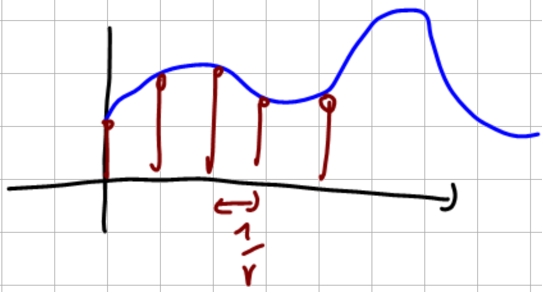
C2, J5, T1

$x(t) = [\cos(7t)]^2$ $\quad r = 10$ kHz $\quad x(n) = ?$ $\quad 0 \le n < 5$



sampling: $x(t) \rightarrow x(n)$ with $t = \frac{n}{r}$

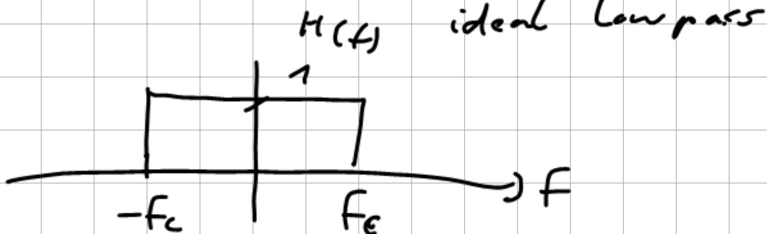
$x(n=0) = \left[\cos\left(7 \cdot \frac{0}{r}\right)\right]^2 = 1^2 = 1$

$x(n=1) = \left[\cos\left(7 \cdot \frac{1}{10000}\right)\right]^2 = 0,9999$

cos in radiant!

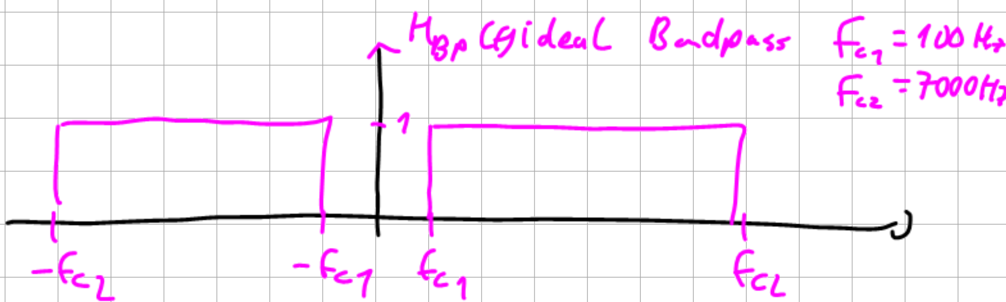
C2 J5 T2

$H(f) = \mathrm{rect}\left(\frac{f}{2f_c}\right)$



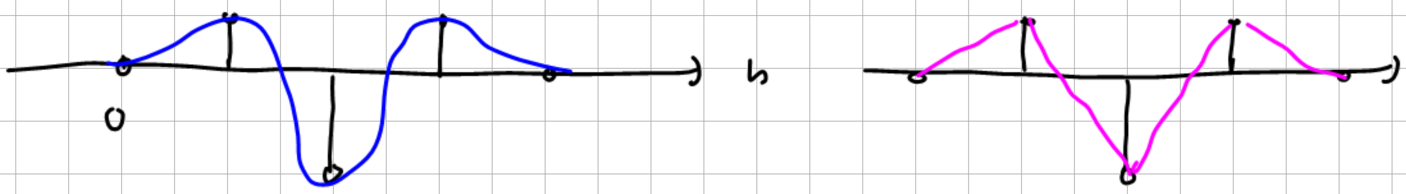
$h(t) = 2f_c \cdot \mathrm{sinc}(2 \cdot f_c \cdot t)$

$= 2 \cdot f_c \cdot si(2 \cdot f_c \cdot \pi \cdot t) = 2 \cdot f_c \cdot \frac{\sin(2f_c \pi t)}{2f_c \pi t}$



$x(n) = [0; 1; -2; 1; 0]$

linear interpolation



C2 J5 T4

has $z(n)$ a DC component? no, DC component has a frequency of $f=0$ Hz which cannot pass the Band-pass.

human voices understandable in $z(n)$?

