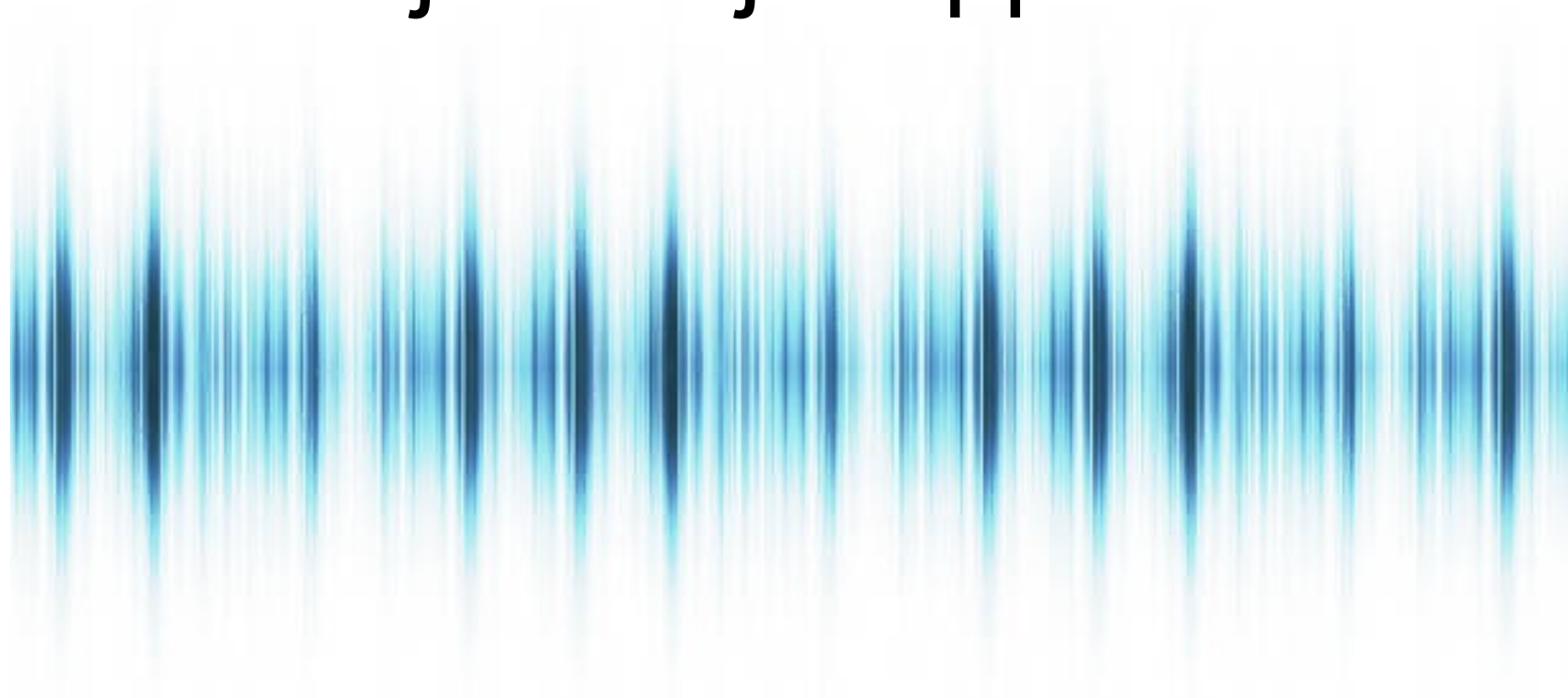


Hljóð - hljóðupptaka

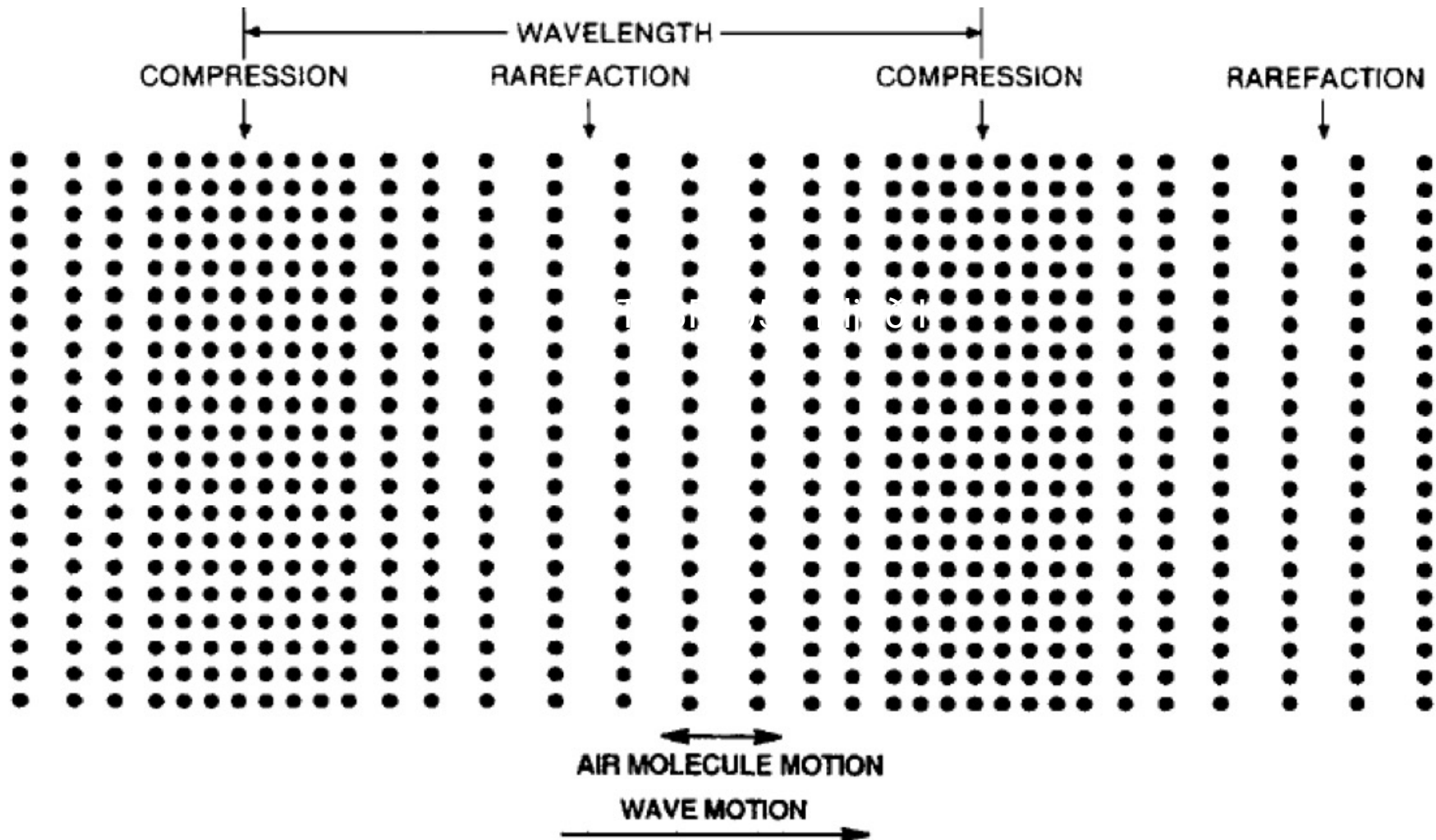


# Hvað er hljóð?

Ekki sama og rafsegulbylgjur! 😊

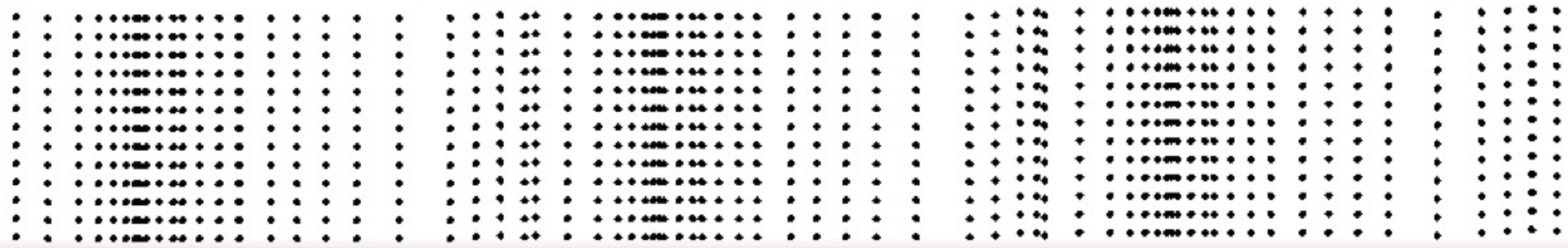
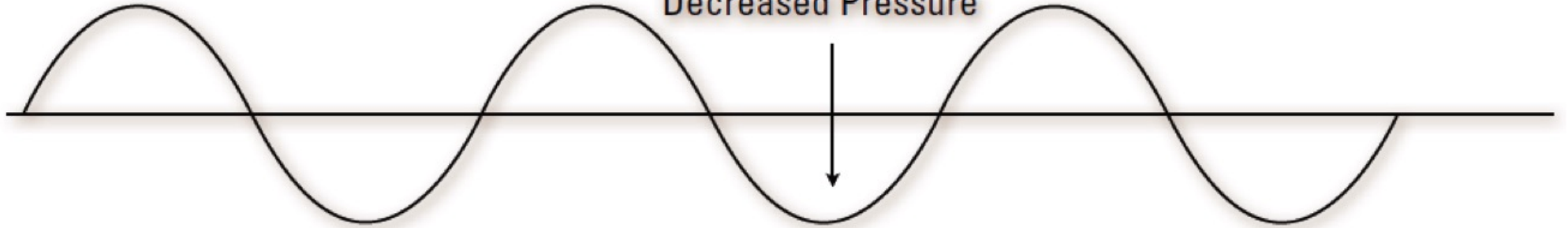
- Við búum í neðsta lagi lofthjúpsins
- Þar er þéttleiki (density) loftsins mikill, loftþrýstingur
- Hljóð myndast við hraðar breytingar á þéttleika loftsins
- Hægar breytingar => Loftþrýstingur, veðurfræði.
- Hraðar breytingar = > Hljóðbylgjur, hljóðhimnur titra.
- Hljóð eða hljóðbylgjur eru þrýstingsbylgjur sem berast í gegnum efni
- Hljóð verður til þegar einhver hlutur kemur af stað titringi í sameindum í loftinu
- (Stök sameind færir fram og til baka en margar saman mynda bylgju)

# Hljóðbylgja titringur sameinda



Increased Pressure

Decreased Pressure



Displacement of Air Molecules



Direction of Propagation



# Hvað er hljóð?

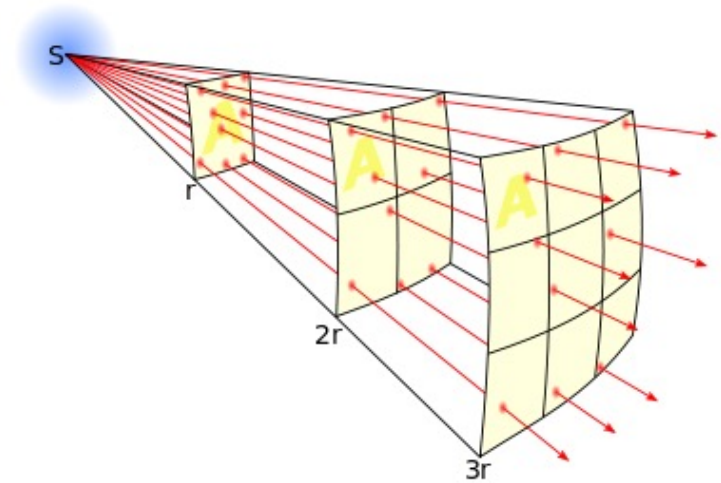
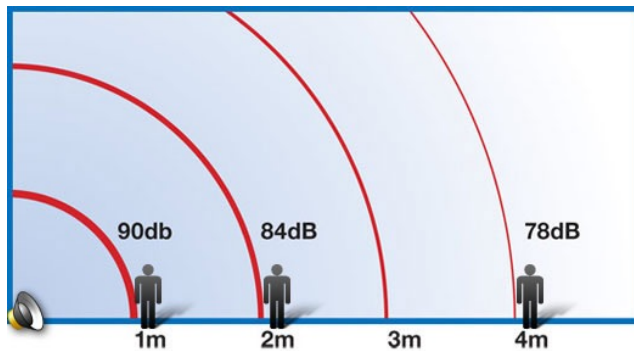
Hljóðhraði er háður efninu sem hljóðið berst um

Hljóðhraði lofts við 21° C er 344 m/s (1238 km/klst)

Hljóðbylgjan dreyfist víðar eftir því sem hún berst frá upphafs þúnkti

