

Lektor 2 for Stovner VGS 12. februar 2013

Sampling av lyd & sånn

Bjørn Aarseth

- Født i Sandefjord for 54 år siden
- Døpt Bjørn Aarseth
- Utdannet elektronikingeniør
- Ansatt i NRK i 1981
- 5. september 2004 = halve livet i NRK!
- Interesser: musikk og elektronikk



Bjørn i NRK



- Begynte med tekn. vedlikehold
- Deretter lab/utvikling av analog/digital lydelektronikk
- Nå: Rådgiver i NRKs Teknologidivisjon
- Medlem i "EBU Pool of Experts"
- På samvittigheten:
 - Alltid Klassisk
 - Alltid Nyheter
 - OBT-studio ("Svisj" mm)
- Akkurat nå:
 - HD og 5.1 i NRK TV
 - Digitalisering av Platearkivet
 - DAB, DVB-T og DRM. Internasjonalt arbeid (EBU)
 - Loudness normalisering av NRKs produksjon



Oppdraget...

- I fysikkpensumet i videregående skole, omhandler ett læreplanmål **lydsampling** i kurset **Fysikk 2** i 3.klasse.
- Læreplanmål lyder: **"Gjøre rede for sampling og digital behandling av lyd"**
- I denne sammenhengen hadde det vært fint om elevene kunne få se dette utført i praksis.

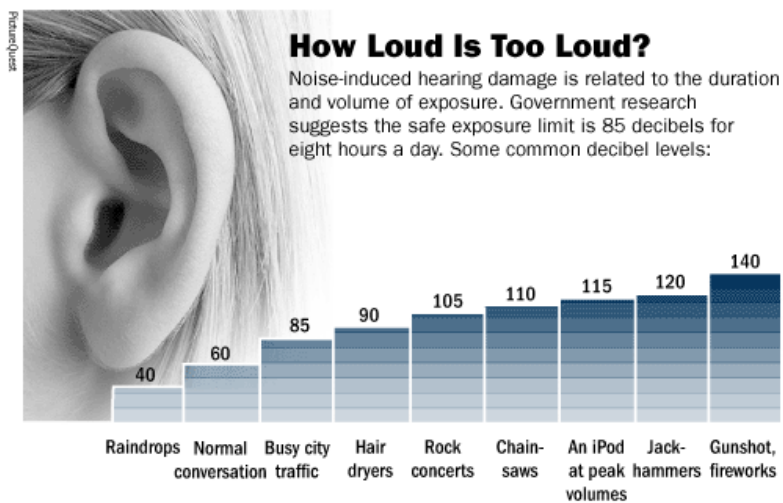
Våre sanser er logaritmiske



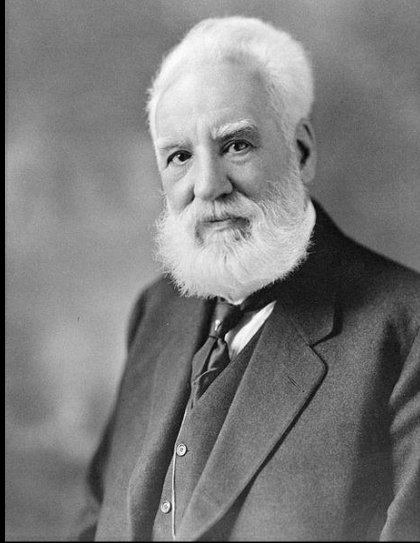
Hvorfor?



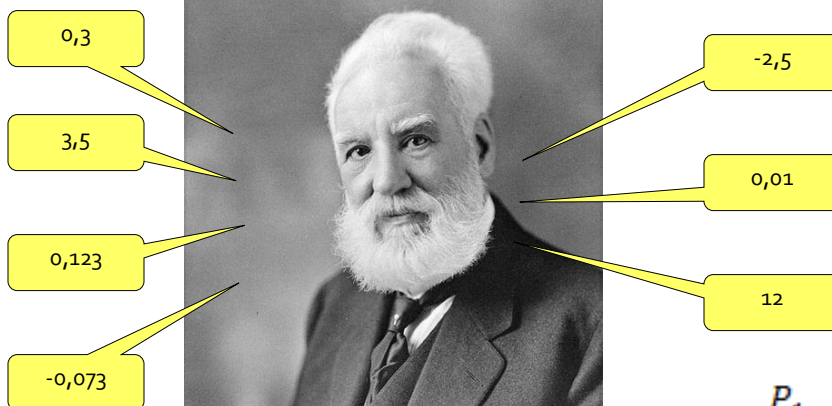
Vi er klar for desibel !!!



Alexander Graham Bell



Måleenhet = Bel??



$$B = \lg \frac{P_1}{P_0}$$

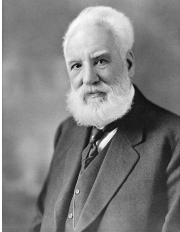
Måleenhet = desiBel!!

3

35

1,23

-0,73

$$dB = 10 \lg \frac{P_1}{P_0}$$


-25

0,1

120

Leksikon, litt på bærtur

- **desibel**, symbol dB, lik 1/10 → bel.
- **bel**, symbol B, enhet for logaritmen (med 10 som grunntall) til forholdet mellom to effekter. Navn etter A.G. Bell. Mest brukt er desibel (dB) (1B = 10 dB), for måling av lydstyrke. Ørets smertegrense er 130 dB.



<http://www.caplex.no/Web/ArticleView.aspx?id=9301704>

Viktig!

dB er en effekt- definisjon!

Desibel

$$dB = 10 \cdot \log_{10} \left(\frac{P_1}{P_0} \right)$$

Eksempler:

$$P_1 = P_0 \rightarrow 0dB$$

$$P_1 = 2 \cdot P_0 \rightarrow 3dB$$

$$P_1 = 0,5 \cdot P_0 \rightarrow \div 3dB$$

$$P_1 = 10 \cdot P_0 \rightarrow 10dB$$



Fra effekt til spenning og strøm

- Ohms lov: $U = R I$
- Effektloven: $P = U I$
- I kombinasjon: $P = U^2/R = I^2 R$
- Altså:

$$dB = 10 \lg \frac{P_1}{P_0} = 10 \lg \frac{U_1^2}{U_0^2} = 20 \lg \frac{U_1}{U_0}$$

Ikke no'
mystisk!

