

Vom Gehör zur MP3: Interdisziplinäre Ansätze zur Akustik im Physikunterricht

Daniel Aichinger¹, Jan-Peter Meyn²

¹ Didaktik der Physik, Westböhmisches Universität in Pilsen, daniel.aichinger@seznam.cz

² Didaktik der Physik, FAU Erlangen-Nürnberg, jan-peter.meyn@physik.uni-erlangen.de

Kurzfassung

Im Schulunterricht wird Akustik meist nur als ein Randgebiet der Lehre von Schwingungen und Wellen behandelt. Dabei gibt es eine Reihe von fachübergreifenden Themen, die in den Lehrplänen anderer Schulfächer stehen, mit den Methoden dieser Fächer aber nicht vollständig erklärt und übermittelt werden können. Beispiele dafür sind die Funktionsweise des menschlichen Gehörs, akustische Alltagsphänomene, musikalische Akustik, Sprachrezeption und Produktion oder die Anwendungen technischer Akustik im Alltag. Im folgenden Text werden fächerübergreifende Arbeits- und Unterrichtsmethoden vorgestellt, die es ermöglichen diese Themen im Unterricht zu behandeln. Die direkte Sinneserfahrung wird dabei in den Vordergrund gestellt. Ausgehend von realen Experimenten zur akustischen Wahrnehmung sollen ihre Eigenschaften und Grenzen erarbeitet werden. Dies ermöglicht den Schülern die behandelten Phänomene selbst zu entdecken, auszuprobieren und anschließend kreativ anzuwenden.

1. Das Konzept

Eine weitreichende curriculare Reform des Bildungssystems ermöglicht Lehrkräften in der Tschechischen Republik ihren Unterricht unabhängig von den früher verbindlichen Lehrplänen zu gestalten. Bereits seit dem 1. 9. 2009 gelten die neuen Rahmenbedingungen für Gymnasien [1], spätestens bis zum 1. 9. 2012 für alle Schultypen von Pre-Primarstufe bis zur Sekundarstufe II.

Das tschechische Schulministerium spezifiziert in einem sogenannten Rahmenbildungsprogramm (RVP) für jede Schulart die wichtigsten Kompetenzen und Inhalte einzelner Fachbereiche, anhand deren Lehrerinnen und Lehrer an Schulen ihre individuellen Schulbildungsprogramme (SVP) zu entwickeln haben. Spezifiziert sind in den Rahmenbildungsprogrammen nur die minimale Anzahl von Unterrichtsstunden für das jeweilige Fachgebiet, dessen allgemeine Unterrichtsziele, durch den Unterricht zu erwerbende Fähigkeiten und Fertigkeiten sowie die für die Schulart obligatorischen Grundinhalte. Nicht mehr vorgegeben sind einzelne Unterrichtsthemen und auch nicht der zeitliche Plan, in dem sie unterrichtet werden sollen. Das ermöglicht den Schulen in Kooperation mehrerer Fächer interdisziplinäre Konzepte zu entwickeln, nach modernen Methoden zu unterrichten und bei Bedarf aktuelle Themen in den Unterricht einzubeziehen.

Das hier vorgestellte Unterrichtskonzept zur Akustik ist Teil eines solchen vom Autor verfassten Schulbildungsprogramms für den Physikunterricht an der Sekundarstufe II. Es entsteht im Rahmen eines Postgradualstudiums in Didaktik der Physik an der Westböhmisches Universität in Pilsen (Tschechien).

Im deutschsprachigen Raum wurden zu dem Thema Akustik im Physikunterricht sowie dem Einsatz von Soundkarte und Audiosoftware in Experimenten bereits mehrere Arbeiten veröffentlicht [2, 3]. In Pilsen haben wir für den Akustikunterricht ein Thema gewählt, das, obwohl es schon über 10 Jahre unser tägliches Leben begleitet, in der Schule immer noch nicht vorkommt: Die MP3-Datenkompression.

Der hier skizzierte Unterricht geht systematisch von der Erprobung und Sensibilisierung der akustischen Wahrnehmung bis hin zur gehörangepassten Audiokodierung bei MP3 vor und trotz aller nötigen Vereinfachung für die Schule bleibt er nahe an der Forschung auf diesem Feld. Damit komplexe akustische Signale in Musik und Sprache analysiert, computer-simuliert oder weiterbearbeitet werden können, beinhaltet das Konzept unter anderem auch Grundlagen der Systemtheorie und digitaler Signalverarbeitung. Alle verwendeten Geräte und Software stammen aus dem Bereich der Studio- und Aufnahmetechnik und ermöglichen eine vielseitige Anwendung bei Messungen, Simulationen und kreativer Arbeit mit Sprache und Musik. Dadurch entsteht ein anwendungsorientierter, fächerübergreifender Akustikunterricht mit starken Bezügen zu Technik, Informatik, Physiologie, Psychologie sowie zum Sprach- und Musikunterricht.

1.1. Lernebenen des Konzeptes

Zielgruppe des Konzeptes sind Schüler der Sekundarstufe II (Alter 15 - 19 Jahre) tschechischer Gymnasien, Fachoberschulen und insbesondere auch von Kunstschulen für Musik, Film und neue Medien. Das Konzept ist so aufgebaut, dass die einzelnen

Lektionen je nach Schulart, Alter und Vorkenntnissen der Schüler und Anzahl der zur Verfügung stehenden Unterrichtsstunden unterschiedlich vertieft in 3 Lernebenen behandelt werden können.

a) Die Grundebene:

In der ersten Lernebene arbeiten Schüler zuerst vorwiegend rezeptiv mit vorgefertigten Hörproben, erleben die wichtigsten akustischen Phänomene und lernen dabei Geräte und Software zur Audibearbeitung kennen. Ziel der ersten Lernebene besteht hauptsächlich darin, akustische Phänomene wiederzuerkennen und richtig zu benennen. Methodisch besteht der Unterricht aus Lernen mit audiovisuellen Lernobjekten, Frontalunterricht mit Demonstrationsexperimenten und einfachen angeleiteten Schülerexperimenten zum „Selberprobieren“. Alle verwendeten Soundbeispiele werden den Lernenden als Wave-Audiodateien frei zur Verfügung gestellt, damit sie diese mit ihrem PC zu Hause, mit einem CD, DVD oder Flash-Speicher basierten Abspielgerät (z.B. MP3-Player) nachhören können.

