

Seminar Sehen und Hören
Abteilung Neuroinformatik
Universität Ulm

Tonhöhenwahrnehmung

Zeitlicher Abriss der Forschung

Ausarbeitung:
Christian Lichtblau

Seminarleitung:
Prof. Heiko Neumann
Prof. Günther Palm
Marcus Borst

Inhaltsverzeichnis

1. Einleitung

1.1. Aufbau des Ohrs

1.2. Erste Experimente zur Tonhöhenwahrnehmung

2. Die fehlende Grundfrequenz, „The missing fundamental“,
und die Ortstheorie

3. Shoutens Residuum Theorie

4. Aktuelle Erkenntnisse zur Tonhöhenwahrnehmung

5. Moderne Theorien zur Tonhöhenwahrnehmung

6. Quellenverzeichnis

1. Einleitung

Ein akustisches System besteht immer aus mehreren Komponenten. Eine Komponente die das akustische Signal erzeugt, indem zum Beispiel eine Saite eines Musikinstruments in Schwingung versetzt wird, eine welche den Schall überträgt und letztlich eine Komponente, welche die ankommenden Schallwellen aufnimmt und verarbeitet.

Eine Schallwelle, eine periodische Kompression von Luft, Wasser oder einem anderen Medium, wird in zwei Bestandteile zerlegt. Der erste Bestandteil ist die Lautstärke oder Intensität, die durch die Amplitude der Schallwelle festgelegt wird. Eine hohe Kompression des Übertragungsmediums erzeugt eine hohe Amplitude und somit einen lauten Ton. Der zweite Bestandteil ist die Frequenz der Schallwelle und damit verbunden die wahrgenommene Tonhöhe. Die Frequenz, mit der Maßeinheit Hertz, bestimmt die Anzahl der Kompressionen in einer Sekunde. Je höher die Anzahl desto höher der erzeugte Ton. Ein erwachsener Mensch kann Frequenzen im zwischen 15/20 Hz bis 20.000 Hz wahrnehmen, wobei die Wahrnehmung von höheren Tonlagen mit dem Alter und der Beanspruchung abnehmen.

Im weiteren wird die Aufmerksamkeit auf der Komponente liegen, welche die Schallwellen aufnimmt und weiterverarbeitet, dem Ohr.

1.1 Aufbau des Ohrs

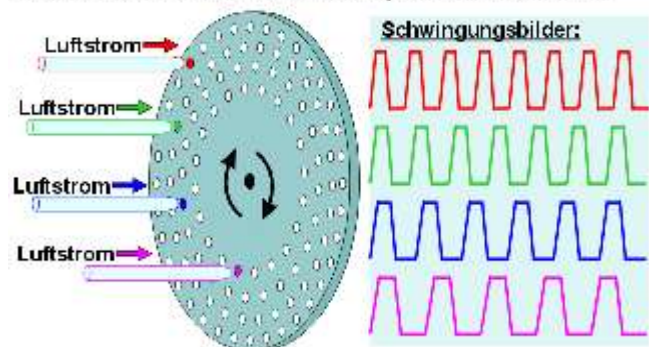
Das Ohr besteht aus drei Bereichen, dem Außenohr, dem Mittelohr und dem Innenohr. Das Außenohr mit der Ohrmuschel nimmt die Schallwellen auf und leitet diese über den Gehörgang an das Trommelfell, einer straff gespannten Membran, welches durch die Druckwelle in der selben Frequenz zum Schwingen gebracht wird. Mit dem Trommelfell sind drei Gehörknöchel verbunden, den Teilen des Mittelohrs. Die Schwingungen werden durch die Knöchelchen verstärkt und an das sogenannte ovale Fenster weitergegeben, hinter dem sich eine Flüssigkeit befindet, die schwerer in Schwingung zu versetzen ist als Luft. Durch die Verstärkung wird der Verlust von ca. 30 dB, der durch das trägere Medium entsteht, kompensiert. Im Innenohr, dem dritten Bereich des Ohrs, befindet sich die für das Hören wichtige Schnecke, oder auch Cochlea. In der Cochlea befindet sich die Basilarmembran, auf der das Cortische Organ sitzt, auf dem wiederum kleine Haarzellen verankert sind. Wird nun eine Schwingung auf diese Membran übertragen, entsteht abhängig von der Höhe der Töne an einer bestimmten Stelle der Basilarmembran eine Welle, welche die Haare der Haarzellen biegt und somit einen Impuls auslöst, der über den Hörnerv an das Gehirn weitergegeben wird.

