mp3 –
ein Beispiel für angewandte Mathematik im Alltag

Anton Schüller[1]
Ulrich Trottenberg[1,2]
Roman Wienands[2]
Felix Frankeser[2]
Gero Stoffels[2]

[1]**Fraunhofer-Institut**
Algorithmen und Wissenschaftliches Rechnen SCAI

[2] **Mathematisches Institut**
der Universität zu Köln

Version 1.1
09.11.2012

Inhaltsverzeichnis

1 Vorbemerkung: mp3 im Schulunterricht

Wer kennt nicht mp3-Player und mp3-Dateien. Mit ihnen ist es möglich, große Mengen von Musik auf kleinstem Raum zu speichern und wieder abzuspielen: denn mit mp3 kann man den Speicherplatz, den man benötigt, um eine Audiodatei zu speichern, auf einen Bruchteil reduzieren. So ist es auf modernen mp3-Playern mit der Größe einer Streichholzschachtel möglich, bis zu 200.000 Minuten Musik (das entspricht 130 Tagen) zu speichern [3].

Dass hinter mp3 jede Menge interessante Mathematik steht, ist in der breiten Öffentlichkeit weitgehend unbekannt geblieben. Wir wollen hier versuchen, einige der Prinzipien, auf denen mp3 basiert, näher zu beleuchten und allgemeinverständlich darzustellen. Dies umfasst biologische, physikalische, mathematische, informatische und musikalische Aspekte.

Da für Schüler/-innen mp3 heute zum selbstverständlichen Alltag gehört, stellt dieses Thema einen direkten Bezug zur Lebenswirklichkeit der Schüler/-innen dar. Daher werden sich mit diesem Thema auch Schüler/-innen angesprochen fühlen, die ansonsten der Mathematik und den Naturwissenschaften eher kritisch gegenüberstehen.

Dieses Skript, das auch eine Reihe von Arbeitsblättern enthält, die im Schulunterricht eingesetzt werden können, wird ergänzt durch eine Reihe beigefügter Materialien, die ebenfalls für den Einsatz im Schulunterricht geeignet sind.

Nach einer kurzen Einführung in Abschnitt 2 behandeln wir in den Abschnitten 3 und 4 die Grundlagen, auf denen mp3 beruht. Abschnitt 5 beschreibt einige Hörbeispiele, die diesem Skript als Audio-Video-Dateien beigefügt sind. Abschnitt 6 ist mathematischer Natur und behandelt das Zerlegen von Tonsignalen/Funktionen in Rechteckschwingungen. In Abschnitt 7 gehen wir auf das Prinzip der Huffman-Codierung ein und erläutern dies an einem Beispiel.

Abschnitt 8 enthält eine Beschreibung der diesem Skript beigefügten Materialien (Audio-Video-Präsentationen, Animationen, Bilder) sowie eine ganze Reihe von Arbeitsblättern für den Unterricht mitsamt Kommentaren und Lösungen.

Dieses Unterrichtsmaterial eignet sich für eine Unterrichtsreihe im Differenzierungsbereich Mathematik/Naturwissenschaft der Mittelstufe (Klassen 8 oder 9) oder im Rahmen einer Projektwoche ab Klasse 8. Es kann auch als Einführung in eine detailliertere Untersuchung zu mp3 in einem Projektkurs der Oberstufe eingesetzt werden.

Bei der Erstellung dieses Skriptes und des beigefügten Unterrichtsmaterials durften wir zurückgreifen auf ausführliche Unterlagen von Prof. Dr. Helmut Neunzert aus seinem Einführungsvortrag auf dem Kongress *Mathematik in der Praxis* im März 2009 [1]. Hierfür bedanken wir uns ganz herzlich.

Ganz herzlich danken wir auch Stefanie Frank und ihren Kolleginnen und Kollegen vom Fraunhofer-Institut für Integrierte Schaltungen (IIS) für die sorgfältige Überprüfung des Manuskripts und viele wertvolle Hinweise, Korrekturen und Verbesserungsvorschläge.

Die Entwicklung dieses Unterrichtsmoduls wurde unterstützt durch eine Projektförderung durch die WestLB-Stiftung Zukunft NRW, für die wir uns ebenfalls ganz herzlich bedanken.

2 Einführung

mp3 ist eine Abkürzung von *MPEG Audio Layer 3*, MPEG eine Abkürzung für die *Moving Picture Experts Group*, die 1988 von der ISO (International Standards Organisation) gegründet wurde, um einen internationalen Standard für die effiziente Kodierung von Videocodes zu entwickeln. Es galt Methoden zu finden, um große Mengen an Video- und Audiodaten so zu komprimieren, dass sie möglichst wenig Speicherplatz benötigen.

Die Anfänge der mp3 Entwicklung gehen auf die 1970er Jahre zurück. Damals begann Prof. Dieter Seitzer an der Friedrich-Alexander Universität Erlangen-Nürnberg mit der Forschung zur Übertragung von Sprache in hoher Qualität über Telefonleitungen. 1979 entwickelte das Team von Prof. Seitzer dann einen ersten digitalen Signalprozessor zur Audiocodierung. 1987 bildeten die Universität Erlangen-Nürnberg und das Fraunhofer-Institut für Integrierte Schaltungen (IIS) eine Forschungsallianz im Rahmen des EU-geförderten Projektes EU147 "EUREKA" für Digital Audio Broadcasting (DAB). Eine ausführliche Darstellung der Geschichte der mp3 Entwicklung findet man in [4].

Seinen heutigen Namen bekam mp3 im Jahr 1995 als Ergebnis einer internen Umfrage der Fraunhofer-Forscher am IIS. 1998 kamen die ersten mp3-Player auf den Markt. Heute gibt es in Deutschland mindestens 9000 Arbeitsplätze im mp3-Bereich (Handel, Hersteller von mp3-Playern) [2].

Die beigefügte Audio-Video-Datei audio_video_kompression.avi demonstriert gleichzeitig an je einem visuellen und auditiven Beispiel, wie sich die Qualität von Bildern und Musik verändert, wenn man die zugehörigen Daten immer weiter komprimiert. Zunächst bleibt die Qualität für die menschliche Wahrnehmung erhalten und man spart große Mengen an Speicherplatz, aber irgendwann werden Qualitätsverluste bemerkbar. Komprimiert man dann immer noch weiter, so verschlechtert sich die Qualität dramatisch.

Im Schulunterricht eignet sich diese Audio-Video-Datei als Demo zur Einführung in das Thema. Die gleichzeitige Kompression von Bild und Ton spricht mit Auge und Gehör zwei Sinne an und macht neugierig auf das, was dahinter steckt.

3 Prinzip von mp3 und Grundlagen

mp3 basiert auf dem Prinzip, dass die Anteile von einem Musikstück, die das menschliche Ohr nicht (explizit) hört, nur sehr viel gröber und ungenauer gespeichert werden müssen als die Anteile, die das menschliche Ohr gut wahrnehmen kann. Dies mag auf den ersten Blick ein wenig überraschend klingen. Hören wir denn nicht alles, was in einem Musikstück enthalten ist? Tatsächlich ist das menschliche Ohr nicht in der Lage, alle Details, die in einem Musikstück enthalten sind, wahrzunehmen. Hierauf gehen wir in Abschnitt 4 näher ein.

Vorher wollen wir uns aber vergegenwärtigen, was Geräusche, Töne und Schall eigentlich sind (vgl. hierzu auch [12], [13]).

Der Schall: Wieso können wir überhaupt das hören, was jemand sagt, der einige Meter von uns entfernt ist? Der Grund hierfür ist das physikalische Phänomen des Schalls. Schall entsteht, weil die Moleküle eines Mediums (z.B. Luft) zum Schwingen gebracht werden. Dadurch stoßen sie an benachbarte Moleküle, bringen auch diese ins

Schwingen usw. Solch eine (mechanische) Schwingung breitet sich in festen, flüssigen oder gasförmigen Stoffen wellenförmig aus [14] (vgl. die hieraus stammenden Animationen Simple_harmonic_motion_animation.gif und Onde_compression_impulsion_1d_30_petit.gif). Die Geschwindigkeit, mit der sich der Schall fortpflanzt, hängt von der Dichte des Mediums ab (zur Erinnerung: Dichte = Masse/Volumen). Für Luft liegt die Schallgeschwindigkeit bei etwa 333 m/s, für Wasser liegt sie bei rund 1400 m/s. Im Vakuum kann sich Schall nicht ausbreiten.

Die Tonhöhe: Entsteht beispielsweise ein Ton dadurch, dass eine Gruppe von Molekülen ganz regelmäßig hin und her schwingt, beispielsweise 400 mal pro Sekunde, so sagen wir auch, der Ton hat eine Frequenz von 400 Hertz, d.h. die Schwingung erfolgt 400 mal pro Sekunde. Das menschliche Ohr kann nur Töne wahrnehmen, die zwischen etwa 16 Hertz und 20.000 Hertz liegen, bei denen die Anzahl der Schwingungen pro Sekunde also zwischen 16 und 20.000 liegt. Dabei reduziert sich die obere Grenze mit dem Alter deutlich. Je höher die Frequenz (also die Anzahl der Schwingungen pro Sekunde) ist, desto höher klingt auch der Ton. Eine Verdopplung der Frequenz entspricht dabei in der Musik genau einer Erhöhung eines Tones um eine Oktave.

Hunde können deutlich höhere Frequenzen hören. Dies machen sich Hundepfeifen zunutze. Mit ihnen kann der Besitzer seinen Hund zurückpfeifen, ohne dass andere Menschen dieses Pfeifen wahrnehmen.

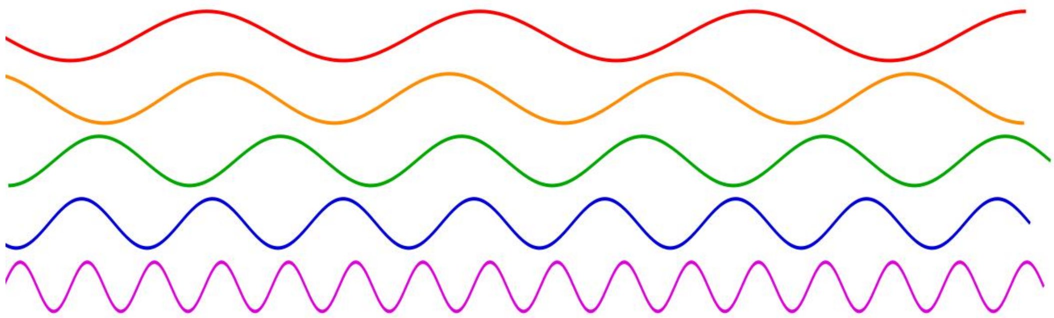


Abbildung 1: Schwingungen unterschiedlicher Frequenz: je kürzer die Länge einer Schwingungsperiode (Wellenlänge), desto höher ist die Frequenz und desto höher klingt der zugehörige Ton.