b) Die Verständnisebene

In der zweiten Ebene lernen Schüler produktiv mit den Geräten und Software zu arbeiten. Zunächst geschieht dies in typisierten Aufgaben mit einer detaillierten Beschreibung und intensiver Lehrerbetreuung. Später arbeiten die Schüler immer selbständiger. In offenen Aufgaben untersuchen sie die Eigenschaften akustischer Phänomene und lernen durch ihr eigenes Handeln wichtige Zusammenhänge zu verstehen. Methodisch wird auf dieser Ebene überwiegend nach der heuristischen und teilweise bereits gemäß der Forschenden Unterrichtsmethode vorgegangen. Bei der Organisationsform handelt es sich um ein Praktikum mit Zweiergruppen.

c) Die Anwendungsebene

Die dritte Lernebene sieht eine erfinderische Anwendung der bis dahin erworbenen Fähigkeiten und Fertigkeiten in offenen Aufgaben ohne Anleitung und in kreativen Schülerprojekten vor. Aus Sicht der Lernpsychologie umfasst die höchste Lernebene des Konzeptes die Ebenen des fernen bis kreativen Wissenstransfers nach der Taxonomie von Haskell [4]. bzw. des nichtspezifischen Transfers nach Niemi-ko [5]. Die Projektideen sollen von Schülern selbst ausgehen und sich thematisch nicht nur auf physikalische Akustik beschränken, sondern die ganze Bandbreite zwischen Naturwissenschaften und Kunst nutzen. Damit werden auch die Herangehensweisen verschiedener Disziplinen von Physik und Informatik über experimentelle Psychologie bis hin zu Musiktheorie oder Phonetik angewandt. Die Organisationsform des Unterrichts in dieser Lernebene ist ein Gruppenunterricht mit Gruppengröße von bis zu 5 Teilnehmern. Die Lehrkraft bleibt dabei im Hintergrund und nimmt die Rolle eines Mentors ein.

1.2. Unterrichtsmittel

Zur Audibearbeitung, Visualisierung der Signaleigenschaften und zum Kontrollhören ist die Nutzung eines Computerraums oder ausreichender Anzahl an Laptops vorgesehen, damit während des Unterrichts möglichst jeder Schüler einen Computer mit einem Kopfhörer zur Verfügung hat. Während an der ersten Lernebene grundsätzlich ein gewöhnlicher PC mit eingebauter Soundkarte noch völlig ausreicht, kommen bei den Aufgaben an der Verständnisebene, wo es zunehmend um Aufnehmen und Messen geht, die Lernenden bereits an die technischen Grenzen der üblichen Hardware und benötigen eine spezielle Ausstattung für Aufnahme und Wiedergabe. Ein leistungsfähiger Rechner spielt dabei allerdings nur die Rolle einer Art von Steuerungseinheit der Audioarbeitsstation.

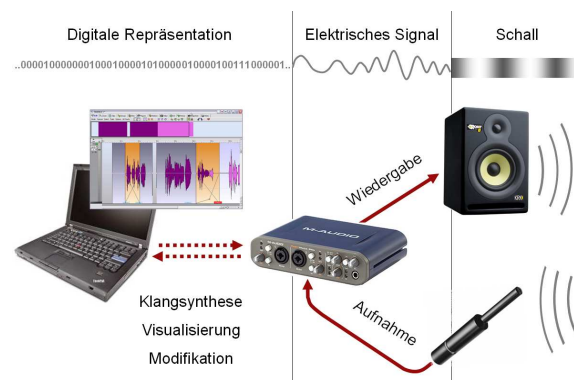


Abb.1: Schematische Darstellung einer Arbeitsstation

Das eigentliche Zentrum der Arbeitsstation stellt eine hochwertige Audioschnittstelle dar. Semiprofessionelle Soundkarten werden zurzeit meistens nicht mehr rechnerintern über PCI, sondern vermehrt als externes Modul über USB 2.0 oder Fire-Wire (IEEE 1394) an den Computer angeschlossen. Diese Bauweise hat besonders in Verbindung mit einem Laptop viele Vorteile. Externe Audioschnittstellen vereinigen so in einem kompakten Gerätgehäuse D/A und A/D Wandler (24 Bit bei 96 / 192 kHz), Analogausgänge für Endverstärker und Kopfhörer, Analogeingänge mit Mikrofonverstärkern (typischerweise 2), Polarisationsspannungsversorgung für Kondensatormikrophone (+48 V) und verfügen im besten Fall über einen separaten Kleinleistungsnetzteil.

Zum Softwarestandard bei computerbasierten Aufnahmesystemen gehören Programme wie Adobe Audition, Steinberg Wavelab, Sony Sound Forge oder Avid Audio Pro Tools, sowie GNU lizenzierte ähnlich zu bedienende und funktionierende freie Software. In unserem Beispiel wird als Hauptsoftware die Vollversion von Wavelab 6.0 in EDU-Lizenz verwendet. Damit Schüler mit verschiedener Software zu arbeiten lernen, ist es allerdings angebracht neben einer für die Schule lizenzierten Software noch zumindest Demoverionen weiterer Audio-Software auf die Arbeitsstationen zu installieren.

Als akustische Sensoren für Messungen und Audioaufnahmen eignen sich am besten kapazitive Kleinmembranmikrophone. Das sind entweder so genannte echte Kondensatormikrophone oder kostengünstigere Elektret-Kondensatormikrophone. Gebraucht werden im Unterricht unterschiedlichen Bauarten mit verschiedenen Richtcharakteristiken, mindestens aber zwei Messmikrophone mit omnidirektionaler Charakteristik (Kugelcharakteristik), regelmäßig nachkalibriert mit einem Schallpegelkalibrator entsprechender Genauigkeitsklasse.

Mobile batteriebetriebene Aufnahmesysteme mit integrierten Mikrofonen bieten eine handliche, kostengünstige und vor allem völlig autonome Alternative zu Computergesteuerten Aufnahmesystemen. Sie können im Unterricht als weitere Arbeitsstationen „zum Mitnehmen“ verwendet werden. Dadurch stehen den Schülern mehrere voneinander unabhängige Aufnahmesysteme zur Verfügung, die nach Absprache mit dem Lehrer für Aufnahmen außerhalb der Schule ausgeliehen werden können.