= > dofnar því lengra sem hún fer.

**“Inverse square law”** þrýstingur lækkar um 6dB við hverja tvöföldun frá upphafspunkti



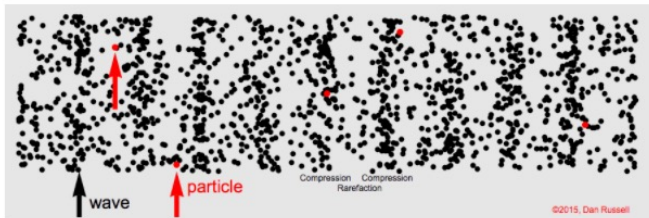
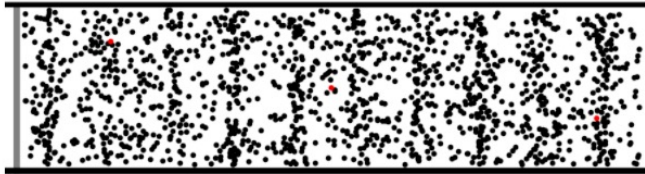
$$\text{Intensity} \propto \frac{1}{\text{distance}^2}$$

$$\frac{\text{Intensity}_1}{\text{Intensity}_2} = \frac{\text{distance}_2^2}{\text{distance}_1^2}$$

# Hvað er hljóð?

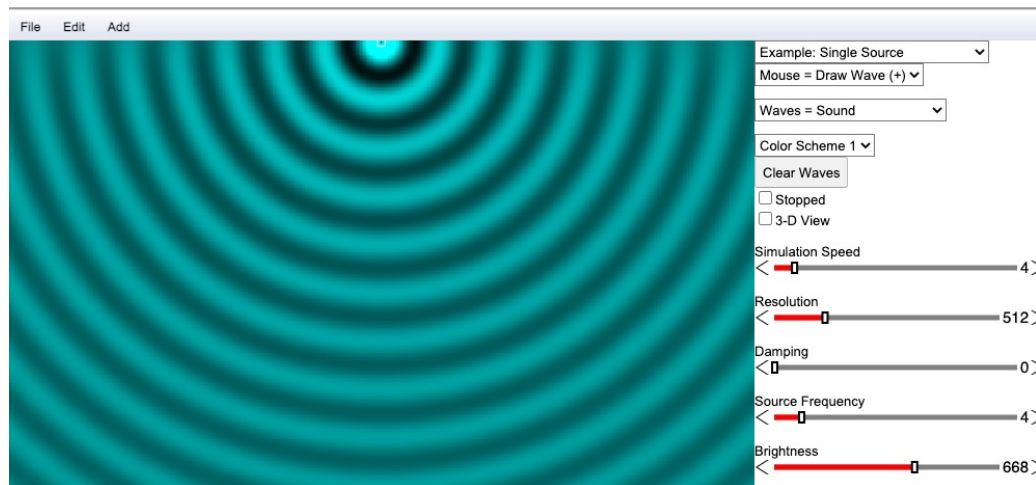
## Longitudinal and Transverse Wave Motion

<http://www.acs.psu.edu/drussell/Demos/waves/wavemotion.html>



## Paul Falstad, applet o.fl

<http://www.falstad.com/ripple/>



# Hljóðbylgjur

- (Wave)Cycle – (Bylgju)Hringur
  - Einn umgangur frá hvíld í plús, hvíld, mínus og aftur hvíld
- Period – Tímabil (t)
  - Tíminn í sec. sem það tekur fyrir hljóðbylgju að fara einn hring.
- Frequency – Tíðni (f)
  - Hversu marga hringi hljóðbylgjan fer á sec.
- **Amplitude – Styrkur (dB)**
  - Hversu mikið færast sameindir frá hvíld
- Wavelength – Bylgjulengd ( $\lambda$ )
  - Hversu langt fer hljóðbylgjan í einum bylgjuhring
- Velocity – Hraði (v)
  - 344m á sec.