Zusammenhang zwischen Wellenlänge, Frequenz und Schallgeschwindigkeit: Ist c die Schallgeschwindigkeit in einem Medium, f die Frequenz einer Schallwelle (d.h. einer sich wellenförmig ausbreitenden Schwingung) und λ (sprich *lambda*) die Wellenlänge, so gilt

$$c = \lambda \cdot f \quad . \tag{1}$$

Sind zwei dieser drei Größen bekannt, so kann man die dritte hiermit berechnen.

Die Lautstärke: Je weiter die Moleküle in der Luft hin und her schwingen, desto lauter ist der Ton. Die Lautstärke beschreibt also den Unterschied zwischen Berg und Tal der Schwingung.

Der Begriff des Tons in der Physik und in der Musik: In der Physik bezeichnet man mit dem Begriff Ton eine regelmäßige (Sinus-)Schwingung mit genau einer Frequenz. Alles, was wir normalerweise hören, ist aber bereits aus einer Vielzahl von Schwingungen zusammengesetzt, die sich überlagern. Auch bei zwei Sängern, die denselben (musikalischen) Ton singen, unterscheiden sich die Stimmen durch die charakteristischen Klangfarben der Sänger; jede Stimme hat ihre eigene Klangfarbe. Physikalisch ist solch ein ge-

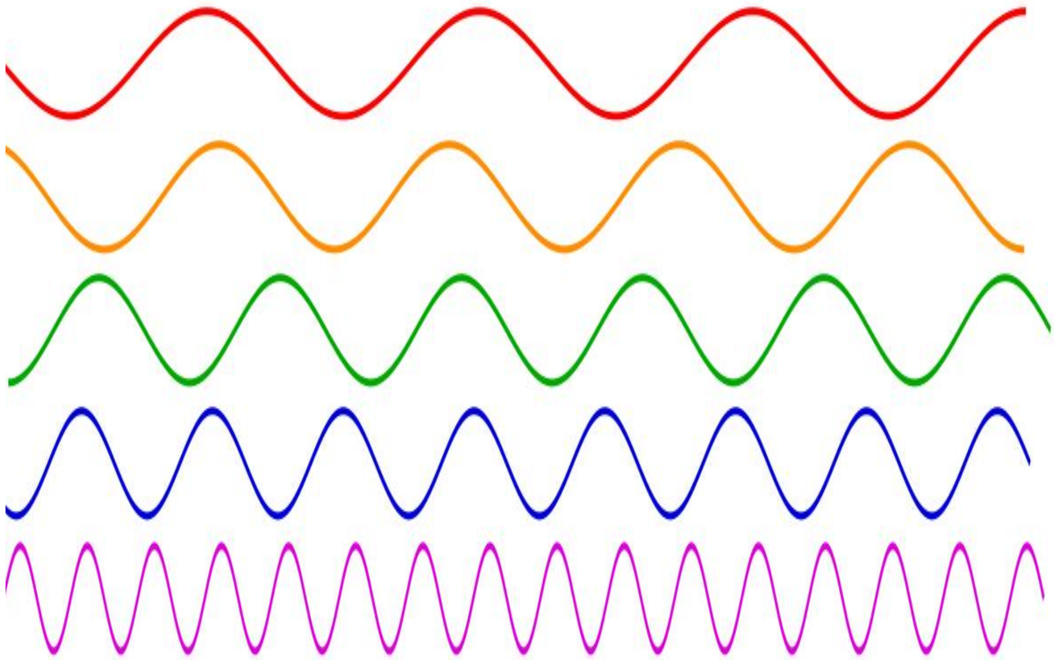


Abbildung 2: Die gleichen Schwingungen wie in Abb. 1, aber mit doppelter Lautstärke (doppelter Amplitude).

sungener Ton bereits kein reiner Ton mehr, sondern eine Überlagerung von regelmäßigen (Sinus-)Schwingungen vieler unterschiedlicher Frequenzen. Trotzdem erkennt man, dass beide Sänger denselben Ton singen. Das menschliche Gehör kann also dem gesungenen (musikalischen) Ton eine Höhe zuordnen. Diese gehörte Tonhöhe entspricht der Frequenz der dominanten (am lautesten klingenden) Frequenz.

Auch unterschiedliche Instrumente haben einen ganz unterschiedlichen Klang, z.B. Gitarre und Klavier, obwohl bei beiden Instrumenten Saiten zum Klingen gebracht werden. Auch hier ist jeder Ton der Instrumente aus vielen Schwingungen unterschiedlicher Frequenzen zusammengesetzt und diese Zusammensetzung unterscheidet sich bei unterschiedlichen Instrumenten. Abb. 3 visualisiert am Beispiel eines Tons eines Saxophons, wie eine solche Schallwelle eines Instruments aussehen kann.

Geräusche: Geräusche haben keine exakt bestimmbare Tonhöhe mehr. Sie sind nichtperiodische Schallereignisse, die durch Überlagerungen vieler Schwingungen unterschiedlicher Frequenz mit rasch wechselnder Amplitude entstehen. Mit anderen Worten: "Der Unterschied von Ton/Klang zu Geräusch ist in der Regelmäßigkeit der Schwingung zu finden. Bei einem Geräusch ist die Schwingbewegung der Luft sehr ungleichmäßig, bei Tönen dagegen handelt es sich um immer wiederkehrende gleichförmige Luftbewegungen" [17].

Anregungen für den Unterricht: Im Internet finden sich eine Reihe von Programmen, die im Unterricht gewinnbringend eingesetzt werden können. Ein Beispiel hierfür ist das Programm *Scope* [6], das einerseits Töne bestimmter Frequenz generieren, andererseits aber auch die Schallwellen unterschiedlicher musikalischer Töne und Klänge sichtbar machen kann:

- So kann man beispielsweise mit dem in Scope eingebauten Signalgenerator Töne unterschiedlicher Frequenz generieren, und damit lassen sich Hörgrenzen im Frequenzbereich individuell feststellen. (Voraussetzung hierfür ist allerdings, dass die verwendeten Lautsprecher diese Frequenzen auch unterstützen; insbesondere im Bereich niedriger Frequenzen erreichen viele Lautsprecher den Grenzbereich des menschli-

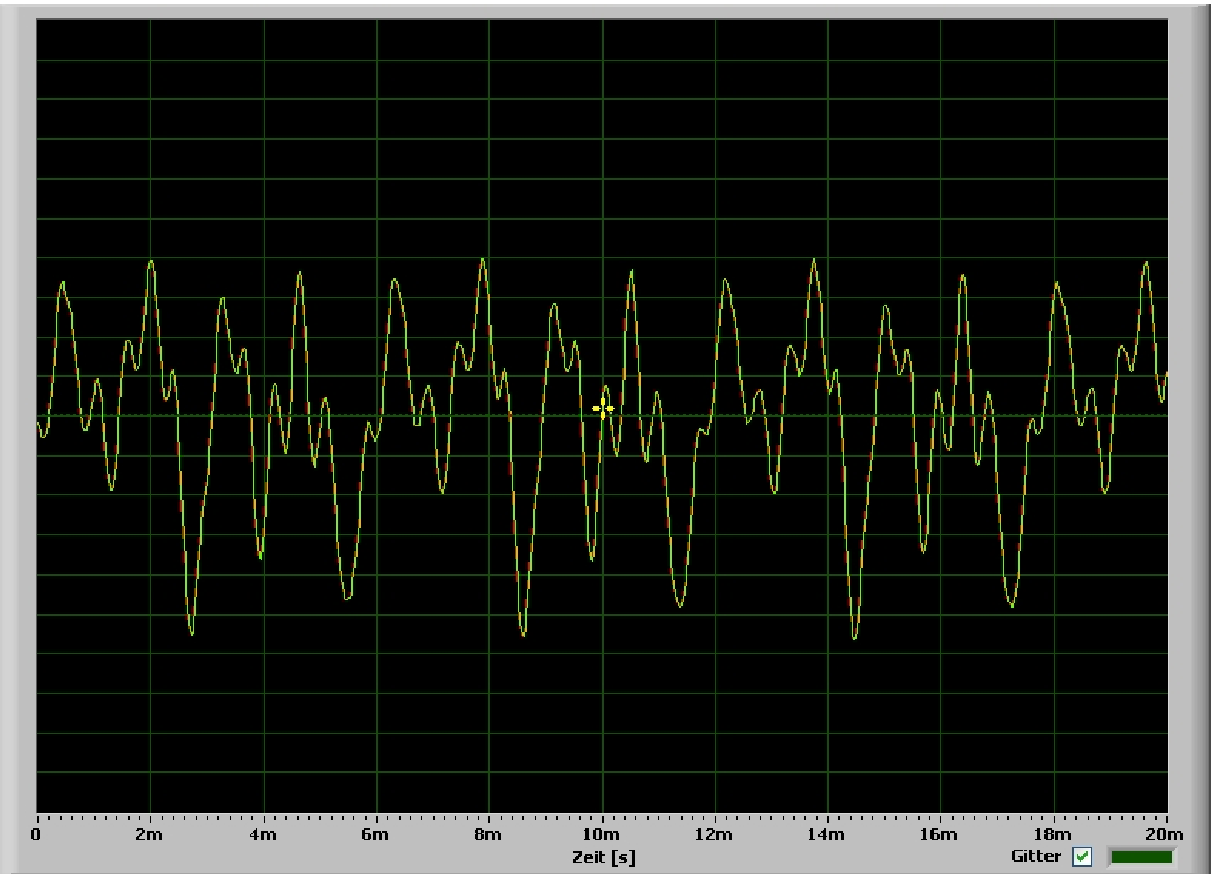


Abbildung 3: Schallwelle eines Tons eines Saxophons (Grafik wurde mit dem Programm Scope [6] erzeugt).

chen Gehörs jedoch nicht.)

- Variiert man den Frequenzbereich der mit dem Signalgenerator erzeugten Töne, so stellt man fest, dass niedrige und sehr hohe Frequenzen leiser klingen als mittlere, auch wenn man die Lautstärke (Amplitude) nicht ändert. Dies entspricht der Tatsache, dass das menschliche Gehör Töne unterschiedlicher Frequenz unterschiedlich gut wahrnehmen kann. Sehr tiefe oder sehr hohe Töne sind für den Menschen viel schlechter hörbar als Töne im mittleren Frequenzbereich (vgl. Abb. 4 und Abschnitt 4).

- Mit dem in Scope eingebauten Oszilloskop kann man Schallwellen sichtbar machen. Interessant ist hier zum Beispiel der Vergleich der von einer Stimmgabel und verschiedenen Instrumenten oder menschlichen Stimmen erzeugten Schallwellen (vgl. Abb. 3). Während die Stimmgabel (bei weichem Anschlag) eine nahezu perfekte Sinuswelle erzeugt, sieht dies bei Instrumenten oder der menschlichen Stimme teilweise anders aus. Die Schüler/-innen können Instrumente mit in den Unterricht bringen, um die Schallwellen, die das Instrument produziert, zu visualisieren. Dabei zeigen sich erhebliche Unterschiede in der Charakteristik und der Komplexität der erzeugten Schallwellen. Dies entspricht den unterschiedlichen Klangfarben der Instrumente.