Spennning: $dB = 20 \cdot \log$

$$U = R \cdot I \quad \text{og} \quad P = U \cdot I \quad \rightarrow \quad P = \frac{U^2}{R}$$

$$dB = 10 \cdot \log_{10} \left(\frac{P_1}{P_0} \right) = 10 \cdot \log_{10} \left(\frac{U_1^2 / R}{U_0^2 / R} \right) = 20 \cdot \log_{10} \left(\frac{U_1}{U_0} \right)$$

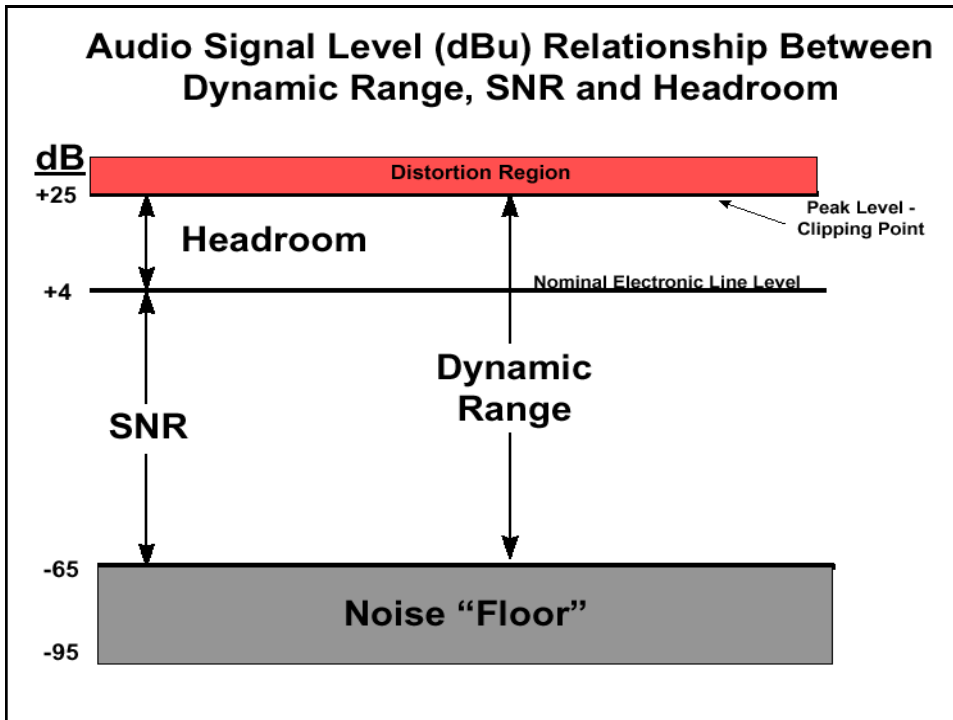
Eksempler:

$$U_1 = U_0 \quad \rightarrow \quad 0dB$$

$$U_1 = 2 \cdot U_0 \quad \rightarrow \quad 6dB$$

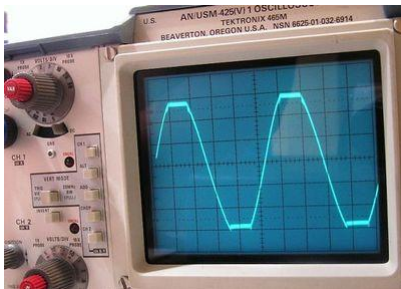
$$U_1 = 0,5 \cdot U_0 \quad \rightarrow \quad \div 6dB$$

$$U_1 = 10 \cdot U_0 \quad \rightarrow \quad 20dB$$



Termisk støy og "klipping"

$$e_n = \sqrt{4 \cdot k \cdot T \cdot B_N \cdot R}$$



Eksempel:

50 ohm dyn. mik.
 $e_n = 0,13\mu\text{V} = 135 \text{ dBu}$
 med 60 dB gain: 75 dBu
 med + 22dBu som
 klippenivå: **DR = 97 dB**
 med +4 dBu som nom.
 maks.: **SNR = 79 dB**

$$k = 1.38 \times 10^{-23} \text{ J/K}$$

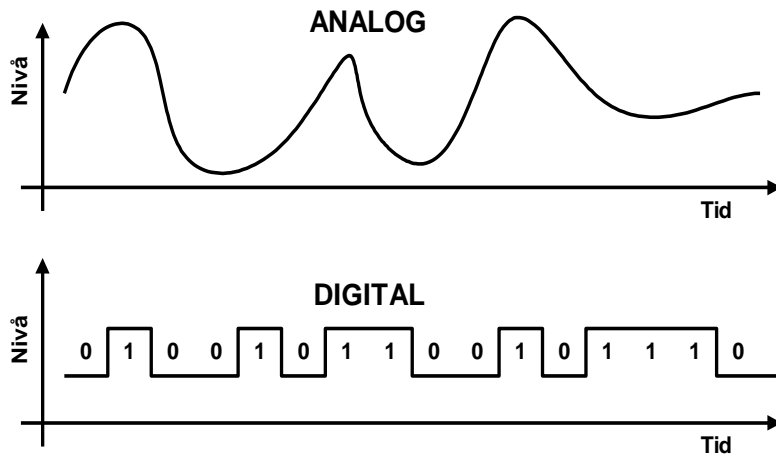
Analog vs. digital:

- Vi lar lyd, bilde og andre signal bli representert av en **analog** størrelse, f.eks. en elektrisk spenning.
- Med **analog** menes en *likeverdig* størrelse.
- Matematisk beskrivelse:
 - Signalet er kontinuerlig i tid og amplitude
 - Signalet eksisterer for alle "tidspunkter" og alle nivåer

Analog vs. digital:

- Vi lar lyd, bilde og andre signal bli representert ved en strøm av tall.
- Disse "tallene" representerer et utsnitt av signalet for noen enkelte tidspunkt og noen enkelte nivåer.
- Matematisk beskrivelse:
 - Signalet er **diskret** i tid og amplitude.
 - Signalet eksisterer kun i enkelte tidspunkter, representert ved et et endelig antall tallverdier.

Analoge og digitale signaler:



Ulemper med analog:

FRAGILE
HANDLE
WITH CARE

- Støy og forvrengning av signalet adderes i hvert ledd signalet passerer:
 - Lagring ("konservering") av signalet reduserer kvaliteten.
 - Kopier blir dårligere enn originalen.
 - Overføring over lange avstander medfører redusert kvalitet. Jo lengre avstand, desto dårligere kvalitet.
- Signalkvaliteten er avhengig av mediet

Digitale signaler

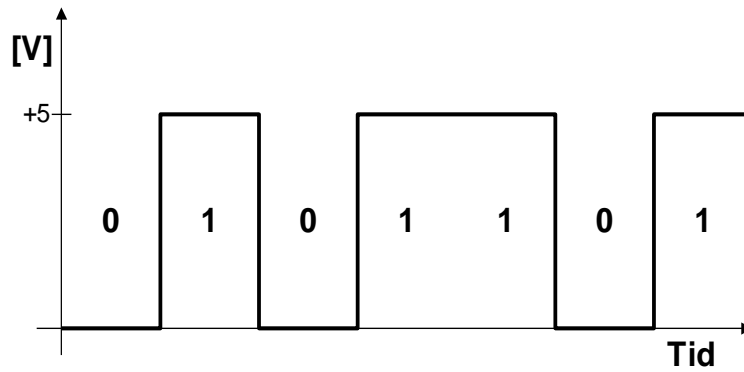
- "Digital" betyr at signalet overføres eller lagres på numerisk form, altså som tall.
- Den "eneste" degraderingen av signalets kvalitet skjer normalt ved konvertering (mellom analog og digital og vice versa).