**Hljóðstyrkur (Amplitude)** er mælanlegur eiginleiki sem segir hve mikið sameindir færast frá hvíld.

Mældur í dB (Decibel: <http://en.wikipedia.org/wiki/Decibel>)

- dB er logaritmískur skali
- **dB SPL – Sound Pressure Level** - Hljóð í andrúmsloftinu.

0 dB SPL = mörkin sem við heyrum eitthvað

Notumst við **plús tölur** til að tala um hljóðstyrk

*Innandyra á samkomustöðum og öðrum stöðum þar sem fólk dvelur í skamman tíma skal jafngildishljóðstig ekki fara yfir 95 dB, háværa hljóðbil ekki fara yfir 110 dB og hæsti hljóðtoppur ekki yfir 130 dB.*

- **dB FS – Full Scale** - Hljóð í tölvunum okkar.

0 dB FS = mesti hljóðstyrkur sem tölvun getur sýnt

Notumst við **mínus tölur** til að tala um hljóðstyrk



# Hljóðhraði

Medium	ft/sec	m/sec	20° C
Air	1130	344	
Water	4920	1500	
Concrete	10,200	3100	
Wood	12,500	3800	
Glass	18,500	5500	
Steel	19,000	5800	

# Hljóðbylgjur

$$\lambda = v/f \quad t = 1/f$$

Frequency (f)	Wavelength ( $\lambda$ )	Period (t)	Speed (v)
250 Hz	137,6 cm	4 ms	344 m/s
500 Hz	68,8 cm	2 ms	344 m/s
1000 Hz	34,4 cm	1 ms	344 m/s

# Purfum að spá í hraða hljóðs í upptökum

## Direct vs upptaka með hljóðnenum

DI upptaka er aðeins á undan hljóðnemaupptöku.

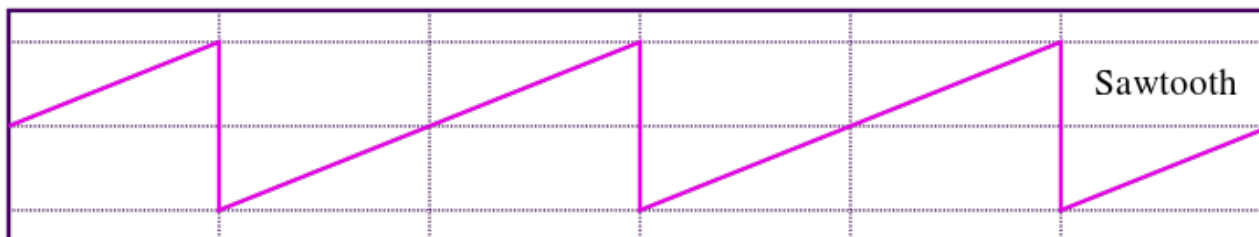
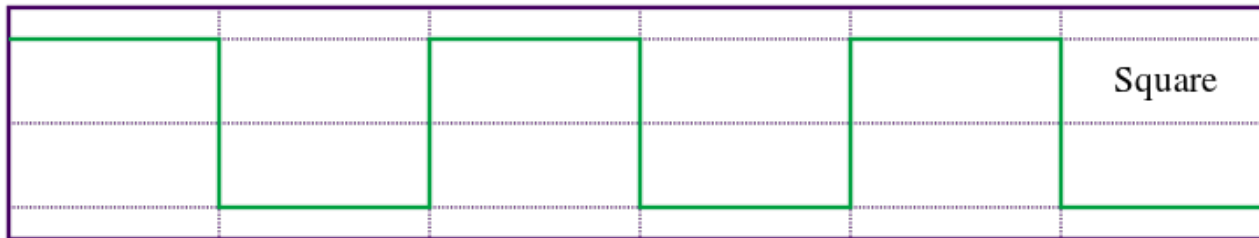
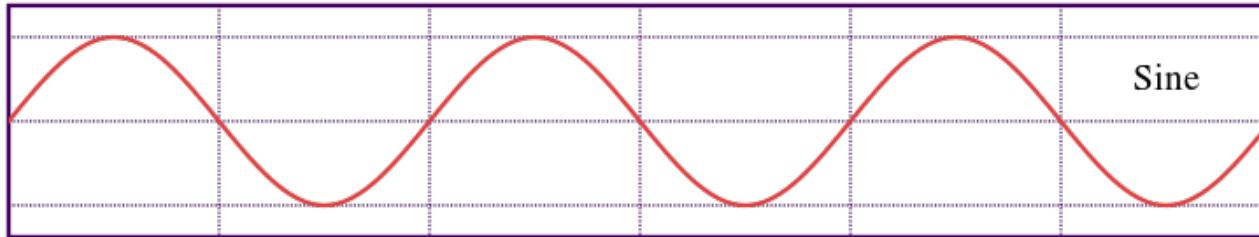
## Upptaka með mörgum hljóðnenum

Mörg merki á mismunandi tímum geta truflað hvert annað. Fasavandamál líkleg

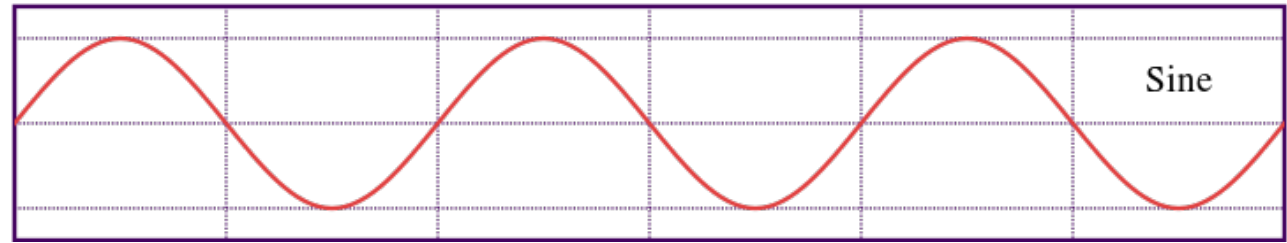
Nálægt: hærra og með minna endurkasti **Upptaka nálægt uppruna vs upptaka lengra frá**

Lengra frá: lægra og með endurkasti frá umhverfinu.

