
















Sensordatenverarbeitung

DATENAUFNAHME (AUDIO)

21.-24.10.2024

Teila

Nr.	Thema	
1	Einleitung; einführende Beispiele	
2	Datenaufnahme; Audio-Datenaufnahme	
3	Bild-Datenaufnahme	
4	Farbe, Segmentierung, Segmentierungsgetr. BV	
5	Koordinatensysteme; Bewegungs-Datenaufnahme	
6	Audiosignal, 1D Frequenzraum, Fouriertransformation	
7	2D Frequenzraum, 2D Filter	
8	Kanten, SdV-Paradigmen, direkte Bildmerkmale	
9	Houghtransformation, Bewegungsmerkmale	
10	Merkmale, Frequenzmerkmale, Sprachmerkmale	
11	Klassifizierungsalgorithmen	
12	Entwicklung und Evaluation sensorbasierter Systeme	
13	Bayes-Schätzung & Bayes-Filter	
14	Anwendungsbeispiele	

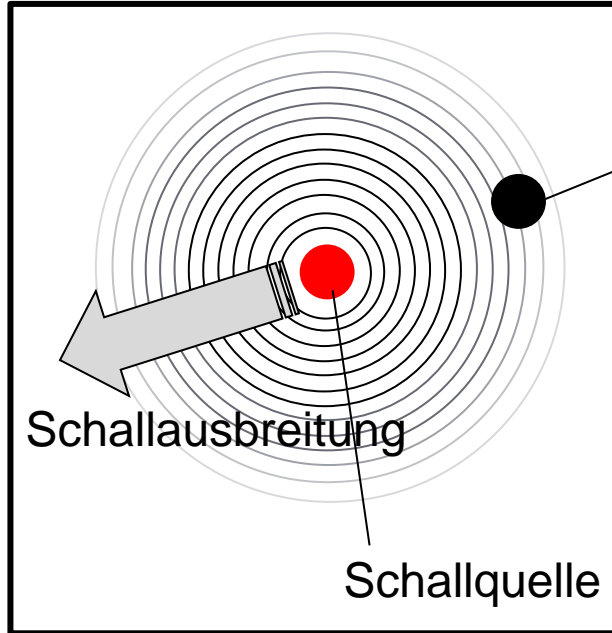
- **Bild und Ton** – in den Medien (Rundfunk, Film, Fernsehen) gebräuchliche Bezeichnung zur Abgrenzung zwischen visuellen und akustischen Signalen
- In der Akustik gelten folgende Definitionen:
 - **Ton**: Signal mit genau einer Frequenz
 - **Klang**: besteht aus mehr als einer Frequenz: Grundschiwingung und Obertöne (ganzzahliges Vielfaches der Grundfrequenz)
 - **Geräusch**: enthält viele verschiedene Frequenzen
- **Audio / Audiodaten / Audiosignale**
 - Aus dem Lateinischen „*ich höre*“ (*audire* = *hören*)
 - Audio wird als Bestandteil von Wörtern verwendet, die mit *Hören* zu tun haben (z. B. Audioanlage, Audiosignal, Auditorium).
 - Menschlicher Hörbereich (Schallwellen 16 Hz bis 20 kHz)

Datenaufnahme allgemein + Audio

- Welche Phänomene spielen bei Audioaufnahmen eine Rolle
- Wie sollte man diese nutzen, um möglichst saubere und gut auszuwertende Audiodaten zu erhalten

- Die klassische Verarbeitungskette zur Messung von Audiosignalen
physikalisches Signal → elektrisches Signal → digitales Signal
- Was ist das Audiosignal: Schalldruckwelle in einem Medium
- Welche Mikrofone gibt es und wie funktionieren sie
 - Wandlerprinzip
 - Akustische Bauform: Richtcharakteristik
 - Sonderformen: Kabellose Mics, Arrays, Körperschall-Mics
- Was muss man bei der Erfassung von Audiosignalen beachten
 - Mikrofontyp
 - Raumakustik
 - Störquellen, Platzierung, Popschutz
- Vom elektrischen zum digitalen Signal: Analog-Digital Wandlung
 - Sampling,
 - Quantisierung,
 - Aliasing

