

Kapitel 3: Schwingung und Tonhöhe

Die Übertragung musikalischer Information spielt sich im Norm fall so ab, dass Energie in Gestalt von Schallwellen an das menschliche Ohr gelangt. („Normalfall“, weil es denkbar ist, dass das Gehörzentrum des Menschen auch direkt elektrisch oder chemisch gereizt werden könnte...) Energie, wie sie in Kapitel 2 abgehandelt wurde, ist im wesentlichen *Träger* der musikalischen Information. Die Information steckt in der spezifischen **Schallstruktur** der Energie. Ohne Energie gibt es keine musikalische Information, aber nicht jede Energie enthält musikalische Information. Die Menge der Energie ist für die transportierte Struktur nicht von großer Bedeutung, wie folgendes Beispiel zeigt:



Abb. 3.1 Energie und Struktur bei der E-Gitarre

Die Schwingungsenergie der abgebildeten Saiten ist nicht sehr groß, sie beträgt einen Bruchteil der Finger-Bewegungsenergie der Spieler. Die an die Luft abgestrahlte Schallenergie der E-Gitarre ist noch viel geringer. Der Tonabnehmer nimmt Bruchteile der Saiten-Schwingungsenergie auf. Im Verstärker „verformt“ dieser kleine Energie-Anteil die erheblich größere Energiemenge, die aus der Steckdose in den Verstärker gelangt: die Gitarren-Energie prägt ihre Struktur der Steckdosen-Energie auf. Die derart „verformte“ Steckdosenenergie erregt die Membran eines Lautsprechers, der nun - im Gegensatz zur E-Gitarre - seine Schwingungsenergie optimal als Schallenergie an die Luft abgeben kann. Für die musikalische Information ist die Zusatzenergie aus der Steckdose nicht von qualitativer Bedeutung, Spielfehler können nicht durch die Elektrizitätswerke korrigiert werden.

3.1. Schwingungen allgemein

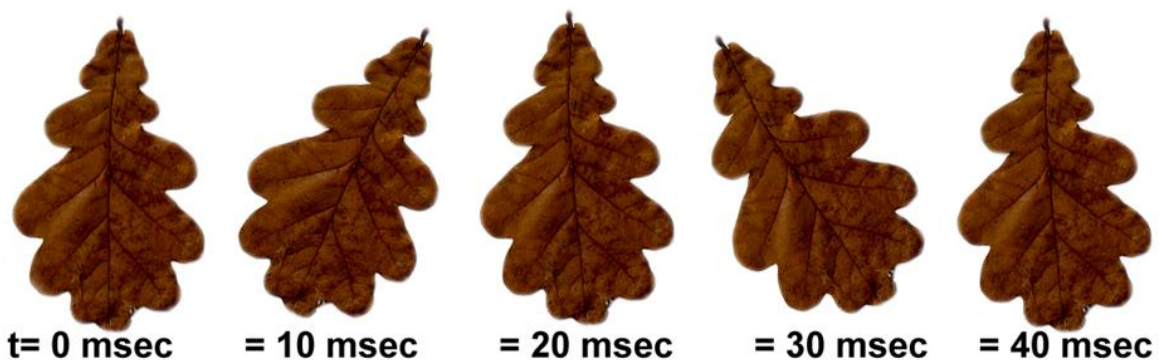
Ziel des Musikmachens ist es, das Trommelfell in mechanische **Schwingung** zu versetzen. Das Trommelfell soll sich so hin- und her bewegen, dass im Gehirn „Musik“ entsteht. Zwischen Trommelfell und Schallquelle, einem Musikinstrument, einem sonstigen Klangerzeuger oder Lautsprecher, befindet sich in der Regel nur Luft. Sobald sich die Schallquelle, die mit der Luft Berührung haben muss, schwingt, wird der chaotischen Bewegung der Luftmoleküle eine gewisse Ordnung aufgezwungen (sofern die Musik ordentlich ist). Diese „Ordnung“ pflanzt sich von der Schallquelle bis zum Trommelfell hin fort (wie in Abschnitt 7.1 genauer untersucht werden wird).

Dabei bleiben die Luftmoleküle, mit denen die Schallquelle Berührung hat, in der Nähe der Schallquelle, und die Luftmoleküle, die sich im Gehörgang befinden, im Gehörgang. Die musikalische Informationsübertragung ist nicht mit einem Transport von Luftmolekülen oder sonstiger Materie verbunden. (Denn sonst entstünde beim Verschicken von Schallwellen ein Wind.) Sehr langsame und starke Schallschwingungen kann man als Vibration spüren. Die meisten musikalisch relevanten Schwingungen sind aber klein und schnell. Man benötigt Hilfsmittel, um sie sichtbar zu machen. Im vorliegenden Kapitel beschäftigen wir uns aber zunächst nur mit Schwingungen und noch nicht mit der Fortpflanzung von Schwingungen im Raum (den Wellen).

Definition

Von **Schwingung** spricht man, wenn sich eine physikalische Größe im Laufe der Zeit periodisch oder nicht-periodisch so ändert, dass sie um einen gewissen Ausgangspunkt („Ruhelage“) oszilliert. Ein am Baum hängendes, im Winde wehendes Blatt „schwingt“, während ein abgerissenes und vom Wind weit weg getragenes Blatt nicht mehr „schwingt“. Beide Male findet zwar eine Veränderung der physikalischen Größe „Ort“ statt, doch nur im ersten Falle gibt es so etwas wie ein Oszillieren.

Bei der Verbildlichung des zeitlichen Vorganges „Schwingung“ bedient man sich in der Regel eines Koordinatensystems, bei dem zeitlich hintereinander vorkommende Ereignisse räumlich nebeneinander dargestellt werden. Eine abendländische Konventionen lautet dabei: „die Zeitachse verläuft von links nach rechts“. Zum Beispiel:



In dieser Darstellung sind 5 Zeitpunkte explizit abgebildet. Einfacher wäre es, nur die physikalische Größe „Auslenkung“ aus der Ruhelage (die bei 20 msec eingenommen wird) in cm oder mm anzugeben, wie das im folgenden Diagramm geschieht:

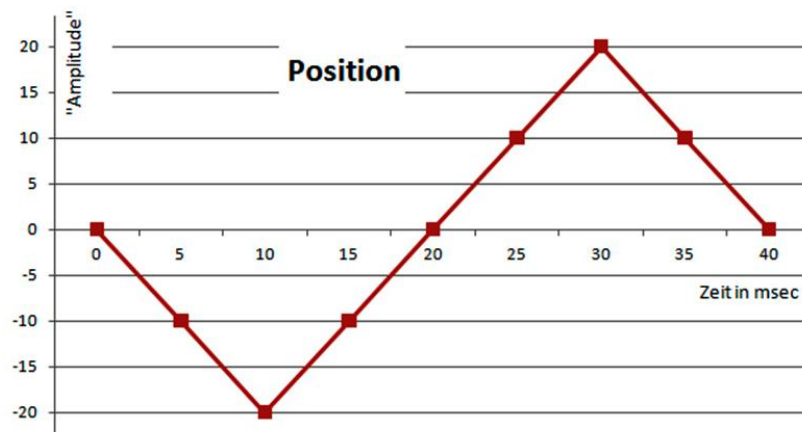


Abb. 3.2 Schwingung des Eichenblattes als Diagramm

Periodisch ist eine Schwingung dann, wenn sich nach immer gleichen Zeiträumen (= „Perioden“) ein Schwingungsablauf genau wiederholt. Ist dies nicht der Fall, so ist die Schwingung **nicht periodisch**. Beispiele periodischer Schwingungen: Bewegung der Erde um die Sonne (Periode = 1 Jahr), Ebbe und Flut (Periode = ein paar Stunden), Kolbenbewegung eines Automotors bei gleichbleibender Geschwindigkeit (Periode = Bruchteile von Sekunden). Nicht-periodische Schwingungen: Bewegung einer Schwingtür in einem Kaufhaus, Bewegungen einer Wetterfahne, Bewegung der Elektronen auf der Sonne. Die letztgenannten Beispiele zeigen, dass ganzheitlich betrachtet nicht-periodische Schwingungen durchaus sinnvolle Information enthalten können („Besucherstrom“ im Kaufhaus, Windböen, „weißes Licht“ der Sonne). Musikalisch sinnvoll sind periodische *und* nicht-periodische Schallschwingungen.

