdot e^{\pm j\omega t} d\omega. \quad (21)$$

Po zastąpieniu zmiennej całkowania ω w (21) zmienną x otrzymuje się:

$$2\pi f(\pm t) = \int_{-\infty}^{\infty} \hat{c}(x) \cdot e^{\pm jxt} dx, \quad (23)$$

zaś zamieniając t na ω

$$2\pi f(\pm\omega) = \int_{-\infty}^{\infty} \hat{c}(x) \cdot e^{\pm jx\omega} dx. \quad (24)$$

Zamieniając z kolei zmienną całkowania x na t , otrzymamy:

$$2\pi f(\pm\omega) = \int_{-\infty}^{\infty} \hat{c}(t) \cdot e^{\pm j\omega t} dt = F[\hat{c}(\mp t)], \quad (25)$$

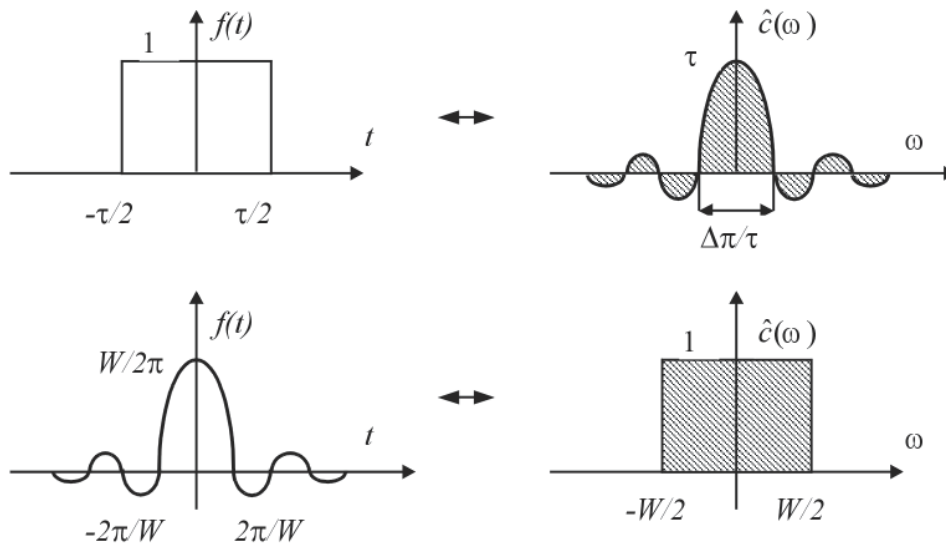
a stąd:

$$\hat{c}(\pm t) = 2\pi f(\mp\omega) \quad (26)$$

Własność ta jest spełniona dla parzystych funkcji $f(t)$ co sprowadza powyższą równość do postaci:

$$\hat{c}(t) \leftrightarrow 2\pi f(\omega) \quad (27)$$

Zależność ta nazywana jest własnością symetrii. Ilustruje ją rys. 1.



Rys. 1. Własność symetrii transformacji Fouriera

Własność liniowości

Dla dowolnych stałych rzeczywistych a_1 i a_2 zachodzi relacja:

$$a_1 f_1(t) + a_2 f_2(t) \leftrightarrow a_1 \hat{c}_1(\omega) + a_2 \hat{c}_2(\omega). \quad (28)$$

Jeśli stałe są równe to z własności liniowości wynika, że widmo sumy funkcji równa się sumie widm tych funkcji.

Zmiana skali w dziedzinie czasu

Dla dodatniej i rzeczywistej stałej a zachodzi:

$$F[f(at)] = \int_{-\infty}^{\infty} f(at) e^{-j\omega t} dt. \quad (29)$$

Po podstawieniu $x = at$, otrzymuje się:

$$F[f(at)] = \frac{1}{a} \int_{-\infty}^{\infty} f(x) e^{-\frac{j\omega x}{a}} dx = \frac{1}{a} \hat{c}\left(\frac{\omega}{a}\right), \quad (30)$$

stąd:

$$f(at) \leftrightarrow \frac{1}{a} \hat{c}\left(\frac{\omega}{a}\right). \quad (31)$$

Podobnie dla ujemnej i rzeczywistej stałej a zachodzi:

$$f(at) \leftrightarrow -\frac{1}{a} \hat{c}\left(\frac{\omega}{a}\right), \quad (32)$$