4 Was hört das menschliche Ohr – und was kann es nicht hören?

Wie hören wir Töne und Musik? Alles, was wir hören, besteht aus Überlagerungen von Schwingungen, die sich in einem Medium wie der Luft wellenförmig ausbreiten. Diese wellenförmige Ausbreitung bedeutet physikalisch gesehen, dass das menschliche Ohr Druckschwankungen wahrnimmt, die aus einer Überlagerung von Schwingungen unterschiedlichster Frequenzen resultieren. Diese Druckschwankungen führen zu einem entsprechenden Schwingen des Trommelfells. Das menschliche Ohr ist wiederum im Stande, dieses Schwingen des Trommelfells über Sinneshaare im Innenohr, die auf unterschiedliche Frequenzen spezialisiert sind, in einzelne Tonfrequenzen zu zerlegen und als Nervenreize an das Gehirn weiterzuleiten [5]. Diese werden dann vom Gehirn als Töne, Klänge und Geräusche interpretiert.

Dem menschlichen Ohr sind hierbei allerdings Grenzen gesetzt:

- Wie bereits in Abschnitt 3 dargelegt, können auch sehr gut hörende Menschen nur Töne zwischen 16 und 20.000 Hertz wahrnehmen. Die obere Grenze nimmt dabei mit dem Alter deutlich ab.

- Abb. 4 zeigt, welche Lautstärke ein Ton einer bestimmten Frequenz haben muss, damit er vom menschlichen Ohr überhaupt wahrgenommen werden kann. Unterhalb dieser *Hörschwelle* (oft auch *Ruhehörschwelle* genannt) ist der Ton für das menschliche Gehör nicht hörbar. Die Abbildung zeigt zudem, dass diese Ruhehörschwelle von der Frequenz des Tones abhängt. Im Bereich von 2 bis 5 kHz (2000 bis 5000 Hertz) reichen schon sehr niedrige Lautstärken aus; hier ist das menschliche Ohr besonders empfindlich. Andererseits muss die Lautstärke bei sehr niedrigen und bei sehr hohen Frequenzen recht hoch sein, damit der Ton für den Menschen überhaupt hörbar ist.

  Musik und Sprache können nur in noch kleineren Frequenz- und Lautstärkebereichen wahrgenommen werden.

- Abb. 5 macht folgende Eigenschaft des menschlichen Gehörs deutlich: Kommt in einem Klang ein Ton mit einer bestimmten Frequenz und einer bestimmten Lautstärke vor (in Abb. 5 ein Ton mit einer Frequenz von 1 kHz = 1000 Hz), so sind benachbarte Frequenzen nur hörbar, wenn sie eine bestimmte Lautstärke (die sogenannte *Mithörschwelle*) überschreiten. Die Hörschwelle wird also in einer Umgebung dieser Frequenz deutlich angehoben. Mit anderen Worten: Eine hohe Lautstärke einer bestimmten Frequenz überdeckt Tonsignale von benachbarten Frequenzen, deren Lautstärke unterhalb der angehobenen Hörschwelle (der Mithörschwelle) liegt.

  Ein Beispiel für diesen *Maskierungseffekt* stellt eine Orgel dar. Wenn sie angeschaltet wird, aber noch keine Töne erklingen, hört man das Atmen der Orgel, d.h. das Gebläse, das für den Luftstrom in der Orgel verantwortlich ist. Wenn die Orgel dann gespielt wird, hört man dieses Geräusch nicht mehr, weil es von der (lauteren) Musik überdeckt (maskiert) wird.

- Vor oder nach sehr lauten Geräuschen kann der Mensch für kurze Zeit leisere Geräusche schlechter oder gar nicht wahrnehmen. Dies kann man mit einem einfachen Versuch überprüfen, wenn man z.B. einen alten Wecker besitzt, dessen Ticken man hören kann. Lässt man den Wecker klingeln, so hört man einerseits während des Klingelns das Ticken des Weckers nicht, obwohl das Ticken weitergeht (dies ist ganz analog zum Beispiel der Orgel wieder eine Maskierung im Frequenzbereich). Hinzu kommt aber, dass man auch kurz nach dem Klingeln des Weckers das Ticken des Weckers nicht hört (obwohl das Ticken ununterbrochen weitergeht). Dies zeigt, dass es neben der oben erläuterten Maskierung durch laute Frequenzen auch eine Verdeckung oder Maskierung im Zeitbereich gibt. Nach lauten Tönen braucht das Ohr eine gewisse Zeit, um leise Geräusche wieder wahrzunehmen.

Die Idee hinter mp3: mp3 macht sich zunutze, dass die akustischen Informationen, die das menschliche Ohr nicht explizit wahrnehmen kann, nicht exakt abgespeichert werden müssen. Ein sogenannter *Audiocodec* speichert unhörbare Teile eines Tonsignals ungenauer als die hörbaren und kann so den erforderlichen Speicherplatz reduzieren.

Um diese Idee ausnutzen zu können, werden die (musikalischen) Töne wieder in die Einzelfrequenzen zerlegt, aus denen sie zusammengesetzt sind. Dabei werden alle Frequenzwerte nicht exakt, sondern ungenauer gespeichert, um Daten einzusparen. Die Genauigkeit wird dabei aber nur so weit reduziert, dass die entstehenden Fehler für das menschliche Ohr nicht wahrnehmbar sind (also unter der Mithörschwelle liegen). Um die notwendige Genauigkeit zu identifizieren, wird ein sogenanntes *psychoakustisches Modell* verwendet, das

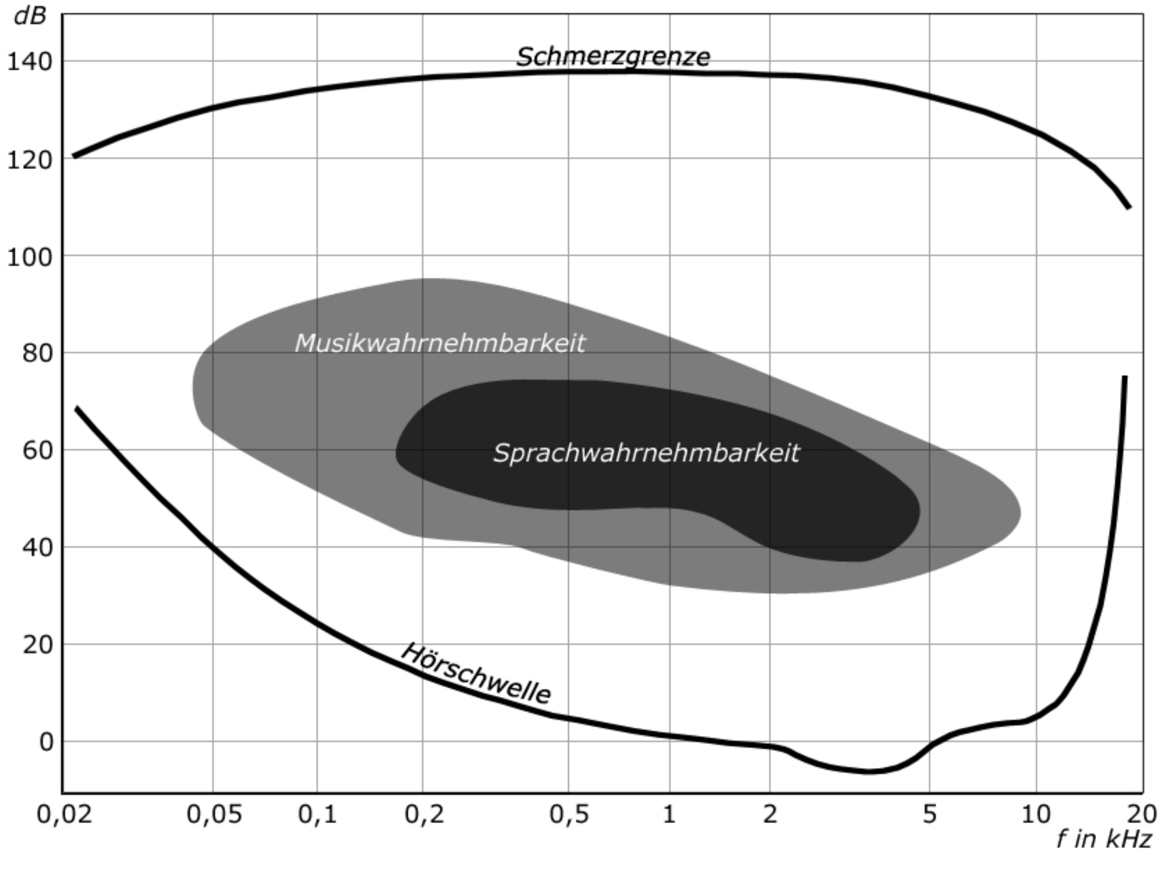


Abbildung 4: Grenzen des menschlichen Gehörs: Hörschwelle, Schmerzgrenze, Musik- und Sprachwahrnehmbarkeit in Abhängigkeit von der Frequenz (nach rechts ist die Frequenz und nach oben die Lautstärke (in der Maßeinheit "Dezibel") aufgetragen. (Man beachte dabei, dass "Dezibel" eine logarithmische Maßeinheit ist. Wegen $\log 1 = 0$ bedeutet 0 Dezibel gerade *nicht*, dass völlige Stille herrscht.)

gerade die Eigenschaften des menschlichen Gehörs widerspiegelt.

Halten wir fest: Für mp3 müssen also die Tonsignale wieder in die einzelnen Frequenzen zerlegt werden, aus denen sie zusammengesetzt sind. Die für das menschliche Ohr hörbaren Anteile müssen nur mit der nötigen Genauigkeit gespeichert werden. Unhörbare Anteile können ungenauer gespeichert werden.

Bemerkung 4.1 *Tatsächlich gibt es noch eine ganze Reihe weiterer Eigenschaften des menschlichen Gehörs, die man nutzen kann, um die Menge der notwendigen Informationen beim Speichern eines Musikstücks weiter zu reduzieren (vgl. hierzu z.B. [8]), ohne dass dies für das menschliche Ohr wahrnehmbar ist. Auch diese Möglichkeiten werden von mp3 genutzt; wir wollen hier aber nicht detaillierter auf diese Eigenschaften eingehen.*

Anregungen für den Unterricht: Auch zu den Inhalten in diesem Abschnitt lässt sich das bereits in Abschnitt 3 erwähnte Programm *Scope* [6] m Unterricht gewinnbringend einsetzen:

- Variiert man den Frequenzbereich der mit dem Signalgenerator erzeugten Töne, so stellt man fest, dass niedrige und sehr hohe Frequenzen leiser klingen als mittlere, auch wenn man die Lautstärke (Amplitude) nicht ändert. Damit ist für die Schüler/-innen unmittelbar erfahrbar, dass das menschliche Gehör Töne unterschiedlicher Frequenz unterschiedlich gut wahrnehmen kann. Sehr tiefe oder sehr hohe Töne sind für den Menschen viel schlechter hörbar als Töne im mittleren Frequenzbereich (vgl. Abb. 4).