Fordeler med digital

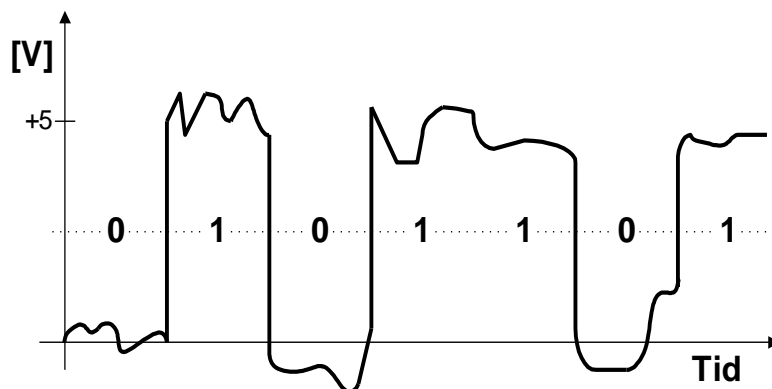


- Signalet er mer robust
 - Vi kan overføre, lagre og kopier så mange generasjoner vi vil uten degradering.
 - Signalkvalitet er uavhengig av medie.
- Digital gir nye muligheter
 - Signalet kan behandles og manipuleres med på helt nye måter.
 - Eks: Ikke-lineær og ikke-destruktiv redigering
 - Vi er i "DATA – VERDEN" !!
 - **Alt** kan lagres
 - Billig billig billig billig.....

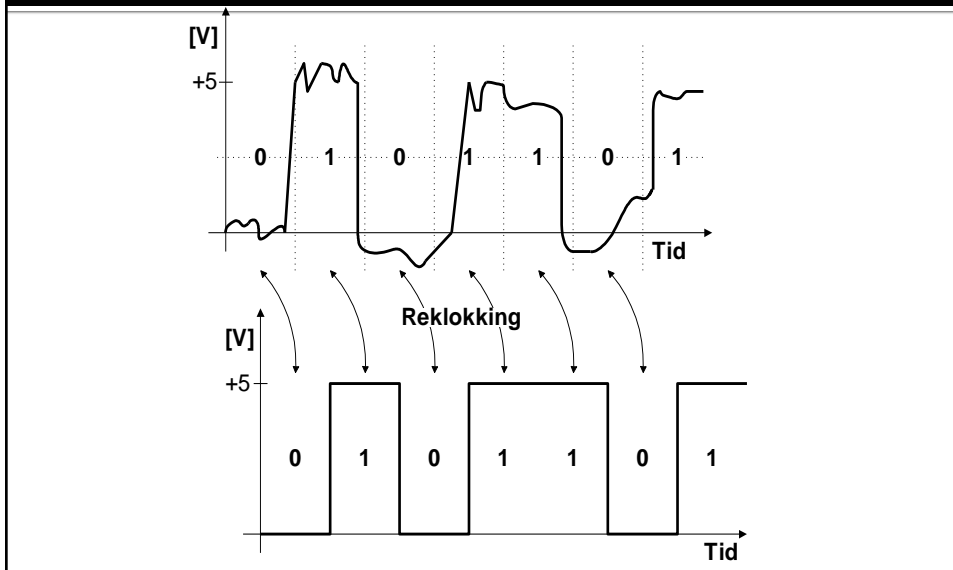
Digitale signaler er robuste (1)



Digitale signaler er robuste (2)



Digitale signaler er robuste (3)



Ulemper med digital:

- Kostbare store bokser med stort effektforbruk?
 - Det var *før*, det..... Nå er det mye billigere!
- Innfører **alltid** forsinkelse
- Må synkroniseres
- Krever ny måleteknikk og nye metoder for evaluering av kvalitet (eks.: MPEG Audio)
- Feil blir mer fatale/uopprettelige
- Ny teknologi krever opplæring

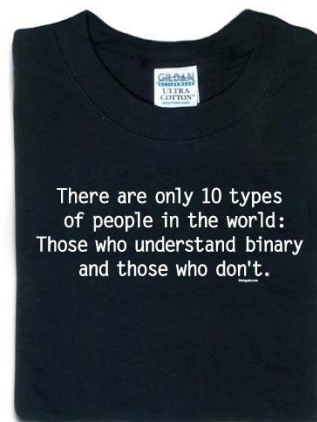
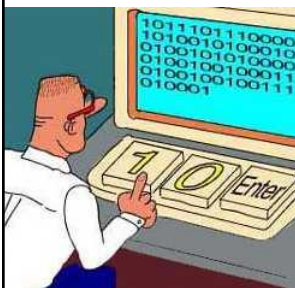
Digital kommer fra latin → finger

- Vi teller på fingrene....
- **Røyksignaler** er en av de tidligste kjente anvendelser av digital ("diskret") teknikk
- **Morse** er den første elektriske anvendelse av digitalteknikk
- I moderne tid assosieres "digital" med data-teknikk og **binære tall**



Binary digiT => BIT

- Et bit er et siffer med binærtall
- Binær = to-tall systemet
- Et bit kan ha verdien 0 eller 1



n bit gir 2^n variasjoner

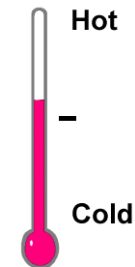
- 1 bit = 2 mulige verdier
- 2 bit = 4 mulige verdier
- 3 bit = 8 mulige verdier

0
0

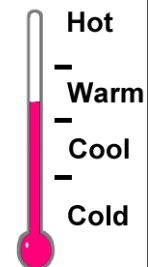
Og dette er riktig?

0	0
0	1
1	0
1	1

0	0	0
0	0	1
0	1	0
0	1	1
1	0	0
1	0	1
1	1	0
1	1	1



1 Bit



2 Bits

"Group by eight" => BYTE

- En BYTE består av 8 bit
- 8 bit = $2^8 = 256$ variasjoner
- En byte kan for eksempel beskrive en bokstav



The char A

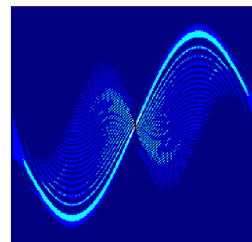
0	0	0	0	0	1	0	0
---	---	---	---	---	---	---	---

Bits & Byte

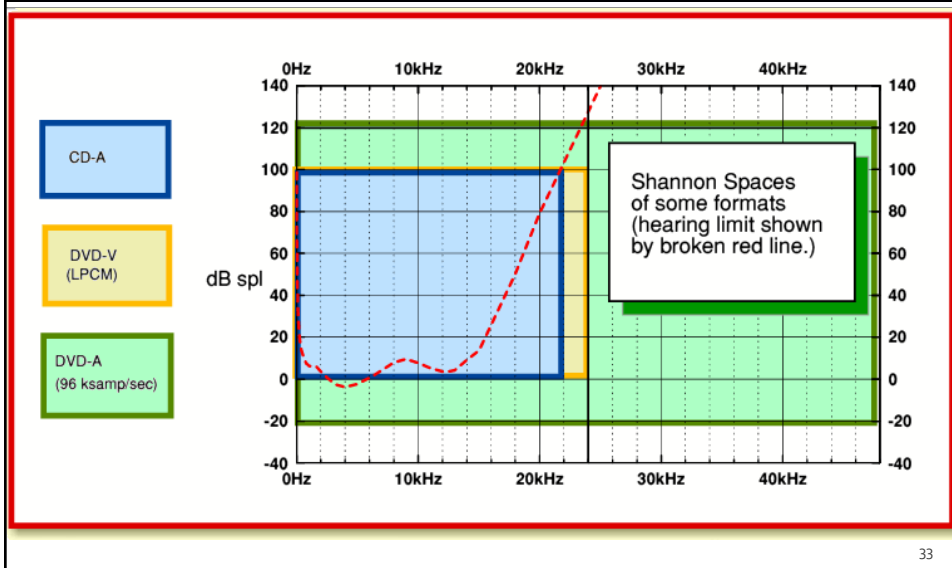
- Overføringshastighet måles normalt i bit pr. sekund
 - ISDN har to 64 kb/s (kilobit pr. sekund) kanaler
 - Stereolyd i CD-kvalitet er ca. 1,5 Mb/s
 - Seriedigital video (SD-SDI) er 270 Mb/s
- Lagringskapasitet måles normalt i antall byte
 - En CD rommer ca. 700 MB (megabyte)
 - En DVD rommer fra 4,7 GB (gigabyte)
 - En normalt stor harddisk rommer 1 TB