old phone system / narrowband : 300 Hz ... 3400 Hz

new -"- / wideband : 50 Hz ... 7000 Hz

$\Rightarrow$ the range from 100 Hz up to 7000 Hz:

human voice can pass

highest frequency in $y(n)$? 24 kHz

-"- in $z(n)$? 7 kHz

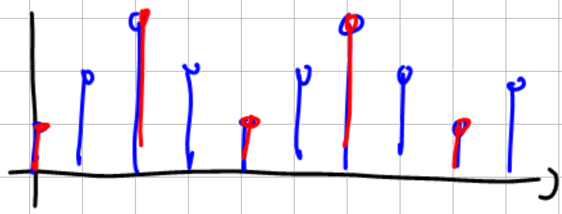
C2 J5 T3

$r = 32$ kHz $f_c = 3400$ Hz

input $\rightarrow$ low pass $\rightarrow$ output $\leftarrow$ wideband or narrowband

it is narrowband

sampling theorem: $f_c = 3400$ Hz $\Rightarrow$ lowest allowed sampling rate $= 6800$ Hz



e.g. $c = 2$

$c = 5 \Rightarrow 5 \cdot 6800\ \text{Hz} = 34000\ \text{Hz}$

$c = 4 \Rightarrow 4 \cdot 6800\ \text{Hz} = 27200\ \text{Hz}$ ✓

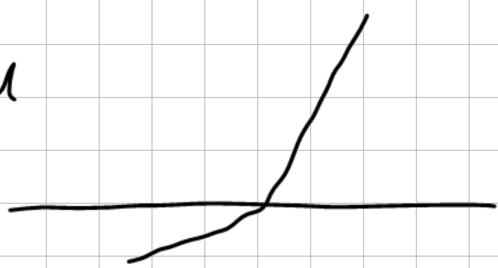
C2, J4, T3

Filter length $= 47 = N$

delay $= 23 \rightarrow \left\lfloor \frac{N}{2} \right\rfloor \leftarrow$ rounded down

C4 J2 T3

Leaky ReLU



$$f(x) = \begin{cases} x, & x \geq 0 \\ \alpha x, & \text{else} \end{cases}$$

$\alpha = 0.3$

counter example for nonlinearity

$a = b = 1 \qquad x_1 = 1 \qquad x_2 = -1$

$$f(a \cdot x_1 + b \cdot x_2) = a \cdot f(x_1) + b \cdot f(x_2)$$

$$f(1 \cdot 1 + 1 \cdot (-1)) = 1 \cdot f(1) + 1 \cdot f(-1)$$

$$f(0) = 1 \cdot 1 + 1 \cdot (-0.3 \cdot 1)$$

$$0 = 1 - 0.3 \quad \lightning$$

Dense Layer: $y_j = \sum_{i=0}^{I-1} w_{ji} x_i + b_j = f(x_i)$

$$f(a \cdot x_1 + b \cdot x_2) = \sum_{i=0}^{I-1} w_{ji} \left(a \cdot x_{1i} + b \cdot x_{2i}\right) + b_j$$

$$= a \cdot \sum_{i=0}^{I-1} w_{ji} \, x_{1i} + b \cdot \sum_{i=0}^{I-1} w_{ji} x_{2i} + b_j$$

$$a f(x_1) + b \cdot f(x_2) = a \cdot \left(\sum_{i=0}^{I-1} w_{ji} x_{1i} + b_j \right) + b \cdot \left(\sum_{i=0}^{I-1} w_{ji} \, x_{2i} + b_j \right)$$

$$= a \sum_{i=0}^{I-1} w_{ji} \, x_{1i} + b \cdot \sum_{i=0}^{I-1} w_{ji} x_{2i} + a \cdot b_j + b \cdot b_j$$

$\neq$

not linear

Fourier Transform (discrete)

$$y(k) = \sum_{n=0}^{N-1} x(n) \, e^{-j 2\pi \frac{nk}{K}} = f(x(n))$$

$$f(a \cdot x_1 + b \cdot x_2) = \sum_{n=0}^{N-1} (a \, x_1(n) + b \cdot x_2(n)) \; e^{-j 2\pi \frac{nk}{K}}$$

$$= \sum_{n=0}^{N-1} a \, x_1(n) \, e^{-j 2\pi \frac{nk}{K}} + b \, x_2(n) \, e^{-j 2\pi \frac{nk}{K}}$$