Zur Wiedergabe von Hörbeispielen und als Schallquelle für akustische Messungen haben sich robust gebaute aktive Studioabhörmonitore bewährt. Sie gewährleisten eine hohe Tonqualität und Lebensdauer und das auch nach mehreren Jahren im Schulsinsatz. Für individuelle Arbeit der Schüler am Rechner eignen sich am Besten geschlossene ohrumschließende Kopfhörer, die Geräusche von Außen dank ihrer Konstruktion effizient dämmen und gleichzeitig nach Außen wenig stören.

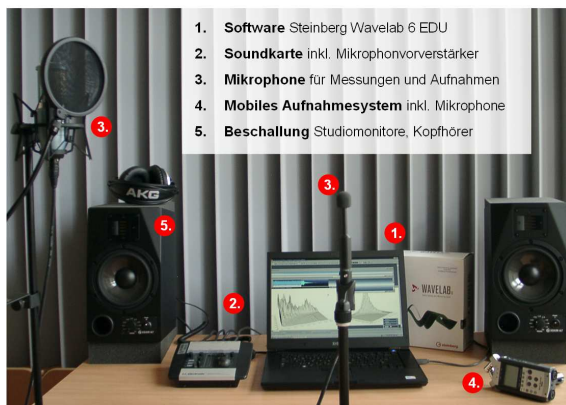


Abb.2: Aufbau einer PC-Audioarbeitsstation (Uni Pilsen)

2. Unterrichtsthemen

In der kompaktesten Variante für Gymnasien umfasst das Konzept 6 Unterrichtsstunden je 45 Minuten. Abhängig von gegebenen Unterrichtsbedingungen und dem Interesse der Schüler können die im Folgenden genannten Themen mit optionalen Praktika bis hin zur zweiten Lernebene vertieft werden. Das bedeutet jeweils mindestens zwei zusätzliche Unterrichtsstunden pro Thema. Zum Abschluss des Kurses wird ein ganztägiger Workshop angeboten, der auf Wunsch der Schüler auf eine Projektwoche erweitert werden kann.

Die Basisthemen im Überblick:

- Akustische Signale und ihre Darstellung (Praktikum: Signale Erzeugen und Auswerten)
- Das menschliche Gehör (Praktikum: Hörtests und Gehörmodelle)
- Komplexe Signale in Sprache und Musik (Praktikum: Mikrophon- und Aufnahmetechnik)
- Digitalisierung der Audioinformation (Praktikum: Digitale Audiotechnik)
- Gehörangepasste Audiokodierung (Praktikum: Kodierungsalgorithmen)
- Kodierungsartefakte und ihre Ursachen (Praktikum: Audioqualitätsbestimmung)

2.1. Akustische Signale und ihre Darstellung

Als Einführung in die Akustik lernen Schüler einfache Signale mit dem Gehör zu erkennen und voneinander zu unterscheiden. Von den zur Verfügung stehenden Visualisierungsmöglichkeiten der verwendeten Audiosoftware beschränkt sich das erste Thema ausschließlich auf die Visualisierung der Signalzeitverläufe (Oszilloskop), damit die Darstellung von Frequenzspektren nicht gleich am Anfang des Kurses zur Verwirrung und Unverständnis führt.

Zunächst werden synthetisch erzeugte Signalmodelle hörbar gemacht und gleichzeitig im Zeitverlauf visualisiert: Einfache Töne verschiedener Frequenzen, komplexe Töne mit wenigen Spektralkomponenten, Breitbandrauschen, Rauschen mit begrenzter Bandbreite, kurze Tonimpulse, Gleichspannungsimpulse und daraus bestehende Impulsreihen. Es werden Eigenschaften von Signalen diskutiert, die zu Kategorisierung führen. Anschließend werden reale Klänge den zugehörigen Modellen und eingeführten Kategorien zugeordnet.

2.2. Die Gehörphysiologie

Menschliche Wahrnehmung von Klängen wird maßgeblich von der Physiologie des Gehörorgans bestimmt. Für den Physikunterricht besonders interessant ist dabei die Funktionsweise des Innenohres mit der sich entlang des spiralförmigen kochlearen Labyrinth erstreckenden Basiliarmembran.

Bereits Hermann von Helmholtz [6] hat Mitte des 19. Jahrhunderts vermutet, dass das menschliche Gehör komplexe Klänge in ihre Spektralkomponenten zerlegt. Er stellte sich den Analysemechanismus ähnlich einer Harfe mit einzelnen verschiedenen gestimmten Seiten vor, die von komplexen Klängen zum Schwingen gebracht werden. Die genaue Rolle der Basiliarmembran beim Analyseprozess hat aber Helmholtz seinerzeit noch nicht durchdrungen. Das ist erst fast 100 Jahre später dem ungarischen Physiker Georg von Bekézszy gelungen und 1961 wurde ihm für die Formulierung der Theorie fortlaufender Wellen an der Basiliarmembran und die damit verbundenen Messungen und Simulationen der Nobel-

preis für Medizin verliehen [7]. In den letzten 50 Jahren wurde seine Theorie durch immer genauer werdende Messtechniken und Modellierung zwar mehrmals kritisch überprüft [8, 9] und schließlich auch in vielen Einzelheiten korrigiert, die wesentlichen Annahmen des Bekézsyschen Innenohrmodells bleiben aber bis heute gültig [10].

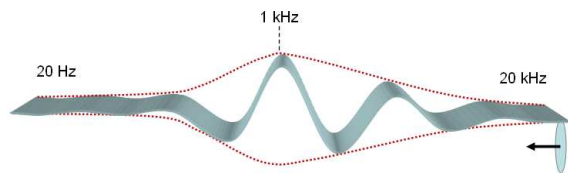


Abb.3: Ein Modell für Wellen an der Basiliarmembran

In der Cochlea kommt es bei lange andauernden tonalen Stimuli zu einer Transformation des zeitlichen Schwingungsverlaufs in ein örtliches Resonanzmuster an der Basiliarmembran. Das stellt die Hüllkurve der fortlaufenden Welle in Abbildung 3 für die Stimulfrequenz 1 kHz (Sinuston) dar. Die Membranauslenkung und Wellenlänge der fortlaufenden Welle auf dem Bild wurden zur Verdeutlichung nicht maßstabsgetreu abgebildet (tatsächliche Amplitude $\sim 10^{-9}$ m, $\lambda_{1\text{kHz}} \gg \text{Membranlänge!}$). Das weist gleichzeitig auf die Grenzen dieses Modells hinsichtlich Amplitude und Wellenlänge hin. Jeder hörbaren Schwingungsfrequenz entspricht ein bestimmter Ort an der Basiliarmembran, an dem die fortlaufende Welle eine maximale Auslenkung erzeugt. Die Basiliarmembran agiert so wie ein mechanischer Frequenzanalysator.