To fundamentale begrensninger

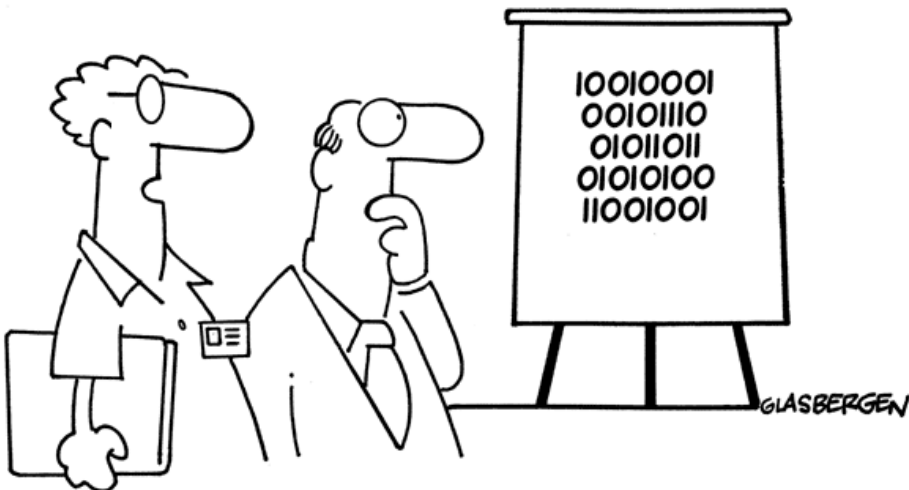
- Digitalisering av lyd innfører to fundamentale begrensninger på lydkvaliteten:
- **Sampling** begrenser oppløsning i *tid/frekvens*
- **Kvantisering** begrenser lydens *dynamikkområde*



"Shannon space"

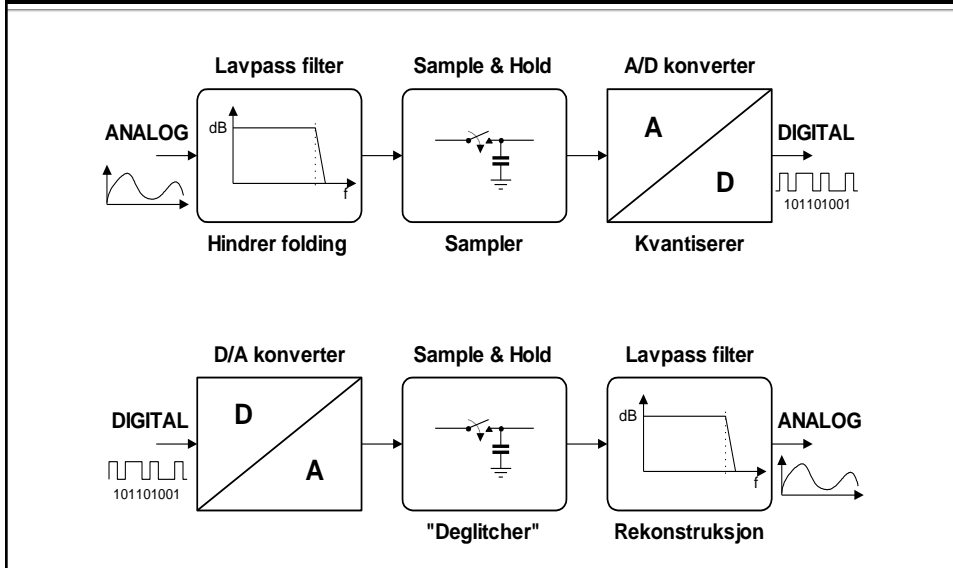


Copyright 2003 by Randy Glasbergen.
www.glasbergen.com



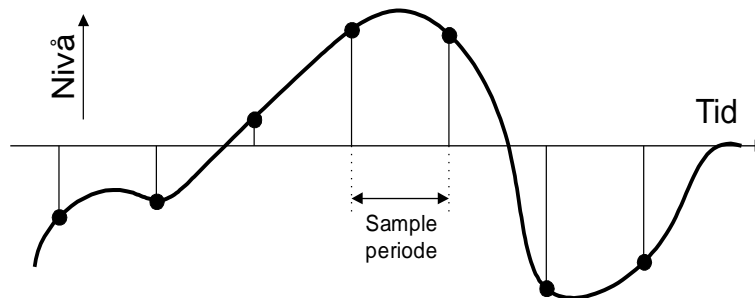
**"We've devised a new security encryption code.
Each digit is printed upside down."**

Analog - PCM digital - analog



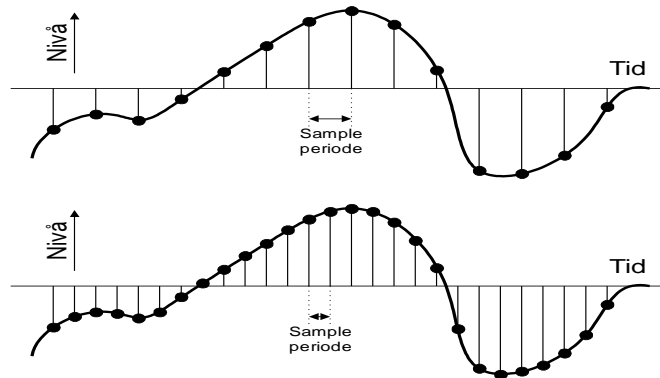
Sampling

- Sampling av et signal betyr å ta ut punktprøver med jevne mellomrom i tid.