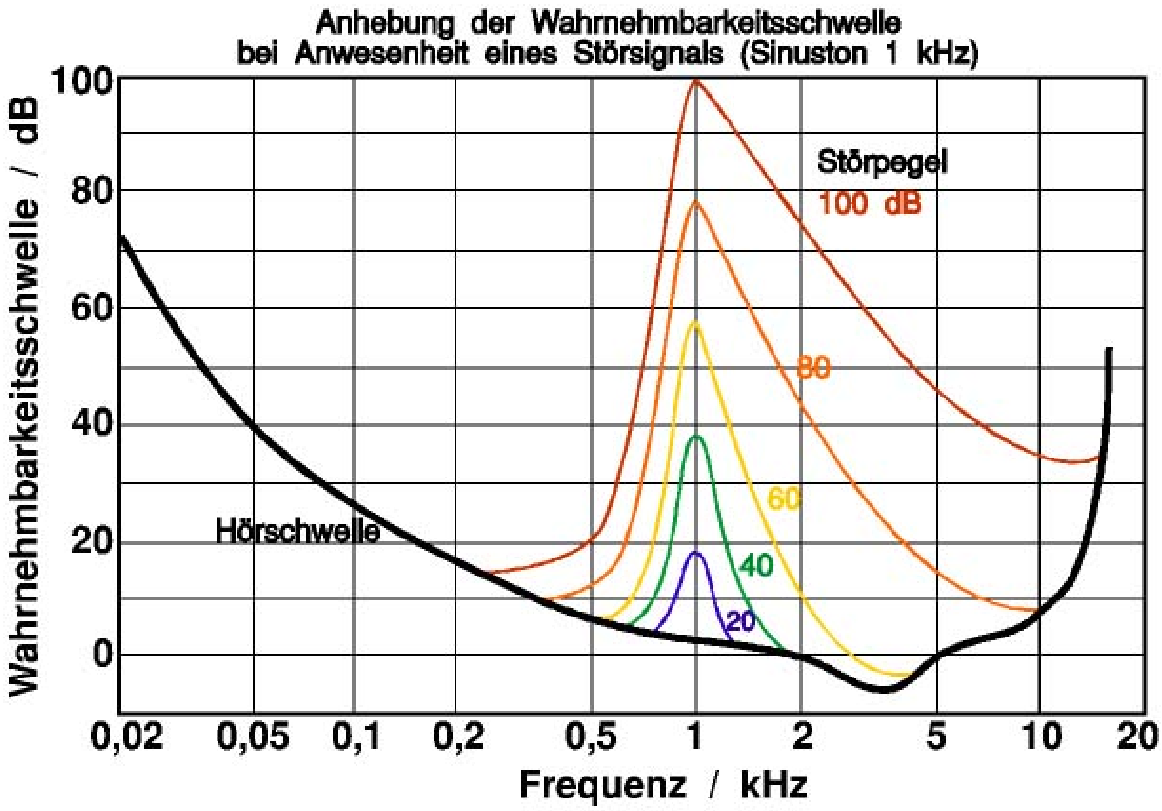


Abbildung 5: Grenzen des menschlichen Gehörs: Anhebung der Hörschwelle durch die Anwesenheit von Tönen mit einer Frequenz von 1 kHz und verschiedenen Lautstärken.

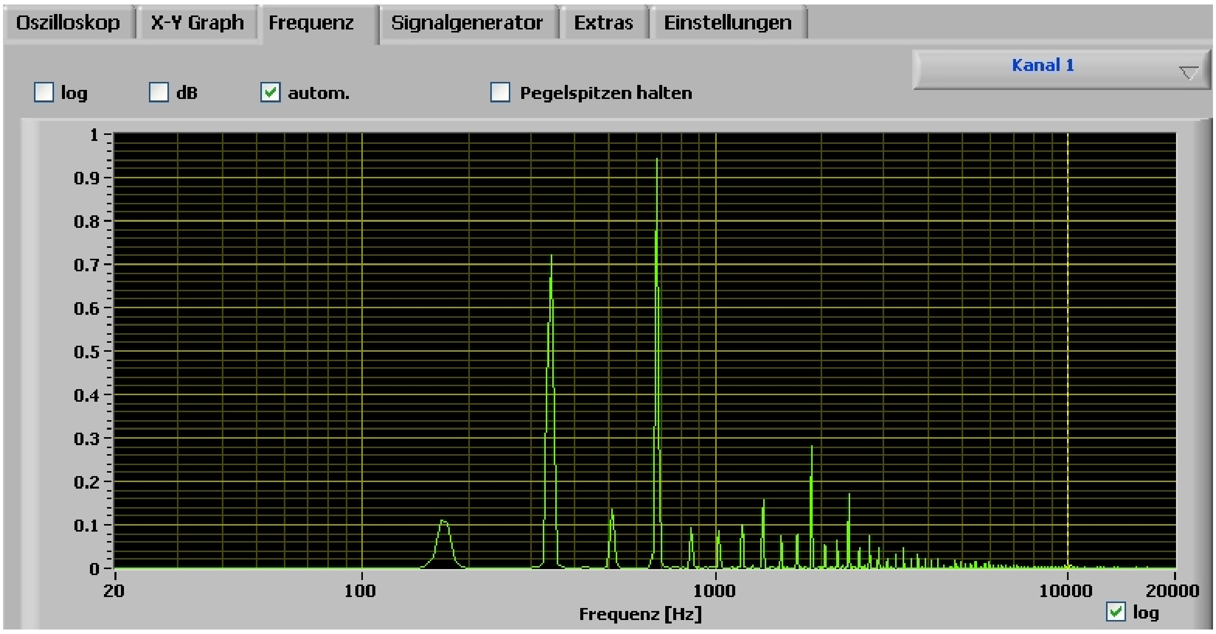


Abbildung 6: Frequenzanalyse eines Tons eines Saxophons (Grafik wurde mit dem Programm Scope [6] erzeugt).

- Das Programm *Scope* besitzt auch eine Frequenzanalyse, mit der man die Frequenzen (d.h. die physikalischen Töne oder Sinus-Schwingungen, die in einem musikalischen Ton enthalten sind) feststellen kann. Abb. 6 zeigt das Frequenzspektrum der in Abb. 3 dargestellten Schallwelle eines Tons eines (Tenor-) Saxophons. Offenbar handelt es sich hier um eine Überlagerung einer Vielzahl physikalischer Töne. Je nach Musikinstrument und Klang kann ein derartiges Frequenzspektrum auch einfacher oder komplizierter aussehen. Auch hier können die Schüler/-innen den Klang ihrer eigenen Instrumente oder ihrer eigenen Stimme analysieren. Bei männlichen Singstimmen ist beispielsweise beim Übergang in die Kopfstimme häufig eine deutliche Veränderung des Frequenzspektrums zu beobachten, wenn die Brustresonanzen wegfallen.

- Interessant ist auch, wie das Frequenzspektrum bei unterschiedlichen auf einem Ton gesprochenen oder gesungenen Vokalen aussieht. Auch diese Unterschiede kann das Ohr offensichtlich wahrnehmen und eindeutig Buchstaben zuordnen.

5 Hörbeispiele für mp3

Die Videoclips musikbeispiel_orig.avi, musikbeispiel_mp3.avi und musikbeispiel_orig_minus_mp3.avi demonstrieren, wie mp3 funktioniert:

- Der Videoclip musikbeispiel_orig.avi enthält ein kurzes Musikstück. Gleichzeitig sind auch die unterschiedlichen Frequenzen, die jeweils zu hören sind (grün gezeichnet), zu sehen.

- Der Videoclip musikbeispiel_mp3.avi enthält dasselbe Musikstück als mp3-Version. Die Menge der in mp3 gespeicherten Informationen ist grün gekennzeichnet. Weiß ist die Menge an Informationen, die mp3 im Gegensatz zur Originalversion nicht zu speichern braucht. Ganz offensichtlich ist das der weit überwiegende Teil der Informationen. Obwohl dieser Teil jetzt fehlt, ist kein Unterschied zur im Videoclip musikbeispiel_orig.avi enthaltenen Originalversion hörbar.

- Der Videoclip musikbeispiel_orig_minus_mp3.avi enthält gewissermaßen die Differenz der beiden vorgenannten Dateien. Er macht den Unterschied beider Versionen deutlich und macht das hörbar, was in der mp3-Version verlorengegangen ist, oder, mit anderen Worten, das, was in der Originalversion enthalten, für das menschliche Ohr aber nicht wahrnehmbar ist, weil es in der Originalversion durch andere dominantere Töne überdeckt wurde. Fallen diese Töne weg, so bleibt etwas Hörbares übrig. Tatsächlich besteht dieser "Rest" aus dem Großteil der in der Originalversion enthaltenen Informationen.

6 Zerlegung von Tonsignalen in Einzelfrequenzen: Multiskalenanalyse

Musikalische Töne bestehen aus einer Überlagerung einer Vielzahl von Schwingungen. Wie in Abschnitt 3 erläutert, sind nur die Schwingungen mit Frequenzen zwischen etwa 20 und 20.000 Hertz für den Menschen hörbar. Der Faktor zwischen den niedrigsten und den höchsten hörbaren Frequenzen beträgt damit immerhin $1.000 = 10^3$, also 3 Zehnerpotenzen. Wenn wir also musikalische Töne wieder in die darin enthaltenen Einzelfrequenzen zerlegen wollen, müssen wir ganz unterschiedliche Frequenz-Skalen betrachten. Da die Frequenzen in einem bestimmten Medium wie der Luft in direktem Zusammenhang mit den zugehörigen Wellenlängen stehen (vgl. Gleichung (1)), können wir ganz analog auch sagen, wir müssen ganz unterschiedliche Skalen von Wellenlängen betrachten.

Eine derartige *Multiskalenanalyse* ist durchaus nicht ungewöhnlich, wenn man die Eigenschaften von Objekten beobachten oder analysieren will.

Bemerkung 6.1 *Multiskalenanalysen sind auch in ganz anderen Anwendungsbereichen oft sehr hilfreich, um Eigenschaften von Objekten zu analysieren oder um bestimmte Problemstellungen zu lösen. Als ein solches Beispiel sei hier nur die Lösung großer Gleichungssysteme mit Millionen von Unbekannten durch sogenannte Multi-Level-Verfahren genannt, in denen auf jedem Level Fehler einer bestimmten Skala separat behandelt werden.*

Man kann (musikalische) Töne, d.h. Tonsignale, die durch Überlagerung von Schwingungen unterschiedlicher Frequenz entstanden sind, durch die mathematische Methode der *Fouriertransformation* wieder in ihre Einzelfrequenzen zerlegen. Das Verständnis der Fouriertransformation setzt jedoch mathematische Hilfsmittel voraus, die in der Schule in aller Regel noch nicht bekannt sind (uneigentliche Integrale, komplexe Zahlen, ...). Für mp3 wird eine Variante der Fouriertransformation eingesetzt, die Modifizierte Diskrete Cosinustransformation (MDCT) und die Polyphasen-Quadratur-Filterbank (PQF).