Die grafische Darstellung von periodischen Schwingungen ist relativ einfach, da es genügt zu wissen, wie sich die interessierende physikalische Größe innerhalb einer Periode verändert:

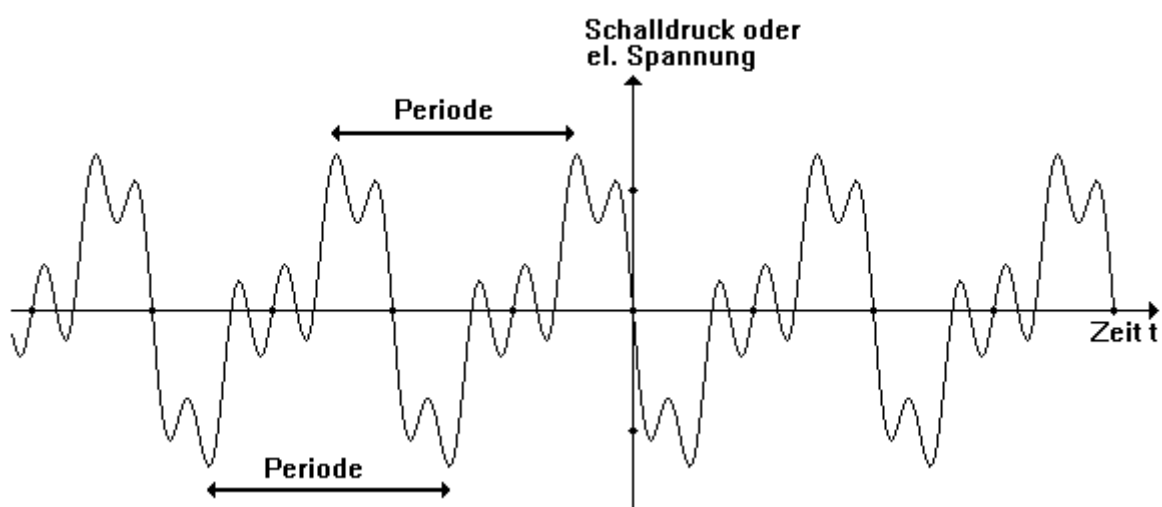


Abb. 3.3 Kontinuierliche Darstellung einer periodischen Schwingung (Posaune)

Während die Zeit („X-Achse“) stets von links nach rechts verläuft, können in der vertikalen Richtung („Y-Achse“) recht unterschiedliche physikalische Größen stehen. Musikalisch relevant sind unter anderem: Luft- oder Schalldruck an einer Stelle, die elektrische Spannung (beispielsweise im Lautsprecherkabel), die Auslenkung einer Saite, die Schwärzung eines Films (beim Lichttonverfahren des Kinos), der Zahlenwert eines Speicherplatzes auf einer CD-ROM usw.

Die Periode wird in Sekunden (sec) gemessen und mit T bezeichnet. Die Anzahl der Perioden pro Zeit heißt **Frequenz** und wird entsprechend in 1/sec gemessen, was auch Hertz („Hz“) heißt.

Frequenz = Anzahl der Perioden pro Zeit

$$f = 1/T \quad (\text{in } 1/\text{sec} = \text{Hz})$$

Nach Abbildung 2.5 kann der Mensch Schallschwingungen zwischen 16 Hz und 12000 bis 20000 Hz hören. Die obere Hörgrenze sinkt mit dem Alter auf ca. 12 000 Hz.

Frequenzmessungen

Heute kann jede Computerbesitzerin einfache Frequenzmessungen im Bereich, den Mikrofon und Soundcard „zulassen“, durchführen (oft zwischen 50 und 8 000 Hz). Man lässt sich das Schwingungsbild grafisch darstellen (siehe auch Kapitel 4) und versucht, die Zeiteinheit auf der Zeitachse zu erkennen, was je nach verwendetem Programm mehr oder minder einfach ist.

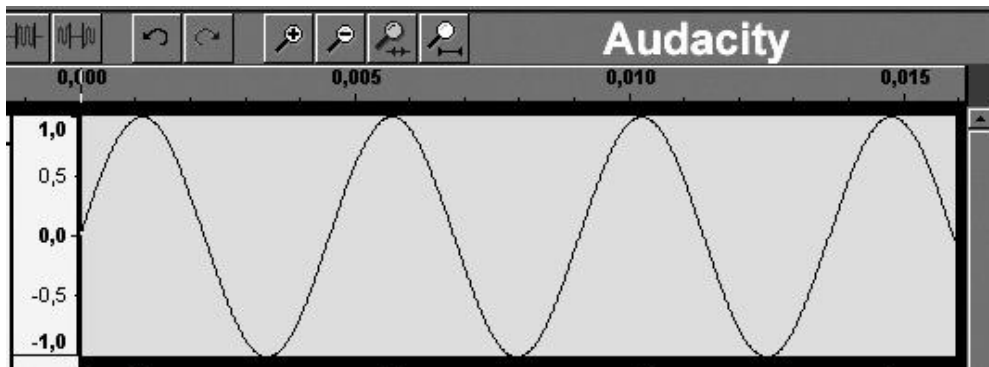


Abb. 3.4 Frequenzbestimmung am Computerbildschirm

Abbildung 3.4 zeigt eine mit dem kostenlosen Audiorecordingprogramm „[Audacity](#)“ aufgenommene und dargestellte Sinusschwingung. Der Zeitmaßstab ist hier explizit mit abgebildet. Eine Periode hat die Dauer von $T = 0.0045$ sec. Daraus folgt für die Frequenz $F = 1/0.0045 = 222$ Hz.

Eine ungewöhnliche, aber durchaus durchführbare Frequenzmessung kann an Vinyl-LP's durchgeführt werden, deren Rillen man mit der Lupe betrachten und die Perioden pro Zentimeter Rillenlänge auszählen kann (Abbildung 4.3). Aus Umdrehungsgeschwindigkeit (33 U/min) und Rillenlänge pro Umdrehung ($2\pi R$) kann die Zeit bestimmt werden, in der das ausgezählte Rillenstück von der Nadel des Plattenspielers überstrichen wird.

Nicht-periodische Schwingungen lassen sich nicht so einfach darstellen wie periodische. Man begnügt sich meist mit zeitlich begrenzten Ausschnitten und einem „qualitativen“ Überblick. Eine „Frequenz“ gibt es definitionsgemäß nicht ([PLAY](#)). Eine nicht periodische Schwingung sieht am Computerbildschirm folgendermaßen aus (der Computer greift aus der analog nicht darstellbaren Schwingung nach einer inneren „Uhr“ Werte heraus, deren Abfolge eine Zufallsverteilung abgeben):

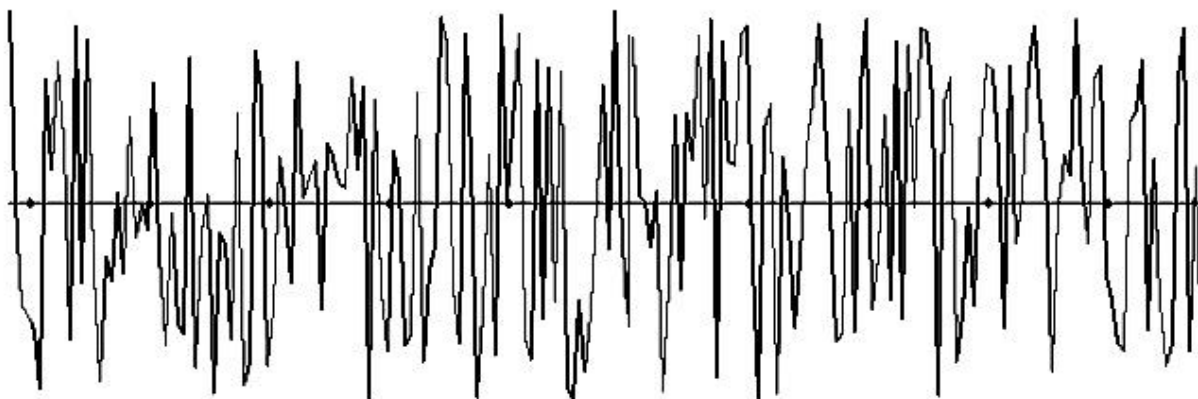


Abb. 3.5 Computerbild einer unperiodischen Schwingung

Der Höreindruck einer Schwingung dieser Art ist eine Art „Rauschen“. Das stimmlose „sch“ erzeugt ein *weißes* Rauschen, bei dem statistisch ausgedrückt alle Frequenzen nach einer Normalverteilung gleich wahrscheinlich sind. Wird diese Wahrscheinlichkeit auf einen Frequenzbereich eingeschränkt, so hört man *farbiges* Rauschen. - Unperiodische Schwingungen kommen beim Musizieren häufig vor. Die deutsche Sprache hat einen sehr hohen Rauschanteil, auf Hawaii dagegen kann es Folgen von 6 oder 8 Vokalen (periodischen Schwingungen) ohne dazwischenliegendes Rauschen (Konsonanten) geben. Weil die menschliche Stimme alle ihre Laute aus periodischen Schwingungen (der Stimmbänder) und Rauschen (des Luftstroms) zusammensetzt, hat bis in die 70er Jahre hinein die elektronische Musik mit der Hypothese gearbeitet, es gäbe keine weiteren Schwingungstypen neben „periodisch“ und „streng unperiodisch“. Indessen bevorzugten 2- und 3-dimensionale Schwingungserzeuger (vgl. Kapitel 9) sog. „quasiperiodische“ Schwingungen, die Regelmäßigkeiten aufweisen, ohne im strengen Sinne periodisch zu sein. Und in den 80er Jahren hat die „FM-Synthese“ für Aufregung gesorgt, deren Schwingungen sich ebenfalls nicht ins bisherige Schema pressen ließen (vgl. Abbildung 10.12). - Hier das Schwingungsbild einer tibetischen Klangschaale ([PLAY](#)):