1



2

Schallempfänger
z.B. Mikrofon



3



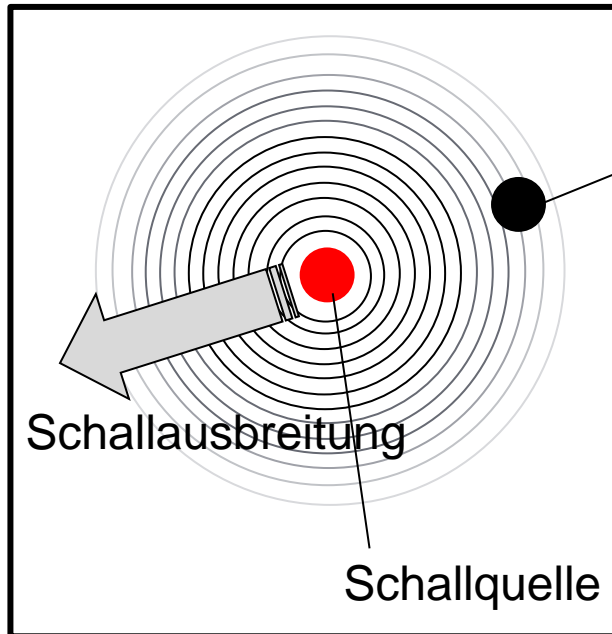
Digitale Verarbeitung

SCHALLDRUCKWELLE

ELEKTRISCHES
SIGNAL

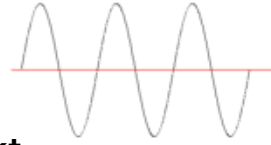
DIGITALES
SIGNAL

1



Schallempfänger

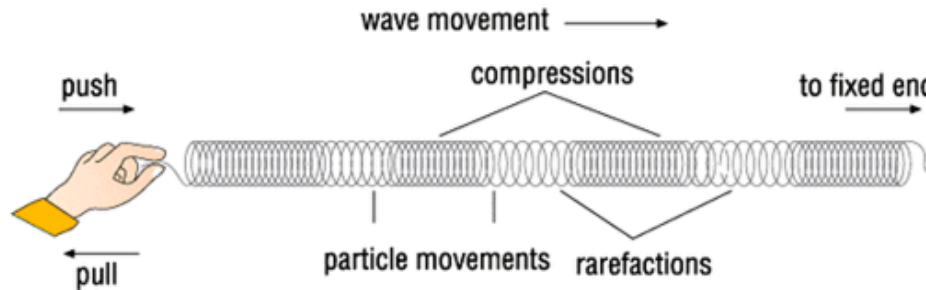
- Was ist Schall, Entstehung
- Wie breitet er sich aus
- Welche Phänomene sind relevant
- Was bedeuten sie für die Messung von Schall mit einem Mikrofon



- Schall wird durch schwingende (vibrierende) Materie erzeugt (= **Schallquelle**, wie z.B. Gitarrensaiten, Lautsprechermembran)
 - Schwingung: eine sich wiederholende Bewegung um den Ruhepunkt
 - Um Materie in Schwingung zu setzen, muss Energie aufgewendet werden
- **Frage: welche Materie schwingt zur Erzeugung von Sprache?**
- Wie wird diese Schwingung für uns Menschen als Schall hörbar?
 - Schwingungen können sich in einem **Medium** ausbreiten (typischerweise Luft, aber auch Wasser, feste Stoffe, oder menschliche Körper)
 - Erreicht die Schwingung das Trommelfell (schwingt!), wird sie als Schall hörbar
- **Frage: kann man im Vakuum Schall wahrnehmen?**
- Sich ausbreitende Schwingungen nennt man Wellen – hier **Schallwellen**
 - Dabei versetzt die Schwingung der Schallquelle benachbarte Partikel des Ausbreitungsmediums (bei Luft sind das Gasmoleküle) in Schwingung
 - Diese Gasmoleküle versetzt wiederum benachbarte Moleküle in Schwingung, usw. ... „Dominoeffekt“
- Die Partikel schwingen **entlang der Ausbreitungsrichtung**
 - solche Wellen nennt man Längswellen oder **Longitudinalwellen**