Hier bietet sich im Unterricht die Einführung von Spektralanalyse als eines effektiven Visualisierungsmittels für komplexe Signale. Ein mechanischer Frequenzanalysator, wie der im Innenohr, ist für Schüler nämlich wesentlich anschaulicher und einfacher zu verstehen als mathematische Transformationen. Um das Verständnis der Gehörphysiologie noch weiter zu vereinfachen, werden im Unterricht 2 Modelle verwendet: Ein hydromechanisches Modell für die Umwandlung einer ins Innenohr eintreffender Longitudinalwelle in eine Transversalwelle an der Basiliarmembran und ein Modell zur Veranschaulichung der Resonanzeffekte bei fortlaufenden Wellen an der Basiliarmembran.

Eine Membran mit kontinuierlich variierender Resonanzfrequenz über ihre ganze Länge macht eine sehr feine Frequenzauflösung einfacher Tonstimuli möglich. Bei mehreren naheliegenden Resonanzmustern an der Basiliarmembran versagt jedoch leicht die Unterscheidung einzelner Frequenzkomponenten. Bei zwei gleichzeitig wirkenden frequenznahen Stimuli unterschiedlicher Intensität dominiert der stärkere Stimulus das Resonanzmuster an der Membran. Dies kann so weit führen, dass es vom Resonanzmuster nur des alleine wirkenden stärkeren Stimulus nicht zu unterscheiden ist. Für das Kompressionsverfahren MP3 stellt diese kleine Unvollkommenheit eine glückliche Konstellation dar.

3. Untersuchung des menschlichen Gehörs

Um die Eigenschaften des eigenen Gehörs zu erproben eignen sich für den Unterricht vereinfachte Gehörtests aus dem Bereich der Audiologie, bei denen Schüler zuerst Beispiele hören und beurteilen, später dann selber zu erzeugen lernen.

3.1. Wahrnehmung von Lautstärke

Anhand von Hörbeispielen lernen Schüler, wie Lautstärke wahrgenommen wird. Dazu wird eine aus kurzen 1 kHz-Tönen bestehende Sequenz mit jeweils steigender Amplitude verwendet. Bei einer Amplitudenänderung in konstanten Schritten wird deutlich hörbar, dass diese nicht zu einer Lautstärkeänderung in konstanten Schritten führt. Während angeleiteter Erzeugung von weiteren Tönequenzen entdecken Schüler, dass sie, um eine hörbare Lautstärkeänderung zu erreichen, bei großen Amplituden einen entsprechend größeren Schritt in der Signalverstärkung machen müssen. Regelmäßige Schritte in der Lautstärke erreichen Schüler in guter Annäherung jeweils durch Verdopplung der Amplitude des Tones (Abbildung 4 unten). Anschließend sollen sie die Gültigkeit ihrer Hypothese für rauschartige Signale und Pulse bestätigen. Mehrfache Verdopplung (x -fach) führt zur Funktion vom Typ $y = 2^x$. So können direkt aus der Sinneserfahrung das Weber-Fechner Gesetz und nach Umrechnung mit Zehnerlogarithmus auch die Lautstärkepegel und die logarithmische Dezibel-Skala eingeführt werden. Mit Hilfe einfacher Berechnungen lernen Schüler die Lautstärkepegel für reale Situationen näherungsweise vorherzusagen.

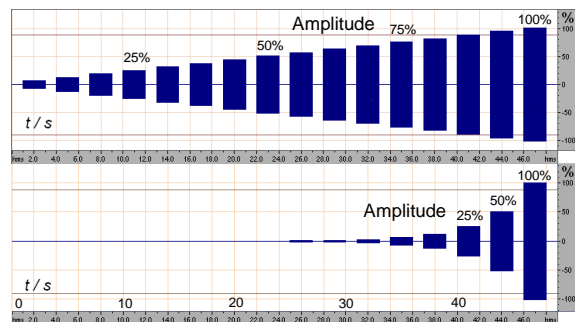


Abb.4: Untersuchung der Lautstärkewahrnehmung

Um die untere Grenze akustischer Wahrnehmung hinsichtlich Lautstärke genauer zu untersuchen, werden Serien von Tönen mit zu- oder abnehmendem Pegel verwendet, zum Beispiel stufenweise jeweils um 2 dB. Schüler suchen in ruhiger Umgebung den leisesten gerade noch hörbaren Ton zuerst in einer solchen Tönequenz für 1 kHz, die so genannte Ruheshwelle bei 1 kHz. Welcher Pegel in der einer relativen Aussteuerungsskala entsprechenden Serie noch hörbar wird, hängt natürlich auch von der Signalverstärkung, Entfernung des Hörers von der Schalquelle und von weiteren Abhörbedingungen ab. In Abbildung 5 liegt die experimentell ermittelte Hörschwelle für 1 kHz Töne am relativen Aussteuerungspegel -72 dB.

Das Finden der Ruhehörschwelle bei 1 kHz ermöglicht für die gegebene Einstellung gemäß der Dezibel-Pegel Definition eine Verbindung zwischen Aussteuerungspegeln des im Rechner generierten Signals (SPL relativ in dB) und den Schalldruckpegeln am Ohr (SPL absolut in dB) herzustellen und damit das Abhörsystem für die konkreten Einstellungen und Abhörbedingungen zu kalibrieren. Gleichzeitig kann bestätigt werden, dass eine Auflösung in Lautstärke von 2 dB im breiten Aussteuerungsbereich an der Grenze der Unterscheidbarkeit liegt [8]. Es ist allerdings auch deutlich hörbar, dass bei hohen Schalldruckpegeln dieses Modell die menschliche Lautstärkewahrnehmung nicht mehr ausreichend beschreibt.