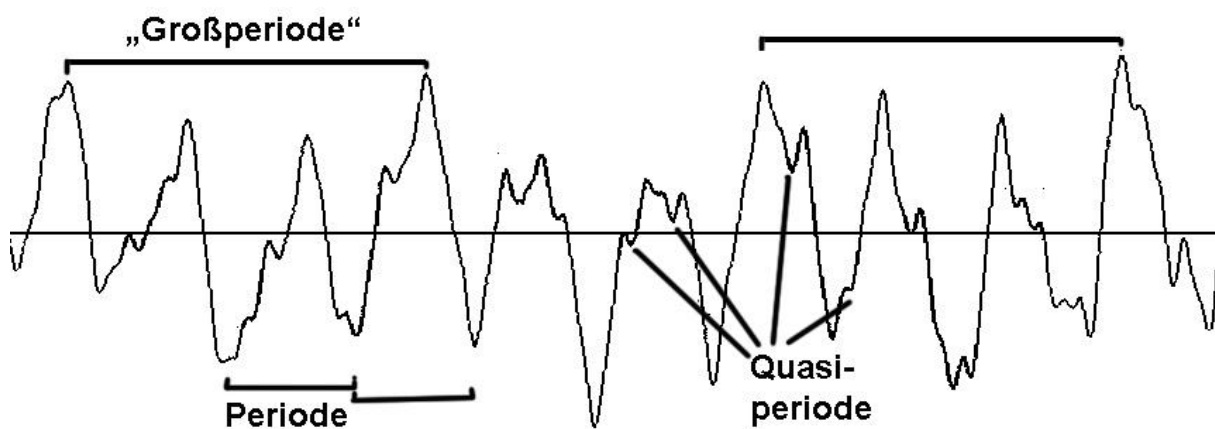


Abb. 3.6 Quasiperiodische Schwingung einer Klangschaale

Weitere Beispiele von Schwingungstypen vor allem bei Sprachlauten und Perkussionsinstrumenten folgen in den jeweiligen Kapiteln. Durch die elektronische Musik sind einige Schwingungsformen besonders „populär“ geworden:

- Die *Sinusschwingung*, definiert als eine Schwingung, bei der die rücktreibende Kraft des schwingenden Partikels proportional der Auslenkung ist.
- Die *Rechteckschwingung*, bei der eine „Klappe“ auf und zu geht („Klarinettenmundstück“).
- Die *Sägezahnsschwingung*, bei der auf eine „langsame“ Auslenkung ein „schnelles“ Zurückschnellen erfolgt („Streichinstrumente“).
- Die *Dreiecksschwingung*, bei der ein „eckiges“ Hin und Her erfolgt („Posaune“).
- Der *Puls*, bei dem das Rechteck unsymmetrisch ist (siehe Kapitel 2!).



Abb. 3.7 Rechteck, Sägezahn, Dreieck, Puls

3.2. Tonhöhe und Frequenz

In der Musikpraxis gibt es merkwürdige Phänomene: So stimmt ein Schlagzeuger seine Toms sorgfältig in der Anordnung HiTom-MidTom-LowTom als Quint-Terz-Grundton, um hernach bedenkenlos in C-Dur *und* D-Dur zu spielen, ohne dass es irgendjemanden stört. Das menschliche Ohr ist offensichtlich in der Lage, die Abfolge „HiTom-MidTom-LowTom“ in unterschiedlichen Tonarten als „passenden“ Dreiklang zu hören. Ähnliches gilt für Instrumente wie Bass Drum, Glocken, Gongs usw.

Experiment 3. 1. „Hat das Schallsignal eine Tonhöhe?“

Es werden **Samples** (= digital gespeicherte Naturklänge) unterschiedlicher Percussionsinstrumente (z.B. die 64 Percussionsinstrumente einer PC-Soundcard in GM-Norm) gespielt und gefragt: „Hat dies Instrument eine Tonhöhe?“. – Sodann werden einzelne Instrumente beschleunigt oder verlangsamt abgespielt und gefragt: „ändert sich die Klangfarbe oder die Tonhöhe?“ ([PLAY: zu Extra-Blatt gehen!](#))

Das Ergebnis dieses Hörexperiments ist:

1. Die Musikpraxis kennt offensichtlich drei Arten von Instrumenten:
 - solche mit genau fixierter Tonhöhe (Klavier, Trompete, Vibraphon),
 - solche mit unterschiedlich interpretierbarer Tonhöhe (Toms, Glockenspiele, Xylophone),
 - solche ohne Tonhöhe (Snare, Triangel, Becken).
2. Bei einigen Instrumenten gehen die (Hör-)Meinungen, ob das Instrument eine Tonhöhe hat oder nicht, auseinander. Die Frage „Hat dieser Klang eine Tonhöhe?“ wird beim Experiment so beantwortet, dass die HörerInnen versuchen, den Klang laut oder leise oder innerlich mitzusingen. Gelingt dies befriedigend, so antworten sie mit „ja“. Die Feststellung „Tonhöhe ja oder nein?“ wird also durch ein individuelles Vergleichsexperiment getroffen.
3. Relativ eindeutig ist die Feststellung, dass gewisse Instrumente beim „Transpoieren“ nur ihre Klangfarbe, andere Tonhöhe und Klangfarbe verändern. „Keine Tonhöhenänderung“ findet selbstverständlich dort statt, wo man auch „keine Tonhöhe“ hört – dennoch sind die Antworten hier eindeutiger als bei Punkt 2.

Experiment 3. 2. Beschreibung der Schwingungsstruktur

Wir sehen die Schwingungsbilder der in Experiment 3.1 verwendeten Klänge an einem **Speicheroszilloskop** oder am Bildschirm eines Computers an.

Ergebnis: Je deutlicher die Periode einer Schwingung ausgeprägt ist, umso klarer erkennt man beim Hören eine Tonhöhe. Alle Instrumente mit klarer Tonhöhe besitzen periodische Schwingungsbilder. Alle ohne Tonhöhe haben strikt unperiodische Schwingungsbilder. Die übrigen Fälle „unklarer“ oder „mehrdeutiger“ Tonhöhe können unklare oder mehrdeutige Periodizität zeigen, zum Beispiel:

- Quasiperiodizitäten (vgl. Abbildung 3.6),
- Periodizitäten, die nicht-periodischen Schwingungen „übergestülpt“ sind (vgl. Abb. 3.8 und 9.5),
- periodische Schwingungen, die so kurz sind, dass das Ohr keine Periodizität feststellen kann (was bei der Bass Drum und stimmhaften Plosivlauten wie „b“ oder „d“ der Fall ist, vgl. Abb. 9.2),
- Mischungen von periodischen und unperiodischen Anteilen.

Diese Aufstellung ist rein deskriptiv. Wir werden später sehen, welche Ursachen deutliche und undeutliche Tonhöhereindrücke haben können: Ursachen, die im Bau der Instrumente oder in der

Obertonstruktur begründet sind. Dass HörerInnen bei einigen Klängen unterschiedlicher Meinung sein können, kann daran liegen, dass sich das „Ohrenmerk“ stärker auf die periodischen Anteile eines Klanges richten kann oder eben nicht.