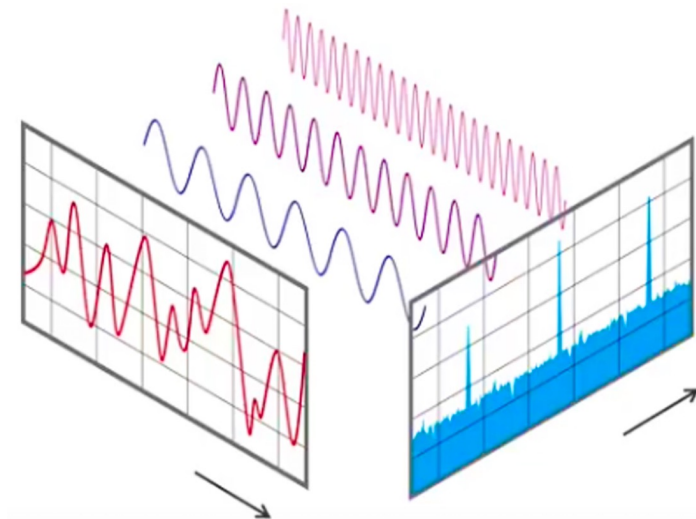
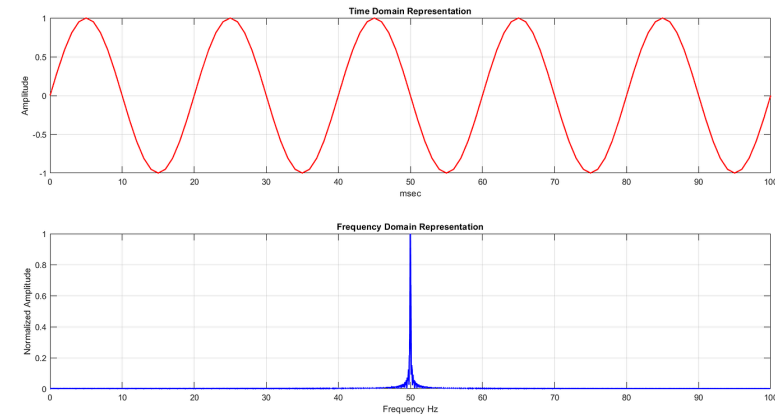
# Waveform grunntegundir af hljóðbylgjum



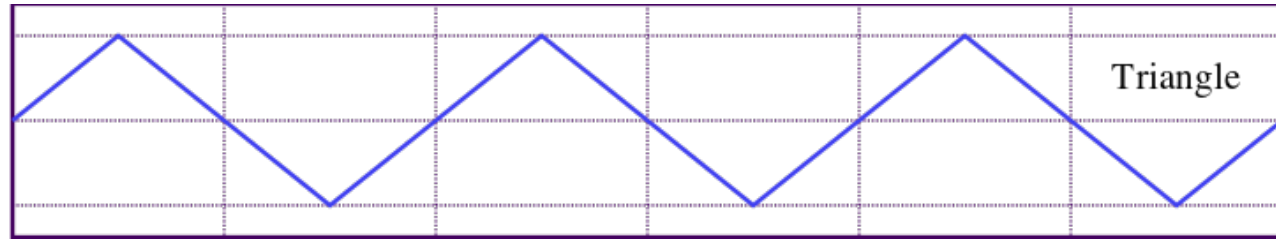
# Sínus



- Hreinn tónn
- Öll orkan er á einni tíðni
- Engir yfirtónar
- Grunn-eining í uppbyggingu hljóðs.
- Allt hljóð er hægt að brjóta niður í mismunandi sínusbylgjur

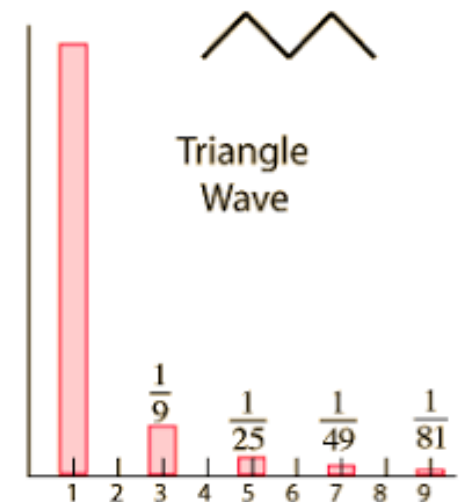


# Þríhyrningsbylja

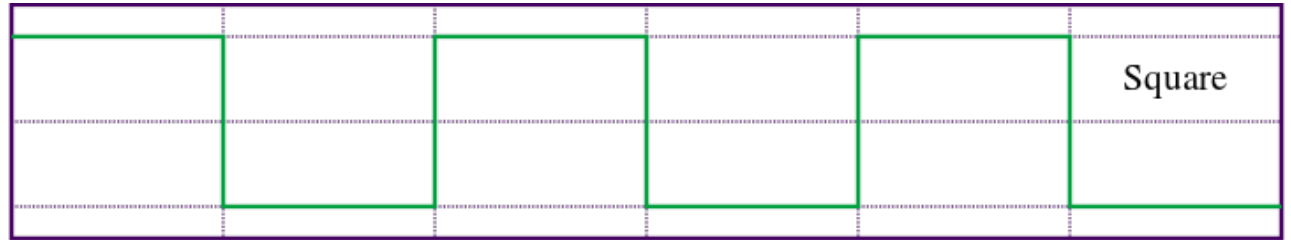


- Oddatölu yfirtónar sem fara minnkandi eftir fjarlægð frá grunntóni
- 3x er sterkur, 5x veikari, o.s.frv.
- (Yfirkeyrt) Segulband hefur gjarnan svona yfirtóna
- (3x grunntónn er áttund plús fimmund)

Styrkur yfirtóna minnkar hratt

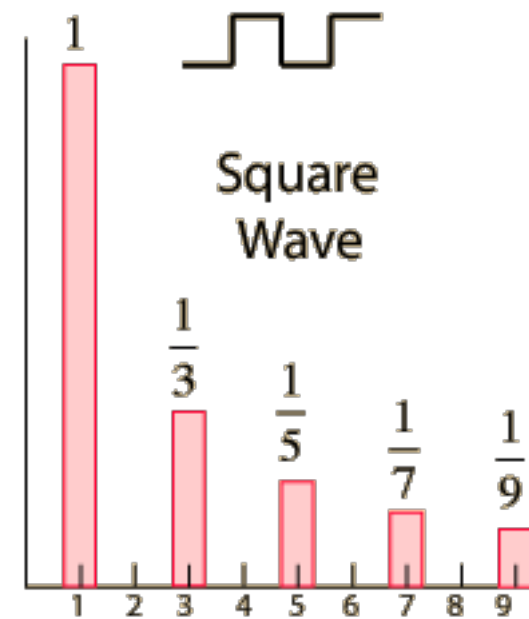


# Kassabylgja

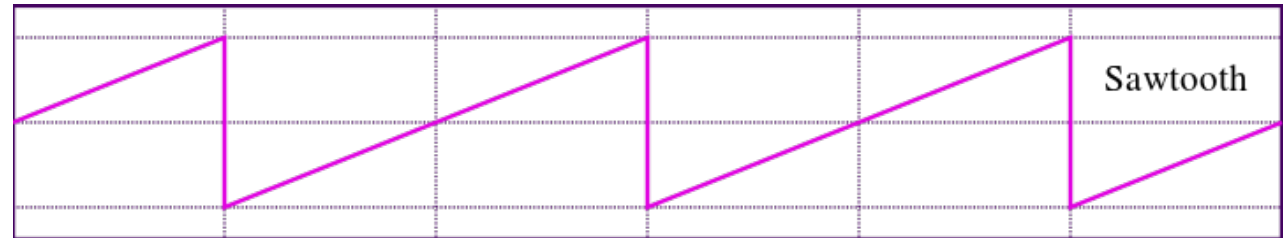


- Óendalega mikið af oddatölu yfirtónum
- Oddatölu yfirtónar, eru yfirtónar sem eru
- $3x$ ,  $5x$ ,  $7x$  o.s.frv. miðað við grunntóninn