1.2 Erste Experimente zur Tonhöhenwahrnehmung

Bereits im 6. Jahrhundert vor Christus stellte Pythagoras fest, dass sich die Tonhöhe genau um eine Oktave erhöht, wenn man die Saite halbiert, mit der dieser Ton erzeugt wurde. Darauf folgerte er, dass die kürzere Saite doppelt so schnell schwingen muss wie die längere. Galileo Galilei beschrieb 1640 schwingende Körper und schlug damals vor anzunehmen, dass die Tonhöhe mit der Zahl der Schwingungen pro Sekunde, also der Frequenz zusammenhängt, womit er aus heutiger Sicht nicht unrecht hatte, sich aber bestimmt nicht im klaren war wie komplex dieser Zusammenhang ist.

Lochsirene:

Der Luftstrom eines Rohres wird durch die rotierende Lochscheibe immer wieder unterbrochen. Die Luftimpulse werden zu Schall.



Die Tonhöhe hängt von der Anzahl der Löcher pro Sekunde ab.

Abbildung 1: Schematische Darstellung einer Lochsirene

der Scheibe immer wieder unterbrochen werden. Die Auftretenden Luftstöße werden als Ton mit einer definierten Tonhöhe wahrgenommen. Die Geschwindigkeit, mit der sich die Scheibe dreht, sowie der Abstand der Löcher auf der Scheibe bestimmen die daraus resultierende Tonhöhe.

Rayleigh veränderte die Periodendauer, indem er jedes zweite Loch auf der Scheibe verschob. Durch diese Verschiebung verdoppelte sich die Periodizität des Signals, woraus Rayleigh folgerte, dass die wahrgenommene Tonhöhe nicht von der Periodizität abhängt, denn die Verschiebung kann so gering gehalten werden, so dass in der Tonhöhe kein Unterschied wahrnehmbar ist. Dieses Experiment machte damals schon deutlich, dass die Tonhöhenwahrnehmung ein Vorgang hoher Komplexität ist und nicht wie vorher angenommen, dass die Tonhöhe lediglich von der Periodendauer abhängt.

Lord Rayleigh erkannte in seinem Experiment mit einer Lochsirene, dass die Tonhöhe nicht ausschließlich von der Periodendauer abhängt. Bei seinem Experiment erstellte er eine Scheibe, auf der Löcher in konzentrischen Kreisen angebracht sind. Wird die Scheibe nun in Rotation versetzt und durch einen Luftstrom angeblasen entstehen Luftstöße, welche durch die Anordnung der Löcher in

2. Die fehlende Grundfrequenz, „The missing fundamental“

Unser Gehör nutzt die Information aus dem Zusammenspiel der einzelnen harmonischen Teiltöne eines Klanggemisches aus, um die Tonhöhe zu bestimmen. Interessanterweise muss diese Grundfrequenz, die für die Tonhöhenempfindung verantwortlich ist, im Klanggemisch selbst nicht vorhanden sein, um diesen Eindruck zu vermitteln. Das Klanggemisch selbst muss nicht ein harmonischer Ton sein, wie er bei Musikinstrumenten vorkommt, sondern kann auch ein komplexes periodisches Signal sein, das physikalisch keine Energie an der Position des fehlenden Grundtons besitzt.

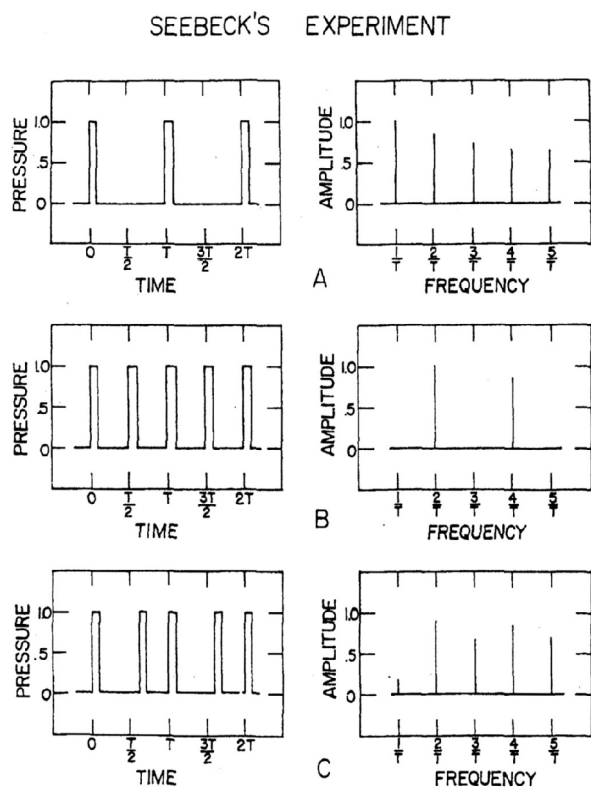


Abbildung 2: Durch eine Lochsirene erzeugter Reiz. Links die ideale Wellenform; rechts die spektrale Repräsentation.