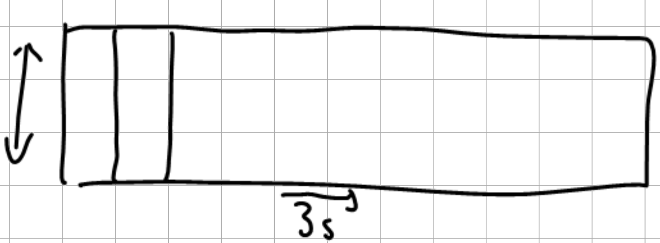
$$a \cdot f(x_1) + b \cdot f(x_2) = a \cdot \sum_{n=0}^{N-1} x_1(n) \, e^{-j 2\pi \frac{nk}{K}} + b \cdot \sum_{n=0}^{N-1} x_2(n) e^{-j 2\pi \frac{nk}{K}}$$

$$= \sum_{n=0}^{N-1} a \, x_1(n) \, e^{-j 2\pi \frac{nk}{K}} + b \cdot x_2(n) \; e^{-j 2\pi \frac{nk}{K}} \quad \checkmark$$

Fourier Transform is linear.

C3, J5, T5

audio 3 seconds, hopsize 10 ms, how many coefficients in the MFCC matrix?

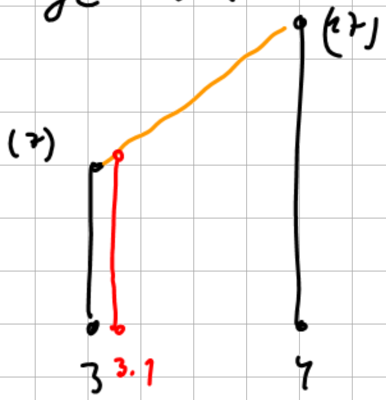


each hopsize introduce a new column

$\Rightarrow$ 300 colums $= \frac{3s}{10ms}$

$\Rightarrow 12 \cdot 300 = 3600$ MFCC Feature

C3 J2 T1



$$f(x) = m x + n$$

$$7 = m \cdot 3 + n$$

$$17 = m \cdot 4 + n$$

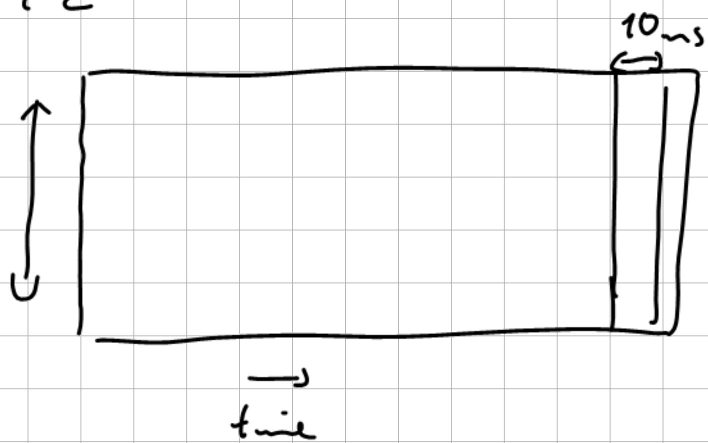
$$m = \frac{17-7}{4-3} = 10$$

$$7 = 10 \cdot 3 + n \Rightarrow n = -23$$

$$f(x) = 10 \cdot x - 23$$

$$f(3.1) = 10 \cdot 3.1 - 23 = 31 - 23 = \underline{\underline{8}}$$

C3, J2, T2



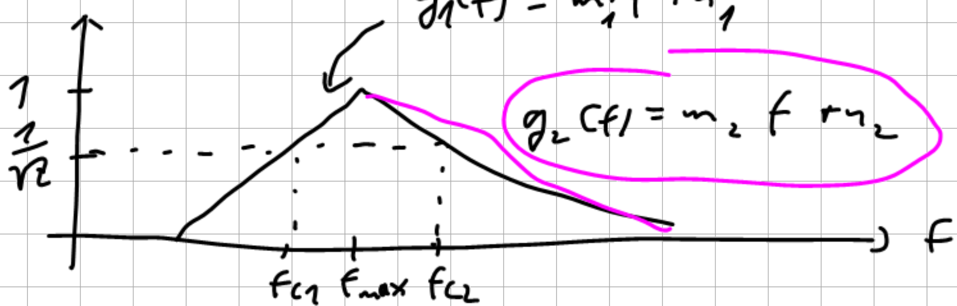
4 Bytes / Bandpass output