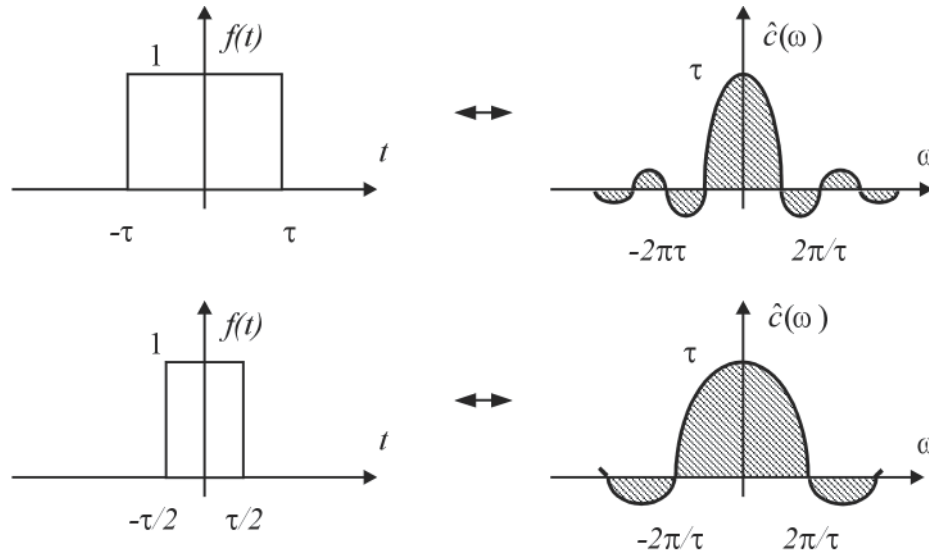
co ogólnie można zapisać

$$f(at) \leftrightarrow \frac{1}{|a|} \hat{c}\left(\frac{\omega}{a}\right). \quad (33)$$

W szczególnym przypadku gdy $a = -1$ zachodzi:

$$f(-t) \leftrightarrow \hat{c}(-\omega) = c^*(\omega) \quad (34)$$

Z własności (33) związanej ze zmianą skali wynika, że skrócenie przebiegu w dziedzinie czasu jest równoważne z rozszerzeniem widma w dziedzinie częstotliwości i odwrotnie. Kompresja podziałki na osi czasu a razy oznacza, że funkcja zmienia się a razy szybciej, częstotliwość jej składowych wzrośnie a razy, zaś jej widmo częstotliwościowe będzie rozszerzone wzdłuż osi częstotliwości a razy. Podobnie, jeżeli funkcja zostanie rozszerzona wzdłuż osi czasu, to znaczy, że będzie zmieniać się wolniej, czyli częstotliwości jej składowych będą niższe. Efekty związane ze zmianą skali ilustruje rys. 2.



Rys. 2. Skrócenie w dziedzinie czasu jest równoważne rozszerzeniu w dziedzinie częstotliwości

Przesunięcie w dziedzinie czasu

Wzór transformacyjny dla funkcji $f(t-t_0)$ będzie miał postać:

$$F[f(t-t_0)] = \int_{-\infty}^{\infty} f(t-t_0)e^{-j\omega t} dt. \quad (35)$$

Po podstawieniu $t-t_0 = x$ otrzymuje się:

$$F[f(t-t_0)] = \int_{-\infty}^{\infty} f(x)e^{-j\omega(x+t_0)} dx = \hat{c}(\omega) \cdot e^{-j\omega t_0} \quad (36)$$

Z (36) wynika, że przesunięcie funkcji w dziedzinie czasu o t_0 sekund nie powoduje zmiany widma amplitudowego $\hat{c}(\omega)$, natomiast powoduje przesunięcie widma fazowego o wartość $-\omega t_0$. Zatem dla składowej o częstotliwości ω przesunięcie o odcinek czasu t_0 jest równoważne przesunięciu fazowemu o $-\omega t_0$ radianów. Możemy więc stwierdzić, że przesunięcie funkcji o t_0 w dziedzinie czasu jest równoważne z pomnożeniem transformaty w dziedzinie częstotliwości przez $e^{-j\omega t_0}$.

Przesunięciem w dziedzinie częstotliwości

Ze wzoru transformacyjnego wynika, że:

$$F[f(t) \cdot e^{j\omega_0 t}] = \int_{-\infty}^{\infty} f(t) e^{j\omega_0 t} e^{-j\omega t} dt = \int_{-\infty}^{\infty} f(t) e^{j(\omega - \omega_0)t} dt = \hat{c}(\omega - \omega_0), \quad (37)$$

zatem:

$$f(t) \cdot e^{j\omega_0 t} \leftrightarrow \hat{c}(\omega - \omega_0). \quad (38)$$

Wynika stąd, że przesunięcie widma funkcji o ω_0 w dziedzinie częstotliwości jest równoważne z pomnożeniem funkcji w dziedzinie czasu przez $e^{j\omega_0 t}$.

Porównując własności transformacji Fouriera związane z przesunięciami w dziedzinie czasu w dziedzinie częstotliwości można zauważyć symetrię:

przesunięcie o odcinek czasu t_0 w dziedzinie czasu jest równoważne przesunięciu fazowemu o $-\omega t_0$ radianów w dziedzinie częstotliwości, zaś przesunięcie widma funkcji o ω_0 w dziedzinie częstotliwości jest równoważne z pomnożeniem funkcji w dziedzinie czasu przez $e^{j\omega_0 t}$.

Różniczkowanie w dziedzinie czasu

Wykonajmy różniczkowanie po czasie wzoru transformacyjnego. Otrzymamy:

$$\frac{d f(t)}{d t} = \frac{1}{2\pi} \frac{d}{d t} \int_{-\infty}^{\infty} \hat{c}(\omega) e^{j\omega t} d\omega, \quad (39)$$

i dalej po przeniesieniu operatora różniczkowania po czasie pod całkę:

$$\frac{d f(t)}{d t} = \frac{1}{2\pi} \int_{-\infty}^{\infty} j\omega \hat{c}(\omega) e^{j\omega t} d\omega, \quad (40)$$

z czego wynika własność:

$$\frac{d f(t)}{d t} \leftrightarrow j\omega \hat{c}(\omega). \quad (41)$$

Wynik ten można uogólnić na n -tą pochodną:

$$\frac{d^n f(t)}{d t^n} \leftrightarrow (j\omega)^n \hat{c}(\omega) \quad (42)$$