regelmäßigen Abständen platziert sind. Das Ergebnis dieses Signals ist eine Tonhöhe mit der Frequenz $2/T$, woraus folgt, dass der Zeitabstand zwischen den Luftstößen um den Faktor zwei halbiert wurde und die resultierende Tonhöhe exakt eine Oktave höher ist, als die Tonhöhe aus dem Signal A. Für die Wellenform des Signals C wurde jedes zweite Loch der Scheibe aus B um einen festen Wert verschoben, so dass die Intervalle der Luftstöße sich regelmäßig abwechseln. (Langes Intervall, kurzes Intervall, ...) Das Ergebnis dieser Wellenform ist, dass eine sehr geringe Energie bei der Frequenz $1/T$ auftritt und die Hauptenergie bei der Frequenz $2/T$, woraus man schließen könnte, dass die Wellenform der Signale B und C die selbe oder eine ähnliche Tonhöhe besitzen. Dem ist allerdings nicht so. Die Signale A und C besitzen die selbe Tonhöhe, aber mit

Seebeck [1] war der erste, der dieses Phänomen in seinem Experiment 1841 entdeckte, bei dem er die Tatsache, dass die Tonhöhe nicht von der Periodendauer abhängt, genauer untersuchte. Dazu benutzte er ebenso wie Rayleigh eine Lochsirene, um Klänge zu erzeugen und zu kontrollieren. In Abbildung 2 A ist die Wellenform von Seebecks erster Sirene dargestellt. Für dieses Signal benutzte er eine Sirene mit regelmäßigen Lochabständen. Die Tonhöhe dieses Signal ist der Grundton mit der Frequenz $1/T$ und ist im ursprünglichem Signal enthalten. Für das zweite Signal wurde eine Lochsirene mit doppelt so vielen Löchern verwendet, die wiederum in

unterschiedlicher Qualität der Töne. Da zu dieser Zeit keine befriedigende Erklärung für dieses Phänomen aufgebracht werden konnte, wurde angenommen, dass es sich hierbei, ähnlich wie bei der optischen Täuschung, um eine akustische Täuschung handeln müsse.

Zwanzig Jahre später schlug Helmholtz [2][3] eine physiologische Erklärung der Fourieranalyse von Schallwellen vor. Dabei nahm er an, dass die Basilarmembran in der Cochlea aus einer Reihe aneinanderhängenden gespannten Saiten besteht, ähnlich wie bei einer Harfe. Er ging davon aus, dass die einzelnen Fasern durch Länge und Spannung eine unterschiedliche Resonanzfrequenz besitzen. Bringt nun eine komplexe Schallwelle die Membran zum Schwingen, so werden nur die Bereiche der Basilarmembran angeregt, die in der selben Frequenz gestimmt sind. Somit wird die Schallwelle in ihre einzelnen Sinusbestandteile zerlegt. Durch diese Theorie wird allerdings noch nicht die akustische Täuschung erklärt. Dafür stellte Helmholtz die Hypothese der nichtlinearen Störung oder Verzerrung im Mittelohr auf. Helmholtz nahm an, dass die Übertragung der Schwingungen der Gehörknöchelchen auf die Cochlea ein nichtlinearer Prozess sei. Diese Schwingungen, die Helmholtz vorschlug und auch gemessen konnte, nennt man Kombinationstöne. Durch die Verzerrung entstehen also neue Frequenzkomponenten im Signal, welche genauso wie die ursprünglichen Frequenzen von der Cochlea wahrgenommen werden. Für nicht komplexe Töne, also harmonische Töne, erzeugt diese nichtlineare Verzerrung Frequenzkomponenten, die den harmonischen Obertönen der Grundfrequenz entsprechen. Bei komplexen Tönen, wie sie Seebeck mit seiner Lochsirene erzeugt hat, würde demnach eine Störfrequenz auftreten, die der Differenz zweier nebeneinander liegenden Frequenzen entspräche. Diese Kombinationstöne werden auch quadratische Komponenten der nichtlinearen Verzerrung genannt. In Seebecks Fall wäre dies die Grundfrequenz $1/T$, da der Abstand je zweier Frequenzen $1/T$ ist, welche auch für die Wahrnehmung der Tonhöhe verantwortlich ist. Die Energie, die dabei zustande kommt, könne, so Helmholtz, zur Energie der Grundfrequenz hinzuaddiert gesehen werden. Da in Seebecks Experiment die meisten spektralen Komponenten den Abstand $1/T$ haben, entsteht eine starke Komponente an dieser Stelle, wodurch die eigentlich nicht vorhandene Grundfrequenz stark hervorgehoben wird. Diese Theorie, die sowohl Helmholtz, Seebeck und auch Ohm unterstützten wird heute auch Ortstheorie oder „place theory“ genannt.