data rate: 24 Bandpasses $\cdot$ 4 Bytes / Bandpass $\cdot$ 100 $\frac{1}{s}$

$$= 24 \cdot 4 \cdot 100 \ \frac{\text{Bytes}}{s}$$

C3 J3, T3

$$g_1(f) = m_1 f + n_1$$



$$g_2(f) = m_2 f + n_2$$

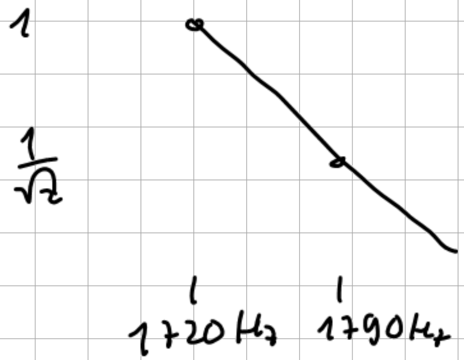
$$-3\,dB = 20 \log_{10} a \Rightarrow a = 10^{\frac{-3}{20}} = 0.707 = \frac{1}{\sqrt{2}}$$

$$\frac{F}{kHz} = 1.96 \cdot \frac{b + 0.53}{26.28 - b}$$

$f_m = 12$ Bark $\rightarrow$ $f_m = 1.72$ kHz $= 1720$ Hz

$f_{c1} = 11.75$ Bark

$f_{c2} = 12.25$ Bark $\rightarrow$ $f_{c2} = 1.79$ kHz $= 1790$ Hz



$$m_2 = \frac{\frac{1}{\sqrt{2}} - 1}{1790\,Hz - 1720\,Hz} = -4.18 \cdot 10^{-3}\,s$$

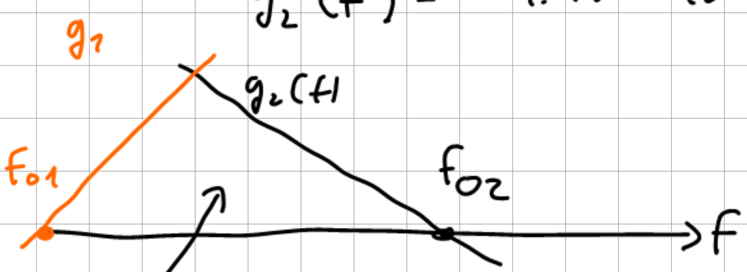
$$1 = m_2 \cdot 1720\,Hz + n_2$$

$$n_2 = 1 - m_2 \cdot 1720\,Hz$$

$$= 1 - (-4.18 \cdot 10^{-3}\,s) \cdot 1720\,Hz$$

$$= 8.20$$

$$g_2(f) = -4.18 \cdot 10^{-3}\,s \cdot f + 8.20$$



Area $= \frac{1}{2} \cdot 1 \cdot (f_{o2} - f_{o1})$

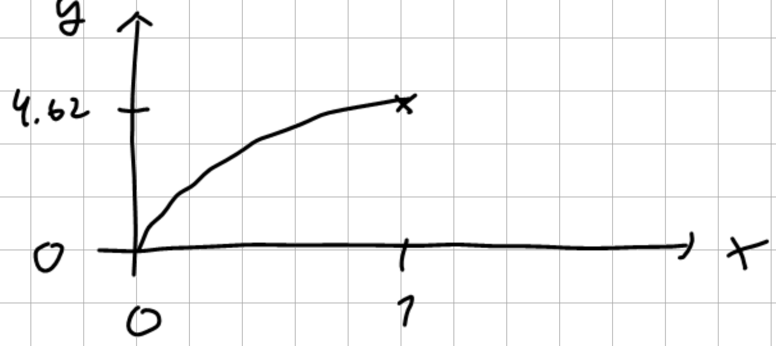
$$g_2(f) = 0 \Leftrightarrow -4.18 \cdot 10^{-3}\,s \cdot f_{o2} + 8.20 = 0$$