Różniczkowanie w dziedzinie częstotliwości

Analogicznie do rozważań na różniczkowaniem w dziedzinie czasu otrzymuje się z dokładnością do stałej dodatniej:

$$-jt f(t) \leftrightarrow \frac{d \hat{c}(\omega)}{d \omega}, \quad (43)$$

oraz

$$\frac{1}{-jt} f(t) \leftrightarrow \int \hat{c}(\omega) d\omega. \quad (44)$$

Uogólnienie tego wyniku na wyższe pochodne funkcji $\hat{c}(\omega)$ prowadzi do relacji:

$$(-jt)^n f(t) \leftrightarrow \frac{d^n \hat{c}(\omega)}{d\omega^n} \quad (45)$$

Różniczkowanie w dziedzinie częstotliwości jest więc równoważne z mnożeniem przez $-jt$ w dziedzinie czasu. Można też wykazać, że z dokładnością do stałej dodatniej funkcja $\int \hat{c}(\omega) d\omega$ jest transformatą Fouriera funkcji $\frac{f(t)}{-jt}$:

$$\frac{1}{-jt} f(t) \leftrightarrow \int \hat{c}(\omega) d\omega. \quad (46)$$

Twierdzenie o splocie

Wzór transformacyjny można zapisać dla iloczynów funkcji:

$$\int_{-\infty}^{\infty} f_1(t) \cdot f_2(t) dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{c}_1(\omega) \cdot \hat{c}_2(-\omega) d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{c}_1(\omega) \cdot \hat{c}_2^*(\omega) d\omega \quad (47)$$

Weźmy funkcję $f(t)$ taką że:

$$f(t) = \int_{-\infty}^{\infty} f_1(\tau) f_2(t - \tau) d\tau \quad (48)$$

i oznaczmy:

$$f(t) = f_1(t) * f_2(t) \quad (49)$$

Funkcję $f(t)$ nazywa się splotem funkcji $f_1(t)$ i $f_2(t)$, a wyrażenie (48) całką splotową.

Twierdzenie o splocie w dziedzinie czasu

Z (47) i (48) wynika relacja:

$$\int_{-\infty}^{\infty} f_1(\tau) f_2(t - \tau) d\tau \leftrightarrow \hat{c}_1(\omega) \cdot \hat{c}_2(\omega), \quad (50)$$

którą symbolicznie można zapisać w postaci:

$$f_1(t) * f_2(t) \leftrightarrow \hat{c}_1(\omega) \cdot \hat{c}_2(\omega), \quad (51)$$

lub podsumować stwierdzeniem, że splot w dziedzinie czasu odpowiada mnożeniu w dziedzinie częstotliwości.

Twierdzenie o splocie w dziedzinie częstotliwości

Analogicznie, biorąc pod uwagę relacje dotyczące splotu można sformułować twierdzenie:

$$f_1(t) \cdot f_2(t) \leftrightarrow \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{c}_1(u) \cdot \hat{c}_2(\omega - u) du, \quad (52)$$

lub symbolicznie:

$$f_1(t) \cdot f_2(t) \leftrightarrow \frac{1}{2\pi} [\hat{c}_1(\omega) * \hat{c}_2(\omega)] \quad (53)$$

Z twierdzenia tego wynika z kolei, że pomnożenie dwóch funkcji w dziedzinie czasu jest równoważne splotowi widm tych funkcji w dziedzinie częstotliwości.

WŁAŚCIWOŚCI SPLOTU

Splot jest szczególnym rodzajem iloczynu ponieważ prawa dotyczące algebry splotu są analogiczne do praw algebraicznego mnożenia. Mamy więc:

prawo przemienności splotu:

$$f_1(t) * f_2(t) = f_2(t) * f_1(t),$$

prawo rozdzielności splotu względem dodawania:

$$f_1(t) * [f_2(t) + f_3(t)] = f_1(t) * f_2(t) + f_1(t) * f_3(t)$$

oraz prawo łączności splotu:

$$f_1(t) * [f_2(t) * f_3(t)] = [f_1(t) * f_2(t)] * f_3(t).$$

Dyskretna transformacja Fouriera

Fizyczna realizacja transformacji Fouriera jest możliwa za pomocą układów analogowych jednak ciągłość przekształceń stoi na przeszkodzie bezpośredniemu wykorzystywaniu komputerów do obliczania transformat. Komputer może bowiem operować wyłącznie na wartościach dyskretnych. Z tego powodu dobrze byłoby funkcje sygnałowe przedstawić w postaci dyskretnej zastępując całkowanie sumowaniem.

Pierwszym krokiem do przystosowania sygnałów o przebiegach ciągłych do obróbki za pomocą komputera jest dyskretyzacja dziedziny czasu, co jest możliwe poprzez zastąpienie przebiegów ciągłych szeregiem Kotelnikowa-Shannona. Czynność taka nazywana jest zwykle próbkowaniem.