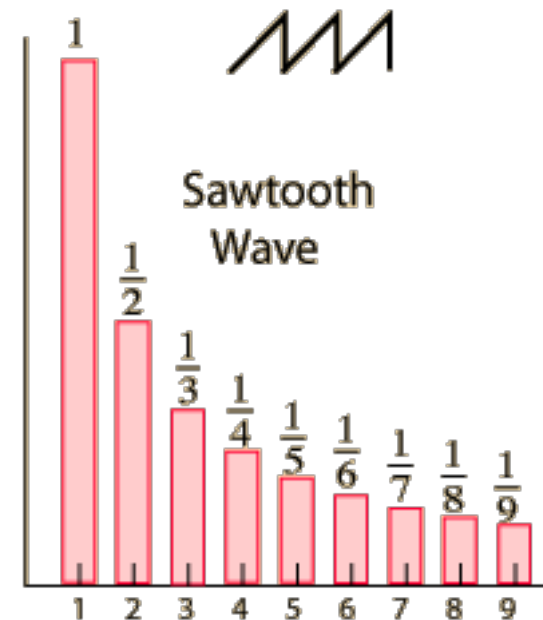
Styrkur yfirtóna minnkar hægar



# Sagtönn



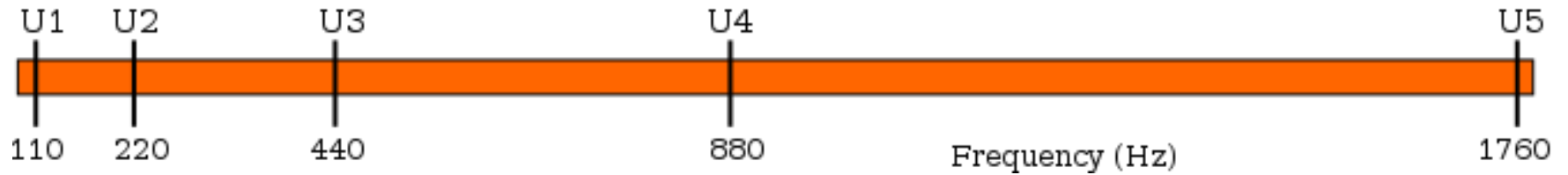
- ✓ Bæði oddatölu og jafntölu yfirtónar
- ✓ Jafntölu yfirtónar eru  $2x$ ,  $4x$ ,  $6x$ ... grunntónn
- ✓ Lampamagnarar hafa jafntölu yfirtóna,
- ✓ aðallega  $2x$  ( $2^{\text{nd}}$  order)



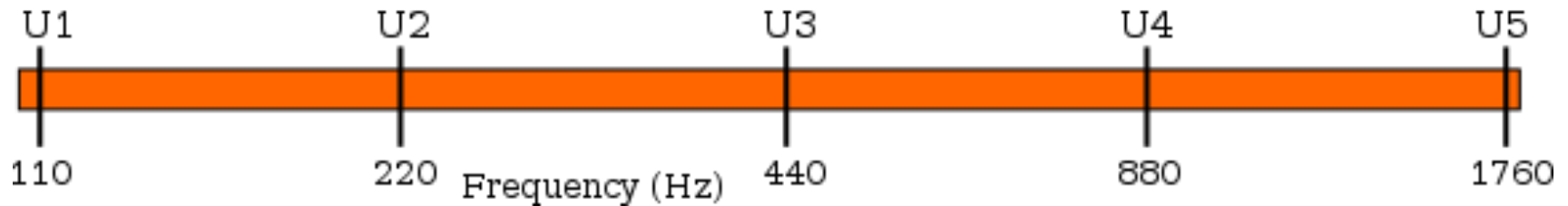


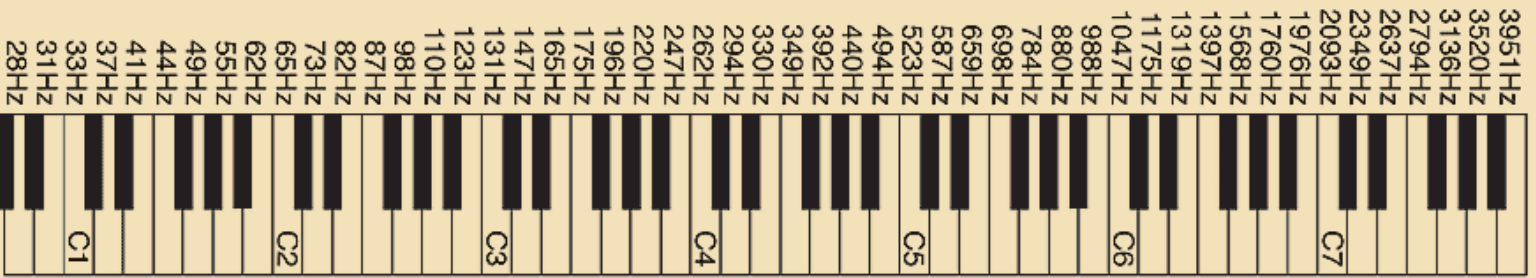
# Eðlisfræðin vs Tónlistinn

## Eðlisfræðin

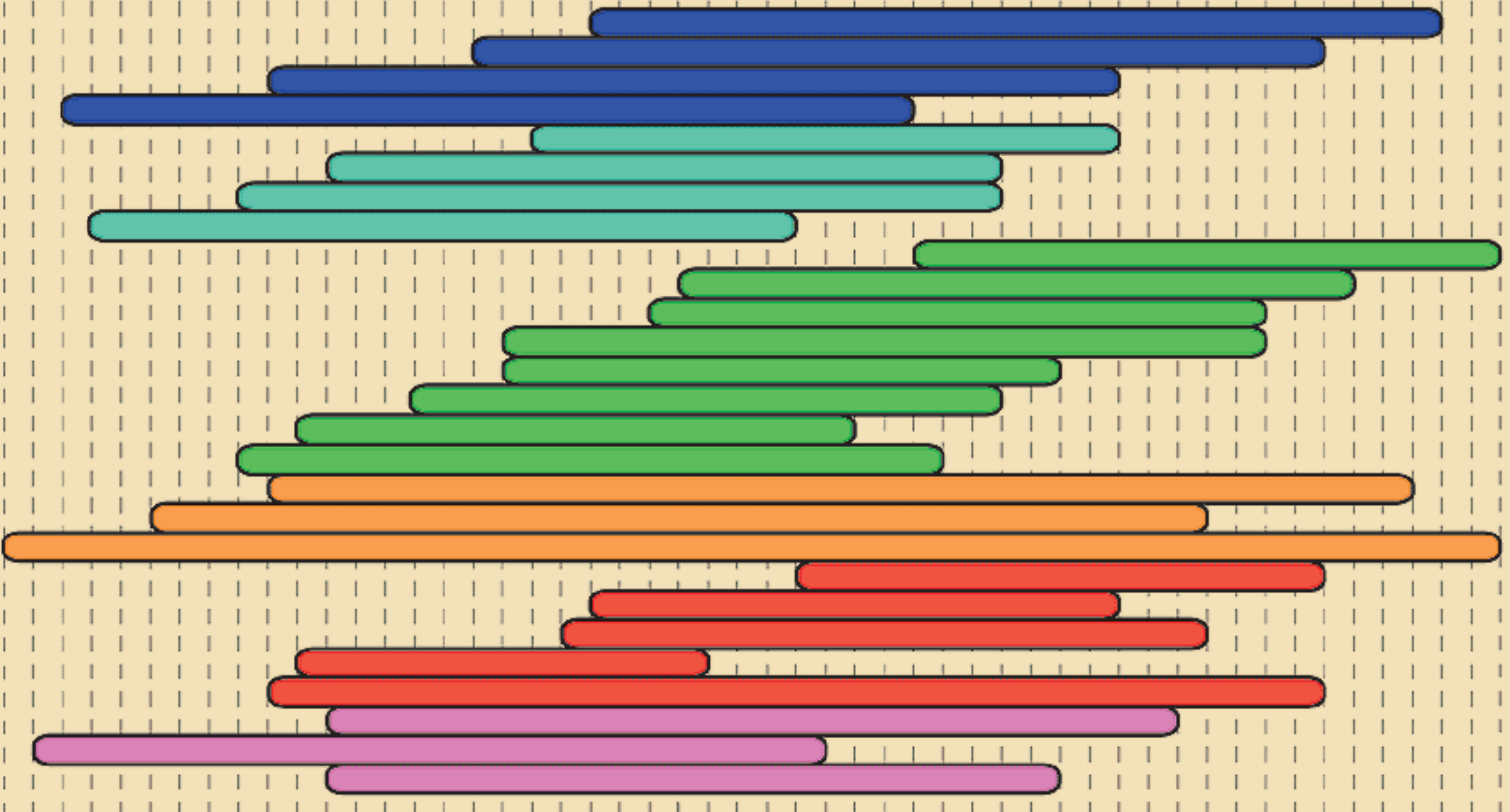


## Tónlistin





- Violin
- Viola
- Cello
- Bass
- Trumpet
- Trombone
- French Horn
- Tuba
- Piccolo
- Flute
- Oboe
- Clarinet
- Alto Sax
- Tenor Sax
- Baritone Sax
- Bassoon
- Harp
- Harpsichord
- Piano
- Xylophone
- Glockenspiel
- Vibraphone
- Timpani
- Marimba
- Guitar
- Bass Guitar
- Voice