3.2. Wahrnehmung der Stimuli-Frequenz

Bei der Untersuchung der Wahrnehmung von Frequenz der Stimuli überprüfen Schüler auch die Frequenzauflösung des Sinnesorgans. Manchen Schülern sind bereits Musikintervalle als Logarithmische Ordnung bekannt. Aufgrund ihrer Einfachheit und guter Annäherung an die menschliche Wahrnehmung im breiten Frequenzbereich wird die logarithmische Verteilung auch in der technischen Praxis sehr oft verwendet. Typische Beispiele dafür sind die logarithmische Frequenzskala in Spektraldiagrammen oder CPB-Filterbanken (Constant Percentage Bandwidth) wie Oktavenfilter und Terzfilter.

Im Praktikum werden anschließend psychoakustische Modelle für die Wahrnehmung der spektralen Distribution akustischer Energie untersucht, die die Eigenschaften des menschlichen Gehörs noch wesentlich genauer approximieren. Das wohl bekannteste Beispiel hierfür ist die Frequenzverteilung nach sogenannter kritischer Bandbreite (Critical Band Rate) und die daraus resultierende Bark-Skala. Sie wurde nach verschiedenen psychoakustischen Untersuchungen entworfen und stellt so indirekt ein Abbild der örtlichen Auflösung der Basiliarmembran dar. Im Unterricht wird auf die Bark Skala und ihre Revision, die ERB-Skala (Equivalent Rectangular Bandwidth) [4] deswegen ein besonderer Wert gelegt.

3.3. Die Ruhe- und Mithörschwelle

Mit ähnlichen Sequenzen wie der bei 1 kHz können Lautstärkepegel für die leisesten gerade noch wahrnehmbaren Töne verschiedener Frequenzen im gesamten übertragenen Frequenzspektrum gefunden werden. Auftragen der ermittelten Pegel über die Frequenz, ergibt die sogenannte Hörschwellenkurve (Abbildung 5). Unter dieser Kurve liegende Töne sind bei gegebenen Abhörbedingungen mit dem Gehör nicht wahrnehmbar. So grenzt die Kurve von unten die sog Hörfläche ab, in der hörbare Töne definierter Frequenz und Lautstärkepegel liegen.

Eine Untersuchung der oberen Lautstärkewahrnehmungsgrenze ist an Schulen ausgeschlossen. Gehörschutz steht hier sowie in weiteren Experimenten immer an erster Stelle. Zu möglichen Folgen der

Exposition gegenüber hohen Schalldruckpegeln und zum Lärmschutz werden Schüler im Unterricht ebenfalls umfassend informiert.

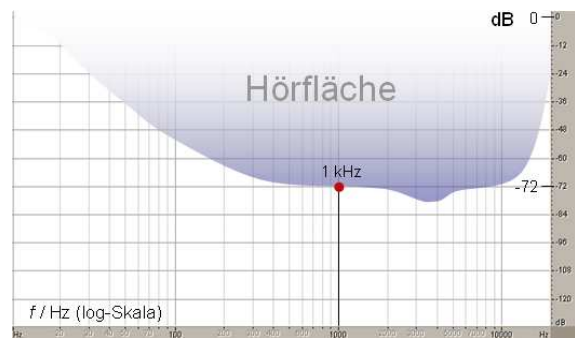


Abb.5: Die Ruhehörschwelle

In Hörtests zur Hörbarkeit von Tönen bei gleichzeitiger Anwesenheit eines starken bandbegrenzten Signals, eines sog. Maskierers, zum Beispiel eines Schmalbandrauschens, können die Lernenden entdecken, dass Töne im naheliegenden Frequenzbereich im Vergleich zur den Ruhehörschwellen für diese Töne erst bei viel höheren Pegeln hörbar werden. Das ist in graphischer Darstellung (Abbildung 6) als eine Anhebung der Hörschwellenkurve gut sichtbar. Die untere Wahrnehmungsgrenze wird in diesem Fall lokal als Maskierungskurve oder insgesamt als Mithörschwelle bezeichnet. Der Bereich mit dem Maskierer gleichzeitig hörbarer Klänge, hier Mithörfläche genannt, wird dadurch verkleinert. Hauptursache dafür ist die bereits beschriebene Funktionsweise der Cochlea. Eine genauere Untersuchung zeigt, wie die Form eines experimentell ermittelten Mithörschwellendiagramms der physiologisch bedingten Hüllkurve einer fortlaufenden Welle an der Basiliarmembran ähnelt (vgl. Abbildung 3 und 6).

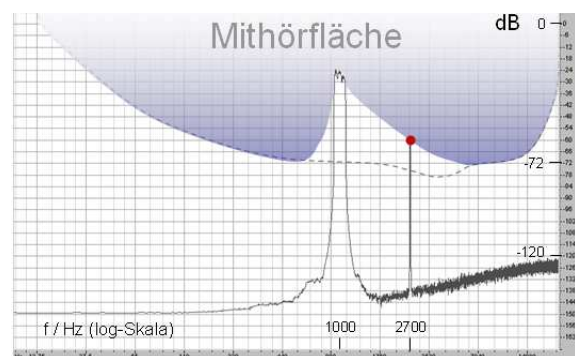


Abb.6: Mithörschwelle beim Schmalbandrauschen

3.4. Die Zeitauflösung

Für weitere Überlegungen sehr wichtig ist die zeitliche Auflösung des menschlichen Gehörs. Schüler lernen diese in verschiedenen Experimenten mit Pulsreihen und Amplitudenschwankungen kennen. In den verwendeten Pulsreihen werden Abstände zwischen einzelnen Pulsen schrittweise verkleinert. Beim Hören soll der geringste zeitliche Abstand der Pulse ermittelt werden, bei dem diese noch als solche gehört werden, bzw. kurz bevor die zeitliche

Auflösung versagt und die Pulsreihe als ein rauer Ton wahrgenommen wird. Hier sollte erwähnt werden, dass es sich um keine scharfe Grenze handelt. Mit guter Wiederholbarkeit kommt man aber auch ohne Training auf Werte zwischen 30 und 25 ms, also Pulsraten zwischen 33 und 50 Pulsen pro Sekunde. Es ist allerdings darauf zu achten, dass Ergebnisse durch das Nachschwingen der Schalgebermembran bzw. allgemein durch die Impulsantwort des Mess-/Abhörsystems nicht verfälscht werden.