Przebiegi o ograniczonym paśmie częstotliwości, dla których transformata Fouriera jest różna od zera tylko w skończonym przedziale częstotliwości można zastąpić następującym szeregiem:

$$x(t) = \sum_{k=-\infty}^{\infty} x(k \cdot \Delta t) \frac{\sin \omega_m (t - k \cdot \Delta t)}{\omega_m (t - k \cdot \Delta t)} = \sum_{k=-\infty}^{\infty} x(k \cdot \Delta t) \operatorname{sinc} 2f_m (t - k \cdot \Delta t), \quad (54)$$

gdzie $\omega_m = 2\pi f_m$ jest graniczną częstotliwością kątową przebiegu $x(t)$, $x(k \cdot \Delta t)$ chwilowymi wartościami przebiegu $x(t)$ (próbki) pobieranymi okresowo w chwilach $k \cdot \Delta t$, Δt odstępem czasowym między próbkami o wartości związanej z graniczną częstotliwością przebiegu $x(t)$ poprzez wyrażenie:

$$\Delta t = \frac{\pi}{\omega_m} = \frac{1}{2f_m}.$$

Jak widać częstotliwość próbkowania nie może być mniejsza od dwukrotnej wartości granicznej częstotliwości przebiegu $x(t)$, czyli

$$f_p = \frac{1}{\Delta t} \geq 2 \cdot f_m \quad (55)$$

Oznacza to, że na składową sygnału o najwyższej częstotliwości równej f_m , powinny przypadać co najmniej dwie próbki. Ograniczenie to nazywane jest kryterium Nyquista. Szereg Kotielnikowa-Shannona (54) przedstawia rozwinięcie przebiegu $x(t)$ w szereg - względem funkcji typu $\frac{\sin \pi x}{\pi x} = \text{sinc } \pi x$, którego współczynniki rozwinięcia są próbkami sygnału. Jeżeli przebieg $x(t)$ jest impulsem o skończonym czasie trwania, to nie może mieć ograniczonego widma. Jednakże, jeżeli funkcja widmowa asymptotycznie i dostatecznie szybko maleje przy $\omega \rightarrow \pm\infty$, to zaniehbując wpływ niskoamplitudowych składowych widma, można również taką funkcję przedstawić w sposób przybliżony szeregiem (54). Ponieważ można przy tym pobrać $\frac{T}{\Delta t}$ próbek, to wynika stąd, że przebieg o ograniczonym widmie i skończonym czasie trwania jest określony ciągiem, zawierającym:

$$N = \frac{T}{\Delta t} = 2 f_m T \quad (56)$$

próbek.

Podobnie można również przedstawić przebieg ciągły, biorąc jego odcinki o długości T .

Przebieg odtworzony z przybliżonego wyrażenia będzie zniekształcony na końcach odcinka T (skutek nieuwzględnienia wpływu niskoamplitudowych składowych widma). Dokładność odtworzenia będzie związana z wyborem granicznej częstotliwości f_m .

Kolejnym krokiem jest dyskretyzacja wartości próbek czyli tzw. kwantyzacja polegająca na przyporządkowywaniu konkretnych wartości próbek do ustalonych poziomów, zwanych poziomami kwantyzacji.

W praktyce, do wykonywania przedstawionych operacji na sygnałach rzeczywistych, wykorzystuje się przetworniki analogowo-cyfrowe. Do takiego przetwornika wprowadza się sygnał, np. w postaci napięcia $u(t)$, a otrzymuje się równoległe lub szeregowo binarne ciągi kodowe, przedstawiające liczbowe wyrażenia próbek sygnału.

Dyskretną transformatą Fouriera (DFT) ciągu N próbek $x(k\Delta t)$, $0 \leq k \leq N-1$ nazywa się ciąg:

$$X(n\Delta f) = \frac{1}{N} \sum_{k=0}^{N-1} x(k\Delta t) e^{-j2\pi\Delta f \cdot k\Delta t \cdot n}, \quad (57)$$

w którym Δt jest odstępem próbek w dziedzinie czasu, $\Delta f = \frac{1}{N\Delta t}$ odstępem próbek w dziedzinie częstotliwości, k numerem próbki w dziedzinie czasu, n numerem próbki w dziedzinie częstotliwości.

W przypadku ogólnym, wartości $x(k\Delta t)$ i $X(n\Delta f)$ mogą być zespolone:

$$X(n\Delta f) = a(n\Delta f) + jb(n\Delta f), \quad (58)$$

gdzie: $a(n\Delta f)$ jest składową kosinusoidalną, zaś $b(n\Delta f)$ składową sinusoidalną.

Składowe te można zapisać w postaci sum:

$$a(n\Delta f) = \frac{1}{N} \sum_{k=0}^{N-1} x(k\Delta t) \cos 2\pi\Delta f \cdot k \cdot \Delta t \cdot n, \quad (59)$$

$$b(n\Delta f) = \frac{1}{N} \sum_{k=0}^{N-1} x(k\Delta t) \sin 2\pi\Delta f \cdot k \cdot \Delta t \cdot n,$$

wynikających bezpośrednio z wzoru (57).

Przy takim określeniu odstepu częstotliwościowego próbek możliwe jest wyróżnienie tylko n ich wartości, $0 \leq n \leq N-1$. Aby się o tym przekonać wyrażmy n w formie:

$$n = rN + n_0 \quad (60)$$

gdzie $r = \frac{n - n_0}{N}$, zaś $n_0 = n$ modulo N (reszta z dzielenia - jeśli liczbę całkowitą n wyrazimy przez $n = n_0 + mN$, gdzie $m = 0, 1, 2 \dots$ to $n_0 \leq N$ nazywa się n modulo N , np. 7 modulo 2 = 1).