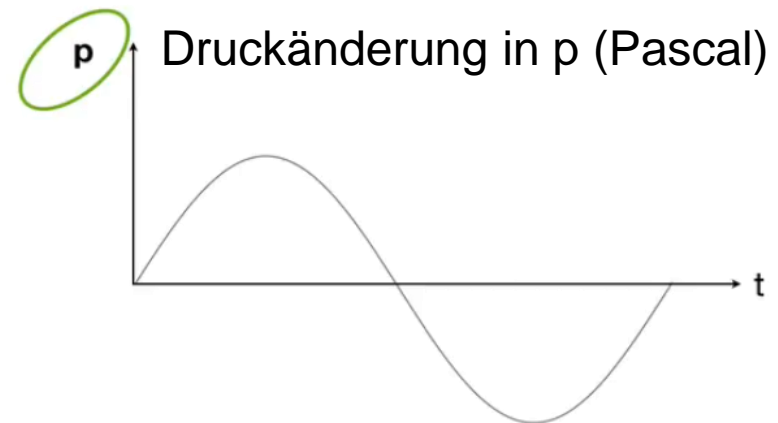


- Längswelle erzeugt Verdichtungen (C) und Verdünnungen (R) im Medium



[media.allrefer.com]

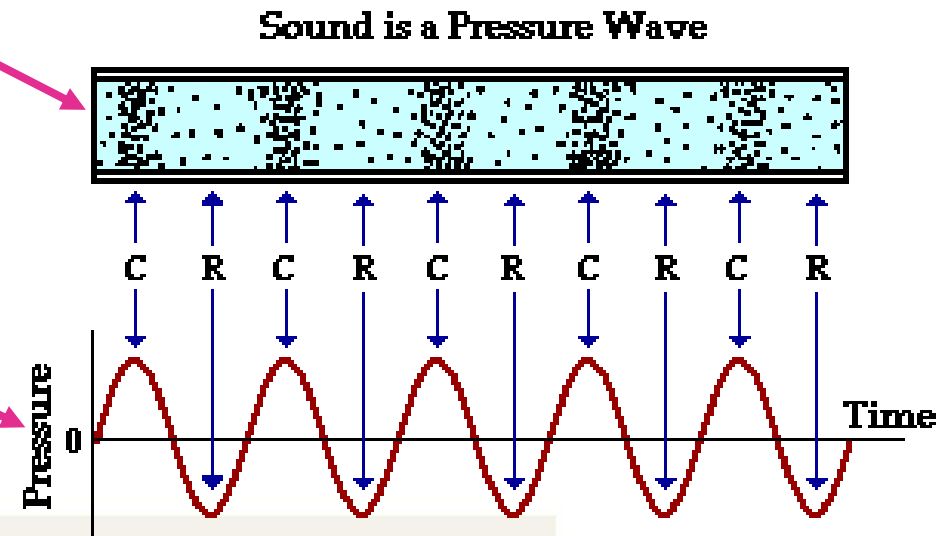
- In Luft als Medium als Überdruck (C) bzw. Unterdruck (R)
- Luftschall: den atmosphärischen Luftdruck überlagerte **Druckänderungen**
- Ruhelage der Gasmoleküle
= atmosphärischer Luftdruck
- Auslenkung um die Ruhelage
= **Luftdruckänderung**
- Richtung der Luftdruckänderung
= Ausbreitungsrichtung des Schalls





<https://www.youtube.com/watch?v=o1H2hdhAgr8>

- Schall ist eine sich ausbreitende, dem atmosphärischen Luftdruck überlagerte Druckschwankung
- Die Ausbreitung erfolgt dadurch, dass die schwingende Schallquelle benachbarte Gasmoleküle in Schwingung versetzt, die wiederum benachbarte Moleküle zum Schwingen bringen (Dominoeffekt)
- Die Gasmoleküle oszillieren dabei um ihre Ruhelage (Vor- und Rückschwingung) entlang der Schallausbreitung – Longitudinalwelle
- Dadurch entstehen Verdichtungen (C) und Verdünnungen (R) (In Luft Über- und Unterdruck)
- Auslenkung um die Ruhelage entspricht der Luftdruckänderung
- Luftschall → Luftdruckänderung
Amplitudenachse Druck p
(gemessen in Pascal)