$$f_{o2} = \frac{-8.20}{-4.18 \cdot 10^{-3}\,s}$$

$$= 1961\ Hz$$

C3 J4 T1

$$y = \log_e(100\,x + 1) \qquad 0 \leq x < 1$$



$\log_e(100 \cdot 1 + 1) = 4,62$

C3 J4 T3

$f(x) = \dots$

sensitivity : $\frac{df}{dx}$ first derivative

e.g. $f(x) = \log x$

sens : $\frac{d \log x}{dx} = \frac{d f(x)}{dx}$

C2 J8 T3

$N = 1000$
$K = 1024$

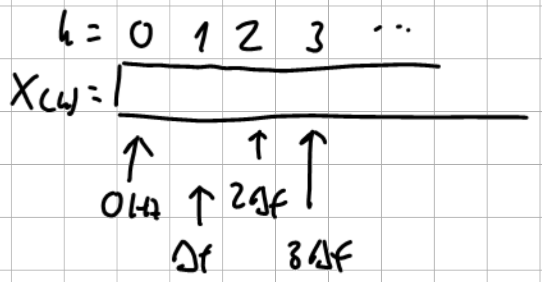
zero padding? yes : $K > N$

no : $K = N$

$K < N$ has no practical application

$$\Delta f = \frac{1}{K \cdot T} \qquad T = \frac{1}{r}$$

$$= \frac{r}{K} = \frac{48\ kHz}{1024} = 46.9\ Hz$$

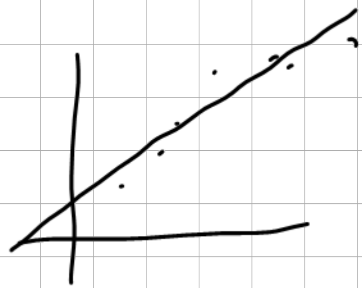


$k = 1 \hat{=} f = k \cdot \Delta f = 46.9\ Hz \approx 50\ Hz$

$$h = \frac{f}{\Delta t}$$

e.g.: $f = 432\ \text{Hz} \Rightarrow h = \frac{432\ \text{Hz}}{46.9\ \text{Hz}} = 9.21 \approx 9$

C3]5 T1



minimize: $\sum_r \left(m \cdot r + n - NG(r)\right)^2$

$$\frac{d}{dm} = \sum_r \cancel{2}\left(mr + n - NG(r)\right) \cdot r = 0$$

$$\frac{d}{dn} = \sum_r \cancel{2}\left(mr + n - NG(r)\right) = 0$$

$$\sum_r m \cdot r^2 + \sum_r n \cdot r = \sum_r r \cdot NG(r)$$

$$\sum_r m \cdot r + \sum_r n = \sum_r NG(r)$$

$$\begin{pmatrix} \sum r^2 & \sum_r r \\ \sum_r r & \sum_r 1 \end{pmatrix} \begin{pmatrix} m \\ n \end{pmatrix} = \begin{pmatrix} \sum_r r \cdot NG(r) \\ \sum_r NG(r) \end{pmatrix}$$

$NG(r) = 0.166,\ 2.385,\ 0.505,\ 3.1,\ 0.836$

$r = 0 \quad 1 \quad 2 \quad 3 \quad 4$

restrict to the first 5 Elements

$$\sum_r r^2 = 0 + 1 + 4 + 9 + 16 = 30$$

$$\sum_r r = 0 + 1 + 2 + 3 + 4 = 10$$

$$\sum_r 1 = 5$$

$$\sum_r r \cdot NG(r) = 0 \cdot 0.166 + 1 \cdot 2.385 + 2 \cdot 0.505 + 3 \cdot 3.1 + 4 \cdot 0.836$$

$$= 16.04$$

$$\sum_r NG(r) = 0.166 + 2.385 + 0.505 + 3.1 + 0.836$$
$$= 6.99$$

$$\begin{pmatrix} 30 & 10 \\ 10 & 5 \end{pmatrix} \begin{pmatrix} m \\ n \end{pmatrix} = \begin{pmatrix} 16.04 \\ 6.99 \end{pmatrix} \quad \cdot 3 \;] \; -$$