Wstawiając (60) do wzoru (57) otrzymuje się:

$$X(n\Delta f) = \frac{1}{N} \sum_{k=0}^{N-1} x(k\Delta t) e^{-j2\pi\Delta f \cdot k\Delta t \cdot (rN + n_0)} \quad (61)$$

Ponieważ $2\pi\Delta f \cdot \Delta t \cdot N = 2\pi$, to człon $2\pi\Delta f \cdot \Delta t \cdot rN$ w wykładniku potęgi można pominąć i wyrażenie (61) przedstawić w formie:

$$X(n\Delta f) = X(n_0\Delta f). \quad (62)$$

Należy zaznaczyć, że dyskretna transformacja Fouriera zachowuje własności transformacji ciągłej.

Odwrotna dyskretna transformacja Fouriera (IDFT- inverse discrete Fourier transform), analogiczna do odwrotnej ciągłej transformaty Fouriera, definiowana jest następująco:

$$x(l \cdot \Delta t) = \sum_{n=0}^{N-1} X(n \cdot \Delta f) e^{j2\pi n \cdot \Delta f \cdot l \Delta t} \quad (63)$$

IDFT wylicza próbki sygnału w dziedzinie czasu $x(l \cdot \Delta t)$, $0 \leq l \leq N-1$ na podstawie znajomości jego widma przedstawionego za pomocą N próbek w dziedzinie częstotliwości. Podobnie jak DFT, IDFT może obliczyć $x(l \cdot \Delta t)$ próbek dla wartości l spoza przedziału $0 \leq l \leq N-1$ będą to jednak powtórzenia wartości $x(l \cdot \Delta t)$ wziętych dla l z wnętrza tego przedziału.

Właściwości ciągłej transformaty Fouriera znajdują swoje ściśle odzwierciedlenie we właściwościach dyskretnej transformaty Fouriera.

Szybka transformacja Fouriera (FFT)

Bezpośrednie obliczenie dyskretnej transformaty Fouriera dla wszystkich składowych widma wymaga wykonania N^2 zespolonych operacji mnożenia i dodawania. Metoda FFT (Fast Fourier Transform) umożliwia redukcję liczby tych operacji do $N \cdot \lg_2 N$, co przynosi znaczne oszczędności czasu potrzebnego na obliczenia. Istnieją dwa algorytmy obliczania FFT: algorytm z podziałem czasowym oraz algorytm z podziałem częstotliwościowym.

FFT Z PODZIAŁEM CZASOWYM

Główną ideą, na której opiera się ten algorytm, jest tzw. dekompozycja N -punktowego DFT na dwa niezależne przekształcenia DFT $N/2$ punktowe. Pierwsze z tych

przekształceń zawiera jedynie próbki parzyste sygnału wejściowego, drugie natomiast jedynie próbki nieparzyste:

$$X(k) = \sum_{n \text{ parzyste}} x(n)W_N^{kn} + \sum_{n \text{ nieparzyste}} x(n)W_N^{kn}. \quad (1)$$

Po podstawieniu $n=2r$ dla n parzystych i $n=2r+1$ dla n nieparzystych otrzymamy:

$$X(k) = \sum_{r=0}^{\frac{N}{2}-1} x(2r)W_N^{k2r} + \sum_{r=0}^{\frac{N}{2}-1} x(2r+1)W_N^{k(2r+1)} = \sum_{r=0}^{\frac{N}{2}-1} x(2r)(W_N^2)^{kr} + W_N^k \sum_{r=0}^{\frac{N}{2}-1} x(2r+1)(W_N^2)^{kr}. \quad (2)$$

Biorąc pod uwagę równość:

$$W_N^2 = e^{-j2\frac{2\pi}{N}} = e^{-j\frac{2\pi}{N/2}} = W_{\frac{N}{2}}, \quad (3)$$

równanie (2) można zapisać w postaci:

$$X(k) = \sum_{r=0}^{\frac{N}{2}-1} x(2r)W_{\frac{N}{2}}^{kr} + W_N^k \sum_{r=0}^{\frac{N}{2}-1} x(2r+1)W_{\frac{N}{2}}^{kr} \quad (4)$$

lub po prostu:

$$X(k) = G(k) + W_N^k H(k), \quad (5)$$

gdzie $G(k)$ jest $N/2$ -punktowym przekształceniem DFT parzystych punktów ciągu $x(n)$, $H(k)$ - $N/2$ -punktowym przekształceniem DFT nieparzystych punktów ciągu $x(n)$, $W_N = e^{-j2\pi/N}$.

FFT Z PODZIAŁEM CZĘSTOTLIWOŚCIOWYM

Weźmy pod uwagę podział ciągu wyjściowego $X(k)$:

$$\begin{aligned} X(k) &= \sum_{n=0}^{\frac{N}{2}-1} x(n)W_N^{kn} + \sum_{n=\frac{N}{2}}^{N-1} x(n)W_N^{kn} = \sum_{n=0}^{\frac{N}{2}-1} x(n)W_N^{kn} + \sum_{n=0}^{\frac{N}{2}-1} x\left(n + \frac{N}{2}\right)W_N^{k\left(n + \frac{N}{2}\right)} = \\ &= \sum_{n=0}^{\frac{N}{2}-1} x(n)W_N^{kn} + W_N^{\frac{N}{2}k} \sum_{n=0}^{\frac{N}{2}-1} x\left(n + \frac{N}{2}\right)W_N^{kn}. \end{aligned} \quad (6)$$