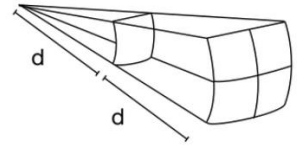


- Schallgeschwindigkeit = Ausbreitungsgeschwindigkeit von Schall im Medium
 - **Dichte des Mediums:** je dichter desto schneller die Ausbreitung
 - **Temperatur des Mediums:** je wärmer desto schneller die Ausbreitung
- Die Schallgeschwindigkeit beträgt ...
- ... in Gasen (Luft) **$331,5 + 0,6T_c$ m/s** (T_c = Lufttemperatur)
 - für $T_c = 20^\circ\text{C} \Rightarrow 343,5 \text{ m/s} = 1236 \text{ km/h}$
- ... in Flüssigkeiten (Wasser 20°C): 1.483 m/s
- ... in Feststoffen:
 - Holz (Eiche): 3.800 m/s , Glas, Eisen: 5.000 m/s
 - Muskel: 1568 m/s , Knochen: 3600 m/s
- Zum Vergleich: Lichtgeschwindigkeit: $299\,792\,458 \text{ m/s} \sim \mathbf{300.000 \text{ km/s}}$
 - Ausbreitung von Licht (im Vakuum)
 - Das gilt auch für alle anderen elektromagnetischen Wellen (Naturkonstante)
- **Frage: fällt euch ein Beispiel des täglichen Lebens ein, in dem die unterschiedliche Geschwindigkeit von Schall und Licht deutlich wird?**

Schall
330 m/s

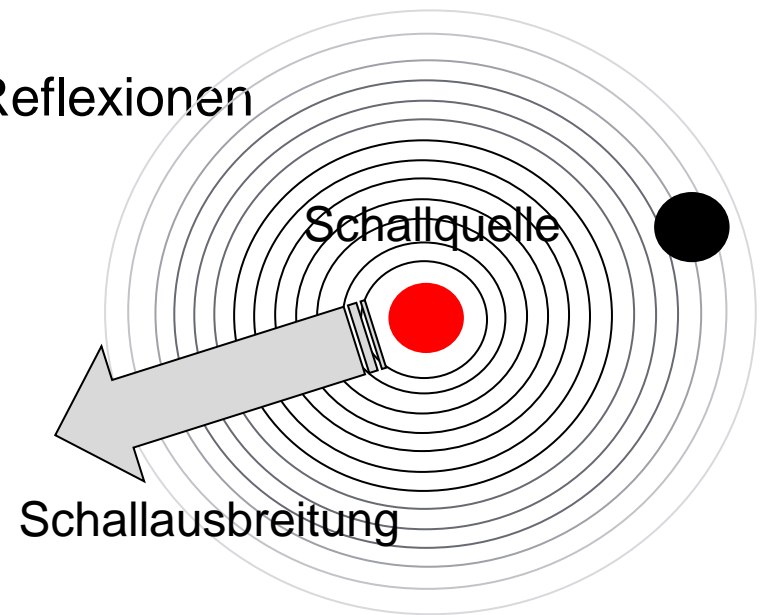
Licht
300.000 km/s

- Schall breitet sich im Medium kugelförmig in alle Richtungen aus
 - Gilt für Nahfeld (1-2m), Punktförmige Schallquelle
 - Abstandsgesetz: die gleiche Energie (Schallintensität Watt/m²) verteilt sich bei einer **Verdopplung des Abstands r** auf das Vierfache der Fläche
 - **Schalldruck** fällt bei Abstand-Verdoppelung auf die Hälfte
 - Entspricht der Abnahme des **Schalldruckpegels um -6dB**

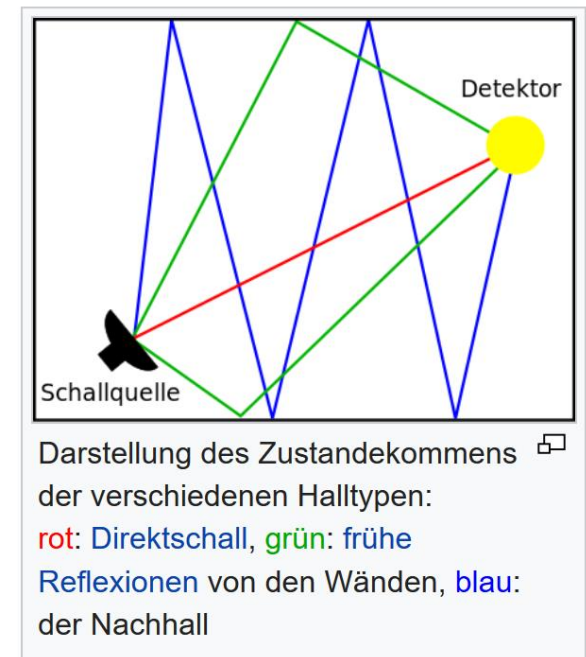


$$-10 \cdot \lg \frac{r_2^2}{r_1^2} = -20 \cdot \lg \frac{r_2}{r_1}$$

- Im geschlossenen Raum: Nachhall durch Reflexionen
 - Detektor nahe bei Schallquelle:
Direktschall überwiegt
 - Detektor weiter weg:
Summe der Reflexionen überwiegt