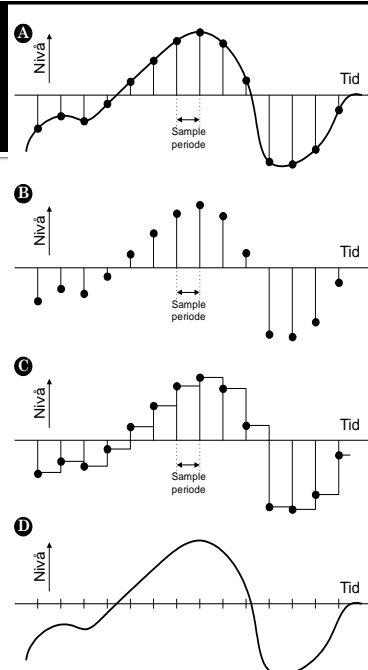
Sampling

- Jo høyere samplingsfrekvens, desto flere "detaljer" kommer med:

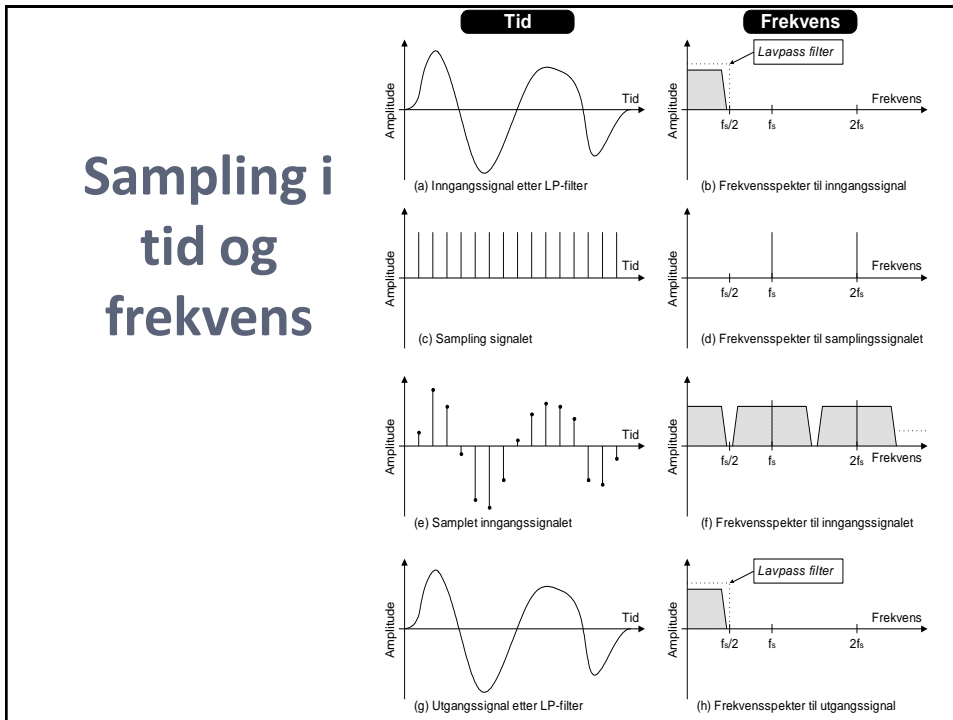


Sampling

- Signalet samples, holdes og til slutt lavpassfiltreres
- Nyquist: ingen informasjon med frekvens under $f_s/2$ har gått tapt
- Eventuell informasjon over $f_s/2$ gir frekvensspeiling



Sampling i tid og frekvens



Nyquist's teorem

“In order to be able to reconstruct a signal, the sampling frequency must be at least twice the frequency of the signal being sampled”

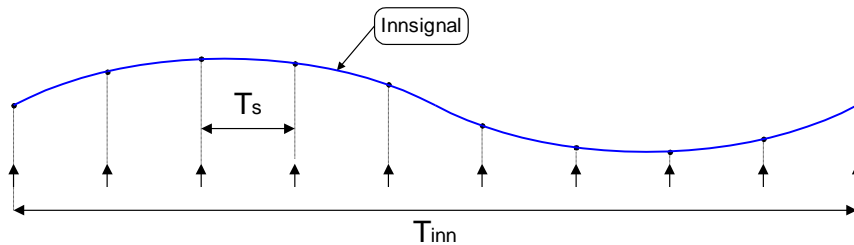
Harold Nyquist (1928)

- Denne minimum samplefrekvensen kalles ***Nyquist frekvensen***

Frekvensspeiling, folding

Nyquist teoremet sier at vi må ha en sample frekvens som er minst det dobbelte av øverste signalfrekvens.

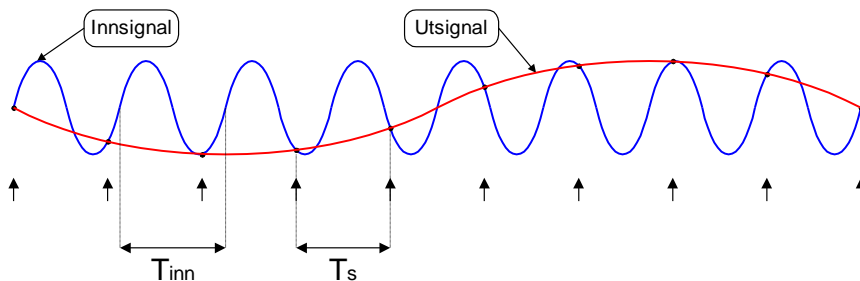
- Eksempel: f_s ca. 10X signalfrekvens:



Frekvensspeiling, folding.

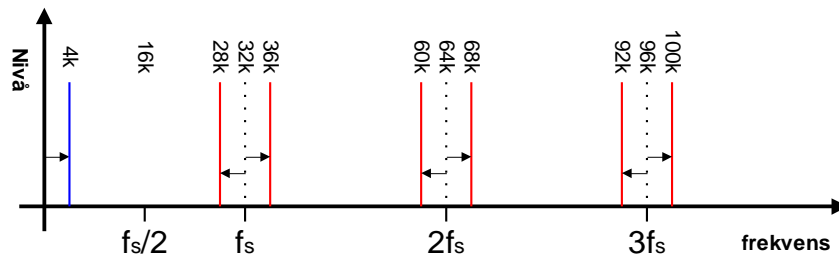
- Når frekvensen til signalet som samples er større enn halve samplingsfrekvensen oppstår frekvensspeiling.

■ Eksempel: signalfrekvens nær samplefrekvens:



Sampling i frekvensplanet

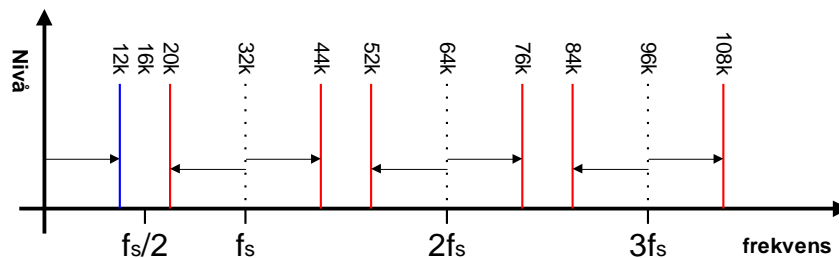
- 4 kHz tone samplet med 32 kHz
 - Pilene markerer "retningen" til signalet, økes innfrekvensen så flytter komponentene seg i pilenes retning.



Sampling i frekvensplanet

12 kHz tone samplet med 32 kHz

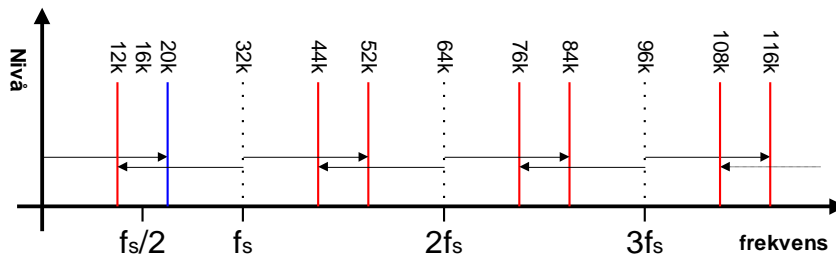
- Pilene markerer "retningen" til signalet, økes innfrekvensen så flytter komponentene seg i pilenes retning.