Wir wollen hier nur das Prinzip einer solchen Multiskalenanalyse deutlich machen. Dabei zerlegen wir die Tonsignale nicht in Sinusschwingungen, sondern in Rechteckschwingungen unterschiedlicher Frequenzen. Hierfür benötigen wir nur elementare Schulmathematik. Dazu gehen wir folgendermaßen vor: In Abschnitt 6.1 führen wir eine Multiskalenanalyse durch fortgesetzte Mittelwertbildung für eine gegebene Zahlenfolge durch. In Abschnitt 6.2 betrachten wir eine Zerlegung eines Tonsignals in sogenannte *Wavelets*, was der Zerlegung des Tonsignals in *Rechteckschwingungen* entspricht (vgl. Abb. 7). Im Unterricht kann man diese Reihenfolge auch umkehren oder den ersten Teil weglassen.

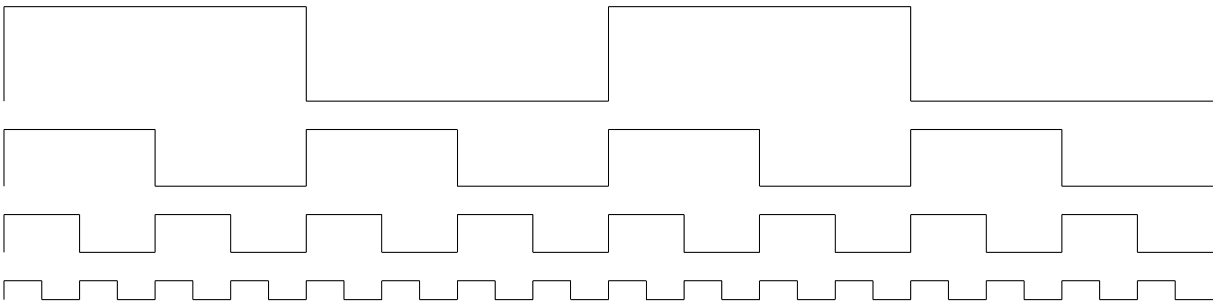


Abbildung 7: Rechteckschwingungen unterschiedlicher Amplitude und Frequenz

6.1 Multiskalenanalyse einer Zahlenfolge durch fortgesetzte Mittelwertbildung

Wir betrachten als Beispiel folgende Zahlenfolge von Quadratzahlen: 0 1 4 9 16 25 36 49. Fassen wir die Zahlen in Paare zusammen und bilden die Mittelwerte dieser Paare, so erhalten wir die Folge $0,5\ \ 6,5\ \ 20,5\ \ 42,5$. Fassen wir diese Zahlen ebenfalls wieder zu Paaren zusammen und bilden die Mittelwerte der Paare, so erhalten wir die Folge $3,5\ \ 31,5$. Für dieses Zahlenpaar haben wir den Mittelwert 17,5. Wir haben jetzt die ursprüngliche Zahlenfolge in mehrere Skalen von Mittelwerten überführt:

Zahlenfolge	0	1	4	9	16	25	36	49
feine Skala von Mittelwerten	0,5		6,5		20,5		42,5	
mittlere Skala von Mittelwerten	3,5				31,5			
grobe Skala von Mittelwerten	17,5							

Um von einer Mittelwertskala wieder zur vorhergehenden zu gelangen, benötigen wir die Abweichungen der Mittelwerte von den zugehörigen Werten auf der vorherigen Skala:

$$17,5 - 14 = 3,5 \ \text{ bzw. } \ 17,5 + 14 = 31,5 \ .$$

Entsprechend auf der nächstgröberen Skala:

$$3,5 - 3 = 0,5 \quad 3,5 + 3 = 6,5 \quad 31,5 - 11 = 20,5 \quad 31,5 + 11 = 42,5 \ .$$

Ganz analog können wir auch von der feinen Skala von Mittelwerten zu unserer ursprünglichen Folge zurückkommen:

$$\begin{array}{cccc} 0,5 - 0,5 = 0 & 0,5 + 0,5 = 1 & 6,5 - 2,5 = 4 & 6,5 + 2,5 = 9 \\ 20,5 - 4,5 = 16 & 20,5 + 4,5 = 25 & 42,5 - 6,5 = 36 & 42,5 + 6,5 = 49 \end{array}$$

Die gröbste Skala von Mittelwerten und diese Abweichungen können wir uns wie in folgendem Schema merken. Wir haben zusätzlich die ursprüngliche Zahlenfolge nochmals dazu geschrieben:

Zahlenfolge	0	1	4	9	16	25	36	49
feinste Skala von Abweichungen	-0,5	+0,5	-2,5	+2,5	-4,5	+4,5	-6,5	+6,5
zweitfeinste Skala von Abweichungen	-3		+3		-11		+11	
zweitgröbste Skala	-14				+14			
gröbste Skala	17,5							

Zu den ursprünglichen Zahlen zurück kommen wir jetzt, indem wir den Mittelwert auf der gröbsten Skala und die entsprechenden gespeicherten Abweichungen auf allen feineren Skalen einfach addieren.

Ein Beispiel:

Zahlenfolge	0	1	4	9	16	25	36	49
feinste Skala von Abweichungen	-0,5	+0,5	-2,5	+2,5	-4,5	+4,5	-6,5	+6,5
zweitfeinste Skala von Abweichungen	-3		+3		-11		+11	
zweitgröbste Skala	-14				+14			
gröbste Skala	17,5							

$\boxed{17,5 + 14 - 11 + 4,5 = 25}$

Dies bedeutet, statt uns eine Zahlenfolge zu merken, können wir uns auch die Mittelwerte auf der gröbsten Skala und die Abweichungen von den Mittelwerten auf allen Skalen merken. Mit diesen Informationen können wir die ursprüngliche Zahlenfolge reproduzieren. Man beachte, dass die Abweichungen von einem Mittelwert sich jeweils nur im Vorzeichen unterscheiden (-3 und $+3$ oder -11 und $+11$ oder $-6,5$ und $+6,5$). Dies entspricht – als Funktion aufgefasst – bereits einer Rechteckschwingung. Dies werden wir im nächsten Abschnitt noch einfacher sehen.

6.2 Zerlegen von Tonsignalen in Rechteckschwingungen unterschiedlicher Frequenzen

Um Tonsignale in Rechteckschwingungen unterschiedlicher Frequenzen zu zerlegen, können wir ganz analog vorgehen. Abb. 8 zeigt links eine Funktion, die wir in Rechteckschwingungen zerlegen wollen. Da wir den Funktionsverlauf in der Praxis oft nicht genau kennen, sondern nur an bestimmten Werten messen, nähern wir die Funktion durch die einzelnen Messwerte an. Diese Messwerte werden durch die gefärbten Balken wiedergegeben.

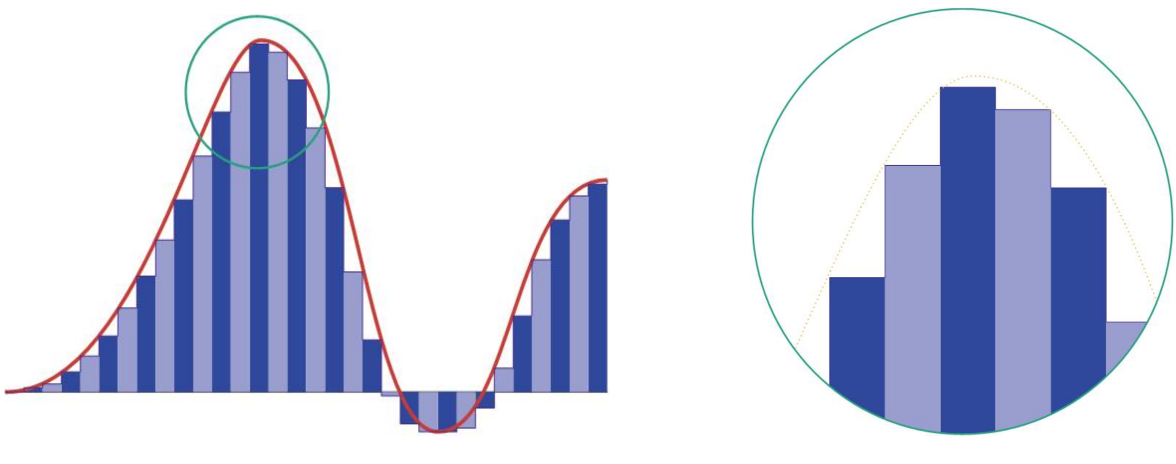


Abbildung 8: Zu analysierende Funktion im Ganzen und ein Ausschnitt in Großansicht

Auf der rechten Seite der Abbildung ist der umkreiste Ausschnitt der Funktion vergrößert dargestellt. Wir erläutern das Prinzip unserer Multiskalenanalyse im Folgenden anhand dieses Ausschnitts.

Die linke Skizze in Abb. 9 zeigt, dass wir wie im vorangegangenen Abschnitt bei der Multiskalenanalyse der Zahlenfolge wieder Mittelwerte der gemessenen Funktionswerte bilden, um auf die nächstgröbere Skala zu kommen. Zurück zur feinen Skala der Funktion können wir wieder kommen, indem wir wieder die Abweichungen zum Mittelwert hinzu addieren.

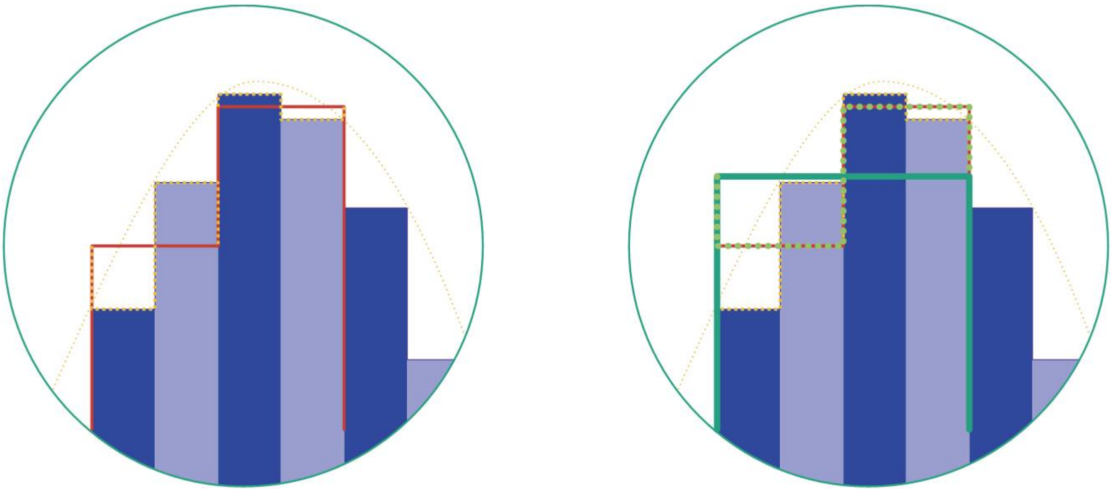


Abbildung 9: Multiskalenanalyse durch Mittelwertbildung: feine Skalen (links) und grobe Skala (rechts)

Dieses Vorgehen können wir wie bei unserer Zahlenfolge im vorangegangenen Abschnitt fortsetzen, um auf gröbere Skalen zu kommen. Die rechte Seite von Abb. 9 zeigt den Mittelwert auf der entsprechenden nächstgröberen Skala.