- **Nachhall** sind kontinuierliche Reflexionen von Schallwellen in einem geschlossenen Raum - Achtung Hall unterscheidet sich von Echo
 - Echo (auch: **Widerhall**) entsteht, wenn die Reflexionen so stark verzögert sind, dass man sie als separates Hörereignis wahrnimmt
- Hier das Wichtigste zu Nachhall:
- Flächen reflektieren Schall in Abhängigkeit ihrer **Oberflächenbeschaffenheit**
- Glatte Flächen reflektieren Schall wie ein Spiegel das Licht, d.h. Einfallswinkel gleich Ausfallswinkel
- Raue Strukturen reflektieren in viele Richtungen, je rauer, desto diffuser ist die Reflexion
- Der Reflexionsgrad ist abhängig von Frequenz und Materialbeschaffenheit
 - Hartes Material absorbiert Schall kaum
 - weiches dagegen stärker
- Das ist der Grund, warum in großen Räumen ohne Absorptionsflächen wie Kirchen, Schwimmhallen oder großen Höhlen Nachhall auftritt



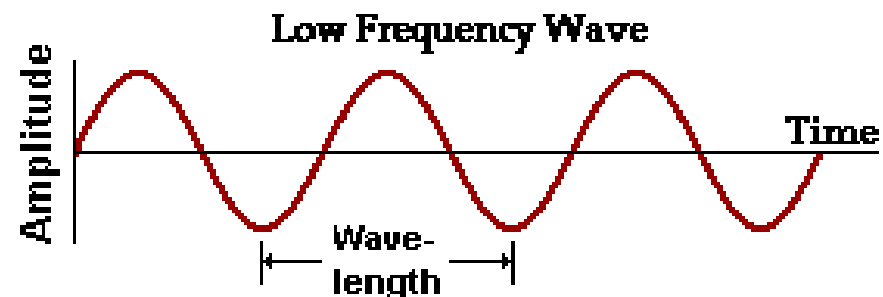
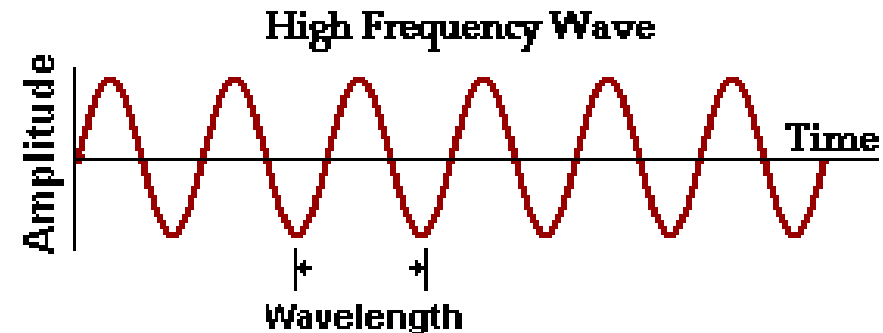
Wikipedia: Nachhall

- Die Luftdruckänderung der Schallwelle ist die **Amplitude A** (in Pascal)
- Die Anzahl der Druckveränderungen pro Sekunde ist die **Frequenz f** der Schallwelle
 - Frequenz wird in der physikalischen Einheit 1/s (Hertz) gemessen.
 - Eine diskrete Frequenz erzeugt einen reinen Sinuston.
 - Geräusche bestehen aus einer Vielzahl von Sinustönen.