Sampling i frekvensplanet

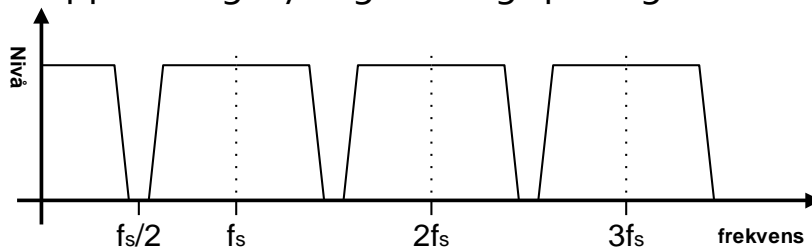
20 kHz tone samplet med 32 kHz

- Pilene markerer "retningen" til signalet, økes innfrekvensen så flytter komponentene seg i pilenes retning.



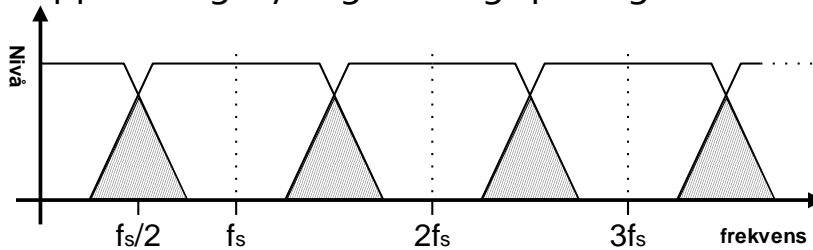
Sampling i frekvensplanet

- Lydsignal hvor frekvensinnholdet er begrenset til under $f_s/2$
- Det er ingen "overlapping" mellom det opprinnelige lydsignalet og speilingene



Sampling i frekvensplanet

- Lydsignal hvor frekvensinnholdet ikke er begrenset til under $f_s/2$
- Det er stor "overlapping" mellom det opprinnelige lydsignalet og speilingene



Valg av samplefrekvens

- Samplefrekvensen må være minst 2X ønsket øvre signalfrekvens
- Synkronisering mot "kanal"
 - Telenettet
 - Video
- Andre hensyn
 - Filtersteilhet
 - "Varispeed"
 - Bitrate (pris!)
 - Kompatibilitet.....

32 kHz = Gammel kringkasting

- Synkron med telenettet (4 x 8 kHz)
- Ønsket båndbredde er 15 kHz (FM)
- Bitrate ønskes så lav som mulig (pris)
- Eksempler
 - NICAM
 - NRK's gamle programlinjenett
 - "Long Play" modus på DAT
 - MP3 på lav bitrate

44,1 kHz = Konsument

- Ønsket båndbredde = 20 kHz
- Synkron med PAL:
 - 3 sampler pr. linje aktiv video
 - PAL: $50 \times 294 \times 3 = 44100$
 - Am. S/H: $60 \times 245 \times 3 = 44100$
- (44,056 kHz p.g.a NTSC)
 - NTSC reduserer delbildefrekvensen med en promille.....
- Eksempler: CD, DAT og MD avspilling

48 kHz = Pro Audio

- Ønsket båndbredde = 20 kHz
- Telesynkron (6 x 8 kHz)
- "Slake" filtre for høy lyd kvalitet
- Mulighet for **varispeed** +/- 12,5 % uten plagsom frekvensspeiling
- Eksempler:
 - Digital video med *embedded audio*
 - DAB, DVB, DVD, BluRay

Lineær kvantisering

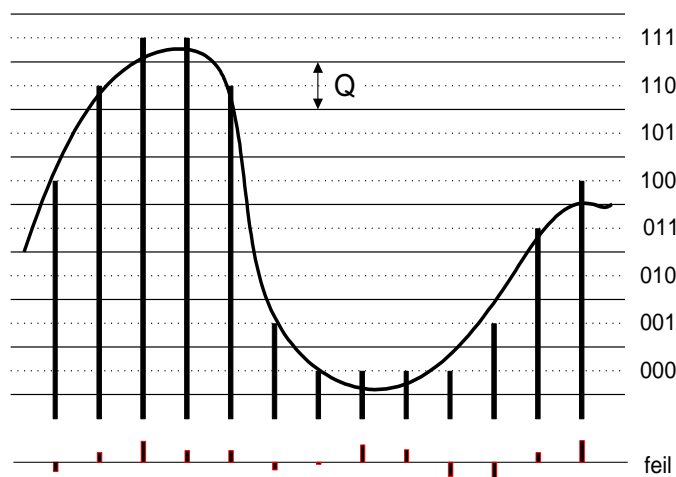
- Lineær kvantisering betyr at alle kvantiseringsintervaller er like store
- Hvert kvantiseringsnivå representeres av en unik digital kode
- Antallet bit bestemmer "oppløsningen"
- Med n bit deles det analoge dynamikk-området opp i 2^n kvantiseringsnivåer:
 - 8 bit gir 256 kvantiseringsnivåer
 - 16 bit gir 65356 kvantiseringsnivåer

Lineær kvantisering

- Q = størrelsen på et kvantiseringsintervall
- Jo flere bit, desto mindre blir Q
- Kvantiseringsfeilen er maks $\pm Q/2$, min 0
- Kvantiseringsfeilen for hvert enkelt sample blir tilfeldig dersom innfrekvens og samplefrekvens er ukorrelert (= alltid!)
- Tilfeldige signal oppfattes som støy. Feilen kalles derfor **kvantiseringsstøy**

Lineær kvantisering

- 3 bit kvantisering av et signal



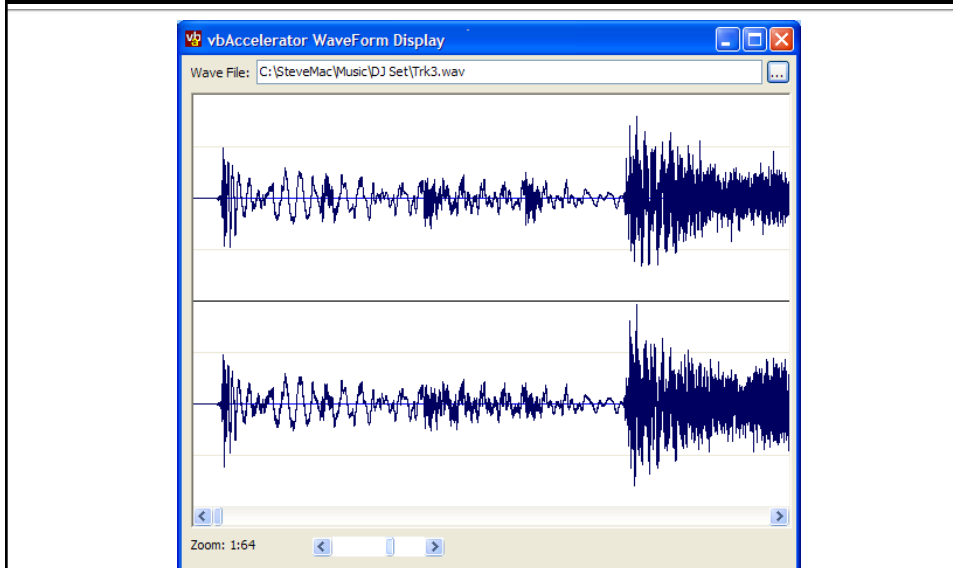
Lineær kvantisering

- Med 1 bit får vi kun to kvantiseringsnivåer.
- Øker vi til 2 bit får vi fire kvantiserings-nivåer, kvantiseringsfeilen er halvert
- Øker vi til 3 bit får vi åtte kvantiserings-nivåer, kvantiseringsfeilen halveres igjen
- Halvering av et nivå = reduksjon på 6 dB
- Altså: hvert bit gir ca. 6 dB dynamikk
- Eksakt: $SNR = 6,02 \times N + 1,78$ dB

Lineær kvantisering

- Uten dither: $SNR = 6,02 \times N + 1,78$ dB
 - 14 bit = 86 dB
 - 16 bit = 98 dB
 - 20 bit = 122 dB
 - 24 bit = 146 dB
- **Dither** reduserer disse tallene noe
- **Noise shaping** gir subjektivt bedre SNR

Hva er "WAV" ????