Ähnlich gibt es bei Schwebungen eine Grenze ab der einzelne Schwebungen nicht mehr unterscheidbar sind, typischerweise bei Schwebungsperioden unter 30 ms. Aus diesen zwei Experimenten kann die Größe der zeitlichen Bearbeitungsblöcke des menschlichen Gehörs ermittelt werden. Sie beträgt je nach Signaltyp in der Größenordnung von 10 Millisekunden. Diesen kleinsten wahrnehmbaren Zeitabschnitt nenne ich bildhaft den „Ohrenblick“.

Grundsätzlich ist die „Ohrenblick“-Länge kürzer bei Transienten als bei lang andauernden tonalen Stimuli. Bei lang andauernden Signalen, wie etwa Tönen oder Rauschen mit einem konstanten Pegel untersuchen Schüler anschließend die Lautstärkewahrnehmung in Abhängigkeit von der Stimulidauer. In einem Hörbeispiel mit verschieden langen Tönen gleicher Frequenz und Amplitude wird die Tondauer gesucht, ab der sich die Lautstärkewahrnehmung nicht mehr ändert. Hier bekommt man je nach Signaltyp Ergebnisse zwischen 100 und 200 ms. Das zeigt, dass die Lautstärkewahrnehmung mit Integrationszeiten arbeitet, die 4- bis 10-mal länger sind als die zeitliche Auflösung.

3.5. Zeitliche Verdeckungseffekte

Mit etwas Übung kann man zwei einzelne nacheinander folgende Pulse auch bei Abständen unter 10 ms unterscheiden. Wenn aber kurz nacheinander zwei unterschiedlich starke Stimuli kommen, dominiert der stärkere (der sog. Maskierer oder Masker) die Wahrnehmung und verdeckt den anderen. Schüler lernen ihr Gehör darauf zu sensibilisieren.

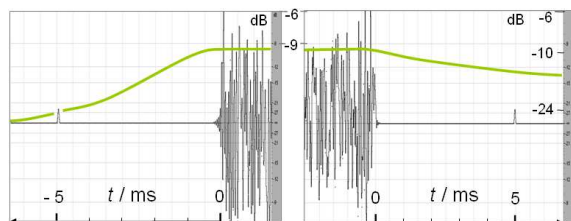


Abb.7: Vor- (links) und Nachverdeckung (rechts)

In der Abbildung 7 ist im Zeitverlauf links ein Puls vor einem Breitbandrauschen zu sehen, der bei gegebener Einstellung an der Grenze der Hörbarkeit liegt, rechts ein vollständig maskierter, d. h. unhörbarer, Puls nach einem Breitbandrauschen. Trägt man die Pegel der Pulse, die beim bestimmten Zeitlichen Abstand nicht gehört werden, graphisch auf, entsteht eine Mithörschwellenkurve im Zeitbereich (Abb. 8).

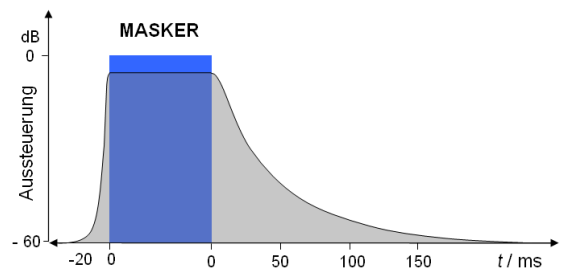


Abb.8: Diagramm der Maskierungskurve im Zeitbereich

4. Komplexe Signale in Sprache und Musik

Im Unterricht bzw. einem Praktikum zu Mikrophon und Aufnahmetechnik lernen Schüler verschiedene Wandlerprinzipie und Mikrophonrichtcharakteristiken anzuwenden, nehmen Sprach und Musiksignale auf, untersuchen deren Zeitverlauf, Frequenzspektrum und gegebenenfalls auch Aussteuerungsstatistiken. Bei der Signalanalyse lernen sie ihre eigenen Erfahrungen aus den vorausgehenden Hörexperimenten umzusetzen. Sie haben zum Beispiel bereits entdeckt, mit wie langen Zeitabschnitten das menschliche Gehör arbeitet, was ihnen bei der Spektralanalyse die Einstellung von Zeitfenstern für verschiedene Signaltypen erleichtert. Fensterlängen zwischen 25 und 100 ms zeigen sich bei Sprache und Musik als eine gute Annäherung an die Akustische Wahrnehmung des Menschen und bieten einen optimalen Kompromiss zwischen Zeit- und Frequenzauflösung.

Die Visualisierung des Frequenzspektrums hilft dominante Frequenzkomponenten in komplexen Signalen zu identifizieren. Für jede Komponente kann nach einem mit einfachen Testsignalen ermittelten Modell dann die Maskierungskurve berechnet werden. Diese entspricht nach Addition für alle Komponenten für das gegebene Zeitfenster einer globalen Mithörschwellekurve. In Abbildung 9 sind die globale Mithörschwelle und die resultierende Mithörfäche für einen obertonreichen komplexen Ton (Kammerton A) angezeigt, der mit einer Geige gespielt wurde.

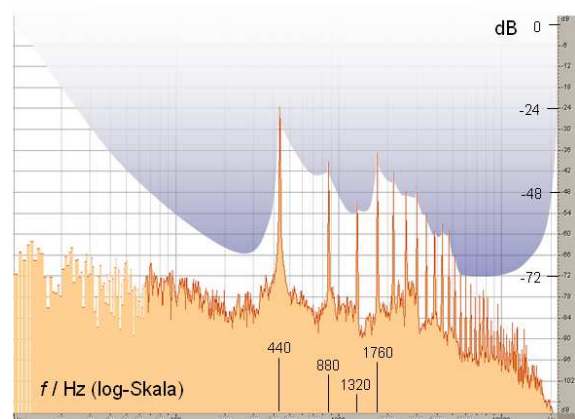


Abb.9: Mithörschwelle eines komplexen Tones

5. Digitalisierung der Audioinformation

Beim Untersuchen komplexer Signale in Sprache und Musik werden alle Signale mit einer hohen Werteauflösung und Abtastrate (24 Bit, 96/192 kHz) aufgenommen oder softwaremäßig generiert. Eine darauf folgende Unterrichtsstunde zur Digitalisierung von Audiosignalen verwendet diese hochauflösten Signale als Ausgangsmaterial für Experimente mit Abtastung und Quantisierung. An Testsignalen wie Tönen, weißem Rauschen, Pulsen und anschließend an den eigenen Aufnahmen erfahren Schüler die Folgen von Abtastung, die hier durch Unterabtastung modelliert wird, und Quantisierung, modelliert durch Requantisierung mit geringerer Bitzahl. Schüler lernen die resultierende Signalveränderung sowohl mit dem Gehör zu erkennen, als auch mit technischen Mitteln auszuwerten.