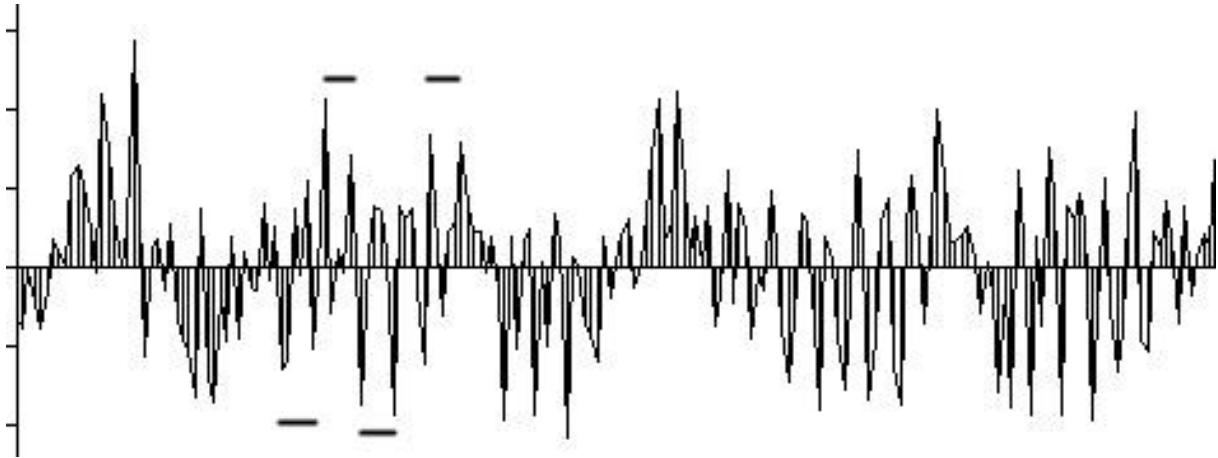


Abb. 3.8 Schwingungsbild eines Tambourins

Die Dimension „Tonhöhe“ spielt bekanntlich in der abendländischen Musik eine große Rolle. Eine klar erkennbare Tonhöhe ist die Voraussetzung für klar erkennbare Melodien und Harmonien. Instrumente mit undeutlichen Tonhöhen haben es daher im Laufe der abendländischen Musiktradition schwer gehabt. Das Erkennen von Tonhöhen ist nicht nur durch die Periodizität der den Höreindruck hervorrufenden Schwingungen, sondern auch durch kulturelle Lernleistungen bedingt. Kinder bezeichnen unterschiedlich hohe Töne oft als „hell - dunkel“ oder „spitz - stumpf“ und verquicken Klangfarbe und Tonhöhe. Dies ist keineswegs „falsch“! Und dennoch muss man in der Grundschule lernen, dass Töne tief und hoch und nicht hell und dunkel sind. - Worin besteht der „Witz“ der folgenden Abänderung eines bekannten Comic von Uli Stein?



Abb. 3.9 Missverständnis bei der Bezeichnung „hoch“ und „tief“

Das Witze-Schema Uli Steins in seinen beiden Bänden „Pisa-Alarm“ besteht darin, dass der Schüler aus seiner Sicht im Recht ist, der Lehrer jedoch etwas anderes meint oder verlangt. Das Pisa-Problem, so Uli Stein, bestehe nicht in der Dummheit der Schüler sondern in der „interkulturellen Kommunikation“ im Klassenzimmer.

Die Metaphorik „hoch - tief“ ist nicht in allen Sprachen des Abendlandes verbreitet: so bezeichnen die Engländer zu hoch oder zu tief gestimmte Instrumente mit „sharp - flat“. Zu hoch oder zu tief singen heißt im Italienischen *crescere* bzw. *calare*.

Nachbemerkung:

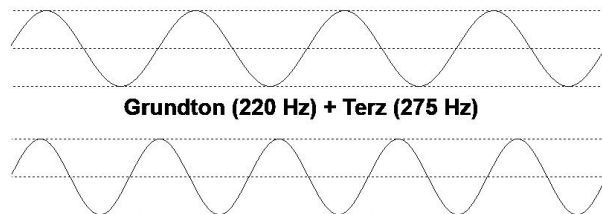
Nach DIN 1320 wird mit „Ton“ die durch eine Sinusschwingung ausgelöste Hörempfindung bezeichnet, also das, was MusikerInnen „**Sinuston**“ nennen. MusikerInnen bezeichnen mit „**Ton**“ jede Empfindung, die eine Tonhöhe hat, d.h. durch eine periodische Schwingung ausgelöst wird. Solch ein (umgangssprachlicher) „Ton“ soll nach DIN 1320 eigentlich als „Klang“ bezeichnet werden. Was MusikerInnen als „**Klang**“ bezeichnen, nennt DIN 1320 „Klanggemisch“.

3.3. Schwebungen

Spielt ein Instrument mehrere Töne oder spielen mehrere Instrumente zusammen, so überlagern sich im Instrument oder in der Luft die einzelnen Schwingungen zu einer Gesamtschwingung. Die Amplitude (z.B. die Abweichung des Luftdrucks vom mittleren Luftdruck) der Gesamtschwingung ist zu jedem Zeitpunkt die Summe der Amplituden aller Einzelschwingungen. In Abbildung 3.10 sind zwei Sinusschwingungen abgebildet, deren Frequenzen sich wie 4:5 verhalten (auf 4 Perioden des oberen Bildes kommen 5 des unteren). Musikalisch gesprochen liegt eine Durterz vor. Die Gesamtschwingung, die durch Addition der Amplituden zu jedem Zeitpunkt entsteht, zeigt das dritte Schwingungsbild. Setzt die Schwingung der Terz um eine Viertel Periode später ein, so ergibt sich das unterste Gesamtschwingungsbild.

Der Hörpsychologie gibt die Überlagerung von Schwingungen einige Rätsel auf:

- Wie erkennt der Mensch, dass bei der „Gesamtschwingung“ zwei unterschiedliche Sinustöne vorliegen?



- Warum merkt der Mensch den Unterschied zwischen dem vorletzten und letzten Schwingungsbild von Abbildung 3.10 nicht? (Mit anderen Worten: Warum spielt die Phasenlage der Einzelschwingungen keine Rolle?)

Und, wir greifen voraus: Warum erkennt der Mensch bisweilen derartige harmonische Sinustonüberlagerungen als „Klangfarbe“ eines *einzigen* Tons?

Experiment 3. 3. Schwebungen

Wir überlagern zwei Sinusschwingungen, halten die Frequenz f_1 der einen Schwingung konstant und verändern die Frequenz f_2 der zweiten Schwingung kontinuierlich. Technisch verwenden wir das „Pitch-Bend“ eines Synthesizers, mit dem üblicherweise glissando-Effekte erzielt werden.

Abb. 3.10 Sinusschwingungen im Terzabstand überlagert

Der Computer spielt auf Synthesizer 1 den Bezugston der Frequenz f_1 und spielt auf Synthesizer 2 denselben Ton, dessen Frequenz f_2 mittels Pitch-Bend leicht verschoben wird. Am Bildschirm kann man die Werte von f_1 und f_2 in Hz ablesen.

Fall 1: Überlagerung in der Luft (d.h. Addition der akustischen Schwingungen) . Die beiden Töne erklingen aus zwei unterschiedlichen Lautsprechern.

Fall 2: Überlagerung im Instrument (d.h. Addition der elektrischen Schwingungen). Die summierten Schwingungen erklingen als Gesamttone aus einem Lautsprecher (Abbildung 3.11).

Fall 3: Überlagerung im Gehirn des Menschen. Die beiden Töne werden stereophon getrennt über Kopfhörer an die beiden Ohren gesendet.

(Unser Stereophones Tonbeispiel kann für alle 3 Fälle verwendet werden: [PLAY!](#)) Ergebnis des Schwebungsexperiments für Fall 1: Der Höreindruck „zerfällt“ in drei recht unterschiedliche Bereiche.

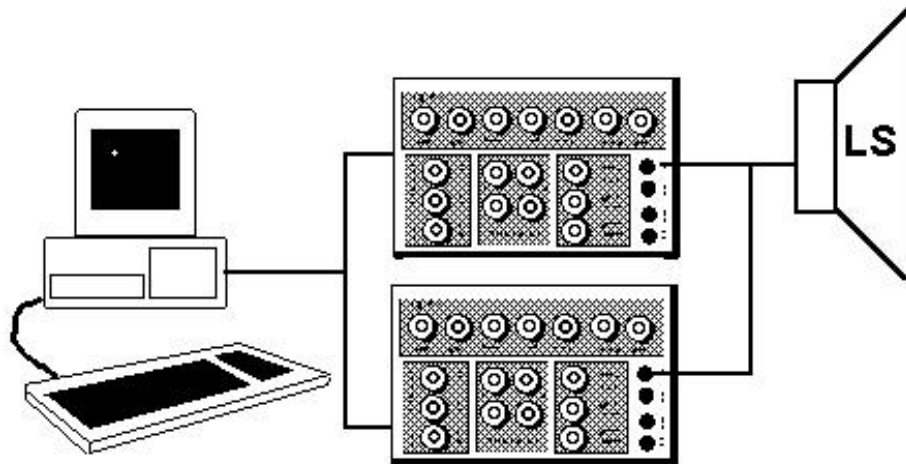


Abb. 3.11 Versuchsanordnung Fall 2

- Bereich 1: Bei geringen Frequenzunterschieden (bis ca. 10 Hz) hört man einen einzigen (Sinus-) Ton, dessen Amplitude sinusförmig schwankt. Dies ist der *Schwebungsbereich* im engen Sinn.
- Bereich 2: Bei etwas größeren Frequenzunterschieden hört man den Übergang zur Zweitonempfindung, wobei die („glatte“) Schwebung in ein raues Knattern übergeht.
- Bereich 3: Vergrößert sich der Frequenzunterschied, sodass die Sinustöne mehr als 20 Hz auseinander liegen, so wird die Zweiton-Empfindung glatt und die Rauigkeit geht in einen „Differenzton“ über. (Ein „Knattern“ von 32 Hz ist hörbar als ein 32 Hz-Ton.)