WAV er kort for "wave"

- Wave filformat = subset av Microsoft's **RIFF** spesifikasjon for lagring av multimedia filer
- **RIFF** = **R**esource **I**nterchange **F**ile **F**ormat
- En RIFF fil starter med en "header" etterfulgt av en rekke "data chunks"



Chunk??????

- A **chunk** is a fragment of information which is used in many multimedia formats, such as [PNG](http://en.wikipedia.org/wiki/Chunk_(information)), [MP3](#), [AVI](#) etc
- Chunk-baserte filformater har "åpen ende"
- Tre ulike chunk'er i WAV:
 - **RIFF** → start på filen
 - **fmt** → beskrivelse av innhold
 - **data** → lyddata

The Canonical WAVE file format

endian	File offset (bytes)	field name	Field Size (bytes)	
big	0	ChunkID	4	The "RIFF" chunk descriptor
little	4	ChunkSize	4	
big	8	Format	4	
big	12	Subchunk1ID	4	The "fmt" sub-chunk describes the format of the sound information in the data sub-chunk
little	16	Subchunk1 Size	4	
little	20	AudioFormat	2	
little	22	NumChannels	2	
little	24	SampleRate	4	
little	28	ByteRate	4	
little	32	BlockAlign	2	
little	34	BitsPerSample	2	
big	36	Subchunk2ID	4	The "data" sub-chunk Indicates the size of the sound information and contains the raw sound data
little	40	Subchunk2 Size	4	
little	44	data	Subchunk2Size	

Introduksjon til digital signalbehandling

Digital signalbehandling

- = DSP
- Nivået til et signal på digital form endres ved å multiplisere alle sampler med en konstant (forsterknings-) faktor.
- Miksing av signaler på digital form foregår ved addisjon.
- Filtrering og alle andre funksjoner gjøres ved å kombinere addisjon og multiplikasjon.

Binær addisjon

- Addisjon på desimal og binær form:

Desimal	Binær
$\begin{array}{r} 15 \\ + 11 \\ \hline = 26 \end{array}$	$\begin{array}{r} 1111 \\ + 1011 \\ \hline = 11010 \end{array}$

Binær addisjon

- Addisjon på desimal, binær og heksadesimal form:

Desimal	Binær	Hexadesimal
$\begin{array}{r} 236 \\ + 727 \\ \hline = 963 \end{array}$	$\begin{array}{r} 11101100 \\ + 1011010111 \\ \hline = 1111000011 \end{array}$	$\begin{array}{r} EC \\ + 2D7 \\ \hline = 3C3 \end{array}$

Binær multiplikasjon

- To tall multipliseres ved "venstre skift" og gjentatt addisjon. Små tall:

Desimal	Binær
12 x 13 =	1 1 0 0 x 1 1 0 1 =
3 6	1 1 0 0
+ 1 2	0 0 0 0
<u> </u>	1 1 0 0
= 1 5 6	+ 1 1 0 0
	<u> </u>
	= 1 0 0 1 1 0 0

Binær multiplikasjon

- To tall multipliseres ved "venstre skift" og gjentatt addisjon. Større tall:

Desimal	Binær
157 x 222 =	1 0 0 1 1 1 0 1 x 1 1 0 1 1 1 1 0 =
3 1 4	0 0 0 0 0 0 0 0
3 1 4	1 0 0 1 1 1 0 1
+ 3 1 4	1 0 0 1 1 1 0 1
<u> </u>	1 0 0 1 1 1 0 1
= 3 4 8 5 4	1 0 0 1 1 1 0 1
	0 0 0 0 0 0 0 0
	1 0 0 1 1 1 0 1
	+ 1 0 0 1 1 1 0 1
	<u> </u>
	= 1 0 0 0 1 0 0 0 0 0 1 0 0 1 1 0

Digital signalbehandling

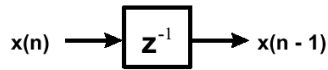
- Addisjon og multiplikasjon av digitale signal gir flere bit ut enn hva vi putter inn.
- Adderes to 16 bits ord fås ett 17 bits ord.
- Adderes mange 16 bits ord blir svaret større.
 - Addisjon av 40 kanaler á 16 bit gir 22 bit til svar
- Multipliseres et n bits ord med en m bits koeffisient fås ett m+n ord.
 - Et 16 bits ord som "fades" med 12 bit forsterkningsfaktor: $16+12=28$ bit til svar

Digital signalbehandling

- Addisjon og multiplikasjon av digitale signal gir flere bit ut enn hva vi putter inn.
- Hvis svaret "trunkeres" vil vi få akkumulert avrundingsfeil.
- Ved "trunkering" forsvinner også dither som er lagt til signalet for å gjøre kvantiseringsprosessen lineær.
- DSP for audio benytter som regel 32 bit ordlengde for å unngå "trunkering".