Abb. 10 zeigt, wie wir durch Addition des Mittelwerts auf der gröbsten Skala und den Abweichungen der Mittelwerte auf den feineren Skalen wieder zu unseren ursprünglichen Messwerten der Funktion zurückkommen. Betrachten wir jetzt die rechte Seite in dieser Abbildung genauer, so stellen wir fest, dass wir tatsächlich unseren Funktionsausschnitt in eine Folge von Rechteckschwingungen zerlegt haben.

Dabei sind wir ganz genauso vorgegangen wie bei der Multiskalenanalyse unserer Zahlenfolge im vorangegangenen Abschnitt. Die Zahlenfolge dort können wir auch auffassen als Messwerte für die Funktion $f(x) = x^2$. Daher haben wir auch dort bereits eine Zerlegung dieser Funktion in Rechteckschwingungen durchgeführt. Dies wird deutlich, wenn wir die Abweichungen auf den einzelnen Skalen nochmals genauer betrachten. Wir stellen dabei fest, dass je zwei dieser Abweichungen den gleichen Betrag haben, sich aber im Vorzeichen unterscheiden; so können z.B. die Werte -3 und $+3$ auf der zweitfeinsten Skala von Abweichungen als eine Rechteckschwingung (der Höhe 3) aufgefasst werden.

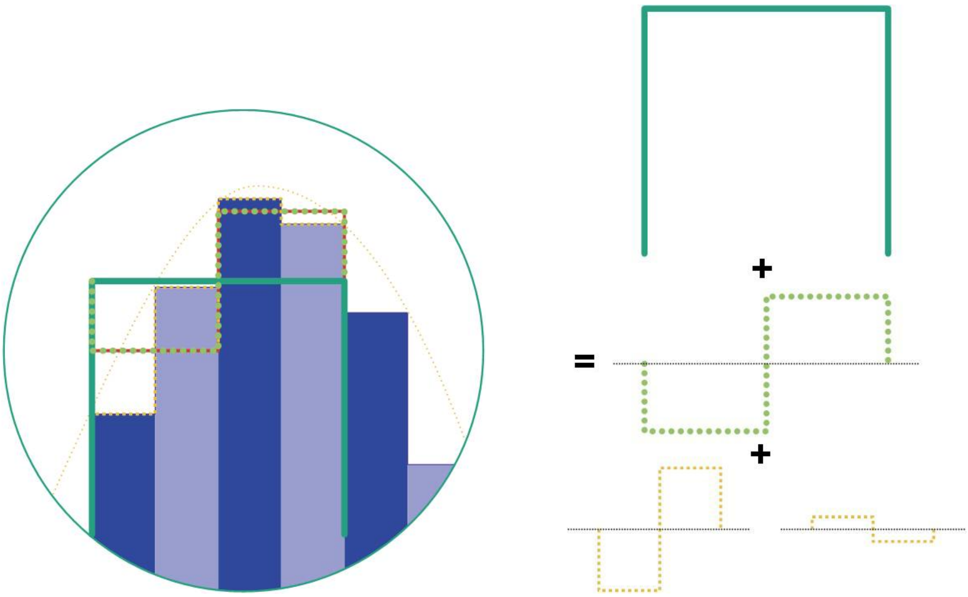


Abbildung 10: Darstellung der diskreten Funktion durch Rechteckschwingungen

7 Effizientes Abspeichern der verbliebenen Informationen in mp3: Huffman-Codierung

Wie wir in Abschnitt 4 diskutiert haben, kann das menschliche Ohr insbesondere in komplexen Klängen (wenn viele Töne gleichzeitig erklingen und sich überlagern) viele Informationen nicht wahrnehmen. Daher werden die unhörbaren Anteile in mp3 nur ungenau gespeichert.

Zusätzlich wird eine weitere Reduktion des zu speichernden Datenvolumens dadurch erreicht, dass man eine sogenannte Huffman-Codierung verwendet. Die Idee der Huffman-Codierung lässt sich am Beispiel der Codierung eines Textes einfach beschreiben: In einem Text kommen Buchstaben unterschiedlich häufig vor, in der deutschen Sprache beispielsweise das e viel häufiger als das y. Deshalb verwendet man bei einer Huffman-Codierung einen nur sehr kurzen Code für häufig vorkommende Buchstaben, längeren Code hingegen für Buchstaben, die nur selten vorkommen [8]. Dies minimiert die Länge des gesamten Codes für den ursprünglichen Text. Gleichzeitig ist aus einem Huffman-Code die ursprüngliche Information schnell, eindeutig und exakt reproduzierbar.

Bemerkung: *Ein anderes Beispiel für eine derartige Codierung ist das Morsealphabet. Ein negatives Beispiel ist hingegen das (klassische) Tippen einer SMS per Handy-Zahlen-Tastatur. Hier muss für häufig verwendete Buchstaben wie z.B. 'e' oder 'n' zweimal gedrückt werden. Daher wird hier die Anzahl der Eingaben nicht minimiert.*

Übertragen auf die Musik bedeutet dies: Meist besteht das ungenau zu speichernde Frequenzspektrum aus wenigen großen und vielen (also häufiger vorkommenden) kleinen Werten (Quantisierungswerte). Die Huffman-Codierung sorgt dann dafür, dass die digitalisierte Darstellung dieses Tons nur sehr wenig Speicherplatz einnimmt. Im Zusammenhang mit mp3 reduziert die Huffman-Codierung den Speicherplatz spürbar.

Im folgenden Abschnitt erläutern wir am Beispiel eines einfachen Textes wie eine Huffman-Codierung im Detail funktioniert. Abschnitt 7.2 enthält einige Bemerkungen zur Effizienz der Huffman-Codierung.

Für weitere Infos zur Huffman-Codierung verweisen wir auf [15, 16].

7.1 Huffman-Codierung am Beispiel eines einfachen Textes

Ausgangspunkt für die Huffman-Codierung ist eine Analyse, welche Buchstaben in dem zu codierenden Text am häufigsten vorkommen, denn diese Buchstaben sollen durch einen möglichst kurzen Code codiert werden, wohingegen Buchstaben, die sehr selten vorkommen, auch durch einen längeren Code beschrieben werden dürfen.

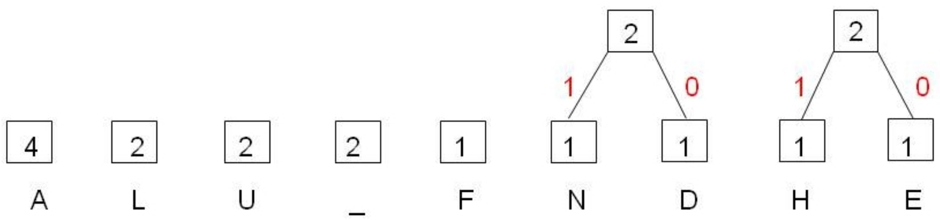
Betrachten wir als Beispiel den rheinischen Schlachtruf "ALAAF UND HELAU". Tabelle 1 gibt das Ergebnis einer Häufigkeitsanalyse an, d.h. sie beschreibt, wie oft welcher Buchstabe in diesem Text vorkommt. Dabei steht ␣ für ein Leerzeichen zwischen zwei Worten.

A	L	U	␣	F	N	D	H	E
4	2	2	2	1	1	1	1	1

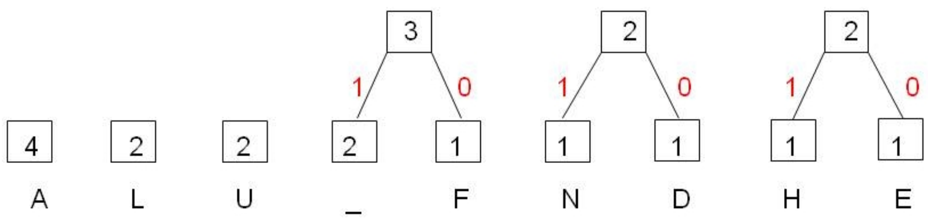
Tabelle 1: Häufigkeitsanalyse des Textes "ALAAF UND HELAU".

Das Vorgehen ist jetzt ganz einfach: Zunächst schreibt man die im Text vorkommenden Buchstaben nebeneinander und zeichnet darüber je ein Kästchen, in dem steht, wie oft die Buchstaben in dem Text jeweils vorkommen. Dann werden jeweils die beiden Kästchen mit den kleinsten vorkommenden Zahlen mit einem neuen Kästchen verbunden, in das die Summme der Zahlen aus den beiden Ursprungskästchen eingetragen wird. Die Linien, die die neuen Kästchen mit seinen Ursprungskästchen verbinden, werden jeweils abwechselnd mit "1" und "0" markiert. Dieses Vorgehen wird solange wiederholt, bis nur noch ein Kästchen übrig ist.

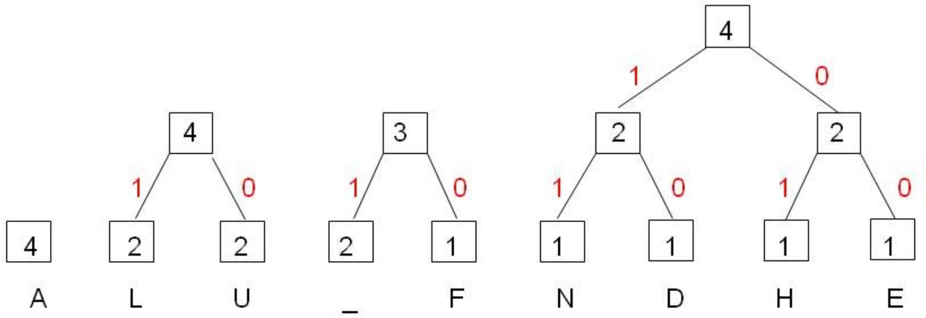
Wendet man dieses Vorgehen auf unser Beispiel "ALAAF UND HELAU" an, so verbinden wir zunächst je zwei Buchstaben, die nur einmal vorkommen, also z.B. N und D sowie H und E zu einem neuen Kästchen, tragen in diese Kästchen die Summen aus den Ursprungskästchen (also 2) ein und markieren die Wege zu den neuen Kästchen jeweils abwechselnd mit 1 und 0 und erhalten so