Doch nicht nur bei komplexen Schallwellen tritt das Phänomen des fehlenden Grundtones auf. Auch bei der menschlichen Stimme muss der Grundton nicht vorhanden sein, die Tonhöhe der Stimme festzustellen. Diese Erkenntnis machte Harvey Fletscher, der aus einem Signal sämtliche tiefen Frequenzen herausfilterte, und trotzdem immer noch die ursprüngliche Tonhöhe wahrnehmen konnte. Auch er konnte für das Phänomen der Fehlenden Grundfrequenz keine Erklärung finden. Als Beispiel für dieses Phänomen ist das Telefon aufzuführen, bei dem nur Frequenzen oberhalb von 300 Hz übertragen werden. Der hörbare Bereich des Menschen liegt allerdings weitaus niedriger. Und trotzdem kann man ohne Probleme eine Grundfrequenz von 100 bis 200 Hz einer Stimme wahrnehmen, obwohl diese Frequenz nicht übertragen wird.

3. Schoutens Residuum Theorie

Schouten [4][5] führte eine ganze Reihe von Experimenten durch, deren Ergebnisse immer größere Zweifel an der Theorie von Helmholtz aufwarfen. Zum einen ist die Tonhöhe der fehlenden Grundfrequenz selbst dann noch existent, wenn die Energie des Reizes so gering ist, dass die Theorie der Verzerrung im Mittelohr diesen Reiz nicht verursachen kann. Zum anderen konnte Schouten keinen Rhythmus oder Schlag feststellen, wenn er einen Sinuston in die Nähe der fehlenden Frequenz platzierte. Daraus folgerte Schouten, dass es höchst unwahrscheinlich sei, dass die Hypothese der nichtlinearen Verzerrung die richtige Erklärung für dieses Phänomen darstellt. Schouten selbst nennt das Phänomen, der Eindruck der fehlenden Grundfrequenz, eine periodische Tonhöhe, oder auch Resttonhöhe (Residuum).

Tonhöhenverschiebung

Eines der elegantesten Experimente, die Schouten in diesem Zusammenhang durchführte, war das Experiment zur Tonhöhenverschiebung. Das Experiment beinhaltet einem Ton, der aus sechs Reizkomponenten besteht, zum Beispiel 1000 Hz, 1200 Hz, 1400 Hz, 1600 Hz, 1800 Hz, und 2000 Hz. Um nun die Grundfrequenz dieses Tones zu bestimmen kann man sowohl die Differenz zweier nebeneinander liegender Komponenten subtrahieren, was eine Grundfrequenz von 200 Hz ergäbe, als auch die Periodizität der Wellenform dieses Signals verwenden, das ebenfalls 200 Hz beträgt. Als zweites Signal verwendete Schouten eine Folge von Komponenten, die exakt den selben Abstand zwischen den Komponenten besitzt, wie das erste Signal, allerdings um 50 Hz verschoben ist. Diese Wellenform hat eine Grundfrequenz von etwa 210 Hz, welche mit der Differenz nicht übereinstimmt.

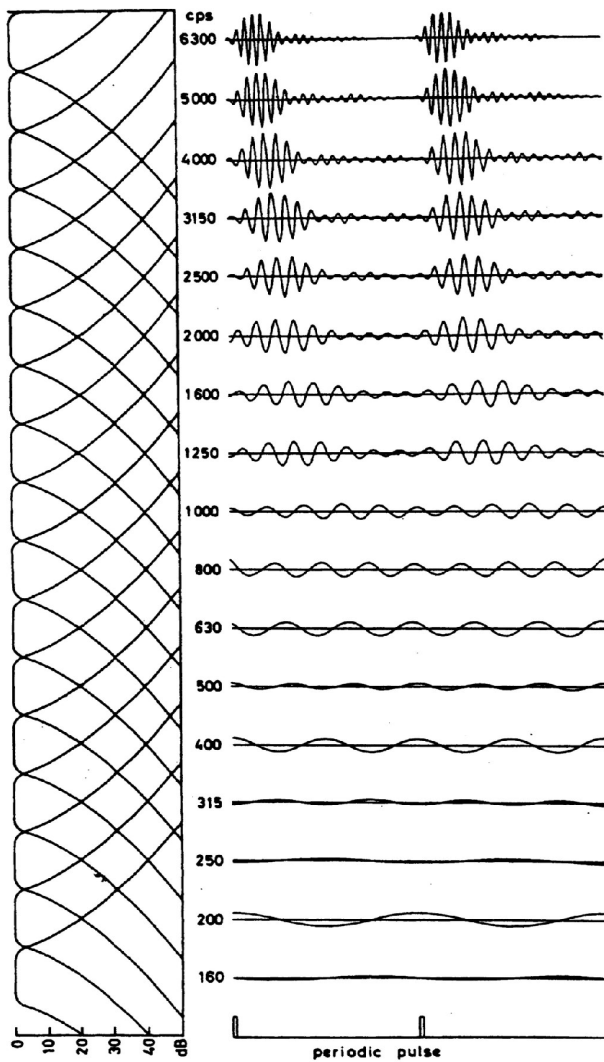


Abbildung 3: Die erste Stufe der auditiven Filterung. Links die Charakteristik des Bandfilters; rechts die Wellenform der einzelnen Komponenten.