Ponieważ występują tu czynniki W_N^{kn} , a nie $W_{N/2}^{kn}$ otrzymane wyrażenia nie są $N/2$ -punktowymi transformacjami Fouriera. Stosując jednak podstawienie $W_N^{Nk/2} = (-1)^k$, możemy równanie (6) zapisać w postaci:

$$X(k) = \sum_{n=0}^{\frac{N}{2}-1} \left[x(n) + (-1)^k x\left(n + \frac{N}{2}\right) \right] W_N^{kn}. \quad (7)$$

Wtedy dla parzystych wskaźników $k=2r$ otrzymamy:

$$X(2r) = \sum_{n=0}^{\frac{N}{2}-1} \left[x(n) + x\left(n + \frac{N}{2}\right) \right] W_N^{2rn}, \quad (8)$$

zaś dla nieparzystych wskaźników $k=2r+1$:

$$X(2r+1) = \sum_{n=0}^{\frac{N}{2}-1} \left[x(n) - x\left(n + \frac{N}{2}\right) \right] W_N^n W_N^{2rn}. \quad (9)$$

Pamiętając, że $W_N^{2rn} = W_{N/2}^{rn}$, równania (8) i (9) można odpowiednio przedstawić w postaci:

$$X(k) = \sum_{n=0}^{\frac{N}{2}-1} \left[x(n) + x\left(n + \frac{N}{2}\right) \right] W_N^{kn} \quad \text{dla } k \text{ parzystego} \quad (10)$$

oraz

$$X(k) = \sum_{n=0}^{\frac{N}{2}-1} \left[x(n) - x\left(n + \frac{N}{2}\right) \right] W_N^n W_N^{kn} \quad \text{dla } k \text{ nieparzystego.} \quad (11)$$

Obliczanie FFT z podziałem częstotliwości polega zatem na wyznaczeniu ciągów

$x(n) + x\left(n + \frac{N}{2}\right)$ oraz $x(n) - x\left(n + \frac{N}{2}\right)$ pomnożonych przez współczynnik W_N^k , a następnie obliczeniu $N/2$ -punktowych transformat obu ciągów.

Podstawy akustoelektronicznej obróbki sygnałów

Kompresja impulsów w dziedzinie czasu polega na skróceniu czasu trwania ograniczonego w czasie impulsu T . Stosunek czasu trwania impulsu wejściowego do czasu trwania impulsu wyjściowego nazywa się współczynnikiem kompresji.

$$\frac{T_{we}}{T_{wy}} = K_k, \quad (1)$$

przy czym charakterystyka fazowo-częstotliwościowa widma sygnału musi spełniać warunek:

$$\Phi_s(\omega) \neq \omega t_0, \quad (2)$$

a jednym z elementów układu obróbki sygnałów jest filtr dopasowany, którego transmitancję wiąże z widmem sygnału następująca zależność:

$$K(j\omega) = A G^*(j\omega) e^{-j\omega t_0} \quad (3)$$

lub też wynikająca z niej para warunków:

$$|K(j\omega)| = A |G(j\omega)|, \quad (4a)$$

$$\Phi_k(\omega) = -[\Phi_s(\omega) + \omega t_0]. \quad (4b)$$

Pierwszy z tych warunków nakazuje aby charakterystyka amplitudowo-częstotliwościowa filtra pokrywała się z modulem widma sygnału (z dokładnością do stałej), co ma wpływ na jakość filtracji dopasowanej (odpowiednią wartość stosunku sygnał/szum).

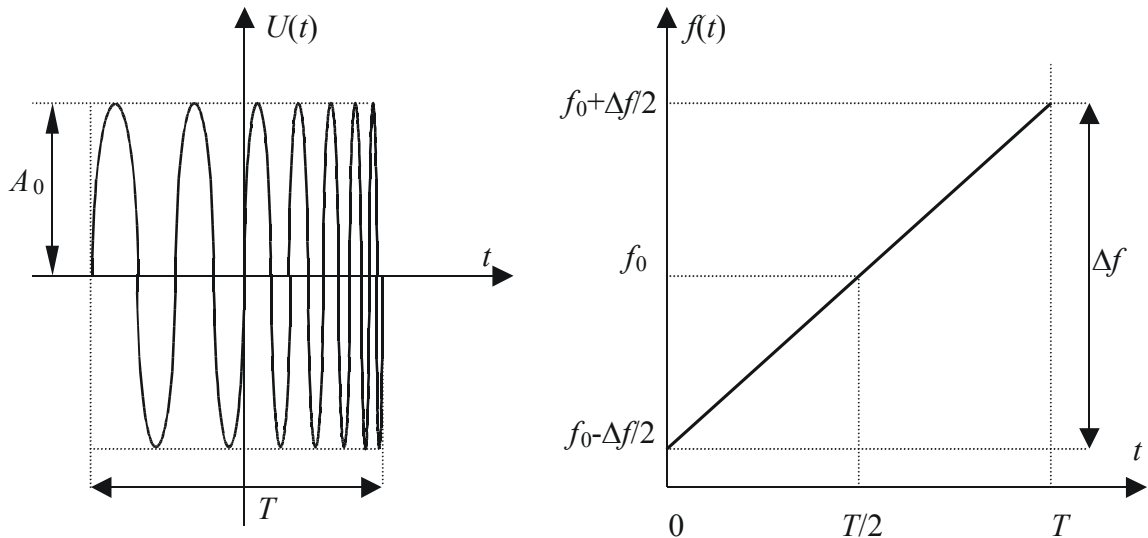
Drugi zaś, nazywany warunkiem kompensacji faz początkowych rozstrzyga o zaistnieniu samej kompresji impulsu. Kompresja zatem jest możliwa wyłącznie dla tzw. sygnałów złożonych czyli takich, dla których iloczyn:

$$m = T\Delta f \gg 1, \quad (5)$$

gdzie wielkość m nazywana jest bazą sygnału, T jest czasem jego trwania, Δf szerokością widma. Dla sygnałów prostych $m=1$.