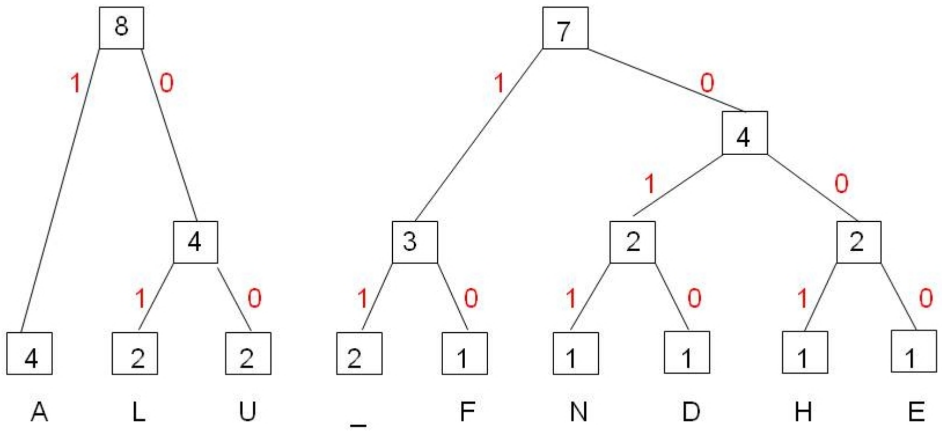
Damit ist nur noch ein Kästchen mit einer 1 übrig geblieben, das wir jetzt mit einem anderen Kästchen, in dem eine möglichst kleine Zahl steht, verbinden müssen. Hierfür könnten wir z.B. das Kästchen über dem Leerzeichen ␣ wählen:



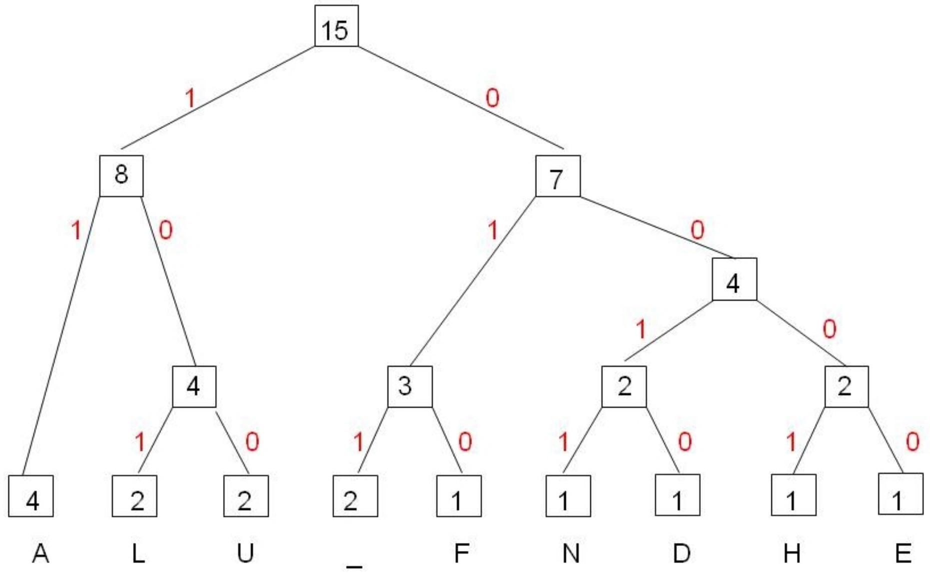
Als nächstes verbinden wir wieder jeweils zwei Kästchen mit den niedrigsten vorkommenden Zahlen, d.h. die Kästchen, in denen eine 2 steht, zu jeweils einem neuen Kästchen, in das wir die Summe (also 4) schreiben, und markieren wieder die Wege abwechselnd mit 1 und 0:



Verbinden wir jetzt das Kästchen mit einer 3 mit einem Kästchen, das eine 4 enthält, sowie die verbleibenden Kästchen, die eine 4 enthalten, so haben wir z.B.



Jetzt können wir nur noch die Kästchen mit der 7 und der 8 verbinden und haben damit folgenden Baum erhalten, in dessen Wurzel die Anzahl der zu codierenden Zeichen steht (in unserem Fall die Zahl 15):



Mit Hilfe dieses Baumes können wir jetzt jeden Buchstaben, der in unserem Text vor-

kommt, codieren. Dazu starten wir bei der Wurzel oben im Baum und wandern von hier aus entlang der verbundenen Kästchen zu dem Buchstaben, den wir codieren möchten. Die Ziffern, die wir an den Verbindungswegen finden, stellen den Code für die Buchstaben dar. Der Buchstabe A wird also durch 11 codiert, der Buchstabe U durch 100, der Buchstabe H durch 0001 und ein Leerzeichen durch 011. “ALAAF UND HELAU” wird damit codiert durch
1110111110100111000011001001100010000010111100 .
Man sieht sofort, dass dies deutlich kürzer ist als der entsprechende binäre ASCII-Code, denn dieser lautet:
01000001011011000110000101100001011001100010000001110101011011100110010000100
000010010000110010101101100011000010111010 1 .

Die Decodierung verläuft entsprechend. Wir starten an der Wurzel des Baumes und folgen gemäß dem Code dem Baum, bis wir an einem Buchstaben ankommen, und wiederholen dieses Vorgehen, bis unsere Ziffernfolge abgearbeitet ist.

Bemerkung: Es ist wichtig, dass wir sowohl bei der Codierung als auch bei der Decodierung oben im Baum (bei der Wurzel) starten und zu den Buchstaben wandern und nicht umgekehrt, weil sonst die Decodierung nicht mehr eindeutig möglich wäre. Würden wir etwa jeweils von den Buchstaben aus nach oben wandern, so könnte etwa ein Code 11001111... sowohl mit AU als auch mit NA beginnen können.

Bemerkung: Entwickelt man wie oben beschrieben einen Huffman-Code, so ist die entstehende Codierung nicht eindeutig, weil nicht vorgeschrieben ist, welche Kästchen miteinander verbunden werden. Im obigen Beispiel haben wir im ersten Schritt N mit D und H mit E verbunden. Wir hätten aber auch N mit H und D mit E verbinden und dann weitergehen können. Daher muss für die Decodierung auch der entstandene Codierungsbaum bekannt sein.

Bemerkung: Wie stark ein Text durch eine Huffman-Codierung komprimiert werden können, hängt stark von der Häufigkeitsverteilung der einzelnen Buchstaben ab. Dies überträgt sich entsprechend auf die Huffman-Codierung bei mp3. Auch hier hängt die Kompressionsrate vom jeweiligen Musikstück ab.

7.2 Bemerkungen zur Optimalität der Huffman-Codierung

Wir haben in Abschnitt 7.1 gesehen, dass wir den Text
ALAAF UND HELAU
mittels einer Huffman-Codierung schreiben können als
1110111110100111000011001001100010000010111100 .

Offensichtlich ist aber diese Ziffernkette doch größer als die Länge des ursprünglichen Textes. Deshalb scheint die Frage berechtigt, wieso denn hier eine kürzere, komprimierte Darstellung erzielt wurde.

Dazu ist zu bemerken, dass in der digitalen Datenverarbeitung jedwede Info letztlich im Dualsystem gespeichert wird, also als Folge der Ziffern 1 und 0. Wir mussten für den obigen Text 9 Zeichen unterscheiden. Für 9 Zeichen sind im Dualsystem bei einer festen Anzahl von Ziffern 4 Stellen erforderlich. Unsere Huffman-Codierung verwendet zwar für einige selten vorkommende Buchstaben auch vier Stellen (für das N z.B. 1100), für die anderen aber nur drei oder zwei Stellen, und benötigt damit insgesamt deutlich weniger

Zeichen, spart also Speicherplatz.

Nun könnten Schüler/-innen auf die Idee kommen, z.B. für die Buchstaben im obigen Text die Zahlen 0 bis 8 zu verwenden und diese ins Dualsystem zu übertragen (vgl. Tabelle 7.2), denn dies müsste doch noch effizienter sein als die Huffman-Codierung, weil dann zwei Buchstaben sogar durch nur eine Ziffer im Binärsystem dargestellt werden könnten:

A:	0
L:	1
U:	10
␣:	11
F:	100
N:	101
D:	110
H:	111
E:	1000

Tabelle 2: Ansatz einer einfachen Codierung im Dualsystem

ALAAF UND HELAU würde dann lauten:
0100100111010111011111110001010
und dies ist offensichtlich kürzer als der oben entwickelte Huffman-Code
1110111110100111000011001001100010000010111100
Dies ist zwar richtig, jedoch kann man die mit Tabelle 7.2 entstandene Ziffernfolge nicht mehr eindeutig decodieren. Sie könnte z.B. auch heißen
AFFHANDH ELAU
und es gibt viele weitere Möglichkeiten einer Decodierung. Man kann auch an kleineren Beispielen leicht feststellen, dass eine Codierung nach Tabelle 7.2 nicht mehr eindeutig entschlüsselbar wäre: So könnte etwa 1010 sowohl LALA als auch UU als auch LAU als auch ULA als auch NA bedeuten.

Eine Codierung, bei der die Decodierung aber nicht in eindeutiger Weise möglich ist, macht jedoch keinen Sinn. Man müsste zusätzlich noch einen binären Code einführen, der eindeutig ist und je zwei Buchstaben trennt, und dies wäre dann weniger effizient als die Huffman-Codierung.

8 Unterrichtsmaterialien

8.1 Audio-Video-Dateien und Animationen

Audio-Video-Datei audio_video_kompression.avi:
Diese Datei eignet sich als Demo zur Einführung in das Thema. Sie demonstriert gleichzeitig an je einem visuellen und auditiven Beispiel, wie sich die Qualität von Bildern und Musik verändert, wenn man die zugehörigen Daten immer weiter komprimiert. Die gleichzeitige Kompression von Bild und Ton spricht mit Auge und Ohr zwei Sinne an. Bei der Kompression bleibt die Qualität für die menschliche Wahrnehmung zunächst erhalten und man spart große Mengen an Speicherplatz, aber irgendwann werden Qualitätsverluste bemerkbar. Komprimiert man dann immer noch weiter, so verschlechtert sich die Qualität dramatisch.

Animationen Simple_harmonic_motion_animation.gif und Onde_compression_impulsion_1d_30_petit.gif:
Diese Animationen erläutern das Entstehen und die räumliche Ausbreitung von Schwingungen.

Videoclips musikbeispiel_orig.avi, musikbeispiel_mp3.avi und musikbeispiel_orig_minus_mp3.avi:
Der Videoclip musikbeispiel_orig.avi enthält ein kurzes Musikstück. Gleichzeitig sind auch die unterschiedlichen Frequenzen, die jeweils zu hören sind (grün gezeichnet), zu sehen.
Der Videoclip musikbeispiel_mp3.avi enthält dasselbe Musikstück als mp3-Version. Die Menge der in mp3 gespeicherten Informationen ist grün gekennzeichnet. Weiß dargestellt ist die Menge an Informationen, die in mp3 nicht gespeichert sind.
Der Videoclip musikbeispiel_orig_minus_mp3.avi enthält gewissermaßen die Differenz dieser beiden Dateien. Er macht den Unterschied beider Versionen hörbar und zeigt das, was in der mp3-Version verlorengegangen ist, oder, mit anderen Worten, das, was in der Originalversion enthalten, für das menschliche Ohr aber im Gesamtzusammenhang des Musikstücks nicht wahrnehmbar ist.