In einem weiterem Experiment stellte Schouten folgende Überlegung an: Die fehlende Grundfrequenz müsse sich genauso verhalten, wie die harmonischer Töne derselben Frequenz, wenn diese aus Nichtlinearitäten besteht. Diese Überlegung legt zu Grunde, dass Kombinationstöne, die Helmholtz vorgeschlagen hatte und auch messen konnte, tatsächlich physikalisch vorhanden sind. Schouten wählte wiederum ein Signal, das eine Periodendauer von 200 Hz besitzt, bei dem die Grundfrequenz von 200 Hz herausgefiltert wurde. Nach dem bisherigen Stand ist klar, dass die wahrgenommene Tonhöhe wiederum der 200 Hz Frequenz entspricht. Zu diesem Signal addierte Schouten eine Frequenz von 206 Hz. Die Idee war, dass wenn eine nichtlineare Verzerrung die Ursache für die wahrgenommene Tonhöhe von 200 Hz ist, dann müsste durch die Addition von 206 Hz eine Indifferenz oder Schwebung von 6 Hz messbar sein. Allerdings konnte Schouten

in keinem seiner Experimente eine solche feststellen.

Sowohl das Ansteigen der Tonhöhe bei der Verschiebung als auch das fehlen einer Schwebung beim Addieren einer Frequenz nahe bei der fehlenden Grundfrequenz lässt für Schouten keinen Zweifel über, dass die Theorie der nichtlinearen Verzerrung von Helmholtz falsch ist. Schouten schlägt seinerseits vor, dass die Tonhöhe aus Epochen der Wellenform bestimmt wird und dass die Periodizität nur dort auftritt, wo der Cochlea Filter es einzelnen Komponenten erlaubt, miteinander zu interagieren.

Schoutens Modell der Tonhöhenwahrnehmung

Schouten schlägt ein Modell der temporalen Theorie vor, das ähnlich wie das Ohr funktioniert und aus zwei Stufen besteht. Als erste Stufe seines Modells verwendet Schouten eine Reihe von Bandpassfiltern, wie sie in Abbildung 3 dargestellt sind, um das Eingangssignal in seine Frequenzkomponenten zu zerlegen. Auf der linken Seite der Abbildung ist die Charakteristik des Bandpassfilters dargestellt, auf der rechten Seite die

Ausgabe die ein solcher Filter erzeugt, wenn er als Eingangssignal einen periodischen Pulse bekommt. Was bei dieser Charakteristik des Bandpassfilters nicht so stark auffällt, ist die zunehmende Bandbreite der einzelnen Filter von den tiefen zu den hohen Frequenzen, entsprechend der Frequenzauflösung der Basilarmembran. Bei den tiefen Bandpassfiltern kann man erkennen, dass das Ergebnis ein Sinuston mit einer einzigen Komponente ist, wie zum Beispiel bei der 400 Hz Komponente. Bei den höheren Frequenzen kann man das Eingangssignal, ein periodisches Pulssignal, sehr deutlich erkennen. Die zweite Stufe besteht aus einem neuronalen Übertragungsmechanismus, wie ihn Schouten nennt. Der Mechanismus funktioniert so, dass die zeitliche Struktur einer Wellenform im Ausgang jedes Filters aus der ersten Stufe im zeitlichen Muster der Nervenfasern erhalten bleibt und somit an höhere Verarbeitungszentren weitergegeben wird.

Probleme bei Schoutens Modell

Das generelle Konzept der Residuums Theorie wurde allgemein akzeptiert, doch kamen immer wieder Versuche zum Vorschein, die mit diesem Modell nicht erklärt werden konnten. Ein Schwachpunkt stellt die These dar, dass der Tonhöhereindruck des Residuums durch die hohen Frequenzbereiche entsteht, was ebenfalls aus der Abbildung 3 ersichtlich ist, da sich die Periodizität in diesen Komponenten am stärksten widerspiegelt. Demzufolge muss dieser hochfrequente Bereich auch hauptsächlich für die Intensität dieses Eindrucks verantwortlich sein, wohingegen niederfrequente Komponenten weniger starken Einfluss nehmen. Diese logische Folgerung kann experimentell widerlegt werden.

Ritsma [6] führte ein Experiment durch, bei dem er feststellen wollte welche Frequenzbereiche oder Komponenten für die Erzeugung der Grundfrequenz verantwortlich sind. Er erzeugte ein Signal das aus zwei Komponentenbereichen bestand. Ritsma stellte durch die Verschiebung dieser Komponenten fest, dass nicht wie zuvor vermutet die hohen Frequenzbereiche die Grundfrequenz beeinflussen, sondern ein Frequenzband, das zwischen dem dritten und fünften Oberton der Grundfrequenz liegt. Für eine Grundfrequenz von 200 Hz ist beispielsweise ein Frequenzband bei etwa 600 bis 1000 Hz verantwortlich. Ritsma nannte demnach diese Region auch spektrale Dominanz.