# Tíðnisvið

- Bassi: 20Hz – 200Hz
- Lág miðja: 200Hz – 800Hz
- Há miðja: 800Hz – 5kHz
- Toppur: 5kHz – 20kHz

# SUÐ

## **Hvít suð (White Noise)**

- Jafn mikill styrkur á öllum tíðnum

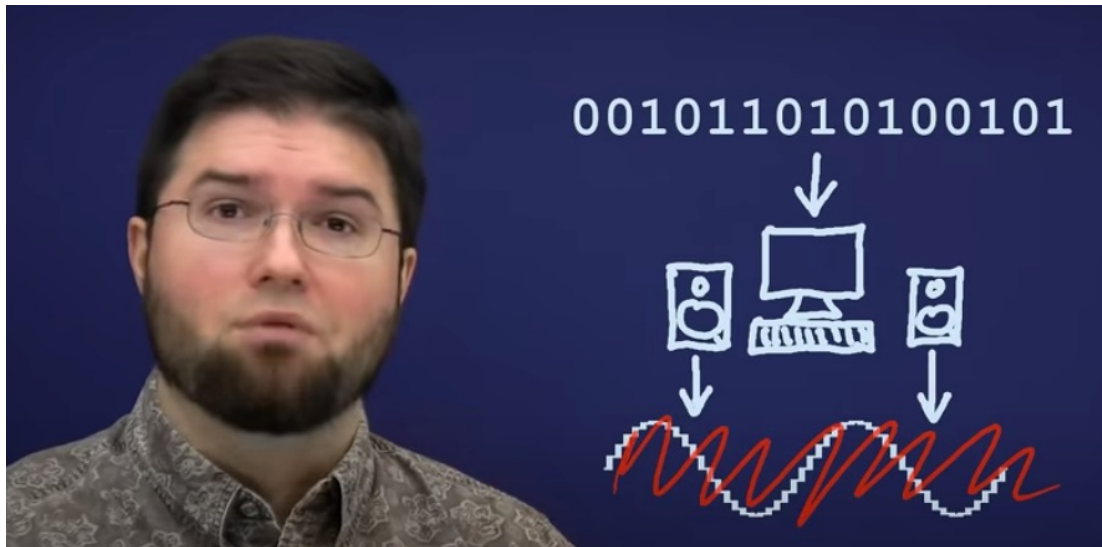
## **Bleikt suð (Pink Noise)**

- Jafn mikill styrkur í öllum áttundum
- Meira eins og eyrað skynjar hljóð

# AD – DA

## Stairsteps – Bits – Dither (24 mín)

<http://www.youtube.com/watch?v=clQ9IXSUzuM>



Monty Montgomery

# Stafræn upptaka

- Hljóð verður til við breytingar á þjöppun sameinda í andrúmsloftinu
- Þrýstingi á hverju augnabliki er gefinn tölulegt gildi.
- Mörgum svona tölum raðað saman og hljóðbylgjan búin til á ný út frá því, nokkurskonar súlurit

# Bitar og Byte

- 1 Byte = 8 Bitar
- 1 KB (KiloByte) = 1.024 Bytes
- 1 MB (MegaByte) = 1.024 KiloBytes
- 1 GB (GigaByte) = 1.024 MegaBytes
- 1 TB (TeraByte) = 1.024 GigaBytes

Af hverju 1.024 en ekki 1.000?  $1024 = 2^{10}$

# Pláss á diskum

Þurfum við að vita c.a. hvað mikið pláss ákveðnar upptökur taka.

Hvað taka 5 mínútur af 16 bita stereo upptöku í 44.1kHz mikið pláss

## Fyrsta skref

Bitafjöldi sinnum "sample rate"

16 bitar á hvert sampl x 44.100 sömpl á sec.  
= 705.600 bitar á sec.



# Pláss á diskum

## Annað skref

Færa yfir í þægilegri tölur

$$705.600 \times 60 = 42.336.000 \text{ bitar á min.}$$

$$42.336.000 / 8 = 5.292.000 \text{ byte á min.}$$

$$5.292.000 / 1.024 = 5.167,97 \text{ KB á min}$$

$$5.167,97 / 1.024 = \underline{\mathbf{5,05 MB á min}}$$

## Þriðja skref

Margfalda með fjölda rása

Stereo er tvær rásir

$$5,05\text{MB} \times 2 = \underline{\mathbf{10,10\text{MB á min}}}$$

## Fjórða skref

Margfalda með mínútufjölda

10,10MB á mín x 5 = **50,50 MB fyrir 5 mín**

Hvað margar mín af 16 bita stereo upptökum komast fyrir á 650MB disk?

**650/10,1 = rúmlega 64 mínútur**

# Hvað taka 60 min af 16 rása 24bita upptökum í 96kHz mikið pláss

- $24 \times 96.000 = 2.304.000$  bitar á sec
- $2.304.000 / 8 = 288.000$  bytes á sec
- $288.000 \times 60 = 172.800.000$  bytes á min
- $172.800.000 / 1.024 = 16.875$  KB á min
- $16.875 / 1.024 = 16.48$  MB á min (1 rás)
- $16.48 \times 16 = 263.68$  MB á min (16 rásir)
- $263.68 \times 60 = 15.820,8$  MB á 60 min (16 rásir)
- $15.820 / 1.024 = \underline{\mathbf{15,45\ GB\ á\ 60\ min\ (16\ rásir)}}$

## Sömplun hljóðs

- Styrkur hljóðs er mældur oft á sec.
- Algengar tíðnir (Hz) eru 44.100, 48.000, 88.200, 96.000, 176.200 og 192.000
- Hver mæling er geymd sem tvíundartala
- Algengur bitafjöldi í tvíundartölunni er 16 og 24 bitar
- (ath, þetta er "fixed point", ekki "floating point" nánar um það síðar)
- Bæði tíðnin og bitafjöldi hafa áhrif á hljómgæði.

# Hverju er hægt að redda í upptökum:

[http://www.youtube.com/watch?v=G2Rhh\\_4GZmU](http://www.youtube.com/watch?v=G2Rhh_4GZmU)



# DSP, Digital Signal Processing

*Hér er hægt að dvelja í nokkur ár, en við sleppum öllum stærðfræði flækjum*

Input  
filter

Sample  
and hold

ADC

DSP

DAC

Output  
filter

Losum okkur við óæskileg merki t.d. eins og of hár tíðnir með LPF	Tekur gildi á skilgreindum tíma og heldur því gildi þar til tekið er aftur gildi.	Analog to digital ADC er í raun fyrsti hlutin af DSP og þar er quantize. Quantize => skráum binary tölur fyrir analog sýni. t.d. 16 bit quantize	Mjög öflug reiknieining Sendir, undirbýr geymslu, nær í . Vinnum með signlaðið á stafrænu formi Coding/Decode Dither er á milli DSP og DAC?	Digital to analog Komum merkinu aftur á analog form.	Lögum merkið til oft með LPF
---	---	--	---	--	------------------------------

# Sampling Theorim

Það er hægt að geyma og endurskapa hljóðmerki með röð af upptökum eða sömplum, sem tekin eru með jöfnu millibili, svo framarlega sem engin tíðni í hljóðinu er hærri en hálf söfnunar tíðnin (e. sample frequency/sampling rate).