$$f = \frac{1}{\text{time}}$$

- Die Geschwindigkeit **v** der Schallwelle und die Frequenz **f** bestimmen die **Wellenlänge l**

$$l = \frac{v}{f}$$



- Der leiseste hörbare Schall verändert den Luftdruck um ca. $\sim 10^{-5}$ Pascal ($1 \text{ Pascal} = 1 \text{ N/m}^2 = 1 \text{ kg/ms}^2$)
 - Definition der Hörschwelle in Luft: $P_0 = 2 \cdot 10^{-5} \text{ Pa} = 20 \text{ } \mu\text{Pa}$ (Mikropascal)
- Der lauteste Schall, den man unter Schmerzen wahrnimmt, verändert den Luftdruck um $100 \text{ Pa} = 10^2 \text{ Pa} = 10.000.000 \text{ } \mu\text{Pa}$
- Aufgrund dieses weiten Bereichs verwendet man statt Schalldruck meist den **Schalldruckpegel** mit logarithmischer Skala in **Dezibel [dB]**
- **Bel** drückt die Quantität einer physikalische Größe in Bezug auf einen Referenzpegel aus, daher hat dB keine physikalische Einheit
- Der **Schalldruckpegel** auch kurz SPL (für Sound Pressure Level) ist definiert als logarithmische Größe und gibt den aktuellen Schalldruck P im Verhältnis zur Hörschwelle P_0 an [in dB]

$$SPL = 10 \log_{10} \left(\frac{P^2}{P_0^2} \right) = 20 \cdot \log_{10} \frac{P}{P_0} \text{ db}$$

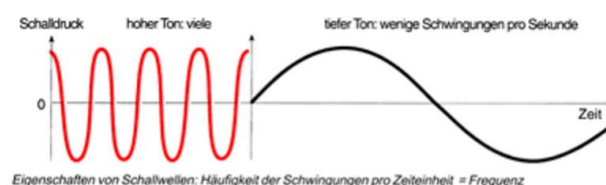
Frage: wieviel dB hat die Hörschwelle?

Beispiel: $P = 100 \text{ Pa}$, $P_0 = 0,00002 \text{ Pa}$
 $P / P_0 = 5.000.000 \text{ db}$
 $\log_{10} 5.000.000 \text{ db} = 6,699 \text{ dB}$
 $20 \cdot 6,699 \text{ db} = 134 \text{ dB}$ (Schmerzgrenze)

Threshold of Hearing P_0	0dB	Leisester hörbarer 1kHz Schall	6dB
Leises Wohnzimmer	20dB	Leises Flüstern	25dB
Kühlschrank (USA)	40dB	Leises Sprechen	50dB
Normale Unterhaltung	60dB	Hauptverkehrsstraße	90dB
Vorbeifahrendes Motorrad	90dB	Schreien	100dB
Presslufthammer	100dB	Hubschrauber	110dB
Rock Konzert	110dB	Luftsirene	130dB
Schmerzschwelle	134dB	Gewehrschuss	140dB
Unmittelbarer Trommelfellriss	160dB	Raketenstart	180dB

Zwischenmenschliche Konversation (50cm Abstand)	60dB
Nahsprechmikrofon ~ 1Pa	94dB

Für mehr Beispiele siehe auch <https://de.wikipedia.org/wiki/Schalldruckpegel>

- **Schall** wird durch schwingende Materie erzeugt
 - Luftschallwellen breiten mit ca. 300m/s als **Longitudinalwellen** aus
 - Die **Amplitude** (Luftdruckänderung in Pascal) bestimmt die Lautstärke
 - Die **Frequenz** (Anzahl der Druckänderungen pro Sekunde in Hertz) bestimmt die Tonhöhe
 - Große Zahlenbereich in Pascal, daher **Schalldruckpegel** definiert, der den aktuellen Schalldruck im Verhältnis zur Hörschwelle in dB angibt
 - Doppelter Abstand von Schallquelle reduziert den Pegel um -6db
 - In geschlossen Räumen sind das Verhältnis von Direktschall und **Reflexionen** (Nachhall) wichtig für die Hör-/Aufnahmequalität
 - Für gute Sprachverständlichkeit:
 - Reflexionsarm
(raue weiche Flächen statt glatte harte)
 - Hohen Anteil an Direktschall
- 
- Das Diagramm zeigt zwei Schallwellen auf einem Koordinatensystem mit der Zeitachse (Zeit) und der Schalldruckachse (Schalldruck). Die linke Welle ist eine rote Sinuskurve mit hoher Frequenz, beschriftet mit 'hoher Ton: viele' und 'Eigenschaften von Schallwellen: Häufigkeit der Schwingungen pro Zeiteinheit = Frequenz'. Die rechte Welle ist eine schwarze Sinuskurve mit niedriger Frequenz, beschriftet mit 'tiefer Ton: wenige Schwingungen pro Sekunde'.

Gute Einführung auf Youtube: [sonic-vision.tv](https://www.youtube.com/channel/UC8v33333333333333333333)