Ein weiteres Problem der temporalen Theorie stellt die Phase der einzelnen Komponenten dar. Eine Phasenverschiebung von einzelnen Komponenten verändert die Wellenform sehr stark, die aus der ersten Stufen resultiert, so dass die Periodizität des Ausgangssignal der hohen Komponenten nicht mehr festgestellt werden kann. Die Phasenverschiebung bewirkt allerdings keine Veränderung der wahrgenommenen Tonhöhe, sondern nur eine Veränderung der Tonqualität. Diese beiden Erkenntnisse,

einerseits, dass die dominante Region für eine periodische Tonhöhe in den niedrigen Frequenzbereichen liegt, und andererseits, dass die Phasenverschiebung eines komplexen Signals die Tonhöhe nicht verändert, führten dazu, dass die temporale Theorie einen ernüchternden Rückschlag einstecken musste.

4. Aktuelle Erkenntnisse zur Tonhöhenwahrnehmung

Der Zentrale Ursprung der Tonhöhe

Bisher wurde bei den Experimenten davon ausgegangen, dass alle Frequenzkomponenten in einem Ohr verarbeitet werden müssen, um den Eindruck der Tonhöhe zu erzeugen. Houtsma und Goldstein [7] führten ein Experiment durch, bei dem sie einen Ton erzeugten, der aus zwei Komponenten bestand. Die beiden Komponenten wurden Testpersonen getrennt auf beide Ohren vorgespielt. Zum Beispiel wurde im linken Ohr eine Frequenz von 1000 Hz und auf dem rechten Ohr 1200 Hz eingespielt. Das Ergebnis war eine zwar schwer wahrnehmbare aber vorhandene Tonhöhe von 200 Hz. Des weiteren wurde die Testpersonen gebeten das Intervall zweier Töne zu identifizieren, die ebenfalls getrennt vorgespielt wurden. Auch hierbei ergab das Experiment, dass eine musikalisch geschulte Testperson keine Schwierigkeiten hatte, dieses Intervall zu bestimmen. Daraus schlossen Houtsma und Goldstein, dass der Mechanismus, der für die Beurteilung zweier getrennter Signale in je einem Ohr verantwortlich ist, an einer zentraleren Stelle liegen muss und nicht irgendein peripherer Analysator ist. Dieses Experiment zeigt ebenfalls, dass die Verzerrung nicht für die Tonhöhe verantwortlich sein kann, da ja nur ein Ton in je ein Ohr vorhanden war.

Verzerrungsprodukte

Wie bereits Helmholtz festgestellt hatte, existieren gewisse nichtlineare Verzerrungen im Ohr. Allerdings sind diese Verzerrungen nicht für die Wahrnehmung der fehlenden Grundtonhöhe verantwortlich. Helmholtz fand in seinem Experiment die sogenannte quadratische Verzerrung als Differenzton zweier nebeneinanderliegenden Komponenten. Dieser Differenzton ist sowohl für moderate als auch für extrem hohe Intensität der Frequenzen unauffällig. Wo hingegen die Kubische Verzerrung sehr von Bedeutung ist. Aus zwei gegebenen Hauptfrequenzen f_1 und f_2 berechnet sich die Verzerrungsfrequenz aus $2 * f_1 - f_2$. Zum Beispiel, wenn zwei Komponenten mit 1000 Hz und 1200 Hz gegeben sind, dann entsteht im Ohr eine starke Verzerrung von 800 Hz. Dabei ist das herausragende dieser Verzerrung, dass sie in ihrer Lautstärke lediglich 20 – 30 dB unter der Lautstärke der Hauptfrequenzen liegt. Damit erzeugt die kubische Verzerrung eine Reihe von einfachen Tönen im Bereich der niedrigen Frequenzen einer periodischen Wellenform. Ist zum Beispiel eine komplexe Wellenform durch $f_1 = 1000$ Hz, $f_2 = 1200$ Hz,

$f_3 = 1400$ Hz und $f_4 = 1600$ Hz gegeben, existiert ebenso eine 800 Hz Komponente aus $2 * f_1 - f_2$, eine 600 Hz Komponente aus $2 * f_1 - f_3$ und eine 1000 Hz Komponente aus $2 * f_2 - f_3$. Diese Verzerrungsprodukte beeinflussen die Qualität der komplexen Wellenform. Eine Bestimmung der dominanten Region der Tonhöhe ist daher nur sehr schwer durchführbar.