Dæmi:

44.100 Hz söfnunar tíðni getur geymt og endurskapað hljóð allt að 22.050 Hz.

# Nyquist tíðnin

Harry Nyquist er annar af höfundum "Sampling Theorim"

Nyquist tíðnin heitir í höfuðið á honum

Nyquist tíðnin er hálf safntíðnin (sample frequency)

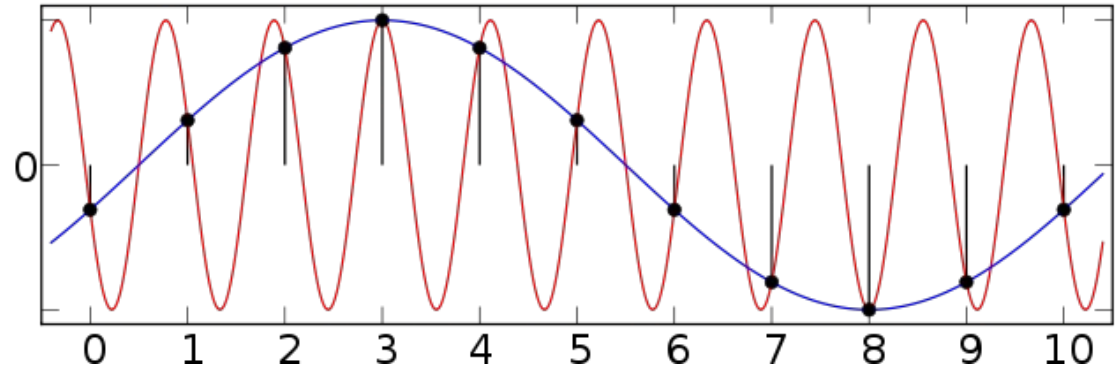
Nyquist tíðni geisladiska er 22.050, ætli það sé tilviljun?

Hæsta tíðni sem við getum heyrt er 20.000



# Aliasing

Við ætlum að **sampla rauða signalið** en þar sem við erum að sampla á of lágrí tíðni verður til Aliasing (**bláa**)

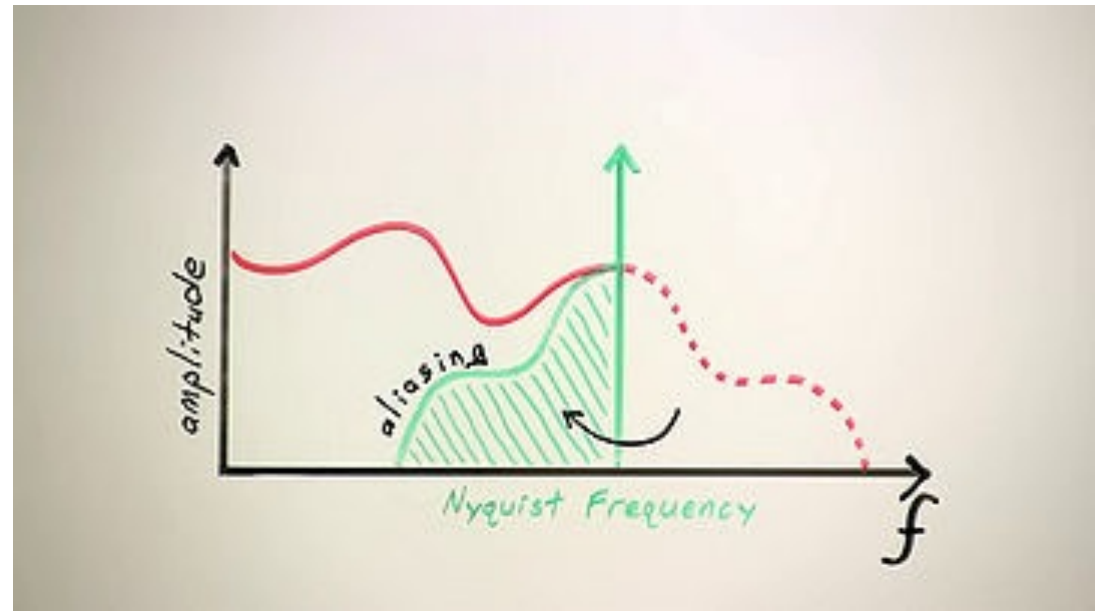
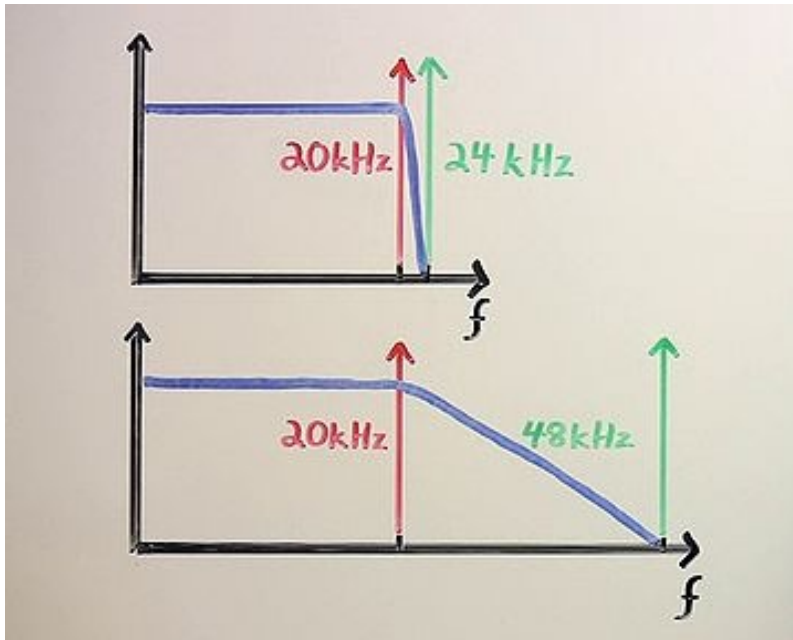


Aliasing er önnur tíðni sem verður til og gefur enga mynd að því merki (signal) sem við viljum taka upp og getur auk þess haft áhrif á önnur merki sem við námum á réttir tíðni.

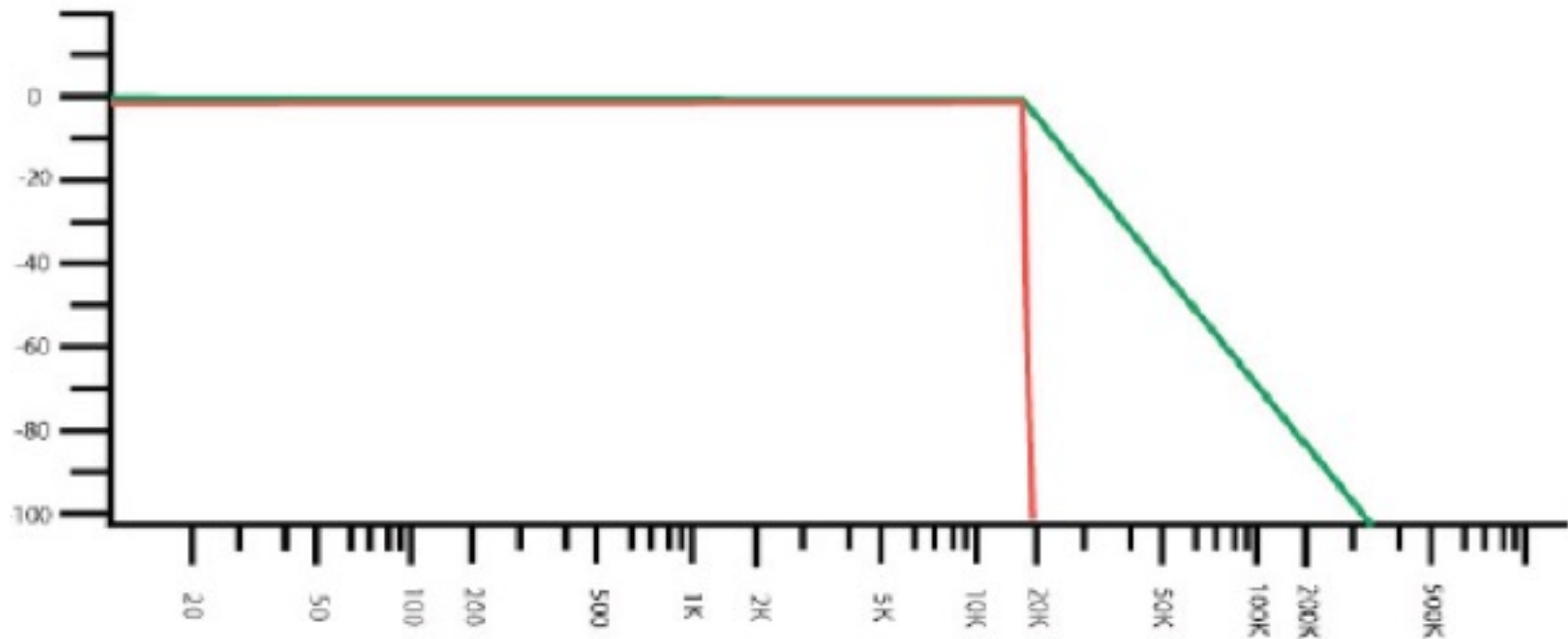
# PCM

## •Monty vinur okkar

<http://xiph.org/video/vid1.shtml> frá 9. mín



# Anti Aliasing Filters



**Figure 2.4**  
A 24 dB/octave filter (green line) and a theoretical brick-wall filter (red line)

# Anti Aliasing Filters

Til að koma í veg fyrir „Aliasing“ verður að takmarka þær tíðnir sem fara í digital formið