Common DSP Building Block Operations

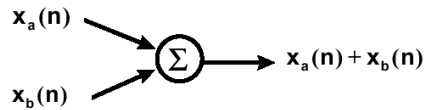
Signal Delay



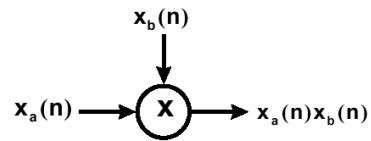
Multiplication a Signal By A Constant



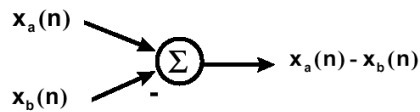
Addition of Two Signals



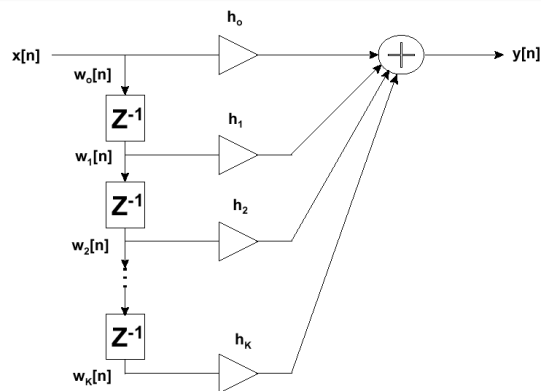
Multiplication of Two Signals



Subtraction of 2 Signals

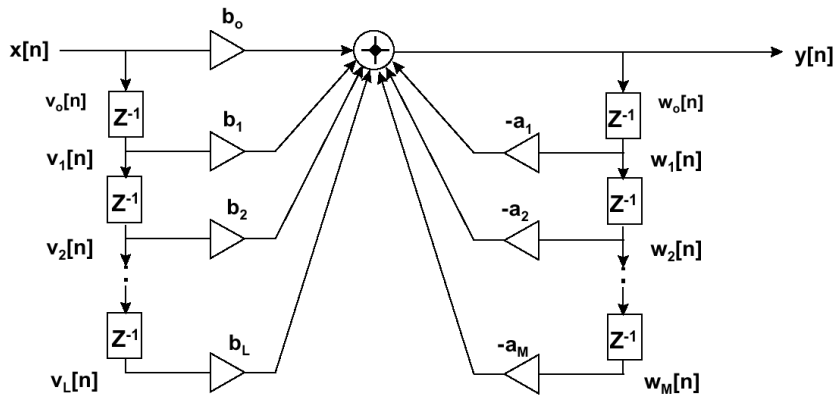


FIR filter



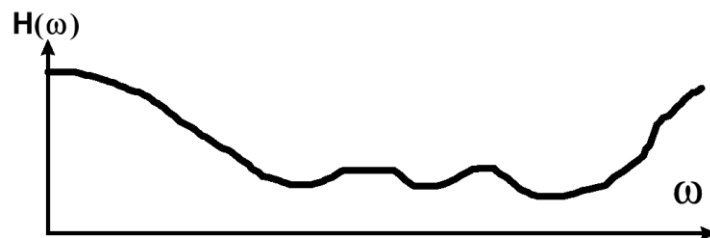
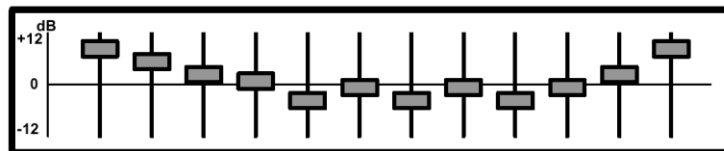
$$y[n] = \sum_{k=0}^{\infty} h[k]x[n-k]$$

IIR filter

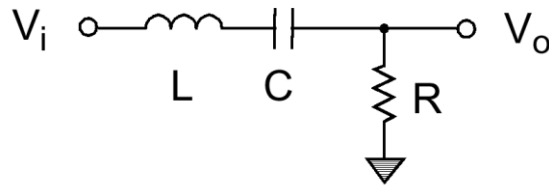


$$y[n] = \sum_{i=1}^M (-a_i y[n-i]) + \sum_{j=0}^L (b_j x[n-j])$$

12 bånd analog grafisk EQ

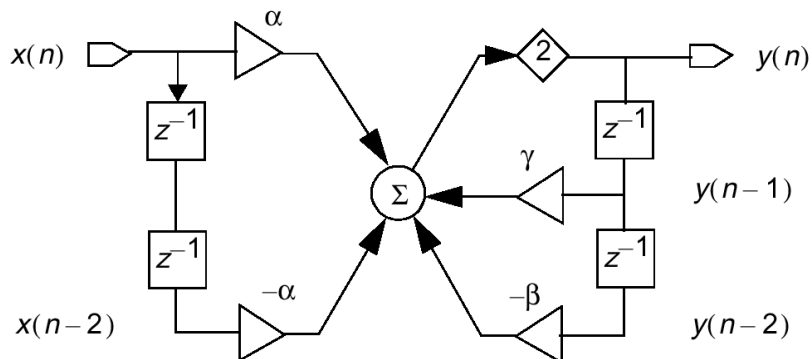


RLC båndpassfilter

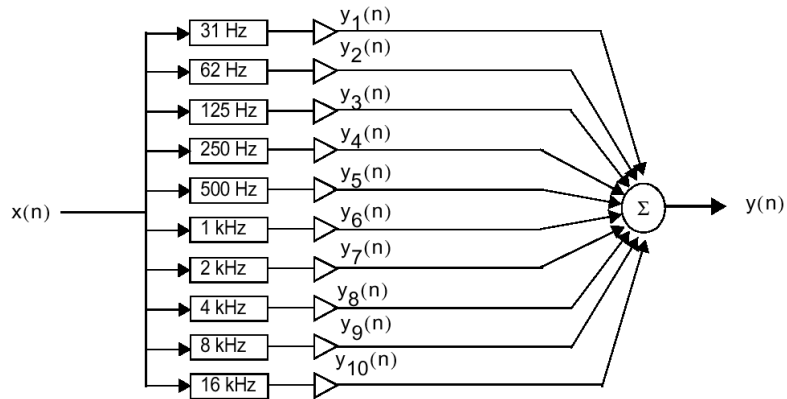


$$\frac{V_o}{V_i} = \frac{R}{R + j(2\pi f)L + \frac{1}{j(2\pi f)C}}$$

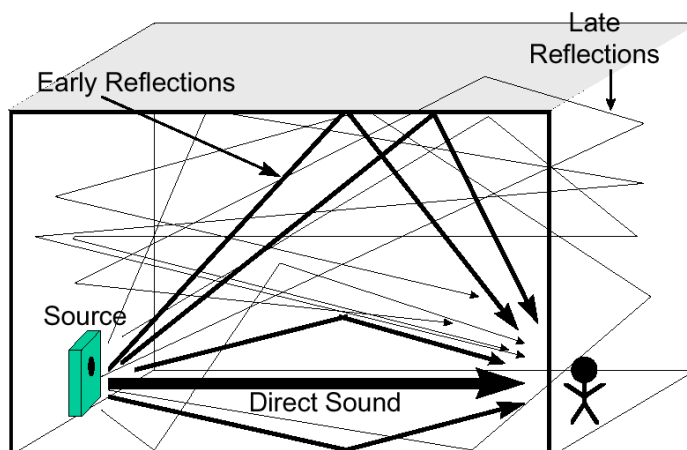
IIR båndpassfilter



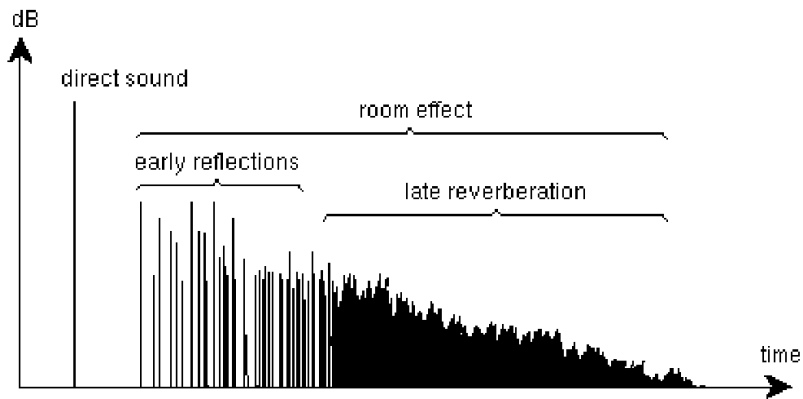
10-bands IIR equalizer



Reverberation of Large Acoustic Spaces



Naturlig etterklang i et rom



James A. Moorer's Digital Reverberation Structure