Bei Fall 2 kann ein Oszilloskop direkt an die Synthesizer angeschlossen und somit die elektrische Überlagerung gezeigt werden. Man hört alle drei Bereiche von Fall 1. Das Bild am Oszillografen ist das einer **amplitudenmodulierten Schwingung** (siehe Abbildung 3.12). Dies Bild ist unabhängig davon, in welchem Schwebungsbereich wir uns befinden. Es bleibt auch dann erhalten, wenn wir bereits zwei „glatte“ Einzeltöne hören. Das Bild sieht bei phasengleichen und gleichstarken Sinustönen, die um ca. 4% gegeneinander verstimmt sind, wie in Abbildung 3.12 aus.

Der Fall 3 („**binaurale Schwebungen**“) ist qualitativ von den Fällen 1 und 2 unterschieden. Bei fast allen Personen ändert sich die Schwebung qualitativ, sie verschwindet jedoch nicht. Der Ton scheint jetzt mit der Schwebungsfrequenz im Kopf zu rotieren. Diese Rotation ist als „Phasing“ bekannt und tritt sogar dann auf, wenn eine minimalistische Tonfolge binaural so dargeboten wird, dass die Tonfolge des einen Kanals um ein paar hundertstel Sekunden gegenüber derjenigen des anderen Kanals verschoben wird. Da im Nervensystem Töne nicht als Schwingungen sondern (digital) als Impulsfolgen kodiert sind, ist es nahe liegend, bei diesen „binauralen Schwebungen“, deren Existenz vor 50 Jahren noch geleugnet worden ist, von einer akustischen Täuschung zu sprechen. Der Mensch kennt diesen künstlichen Effekt nur von der Raumbewegung her, also projiziert er ihn auf eine solche...

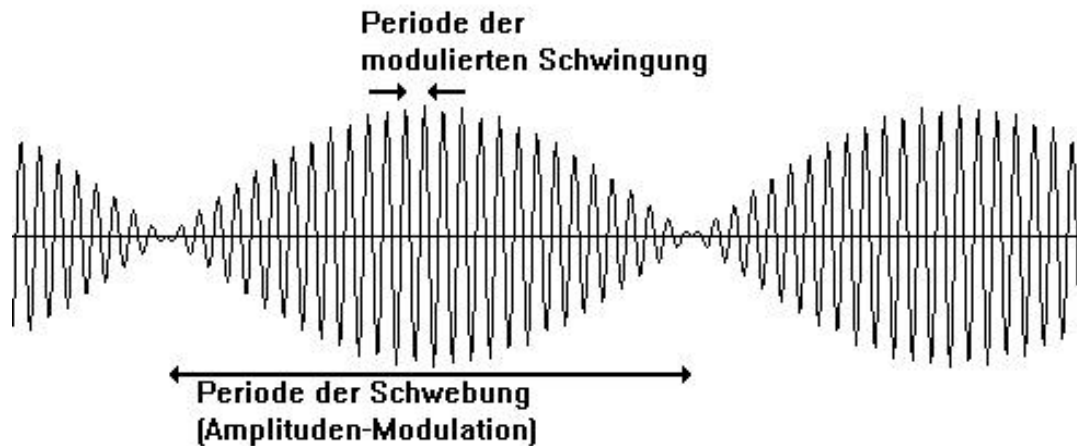


Abb. 3.12 Schwebungsbild: Amplitudenmodulation

In allen drei Fällen ist eindeutig festzustellen, dass die Schwebung im Bereich 1 (Frequenzdifferenz kleiner als 10 Hz) eine Frequenz hat, die genau der Differenz der beiden Sinustöne entspricht, also

$$f_{\text{Schwebung}} = f_1 - f_2$$

(falls f_1 größer als f_2) beträgt (PLAY: [440/441 Hz](#), [440/442 Hz](#), [440/443 Hz](#)). Die Frequenz des gehörten Tons, der schwebt, ist

$$f_{\text{Träger}} = (f_1 + f_2)/2$$

(d.h. das arithmetische Mittel aus den beiden überlagernden Frequenzen).

Anwendung der Schwebungen:

In **Stimmgeräten** wird eine eintreffende Schwingung der Frequenz (f_2) einer im Stimmgerät einprogrammierten Schwingung der Frequenz (f_1) überlagert. Wenn die Schwebung verschwindet, sind beide Frequenzen gleich groß.

Schwebungen zwischen zwei **Blockflöten** führen dazu, dass diese Instrumente im Duo sehr penetrant „flattern“ können. Fröhliche Weihnacht!

Die Schwebungsfrequenz kann auch in den hörbaren Bereich fallen. Es treten in diesem Falle „Kombinationstöne“ auf. Beim Stimmen einer Geige hat Tartini entdeckt, dass bei einer reinen Quint dieser „Kombinationston“ mit einem der beiden zu stimmenden Saitentönen übereinstimmt. (Allerdings ist „Kombinationston“ eine allgemeinere Bezeichnung, Schwebungstöne sind die einfachsten „Kombinationstöne“.) Als Ersatz tiefer Orgelpfeifen können ebenfalls zwei höhere (und billigere) zur Schwebung gebracht werden.

In elektronischen Musikinstrumenten der 20er Jahre (Ondes Martenot, **Theremins Ätherophon** [siehe Abbildung 3.13], Magers Sphärophon) wurden zwei Hochfrequenzschwingungen von ca. 100 000 000 Hz (= 100 MHz = UKW-Wellenbereich) erzeugt. Die eine der beiden Schwingungen wurde ganz geringfügig durch Handbewegungen in der Luft verändert, zum Beispiel um 0,001% auf 100,001 MHz. Wurden beide hochfrequente Schwingungen einander überlagert, ergab sich eine Schwebungsfrequenz im hörbaren Bereich. Im Beispiel: 100 001 000 Hz - 100 000 000 Hz = 1000 Hz.

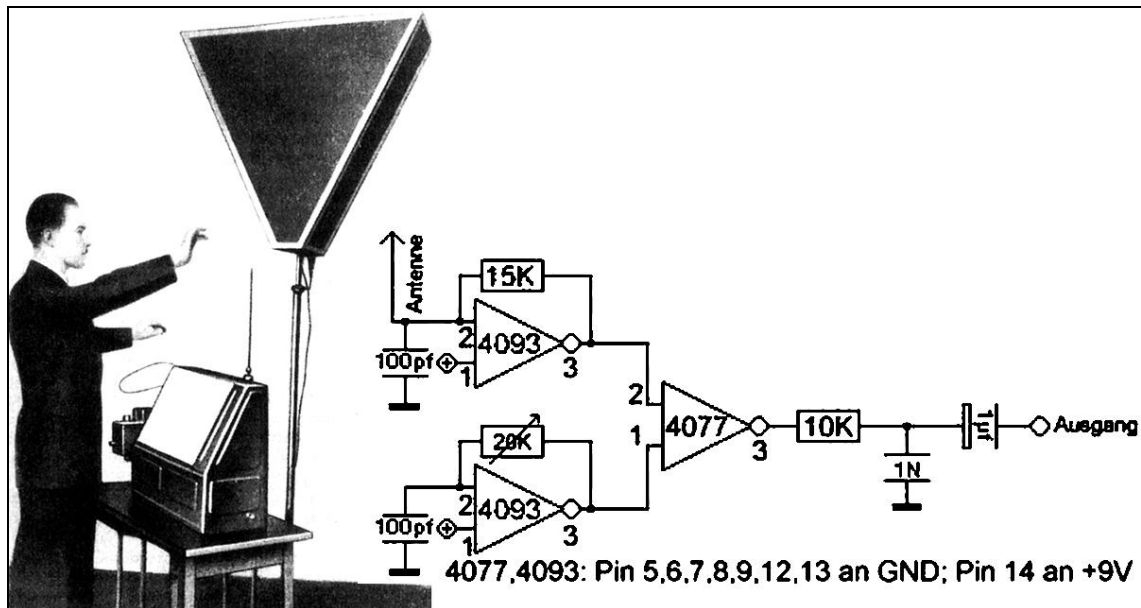


Abb. 3.13 Hinter dem Theremin (links dessen Erfinder) verbirgt sich eine einfache elektronische Schaltung, die man mit wenigen Bauteilen eines Elektronikshops nachahmen kann

Mathematisch sind Schwebungen leicht erklärbar: Addiert man zwei Sinusfunktionen unterschiedlicher Frequenz, so erhält man nach dem „Additionstheorem“ der Trigonometrie eine einzige, amplitudenmodulierte Sinusfunktion.