Tíðnir fyrir ofan 20kHz geta búið til alias tíðnir fyrir neðan 20kHz

Filter fjarlægir tíðnir fyrir ofan Nyquist tíðnina áður en AD breyting á sér stað Low Pass Filter

AD => analog to digital => breytum hliðrænu merki í stafrænt merki

## Anti Aliasing Filters

***Ekki eins einfalt og það lítur út fyrir að vera***

Cutoff band – 3dB minnkun á styrk

dB/Octave – mæling á því hve mikið styrkur minnkar per áttund þar til enginn styrkur er

6dB = 1<sup>st</sup> order, 12dB = 2<sup>nd</sup> order etc.

Eftir því sem filterinn er brattari, því meiri líkur á fasatruflunum, og tíðnitruflunum

# Anti Aliasing Filters

## Fullkominn Anti-Aliasing Filter

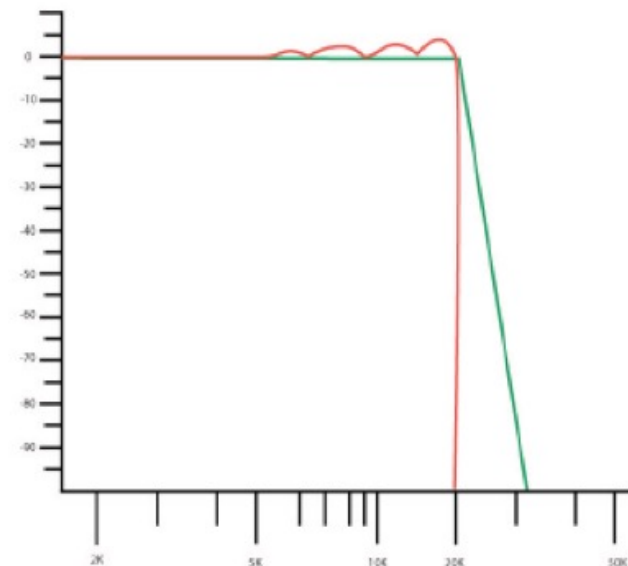
Hleypur öllum æskilegum tíðnum í gegn

Hleypur engum óæskilegum tíðnum í gegn

Hefur enga fasa- eða tíðnitruflun

Svona filter, analog, er ekki hægt að gera fyrir 44.1kHz sampling rate  
Hefur ekki nema bilið á milli 20kHz og 22.05kHz til að fara úr fullum styrk í engan

Hvað er til bragðs?



# Anti Aliasing Filters

## Málamiðlun: (heimadót)

Byrja undir 20kHz

Fæst heyrum við mikið fyrir ofan 16kHz

Fara ekki alveg í núll við 22.05kHz

Ef förum það langt niður að aliasing tíðnir heyrast lítið sem ekkert

Sættum okkur við lítilsháttar fasa og tíðnitruflanir

Eru fyrir ofan 10kHz og breyta ekki öllu

**HVERNIG ÞETTA ER GERT BREYTIÐ MIKLU VARÐANDI GÆÐI CONVERTORA**

# Anti Aliasing Filters

Gagnrýni á stafrænt hljóð

Harður toppur – Yfirleitt vegna þess að Anti Aliasing filterar eru ekki nógu góðir

Ástæða fyrir hærri sampling rate

Ef við sömplum á 96kHz er Nyquist tíðnin 48kHz

Anti Aliasing Filter getur verið 20kHz – 48kHz

Miklu auðveldara að gera góðan svoleiðis filter

Tíðni og fasatruflanir allar fyrir ofan 20kHz – heyrast ekki



# Bitafjöldi

Tíðni sampla er annar helmingurinn af jöfnunni.

Hinn er nákvæmni sampla

Því fleiri bitar sem eru notaðir í hvert sampl, því nákvæmari er upptakan

1 biti = tvær tölur (0 eða 1)

2 bitar = fjórar tölur (00, 01, 10, 11)

4 bitar = átta tölur

N bitar =  $2^N$  tölur

16 bitar =  $2^{16}$  tölur = 65.536

24 bitar =  $2^{24}$  tölur = 16.777.216

# Bitafjöldi - Dynamic Range

Hvert sampl mælir hljóðstyrk á þeim tímapunkti sem samplið er tekið

Munur á hæsta mögulega styrk og minnsta mögulega styrk er „dynamic range“

Bakgrunns suð c.a. 25 dB SPL

Sterkustu tónar í sinfoníu c.a 105 dB SPL

Dynamic range c.a 80 dB

# Bitafjöldi

Hver biti tvöfaldar fjölda talna sem við höfum

Tvöföldun á hljóðstyrk í SPL er 6dB

Hver biti gefur okkur því 6dB í aukinn „dynamic range“

$80\text{dB} / 6 = 14$  bitar (upprunalega áttu geisladiskar að vera það)

16 bitar = 96dB dynamic range

„Perfect Sound Forever“ auglýstu Sony þegar þeir settu fyrstu geisladiskana á markað

# Bitafjöldi

Af hverju meira?

Bakgrunns suðið geymir líka upplýsingar

Efri mörkin (105dB) eru RMS eða meðaltal, Peak value getur farið töluvert hærra

20 bitar gefa 120dB dynamic range.

Það dekkar frá þröskuldi þess sem við heyrum og upp að sársaukamörkum

24 bitar = 144dB

Gott að hafa auka þegar við erum að vinna hljóðið (stærðfræði)

Notum yfirleitt ekki alla bitana

Ekki nema upptakan fari upp í 0dB

# Dithering

Hvert sampl er rúnnað af í næstu tölu sem er til umráða.

Jafnvel þótt við höfum margar tölur er alltaf einhver afrúnnun í gangi

Skömmunarvilla = Quantization Error

Aldrei meiri en helmingur af bili tveggja talna

Alvarlegri þegar við notum færri bita

Alvarlegri á lægri hljóðstyrk

(Við notum færri bita á lægri hljóðstyrk)

# Dithering

Ef við höfum tölurnar 0-9 til umráða og mælingin gefur okkur 4.6 þá er það alltaf rúnnað af upp í 5.

Ef þetta gerist hundrað sinnum í röð er meðaltalið 500 í stað 460.

**Dithering bætir við lítilli Random tölu þannig að afrúnnunin verður ekki alltaf eins**

Ef random talan er t.d á bilinu 0-0.2 þá gæti 4.6 verið rúnnað upp í 5 í c.a. 60% tilfella og niður í 4 í 40% tilfella.

100 skipti, meðaltal = 460

**Hljóðið bjagast ekki (engir nýir yfirtónar verða til) en við fáum suð í staðinn.**

# Dithering Noise Shaping

Í stað þess að bæta við algjörlega random tölu sem gerir jafnt suð yfir allt tíðnisviðið er séð til þess að suðið sé fyrst og fremst á efsta tíðnisviðinu.

Þá verður það enn þá minna áberandi

<https://www.youtube.com/watch?v=IRlohQw-1DY>