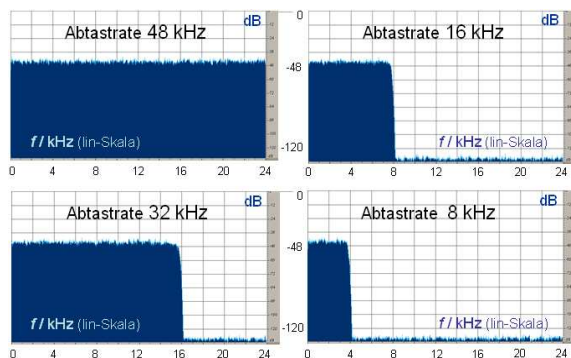


Abb.10: Abtastung modelliert mit Unterabtastung

Abbildung 10 zeigt die übertragene Bandbreite eines weißen Rauschens bei Verringerung der Abtastrate am Beispiel von 48, 32, 16 und 8 kHz Abtastung. Zunächst wird die Klangänderung mit dem Gehör beurteilt. Danach sollen die Schüler die übertragene Bandbreite quantitativ bestimmen und zusammen mit der Dateigröße des Signals nach Unterabtastung für alle verwendeten Abtastraten notieren. Aus der Erfahrung heraus wird das Abtasttheorem abgeleitet. Dabei wird auch näher auf den Grund zur Tiefpassfilterung, also die Verhinderung der sonst auftretenden spektralen Spiegelung (Aliasing), eingegangen.

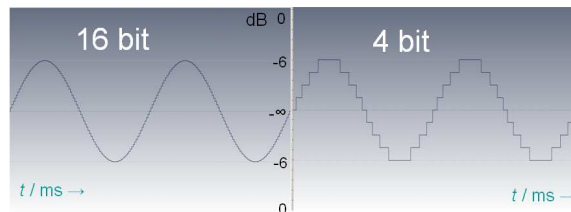


Abb.11: Quantisierung modelliert mit Requantisierung

Im nächsten Experiment lernen Schüler, dass Verwendung diskreter Aussteuerungswerte mit sich immer die Entstehung eines Rundungsfehlers trägt. Sie quantisieren die mit hoher Auflösung generierten und aufgenommenen Signale schrittweise neu mit 16, 12, 10, 8, 6, 4 und 2 Bit. Zu dem untersuchen sie bei gleichzeitiger Kontrolle mit dem Gehör sowohl den Zeitverlauf und Spektrum der requantisierten

Signale, als auch die resultierenden Dateigrößen. Im Zeitverlauf eines Tones in Abbildung 11 erkennt man die durch Requantisierung entstandenen diskreten Werte als äquidistant verteilte Stufen. Je weniger Stufen zur Verfügung stehen, desto größer ist der Rundungsfehler. Subtrahiert man den Zeitverlauf des requantisierten Signals von dem des Originals, bleibt der Zeitverlauf des Quantisierungsfehlers übrig. Das dazugehörige Spektrum weist bei harmonischen Signalen, wie dem in Abbildung 11, dann stark ausgeprägte, sich periodisch wiederholende Maxima auf. Harmonische Anteile in Sprache und Musik sind aber meistens nur quasiperiodisch und besitzen deswegen eine rauschartige Zufallskomponente. Und natürlich besteht Sprache nicht nur aus Vokalen und Musik nicht nur aus langen Tönen.

Konsonanten, Einschwingvorgänge, perkussive Klänge, die alle sind im Zeitverlauf stark nichtdeterministisch. In dem Falle hat auch der Zeitverlauf des Quantisierungsfehlers inhärent einen Zufallscharakter und ist im Spektrum eher gleichverteilt. Deswegen wird in der technischen Praxis oft anstatt von Quantisierungsfehler der Begriff Quantisierungsrauschen verwendet. In Experimenten entdecken Schüler, dass der Abstand zwischen dem Pegel des Signals und dem des entstandenen Quantisierungsfehlers näherungsweise 6 dB pro Bit beträgt. Dadurch lässt sich dieser für eine bestimmte Quantisierung mit hoher Zuverlässigkeit vorhersagen.

6. Gehörangepasste Audiokodierung

Schüler haben in vorheriger Lektion bereits erfahren, dass mit Unterabtastung und Requantisierung Audiodateien kleiner gemacht werden können, dafür aber hörbar "anders" als das Original klingen. Bei einer Dateigrößenreduktion um den Faktor 10 gegenüber dem CD-Standard (16 Bit, 44,1 kHz), beispielsweise bei 8 Bit, 8 kHz ist die resultierende Audioqualität dann bereits unterhalb der heutigen Anforderungen an Sprach- und Musikübertragung.

Wie macht man das also bei MP3, auf DVDs, beim digitalen DVB-T Fernsehen oder bei DAB Radio?

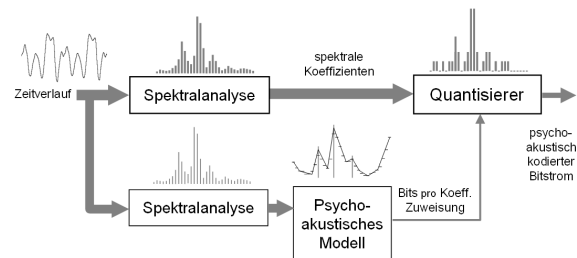


Abb.12: Blockdiagramm gehörangepasster Kodierung

Es ist weiterhin nichts anderes als Unterabtastung und Neuquantisierung, was in dem Koder geschieht, nur wird sie klüger angewandt. Der Trick liegt darin, das Signal gemäß seinen Eigenschaften selektiv im Spektrum zu quantisieren. Ein solcher sogenannter Sub-Band Koder ist in der Abbildung 12 schematisch dargestellt. Zuerst werden hier Zeitabschnitte

des Signals in ihre Spektralkomponenten zerlegt, genauso wie im Unterricht öfters zuvor bei Signalanalyse. Die Länge der Zeitabschnitte muss dabei die „Ohrenblick“-Länge respektieren und ebenso die auftretenden Vor- und Nachverdeckungsphänomene. Das resultierende Spektrum wird in schmalen Bändern weiterbearbeitet (daher Sub-Band). Um die Datenrate insgesamt zu bewahren, wird das Signal in den Teilbändern jeweils im reziproken Verhältnis zu ihrer Anzahl unterabgetastet. In jedem Sub-Band wird die Aussteuerungsskala auf die stärkste darin auftretende Spektralkomponente skaliert (Quasi-Float Darstellung). So wird in jedem Band die volle Aussteuerung genutzt. Der zugehörige Bandverstärkungsfaktor (Scale Factor) muss aber immer mitübertragen werden. Bis zu diesem Schritt wird das Signal nur mathematisch transformiert und kann ohne jeglichen Informationsverlust in seiner ursprünglichen Form wiederhergestellt werden.