$$0\, m + (-5)\cdot n = 16.04 - 3 \cdot 6.99$$

$$n = \frac{16.04 - 3 \cdot 6.99}{-5} = 0.986$$

$$30 \cdot m + 10 \cdot 0.986 = 16.04$$

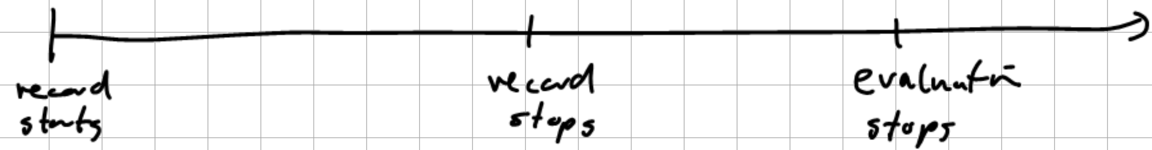
$$m = \frac{16.04 - 9.86}{30} = 0.206$$

$$\begin{pmatrix} \sum_r r^2 & \sum_r r \\ \sum_r r & \sum_r 1 \end{pmatrix} \quad \ldots$$

```
r = np.arang(10)

[ ← r = np.zeros((10))
  for i in range(r.shape[0]):
      r[i] = i ]  ←

a = np.sum(r*r)    or   np.sum(r**2)

b = np.sum(r)

c = r.shape[0]     or   np.sum(r**0)

A = np.array([[a, b], [b, c]])
```

C4]6 T2

audio 30 s evaluatin 28 s



Latency = 58 s

real time factor $= \frac{28}{30} < 1$

$\Rightarrow$ in principle capable of real time evaluation.

C4 J6 T3

not capable of real time application $\Rightarrow$ whisper takes much larger than 500ms

| user experience |

C4 J6 T1 Levenshtein distance

evaluated by each letter
- each insertion : distance + 1 — 2
- each missing : distance + 1 — 7
- each replacing : distance + 1 — 2

11 = Levenshtein distance

correct sentence :

T o d a y i t i s m u c h t o o c o l d .

T w o d a y s i t i s 2 c o l d .

C4 J7 ~~T~~ T2

$$p(n) = \frac{1}{n \cdot \ln(1.78 \cdot 10^5)}$$

$$p(1) = \frac{1}{1 \cdot \ln(1.78 \cdot 10^5)} = \dots$$

$$p(2) = \frac{1}{2 \cdot \ln(1.78 \cdot 10^5)} = \dots$$

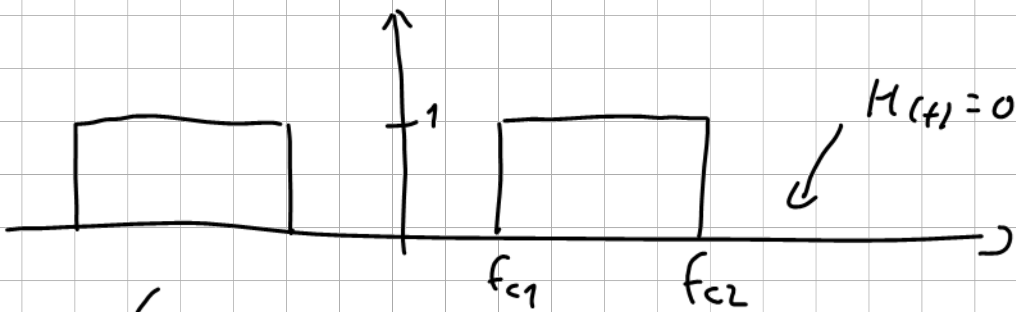
C4 J5 T1

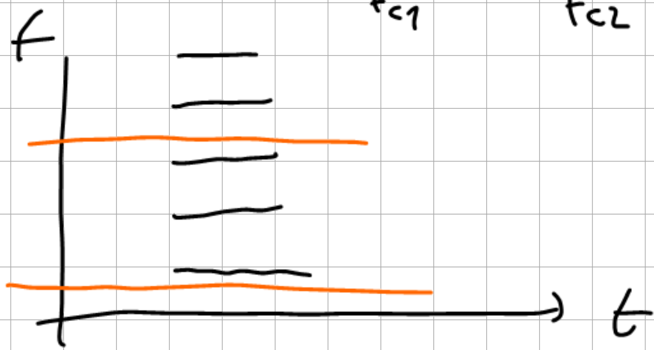
$x \longrightarrow \boxed{h} \longrightarrow y = h * x$

$Y(f) = H(f) \cdot X(f)$

$X(f) = \frac{1}{H(f)} \cdot Y(f)$

$H(f)$ ideal Bandpass





C4 J4