5. Moderne Theorien zur Tonhöhenwahrnehmung

Die zwei jüngsten Theorien zur Tonhöhenwahrnehmung kommen aus dem Bereich der Ortstheorien. Beide Theorien führen in der ersten Stufen eine spektrale Analyse der komplexen Wellenform des Eingangssignals durch und zerlegen das Signal auf diese Weise in Teilkomponenten, wie es in Abbildung 3 der Fall ist. Aus dieser Repräsentation wird dann mittels einer Transformation die Information für die Bestimmung der Tonhöhe gewonnen. Beide Theorien versuchen die unveränderte Wahrnehmung der Tonhöhe zweier unterschiedlicher Reize zu erklären. Zum Beispiel besitzt ein Signal, das die Komponenten 1000, 1200 und 1400 Hz besitzt, ebenso wie ein Signal mit den Komponenten 1600, 1800 und 2000 Hz eine sehr starke 200 Hz Komponente. Beide Theorien erkennen, dass es sich um zwei unterschiedliche Komplexe handelt, wobei das niedriger frequente Signal eine stärker wirkende Tonhöhe erzeugt als das höher frequente. Daraus folgt, dass die Theorie definieren muss, was mit dem Begriff der „Stärke einer Tonhöhe“ gemeint ist, und wie diese Stärke gemessen werden kann.

Es ist offensichtlich, dass die Ergebnisse der aktuellen Forschung zur Aufdeckung der neuen Dimension in der Tonhöhenwahrnehmung geführt haben und dass diese neue Dimension in die einzelnen Theorien aufgenommen wurde. Während bei den praktischen Studien des letzten Jahrhunderts die Tonhöhenwahrnehmung immer als die Bestimmung der durchschnittlichen Tonhöhe gesehen wurde, muss jetzt auch die Stärke des Tonhöhenindrucks erkannt werden. Des weiteren muss auch die Zweideutigkeit und die Anzahl von verschiedenen alternativen Tonhöhen untersucht und quantifiziert werden.

Die Muster – Transformationstheorie von Wightman

In der ersten Stufe dieser Theorie [8] wird, wie bei den meisten aktuellen Ortstheorien, eine Frequenzanalyse mit einem Bandpassfilter durchgeführt, wie er in Abbildung 3 dargestellt ist. Das Filtern liefert ein Maß über die Stärke der Erregung der einzelnen Frequenzen, wobei die Feinheit des Filters eine herausragende Rolle spielt. Dieses Muster der Aktivität kann auch als Funktion $f(x)$ dargestellt werden, wobei f als Stärke der Aktivität an den verschiedenen Stellen x der Basilmembran gesehen werden kann. Wightman nimmt an, dass eine Transformation dieser Funktion auf die selbe Weise bewerkstelligt werden kann, wie die Berechnung einer Fouriertransformation des Spektrums

der originalen Wellenform. Die Transformation liefert als Ergebnis eine Annäherung an die Autokorrelationsfunktion der Wellenform. Bei der Autokorrelationsfunktion wird ein Signal mit sich selbst verglichen, indem es unter allen möglichen Verschiebungen mit sich selbst multipliziert wird, also mit sich selbst gefaltet. Bei einer Verschiebung von Null ergibt sich somit ein absolutes Maximum. Existiert im Signal eine Periodizität, wie es bei Lauten der Fall ist, entstehen bei der Verschiebung um das Vielfache der Periodendauer lokale Maxima. Daraus folgt bei der Tonhöhenwahrnehmung, dass ein zweites lokal Maximum entsprechend der wahrgenommenen Tonhöhe existiert. Im Falle eines Pulssignals enthält das Spektrum alle harmonischen von $n * (1 / T)$, wobei n die Werte $1, 2, \dots$ annimmt. Dieses Spektrum enthält eine starke Periodizität aufgrund der regelmäßigen Platzierung der einzelnen Komponenten. Die Transformation dieses Spektrums liefert eine starke Komponente bei 5 msec , also $1 / 200 \text{ Hz}$. Die Stärke dieser Komponente ist abhängig von der Isolation einzelner Komponenten, welche durch die Filterung des komplexen Eingangssignals bewerkstelligt wird. Bei höheren Frequenzen, die aufgrund der Anpassung an die Basilarmembran nicht so stark aufgelöst werden, ist die Periodizität weniger offensichtlich und somit wird laut Theorie keine Tonhöhe wahrgenommen. Daraus folgt, dass die Theorie im Prinzip eine dominante Region vorherzusagen vermag. Der einzige freie Parameter in diesem Modell ist die Feinheit des Filters für die Frequenzanalyse. Von diesem Parameter ist im wesentlichen die Autokorrelationsfunktion abhängig, die je genauer die einzelnen Komponenten erkannt werden, das zweite lokale Maximum umso deutlicher aufzeigt.