Ein parallel arbeitendes mathematisches Modell des menschlichen Gehörs berechnet für jedes verwendete Frequenzband eine Mithörschwelle und teilt dem Quantisierer zu, mit wie viel Bits die im Band beinhalteten Spektralkomponenten quantisiert werden sollen, damit der Quantisierungsfehler in diesem Band nicht über die Mithörschwelle kommt und somit nicht hörbar würde. Auch für die Quantisierung im Spektrum gilt, dass bei einer Verringerung der Quantisierungsgenauigkeit um 1 Bit der Quantisierungsfehler um 6 dB ansteigt. Bei stark auftretender Maskierung reicht dann in den betroffenen Bändern eine Quantisierung der Spektrallinien mit 1 oder 2 Bit und sogar in wenig maskierten Teilen des Spektrums braucht man nur selten mehr als 6 Bit statt 16 Bit (CD-Audio) oder 24 Bit (Studiotonspur). Dazu kommt noch ein Ersparnis an Datenmenge durch die unter der Mithörschwelle liegenden Spektralkomponenten, die überhaupt nicht kodiert werden, ohne dass sich ihre Absenz bemerkbar macht.

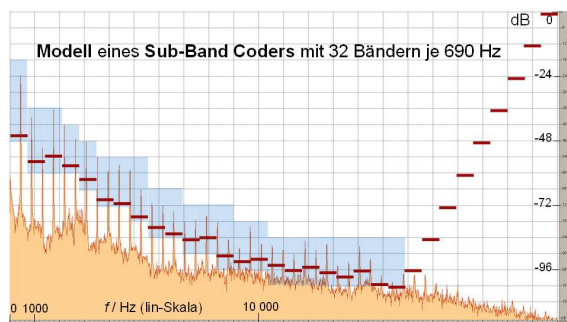


Abb.13: Selektive Quantisierung im Spektrum (Modell)

6.1. Kodierungsartefakte und ihre Ursachen

Zum Abschluss des Kurses untersuchen Schüler die Hörbarkeit von Kodierungsartefakten bei verschiedenen Einstellungen eines MPEG Koders. Sie verwenden als Quellensignale ihre eigenen Musik- und Sprachaufnahmen und lernen bei Hörtests verschiedene Arten auftretender Artefakte herauszuhören.

7. Literatur

- [1] *Organisation of the education system in the Czech Republic 2009/10*. European Commission - Education, Audiovisual & Culture Executive Agency, EURYDICE data base, URL: <http://eacea.ec.europa.eu/education/eurydice/eurybase_en.php> (Stand 3/2011)
- [2] MATHELITSCH, LEOPOLD; VEROVNIK, IVO: *Akustische Phänomene*. Praxis Schriftenreihe Physik, Aulis Verlag, Deubner&Co, Köln 2004, ISBN: 978-3-7614-2532-9
- [3] VOBKÜHLER, ADRIAN; NORDMEIER, VOLKHARD: *SOUNDS - Akustik mit der Soundkarte*. In: *Lehren und Lernen mit neuen Medien*. LIT Verlag, Münster 2006, Band 26, S. 102-104, ISBN 3-8258-9561-0
- [4] HASKELL, ROBERT E.: *Transfer of Learning: Cognition and Instruction*. Academic Press, New York 2001, ISBN: 978-0-1233-0595-4
- [5] NIEMIERKO, BOLESŁAW: *Taxonomies of Educational Goals as a Lead into Creative Teacher Training*. In: *Polish Journal of Social Science*, Vol. IV No. 1, Herausgeber: Mięczysław Plopa, Elbląg University of Humanities and Economy, Elbląg 2009, S.93-106, ISSN 978-8-3891-1319-8
- [6] VON HELMHOLTZ, HERMANN: *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik*. Verlag F. Vieweg, Braunschweig 1862. On-line auf: *The Virtual Laboratory*, Max-Planck-Institute for the History of Science, Berlin, ISSN 1866-4784, URL: <<http://vlp.mpiwg-berlin.mpg.de/>> (Stand 3/2011)
- [7] *The Official Web Site of the Nobel Prize*. © Nobel Media AB 2011, URL: <http://nobelprize.org/nobel_prizes/medicine/laureates/1961/> (Stand 3/2011)
- [8] ZWICKER EBERHARD, FASTL HUGO: *„Psychoacoustics: Facts and Models“*. Springer Verlag, Berlin, Heidelberg, New York 1999, ISBN 3-5406-5063-6
- [9] MOORE, BRIAN C. J.: *An Introduction to the Psychology of Hearing*. Academic Press, London, New York 1982, ISBN 0-1250-5622-2
- [10] PUJOL, RÉMY ET AL.: *Promenade around the cochlea*. Projektwebseite, Montpellier, France, URL: <www.cochlea.org> (Stand 3/2011)
- [11] BOSI, MARINA; GOLDBERG, RICHARD E.: *Introduction to Digital Audio Coding and Standards*. Springer Science + Business Media, New York 2003, ISBN 1-4020-7357-7
- [12] WATKINSON, JOHN: *The Art of Digital Audio*. Focal Press, Oxford 2004, ISBN 0-4020-7357-7
- [13] AICHINGER, DANIEL; ADAMI, ALEXANDER; BELGHAOUTI, ISABELLE: *From EUREKA dream to DAB reality*. In: *10th Workshop Digital Broadcasting*, Proceedings, Fraunhofer Verlag, Stuttgart 2009, S. 34-42, ISBN: 978-3-8396-0054-2