Zwiększenie bazy sygnału uzyskuje się przez wprowadzenie modulacji wewnątrzimpulsowej, czyli spełnienie zależności (2). Zmiana wewnętrznej struktury impulsu może dotyczyć częstotliwości (fazy) lub amplitudy, bądź też kombinacji tych wielkości i może przebiegać w sposób ciągły (modulacja) lub dyskretny (manipulacja).

Jednym z najważniejszych z praktycznego punktu widzenia sygnałów złożonych jest sygnał z liniową modulacją częstotliwości (LFM) przedstawiony na rys. 1.



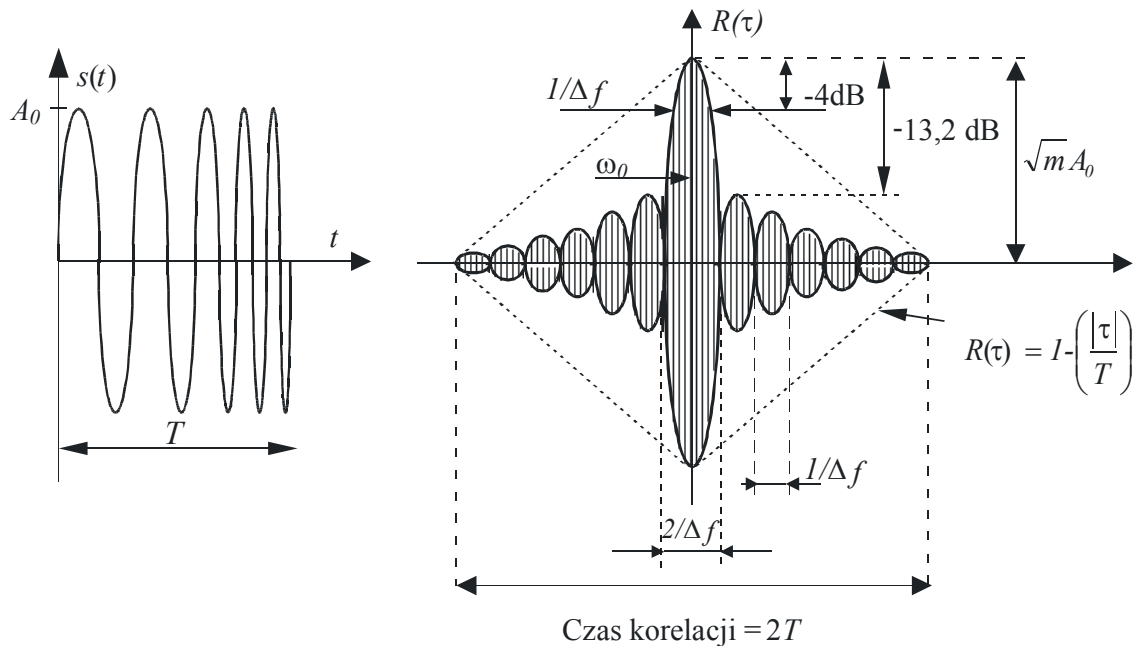
Rys. 1. Sygnał z LFM

Sygnał taki opisany jest zależnością:

$$u(t) = A_0 \cos\left(\omega_0 t + \frac{\mu t^2}{2}\right), \quad (6)$$

gdzie A_0 jest amplitudą, a $\mu = 2\pi\Delta f/T$.

Impuls z LFM na wejściu i wyjściu filtru dopasowanego przedstawiony jest na rys. 2.



Rys. 2. Sygnał z LFM. Postać sygnału na wejściu (po lewej) oraz na wyjściu (po prawej) filtru dopasowanego

Widmo tego sygnału ma w przybliżeniu postać prostokątną i jest tym bardziej prostokątne im większa jest baza sygnału. Na wyjściu filtru dopasowanego do sygnału z LFM napięcie wyjściowe przyjmuje postać:

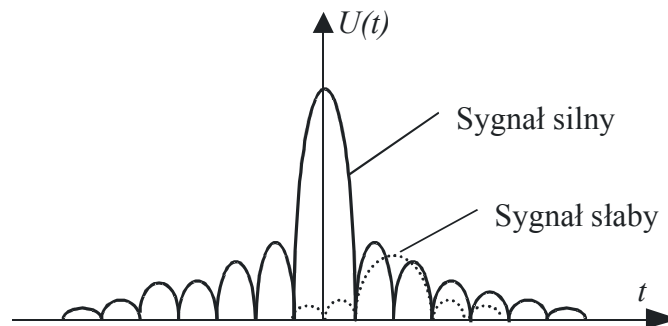
$$y(t) = \sqrt{m} A_0 \frac{\sin \frac{\Delta\omega}{2}(t-T)}{\frac{\Delta\omega}{2}(t-T)} \cos \omega_0(t-T) \quad (7)$$

i pokrywa się z funkcją autokorelacji sygnału wejściowego $r(t)$, która wyraża się ogólną zależnością:

$$R_x(t_1, t_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} X_1 X_2 f(x_1, x_2, t_1, t_2) dx_1 dx_2 \quad (8)$$