Die Schwingungen haben die Gleichungen $Y_1 = A \sin(2\pi f_1 t)$ und $Y_2 = A \sin(2\pi f_2 t)$. $Y_1 + Y_2 = A(\sin(2\pi f_1 t) + \sin(2\pi f_2 t)) = 2A \cos(\pi(f_1 - f_2)t) \times \sin(\pi(f_1 + f_2)t)$. [Additionstheorem: $\sin(\alpha) + \sin(\beta) = 2\cos((\alpha - \beta)/2) \times \sin((\alpha + \beta)/2)$.] Dies ist eine Sinusschwingung der Frequenz $(f_1 + f_2)/2$, deren Amplitude mit der Frequenz $(f_1 - f_2)/2$ cosinusförmig schwankt.

Das Oszilloskop gibt in Fall 2 das mathematische „Additionstheorem“ wieder, es macht keinen Unterschied zwischen Schwebungen kleiner oder großer Frequenz. Die Mathematik kann daher nicht erklären, warum der Mensch einmal die Schwebungen mit „Mittelton“, das andere Mal aber zwei getrennte Töne hört. Dies ist Aufgabe der „Hörtheorien“. (Tonbeispiel: [PLAY!](#))

3.4. Die Tonhöhenunterscheidungsfähigkeit des Ohrs

Gleichzeitig erklingende und leicht gegeneinander verstimmte Töne oder Intervalle erkennt man an den Schwebungen. Wie steht es aber mit hintereinander erklingenden Tönen? Zur Feinuntersuchung verwenden wir das Pitch-Bend eines Synthesizers oder einer Soundcard. Pitch-Bend unterteilt den temperierten Ganzton in 64 Schritte. Dies entspricht einem Frequenzverhältnis (Intervall) von ca. 0,3827%. Wir nennen ein solches Intervall im folgenden „1 Bend“.

1 Bend = 0,3827% ist die „Default“- oder Standardeinstellung von GM-Soundcards. Meist lässt sich Pitch-Bend noch feiner (kleiner) einstellen. Die Standardeinstellung rührt daher, dass mit dem Bend-Rad am Keyboard aus der Mittelstellung heraus nach oben und unten ein glissando von maximal einem Ganzton möglich sein soll. Dies reicht für übliche diatonische Skalen ja aus. Der maximale Umfang der halben Radbewegung kann aber in der Regel auf bis zu 2 Oktaven vergrößert (Bend also vergrößert) oder auf einen Halbton verkleinert (Bend also verfeinert) werden. Im letztgenannten Fall wäre 1 Bend = 0,19135%. Hi-End-Cards oder -Synthesizer unterteilen die gesamte Bend-Rad-Bewegung nicht in 128, sondern $128 \times 128 = 16\,384$ Teile. Dann wird 1 Bend = 0,00299%.

Experiment 3.4. Mini-Intervalle

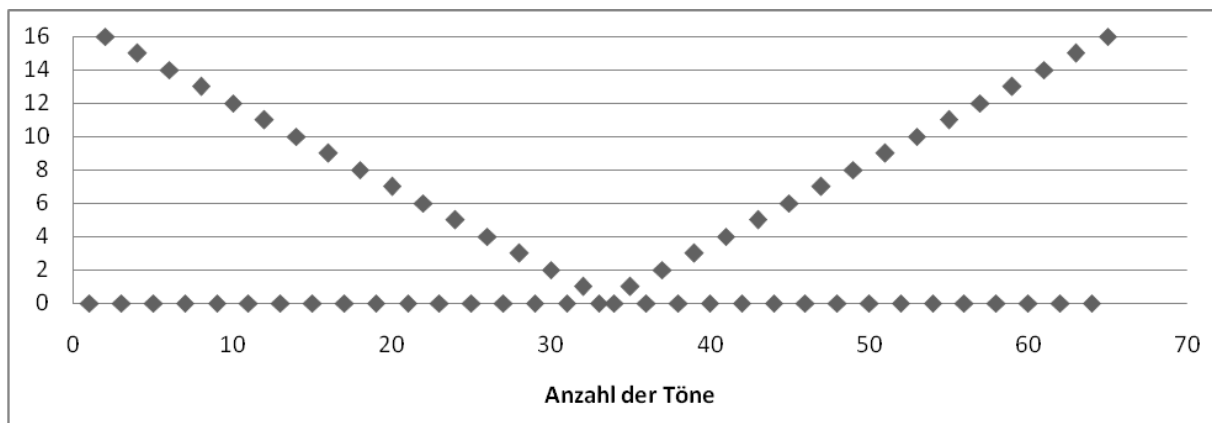
- Wir spielen eine Abfolge von Sinustönen, deren Frequenz sich um 1 Bend voneinander unterscheiden. Nach 64 Tönen haben wir einen Ganzton erreicht ([PLAY!](#)).
- Wir spielen Zweitonintervalle (z.B. Quinten) hintereinander, deren Größe um jeweils 1 Bend zunimmt.
- Wir spielen einen Sinuston, dessen Frequenz um 1, 2, 3 usw. Bends moduliert wird. Diese Frequenzmodulation ist ein echtes „Vibrato“ im Gegensatz zur Amplitudenmodulation der Schwebung. Die Frequenzänderung findet 1 mal pro sec, d.h. mit 1 Hz statt.

Außer im Falle des Vibratos scheinen die Bends unterhalb der Hörbarkeitsgrenze zu liegen. Gefühlsmäßig nimmt man noch ein Ansteigen wahr, rational jedoch lässt sich dies nicht feststellen.

Experiment 3.4. Quantitativer Teil

Teil 1: Es werden hintereinander zwei Instrumentaltöne (z.B. Klavier) gespielt, die um wenige Bends gegeneinander verstimmt sind. Es soll festgestellt werden, ob der 2. Ton tiefer, gleich oder höher als der 1. Ton ist ([PLAY!](#)).

Teil 2: Es wird eine „Pitchbend-Folge“ im Wechsel mit einem konstanten Vergleichston gespielt. Die Tonhöhe der „Pitchbend-Töne“ nimmt von $\frac{1}{4}$ Ton-Intervall kontinuierlich ab und dann wieder zu. In einem gewissen Bereich sind keine Tonhöhenunterschiede mehr zu hören. Da in der Mitte des Experiments die Abfolge „Vergleichston/Pitchbendton“ zu „Pitchbendton/Vergleichston“ vertauscht wird, man aber zunächst hiervon nichts bemerkt, ist der Augenblick, in dem der Hörer das „Umkippen“ der Abfolge wahrnimmt, der Augenblick, ab dem ein Tonhöhenunterschied wahrgenommen wird. ([PLAY!](#))



Das kleinste Intervall, bei dem gerade noch ein Tonhöhenunterschied wahrgenommen wird, heißt „Just Noticeable Difference“ (JND). Für diese Pitch-JND kann festgestellt werden:

- Die Pitch-JND ist von der Tonhöhe und Lautstärke abhängig: sie ist bei ca. 1000 Hz am kleinsten, ist bei 1720 Hz kleiner als bei 440 Hz; sie ist bei lauterem Tönen kleiner als bei leiseren.
- Die Pitch-JND liegt im optimalen 1000 Hz-Bereich bei ca. 0,7%. Bei einigen 100 Versuchen mit MusikstudentInnen hat sich ein Schwellenwert von 2 bis 4 Bend herausgestellt.