Goldsteins Theorie zur Berechnung der Tonhöhe

Goldsteins [9] Theorie teilt einige Merkmale der Theorie von Wightman. Beispielsweise benutzt er in der ersten Stufe seines Modell ebenfalls einen Bandpassfilter für die Frequenzanalyse. Weiterhin benutzt Goldstein ein Kriterium, um die Komponenten in aufgelöst und nicht aufgelöst einzuteilen. Ist eine Komponente aufgelöst, so berechnet er eine Frequenz f_i . Somit ergibt sich bei einer periodischen Welle eine sortierte Reihe von Frequenzen von $f_1, f_2, f_3, \dots, f_n$, wobei die höchste Frequenz bei f_n liegt. In der nächsten Stufe beschreibt Goldstein einen „optimalen Prozessor“, der eine Analyse der erhaltenen spektralen Komponenten vornimmt. Dabei wird versucht zwei Zahlen f_0 für die Grundfrequenz und k für das niedrigste harmonische Intervall so festzulegen, dass $f_0 * k \rightarrow f_1, f_0 * (k + 1) \rightarrow f_2, \dots, f_0 * (k + m) \rightarrow f_n$ gilt, wobei „ \rightarrow “ bedeutet, dass der geschätzte Wert auf der linken Seite mit maximaler Wahrscheinlichkeit der selbe Wert ist, wie der auf der rechten Seite. Zum Beispiel, wenn $f_1 = 1003 \text{ Hz}$, $f_2 = 1205 \text{ Hz}$ und $f_3 = 1390 \text{ Hz}$ ist, dann ist eventuell $f_0 = 200$ und $k = 5$. Das bedeutet, $1000 \rightarrow 1003$, $1200 \rightarrow 1205$ und $1400 \rightarrow 1390$. Die Maximierung geschieht hierbei im Bezug auf f_0 , dem Frequenzwert der

Tonhöhe und k dem Wert der niedrigsten harmonischen Frequenz. Die Differenz zwischen dem gemessenen und dem vorhergesagten Wert ist ein Maß für die Qualität oder die Stärke der wahrgenommenen Tonhöhe. Die Stärke dieses Modells ist, dass einige der Schätzformeln ausgearbeitet wurden und in einer Vielzahl von Versuchsreihen getestet werden konnten. Allerdings kann die genau Frequenz einer Komponente nicht bestimmt werden, wodurch eine Varianz der Frequenz entsteht. Diese Tatsache und die Tatsache, dass der Einfluss einer Komponenten auf eine andere völlig verloren geht, sind weniger befriedigende Aspekte dieser Theorie.

6. Quellenverzeichnis

Abbildungsverzeichnis

Abbildung 1: Zentrale für Unterrichtsmedien im Internet e.V.

<http://www.zum.de/dwu/pas008vs.htm>

Abbildung 2: The Nervous System, Donald b. Tower Editor in Chief, Vol. 3
Human Communication and its Disorders,
Pitch Perception, Davic M. Green, Seite 148.

Abbildung 3: The Nervous System, Donald b. Tower Editor in Chief, Vol. 3
Human Communication and its Disorders,
Pitch Perception, Davic M. Green, Seite 150.

Qellenverzeichnis

The Nervous System, Donald b. Tower Editor in Chief,
Vol. 3: Human Communication and its Disorders,
Pitch Perception, Davic M. Green.

Psychophysikalische Untersuchung von spektralen und zeitlichen Mechanismen
des auditorischen Systems anhand harmonischer und unharmonischer
Amplitudenmodulationen: relatives und absolutes Gehör
Diplomarbeit von Stefan Bleeck
<http://www.tonhoehe.de/diplom.html>

Literaturverweise

- [1]: Seebeck , A. (1841),
Beobachtungen über einige Bedingungen der Entstehung von Tönen,
Ann. Phys. Chem., 53: 417 – 436.
- [2]: Ohm, G. S. (1843),
Über die Definition des Tones, nebst tonbildender Vorrichtungen,
Ann. Phys. Chem., 59: 513 – 565.

- [3]: von Helmholtz, H. L. F. (1863),
Die Lehre von den Tonempfindungen als physikalische Grundlage für die
Theorie der Musik,
F. Vieweg und Sohn, Braunschweig.
- [4]: Schouten, J. F. (1940),
Five Articles on the Perception of Sound,
Institute for Perception, Eindhoven.
- [5]: Schouten, J. F., Ritsma, R. J., Cardazo, B. L., (1962),
Pitch of residue,
J. Acoust. Soc. Am., 34: 1418 – 1424.
- [6]: Ritsma, R. J. (1967),
Frequencies dominant in perception of the pitch of complex sounds,
J. Acoust. Soc. Am., 42: 191 – 198.
- [7]: Houtsma, A. J. M., and Goldstein, J. L. (1972),
The central origin of the pitch of complex tones: Evidence from the musical
interval recognition,
J. Acoust. Soc. Am., 51: 520 – 529.
- [8]: Wightman, F. L. (1973),
The pattern-transformation model of pitch,
J. Acoust. Soc. Am., 54: 407 – 416.
- [9]: Goldstein, J. L. (1973),
An optimum processor theory for the central formation of the pitch of
complex tones,
J. Acoust. Soc. Am., 54: 1496 – 1516.