Zależność ta opisuje statystycznie uśrednioną liniową współzależność między wartościami sygnału w przekrojach czasowych t_1 i t_2 , ma postać impulsu o obwiedni typu sincx wypełnionego stałą częstotliwością nośną f_0 . Dla tego typu sygnałów charakterystyczne są tzw. listki boczne, występujące symetrycznie po obu stronach listka głównego. Z rys. 2 wynika również, że listek główny tego napięcia, stanowiący sygnał użyteczny, uległ znacznemu skróceniu w stosunku do sygnału wejściowego i czas jego trwania, mierzony na poziomie -4dB , jest równy $1/\Delta f$, zaś amplituda wzrosła \sqrt{m} razy.

Stałość częstotliwości nośnej impulsu po kompresji (skomprimowanego) oraz skrócenie czasu jego trwania wynika z warunku (4b). Warunek (4a) wpływa natomiast na amplitudę listka głównego, jego szerokość oraz poziom listków bocznych (stosunek amplitudy najwyższego z listków bocznych do amplitudy listka głównego). Dla sygnału z LFM poddawanego obróbce w filtrze dopasowanym poziom ten jest wysoki, co stwarza niebezpieczeństwo maskowania sygnału słabszego przez listki boczne sygnału silniejszego przy przetwarzaniu sygnałów znacznie różniących się poziomami (rys. 3).



Rys. 3. Maskujące działanie listków bocznych sygnału silnego

Z powodów widocznych na rys. 3 poziom listków bocznych jest istotny z punktu widzenia obróbki sygnałów złożonych. W celu jego obniżenia stosuje się tzw. ważenie. Polega ono na celowym wprowadzeniu niedopasowania pomiędzy widmem sygnału a funkcją przenoszenia filtru odbiorczego:

$$|K(j\omega)| \neq A|G(j\omega)|. \quad (9)$$

Funkcję przenoszenia filtru dobiera się oczywiście w taki sposób aby uzyskać maksymalne stłumienie listków bocznych przy równocześnie najmniejszym poszerzeniu listka głównego i najmniejszych stratach dynamiki. Do najpopularniejszych funkcji wagowych należy niewątpliwie funkcja Hamminga, zgodnie z którą charakterystyka amplitudowo-częstotliwościowa filtru odbiorczego przyjmuje postać:

$$K_H(\omega) = 0,08 + 0,92 \cos \frac{2\pi(\omega - \omega_0)}{\Delta\omega}. \quad (10)$$

Zastosowanie tej funkcji pozwala uzyskać znaczne obniżenie poziomu listków bocznych przy ok. 1,5-krotnym poszerzeniu listka głównego oraz niewielkich stratach dynamiki. Niekorzystne efekty towarzyszące obróbce wagowej wynikają z wprowadzonego niedopasowania.

Dla sygnałów z LFM o dużej bazie ($m \geq 30$) synteza filtra dopasowanego sprowadza się do wykonania filtra o maksymalnie prostokątnej charakterystyce amplitudowo-częstotliwościowej i charakterystyce fazowo-częstotliwościowej o postaci:

$$\Phi_k(\omega) = -[\Phi_s(\omega) + \omega t_0] = -\frac{(\omega - \omega_0)^2}{2\mu} - \omega t_0. \quad (11)$$

Kwadratowi zmian fazy odpowiadają zatem liniowe zmiany opóźnienia czasowego w funkcji częstotliwości:

$$t_d(\omega) = \frac{d\Phi_k(\omega)}{d\omega} = -\frac{(\omega - \omega_0)}{\mu} - t_0, \quad (12)$$

gdzie t_0 jest stałym opóźnieniem spełniającym warunek $t_0 \geq T$ będący warunkiem realizowalności filtra.

Funkcja $t_d(\omega)$ odzwierciedlająca zależność opóźnienia czasowego harmonicznego widma sygnału od częstotliwości nazywa się funkcją dyspersyjną. Dla sygnału z LFM ma ona postać

$$t_d'(\omega) = \frac{d\Phi_s(\omega)}{d\omega} = \frac{\omega - \omega_0}{\mu}. \quad (13)$$

Istnieją zasadniczo dwie metody realizacji takich charakterystyk dyspersyjnych sygnałów. Pierwsza z nich - aktywna polega na formowaniu sygnału z liniową charakterystyką dyspersyjną w dziedzinie czasu, natomiast druga - pasywna polega na formowaniu analogicznego sygnału w dziedzinie częstotliwości i realizowana przez układ, którego głównym ogniwem jest filtr dyspersyjny o charakterystyce określonej przez (12). Filtr dyspersyjny spełnia w tej metodzie funkcję elementu „rozciągającego” impuls w dziedzinie czasu, jak i nadającego mu zadaną strukturę wewnętrzną. Wspólną cechą tych metod jest formowanie sygnału na niskim poziomie mocy i przy niskiej częstotliwości. Z racji występowania w metodzie aktywnej wielu problemów związanych głównie z niedopasowaniem w praktyce wykorzystuje się raczej metodę pasywną.