Die Tatsache, dass die Pitch-JND bei ca. 0,7% liegt, bedeutet, dass man die Stufen des Pitch-Bend-Rades am Synthesizer bzw. an der Soundcard (0,38%) gerade nicht mehr erkennt, dass also die stufenförmige Bewegung eines Bend-Rades als kontinuierliches glissando erlebt wird. Dies stimmt gut mit den Erfahrungen überein. - Ergebnisse von Untersuchungen nach der „Vibrato“-Methode von Eberhard Zwicker 1982 ergeben ein noch differenzierteres Bild:

- Bis ca. 500 Hz ist die Pitch-JND konstant 3,6 Hz (also nicht als Intervall in Prozent ausdrückbar).
- Ab ca. 1500 Hz ist die Pitch-JND ziemlich konstant 0,71%.
- Bei 1000 Hz ist die Pitch-JND 5 Hz entsprechend 0,5%.

Dies kann bedeuten, dass der Mensch zwei „Strategien“ beim Tonhöhenunterscheiden hat, eine für niedrige, eine für höhere Frequenzen. Diese Arbeitsteilung ist in der Tat neurophysiologisch zu finden. Dennoch stellen die sehr kleinen Pitch-JND-Werte eine Herausforderung an die Hörpsychologie dar: Wie schafft es das kleine und mit Sinneszellen grob bestückte Ohr, solch feine Unterschiede wahrzunehmen?

3.5. Hörtheorien

Die Antwort auf diese Fragen versuchen die sog. „Hörtheorien“ zu geben, die sich seit Mitte des 19. Jahrhunderts immer weiter verfeinert haben. Zunächst zur **„Anatomie“ des Ohres**. Wenn der Schall wie in Abbildung 2.5 dargestellt, aufs Trommelfell trifft, dieses in Bewegung setzt und diese Bewegungen dann über die Knöchelchen ans ovale Fenster des Innenohrs, das wie eine Schnecke aussieht, gelangen, dann ist das Terrain der Hörtheorien erreicht. Das Innenohr ist mit einer Flüssigkeit gefüllt, die wie jede Flüssigkeit (Stichwort hydraulisches Bremssystem beim Auto) weitgehend nicht komprimierbar ist. Das ist ein großer Unterschied zur Luft!

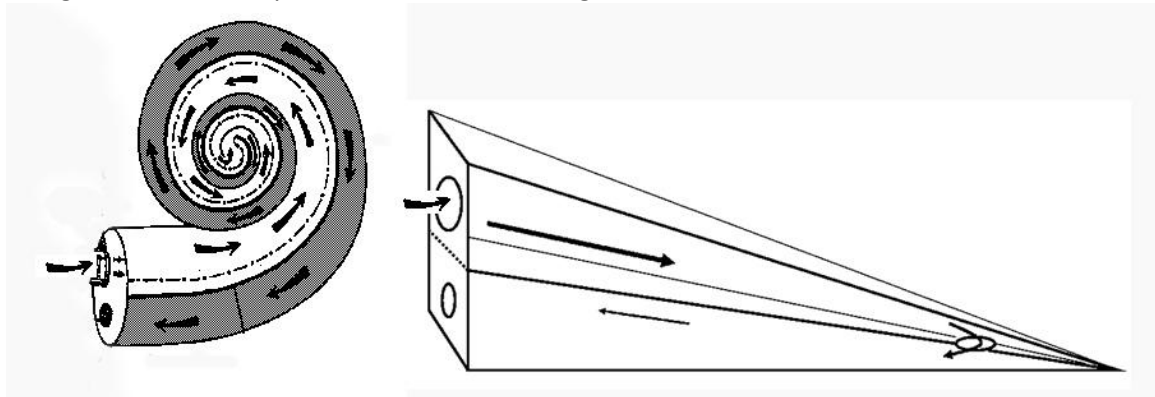


Abb. 3.14 Innenohr schematisch (eingerollt und aufgerollt)

Die Länge der (aufgerollten) Schnecke ist ungefähr 3,2 cm. In der Mitte ist sie unterteilt durch das Corti'sche Organ, das aus diversen Membranen besteht, zwischen denen sich Härchen unterschiedlicher Stärke befinden, an denen die Hörnerven sitzen. Mechanisches Basis des Hörvorgangs ist, dass die Flüssigkeit im Innenohr bewegt wird (Pfeilrichtung) und dabei die Härchen „irgendwie“ in Bewegung geraten, was die Nerven reizt. Über die Details dieses Vorganges gibt es folgende Theorien:

1. Orts- und Wanderwellentheorie im Innenohr

Nach der Ortstheorie ist die Frequenzbestimmung eine Frage des Orts maximaler Erregung der Basilarmembran. Bei der Wanderwellentheorie kommen weitere Faktoren hinzu, die eine raffinierte (aber erklärbare) Arbeitsweise des Gehörnervensystems implizieren. Der eintreffende Schall löst eine Wanderwelle aus, die ein raum-zeitliches Erregungsmuster entlang der Basilarmembran erzeugt und dabei die Haare (unterschiedlich stark und zeitlich versetzt) bewegt. Diese Wanderwelle gelangt nicht bis ans Ende des Innenohrs, weil sie unterwegs „absorbiert“ wird. Folgende Parameter werden von den Sinneszellen „abgegriffen“ und zwecks Weiterverarbeitung ans Nervensystem geleitet:

- Ort der maximalen Erregung der Basilarmembran (Ortstheorie nach **Helmholtz**).
- Ort, an dem die Wanderwelle durch Absorption zum Erliegen kommt (Ortstheorie nach **Békésy**).

- Wellengeschwindigkeit entlang der Basilarmembran (Sinneszellen werden hintereinander gereizt). Physikalisch: Wellengeschwindigkeit ist frequenzabhängig.
- Erregungsstärke und Größe der erregten Fläche (ist nicht nur von der Lautstärke abhängig, sondern z.B. auch vom Absorptionsvermögen der Innenohrflüssigkeit). Physikalisch: Absorption ist frequenzabhängig.

Die Orts- und Wanderwellentheorie wird durch folgende Modell-Bilder veranschaulicht:

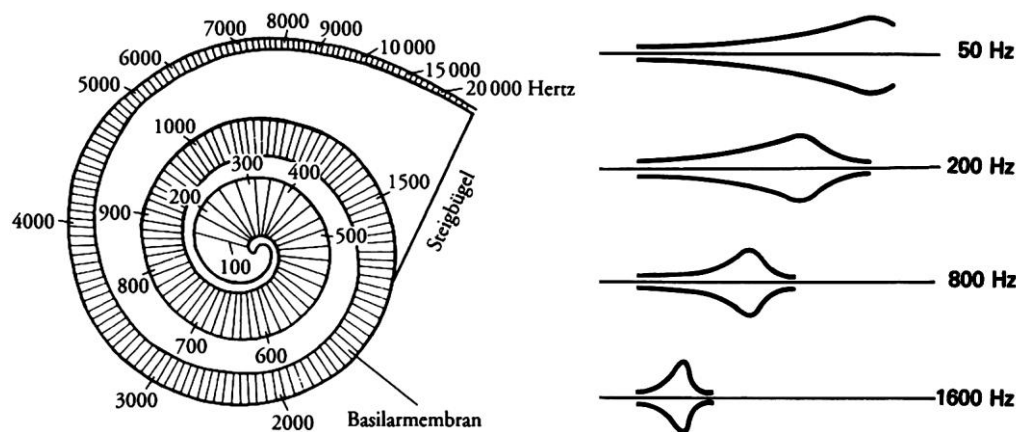


Abb. 3.15 Veranschaulichung der Orts- und Wanderwellentheorie

2. Kodierung durch verschiedene Neuronen

Physiologische Basis: In den Neuronen können Impulse unterschiedlicher Frequenz abgefeuert werden. Neben einer statistischen Grunderregung gibt es signifikante Impuls-Frequenzen. Folgende Parameter könnten eine Information über Frequenzen ans Gehirn weiterleiten:

- Änderung der Neuronenimpulsfrequenz („**Frequenztheorie**“).
- Arbeitsteilung mehrerer Neuronen bei der Kodierung höherfrequenter Schwingungen („**Volley-Prinzip**“).
- Anzahl und Ort der erregten Neuronen/ der nicht erregten Neuronen („**Ortstheorie**“).
- Faserlänge („Empfindlichkeit“) oder frequenzabhängige Reizschwellenwerte der Neuronen (sog. „**Bestschalltheorie**“).

Die beiden Theorien 1 und 2 können „arbeitsteilig“ eingesetzt werden. Für niedrige Frequenzen kann eher Theorie 2, für hohe Frequenzen Theorie 1 eingesetzt werden.

3. Resonanz und Beweglichkeit der Haarzellen

In jüngster Zeit wird eine Theorie diskutiert, wonach die Haarzellen je nach Beweglichkeit in Eigenschwingung (Resonanz) geraten und dadurch auf die Basilarmembran und deren Schwingung rückwirken. Die Beweglichkeit der Haarzellen und damit die Resonanzfähigkeit ist vom psychischen Zustand (Stress!) von endogenen oder exogenen Drogen, von der akustischen Belastung usw. abhängig. Die Resonanzfrequenz ist sehr genau, während der Ort maximaler Erregung ungenau ist.

4. Verarbeitung durch die synaptische Struktur der Nervenleitung

Die Nervenfasern (Neuronen) sind an „Synapsen“ miteinander verbunden. Diese Synapsen sind als logische Netze konstruiert und arbeiten (via Botenstoffe, Transmitter) chemisch abhängig. Es wird festgestellt, dass synaptische Schaltungen beispielsweise den Ort der Basilarmembran „fokussieren“

und damit die Frequenzbestimmung präzisieren. Die synaptische Struktur der Nerven zwischen Innenohr und Gehirn ist in mehreren Stufen organisiert. So ist zum Beispiel jene Stufe wichtig, auf der die Nerven der beiden Ohren miteinander verbunden sind, weil frühestens hier ein Richtungshören möglich ist.

Die hier zitierten Hörtheorien geben auf die Fragen nach dem Schwebungsphänomen (Abschnitt 3.3) und nach dem guten Auflösungsvermögen (Pitch-JND) teilweise Auskunft. Wir nehmen noch ein weiteres Phänomen hinzu, das das Ohr auch qualitativ vom Auge unterscheidet: das (geschulte) Ohr kann sehr genau Intervalle hören, also Frequenzverhältnisse „messen“, während das Auge nicht in der Lage ist, Längen- oder Raumverhältnisse exakt zu messen. Alle Hörtheorien, die Aussagen darüber machen, dass eine frequenzabhängige Verarbeitung oder Zubereitung des akustischen Signals im Ohr stattfindet, helfen zu erklären, warum ein zusammengesetzter Klang im Ohr quasi wieder zerlegt werden kann. Für die Ortstheorien spricht, dass dabei Phasenunterschiede keine Rolle spielen. Das Schwebungsphänomen besagt, dass erst ab ca. 10 Hz unterschiedliche und voneinander unabhängige Innenohrbereiche Frequenzunterschiede registrieren können.

Das gute Auflösungsvermögen des Ohrs ist durch die Menge der Sinneszellen und die Mechanik des Innenohrs nicht erklärbar, denn es gibt 12 bis 19 000 „innere“ und 3 500 „äußere“ Härchen. In dem 2-Oktavenraum zwischen a' und a^3 gibt es ca. 200 gerade noch unterscheidbare Tonhöhenstufen (JND's) und etwa 1350 Härchen, sodass für jede JND ungefähr 7 Härchen „zuständig“ wären – ein absolut undenkbares Phänomen bei einem derart „wogenden“ Haarenfeld. Erst durch die „Frequenztheorie“, die „Resonanztheorie“ und/oder die synaptische Präzisierung des akustischen Signals wird diese Fähigkeit tendenziell erklärbar.

3.6. Der Verdeckungseffekt

Der gewisse Grundkonsens der Hörtheorien kann einige Effekte erklären, die für die Musik wichtig sind. Dazu gehört der Verdeckungseffekt, wonach man ein Schallsignal dann nicht mehr hören kann, wenn ein anderes intensiver ist und „in seiner Nähe“ liegt.

Experiment 3.5. Verdeckungseffekt

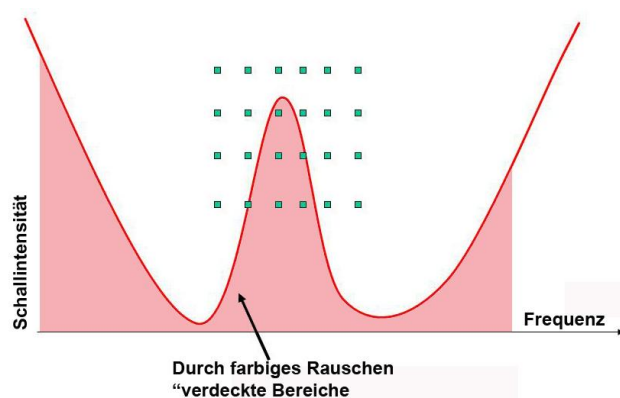


Abb. 3.16 Experiment zum Verdeckungseffekt

Ein Synthesizer erzeugt ein schmalbandiges Rauschen im Bereich um 1000 Hz (grau schraffiert in der Abbildung). Es werden Sinustöne von gleicher Schallintensität und unterschiedlicher Frequenz gespielt. In Abbildung 3.15 stehen die kleinen Quadrate für je einen Sinuston. - Ergebnis des Hörexperimentes ist, dass die Sinustöne innerhalb des Bereichs des Bandrauschens nicht hörbar sind. Das Rauschen „verdeckt“ den Sinuston. Erklärt wird der Effekt dadurch, dass das „farbige Rauschen“ (= des verdeckende Schallsignal) jenen Bereich des Innenohrs betäubt („beschäftigt“), in dem

der Sinuston (= das verdeckte Schallsignal) gehört werden müsste. ([PLAY!](#))

Der Verdeckungseffekt ist in der musikalischen Praxis nicht unwichtig:

Spielen **zwei Musikinstrumente** gleichzeitig, so kann das eine das andere so verdecken, dass letzteres nicht hörbar ist. Verdeckt wird dann, wenn beide Instrumente im gleichen Frequenzbereich operieren, soll andererseits eine SängerIn ein Opernorchester übertönen, so muss die Stimme der SängerIn in einem Frequenzbereich „stark“ sein, in dem das Orchester „schwach“ ist (Abbildung

3.17). Ist dies nicht der Fall, ist die SängerIn unhörbar. In der Tat trainieren OpernsängerInnen sog. **Formantbereiche**, die bei den Orchesterinstrumenten nur schwach ausgeprägt sind. Dasselbe gilt für Soloinstrumente in einem Sinfonieorchester. (Tonbeispiel: ein Mann und eine Frau „joiken“ bei starker Meeresbrandung. [PLAY!](#))

Bei der Datenreduktion durch „mp3“ wird der Verdeckungseffekt ebenfalls ausgenutzt. Kurz gesagt: alles, was ohnedies nicht gehört weil verdeckt wird, wird einfach weggelassen. (Die Kunst ist aber, wie dies technisch geschehen soll!) Siehe hierzu ausführlich Kapitel 4.3!

Während der Verdeckungseffekt durch die Ortstheorie des Hörens erklärbar ist und offensichtlich ein Effekt des Innenohrs ist, ist der **Cocktailparty-Effekt** eine Leistung des Gehirns: sprechen zwei Menschen gleichzeitig etwa gleich laut, so kann sich der Hörer auf einen der beiden konzentrieren, verliert aber bei dem, was der andere sagt, den Faden. ([PLAY!](#))

Der Pavarotti-Effekt

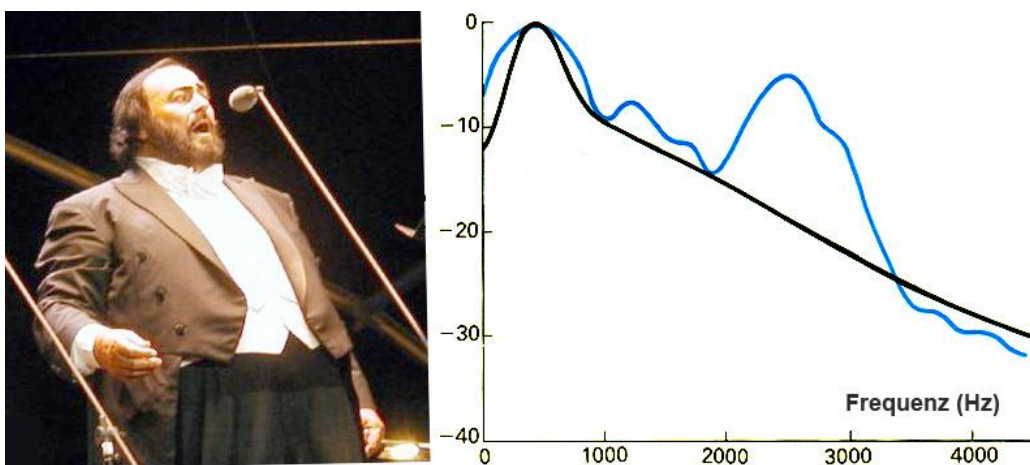


Abb. 3.17 Der Pavarotti-Effekt (nach John R. Pierce): Was bedeuten die Kurven im rechten Diagramm? Welche Art Stimmbildung hat Pavarotti genossen?