Bild ruhehoerschwelle.jpg:
Dieses Bild zeigt (Ruhe-)Hörschwelle, Schmerzgrenze, Musik- und Sprachwahrnehmbarkeit in Abhängigkeit von der Frequenz eines Tones. Die Hörschwelle kennzeichnet die Lautstärke, die ein Ton einer bestimmten Frequenz haben muss, damit er vom menschlichen Ohr wahrgenommen werden kann. Unterhalb dieser *Hörschwelle* wird der Ton vom menschlichen Gehör nicht wahrgenommen. Dieses Bild zeigt zudem, dass diese Hörschwelle von der Frequenz des Tones abhängt. Im Bereich von 2 bis 5 kHz (2000 bis 5000 Hertz) reichen schon sehr niedrige Lautstärken aus; hier ist das menschliche Ohr besonders empfindlich. Andererseits muss die Lautstärke bei sehr niedrigen und bei sehr hohen Frequenzen recht hoch sein, damit der Ton für den Menschen überhaupt hörbar ist.

Musik und Sprache können nur in noch kleineren Frequenz- und Lautstärkebereichen wahrgenommen werden.

Gut beobachtende Schüler werden feststellen, dass in manchen Frequenzbereichen weniger als 0 (Dezibel) erforderlich sind, und darüber verwundert sein. Man beachte aber, dass die Lautstärke hier in Dezibel gemessen wird und dass Dezibel eine logarithmische Maßeinheit ist. Wegen $\log 1 = 0$ bedeutet 0 Dezibel gerade *nicht*, dass völlige Stille herrscht.

Bild mithoerschwelle.jpg:
Dieses Bild macht folgende Eigenschaft des menschlichen Gehörs deutlich. Kommt in einem Klang eine bestimmte Frequenz mit einer bestimmten Lautstärke vor (in diesem Bild ein Ton mit einer Frequenz von 1 kHz = 1000 Hz), so sind benachbarte Frequenzen nur hörbar, wenn sie eine bestimmte Lautstärke (die sogenannte *Mithörschwelle*) überschreiten. Die Hörschwelle wird also in einer Umgebung dieser Frequenz deutlich angehoben. Mit anderen Worten: Eine hohe Lautstärke einer bestimmten Frequenz überdeckt Tonsignale von benachbarten Frequenzen, deren Lautstärke unterhalb der angehobenen Hörschwelle (der Mithörschwelle) liegt.

Entsprechende Kurven für die Mithörschwelle kann man auch für dominante Töne anderer Frequenzen darstellen.

8.2 Arbeitsblätter

Arbeitsblatt: Experiment zum menschlichen Gehör

Materialien:
hörbar tickender Wecker oder hörbar tickende Eieruhr, Stoppuhr, Stift, Papier

Durchführung:
Sucht einen Platz, an dem ihr möglichst weit von anderen Gruppen entfernt seid.

Dieser Versuch wird genau so oft durchgeführt, wie Mitglieder in der Gruppe sind. Jedes Mitglied ist einmal Versuchsperson. Ferner wird jeweils ein "Zeitmesser" bestimmt.

Stellt nun in jedem Versuch den Wecker so, dass er in einer Minute klingeln wird. Verhaltet euch nun ganz ruhig. Die Versuchsperson konzentriert sich auf das Ticken des Weckers. Klingelt nun der Wecker, beginnt die Zeitmessung. Die Zeit soll vom Beginn des Weckerklingelns bis zu dem Zeitpunkt, an dem die Versuchsperson das Ticken des Weckers wieder wahrnimmt, gestoppt werden. Haltet diese Zeit in Sekunden fest und tragt die Ergebnisse für alle Versuchspersonen in folgende Tabelle ein.

Zum Abschluss messt ihr einmal, wie lange das Klingeln des Weckers dauert und tragt auch dies in die Tabelle ein.

Versuchsperson	**Zeit (s)**
Dauer des Klingelns des Weckers	

Aufgaben:

1. Vergleicht die Ergebnisse der verschiedenen Versuchspersonen miteinander und mit der Dauer des Klingelns des Weckers.
2. Überlegt in der Gruppe, welche Ursachen für die Ergebnisse verantwortlich sein können. Berücksichtigt dabei, dass das Ticken des Weckers die ganze Zeit über weitergeht.

Arbeitsblatt: Experiment zum menschlichen Gehör

Infos für Lehrkräfte:
Der Versuch sollte in Gruppen zu mindestens drei Schüler/-innen durchgeführt werden. Es sollten nicht zu viele Gruppen sein, da sich diese gegenseitig durch das Klingeln des Weckers sonst stören würden.

Ergebnis der Gruppenarbeit sollte sein, dass es nach dem Klingeln des Weckers einige Zeit dauert, bis die Versuchspersonen das Ticken des Weckers wieder wahrnehmen können: Nach sehr lauten Geräuschen kann das menschliche Ohr für kurze Zeit leisere Geräusche schlechter oder gar nicht wahrnehmen (vgl. Abschnitt 4).

Da die Zeiträume, bis das menschliche Ohr leisere Geräusche wieder gut wahrnehmen kann, recht kurz sind, kann es sein, dass mehrere Messreihen erforderlich sind und man die Mittelwerte dieser Messreihen verwenden muss, um Aussagen treffen zu können. Bei einer einzelnen Messung können z.B. durch Abweichungen in der Reaktionszeit des Zeitnehmers mitunter Messfehler entstehen, die das Ergebnis signifikant beeinträchtigen.

Arbeitsblatt: Hörschwelle

Das menschliche Gehör kann nicht alle Töne wahrnehmen, sondern nur diejenigen, deren Lautstärke oberhalb der sogenannten Hörschwelle liegen. Abb. 1 zeigt u.a. die Hörschwelle in Abhängigkeit von der Frequenz eines Tones. Nach rechts ist die Frequenz des Tones aufgetragen, nach oben die Lautstärke, die in der Einheit Dezibel (dB) gemessen wird.

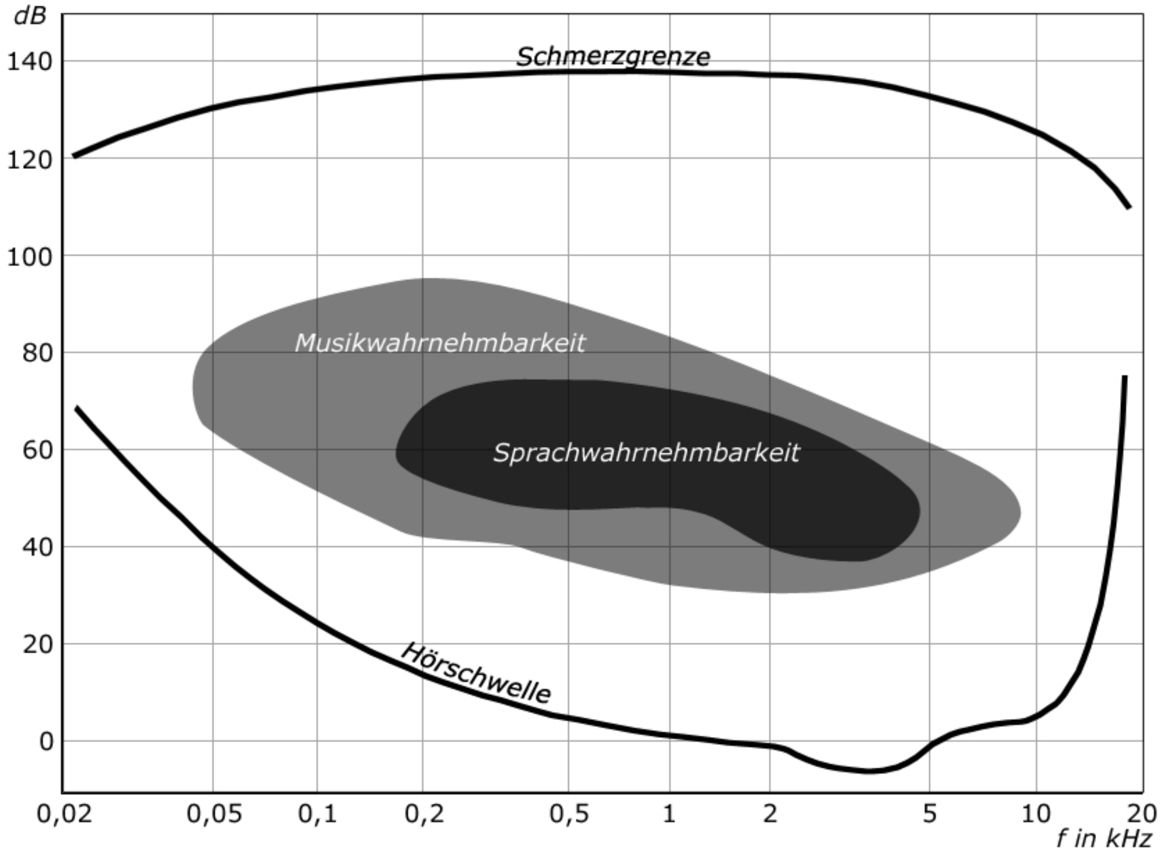


Abbildung 1: Grenzen des menschlichen Gehörs: Hörschwelle, Schmerzgrenze, Musik- und Sprachwahrnehmbarkeit in Abhängigkeit von der Frequenz.

Aufgaben:

1. Beschreibe den Verlauf der in Abb. 1 als Hörschwelle bezeichneten Kurve.
2. Welche Bereiche werden durch die Hörschwelle voneinander getrennt?
   Gib in diesem Zusammenhang an, ob folgende Töne für das menschliche Ohr hörbar sind: a) 10 kHz mit 0,1 dB b) 0,5 kHz mit 20 dB
   c) 0,05 kHz mit 30 dB d) 20 Hz mit 60 dB
3. In welchem Frequenzbereich sind Töne von sehr geringer Lautstärke für das menschliche Ohr hörbar?
4. In welchen Frequenzbereichen ist eine Lautstärke von mehr als 40 dB erforderlich, damit der Ton vom menschlichen Ohr wahrgenommen werden kann?
5. Das menschliche Ohr kann Töne mit Frequenzen über 20 kHz nicht wahrnehmen. Erläutere dies in Zusammenhang mit Abb. 1.

Arbeitsblatt: Hörschwelle

Lösungen der Aufgaben:

1. Die Abbildung zeigt die Hörschwelle in Abhängigkeit von der Frequenz eines Tones. Die Kurve beginnt bei 0,02 kHz und 70 dB. Sie fällt dann parabelförmig bis zu etwa 2 kHz und etwas unter 0 dB. Zwischen 2 und 5 dB ist die Lautstärke dann nochmals niedriger. Danach steigt die Kurve bis zu etwa 11 bis 12 kHz und etwa 10 dB moderat an. Jenseits von 12 kHz ist ein extrem starkes Wachstum zu verzeichnen. Bei etwa 18 kHz sind bereits 80 dB erreicht.

   Zusatzinfo für Lehrkräfte: Besonders genau beobachtende Schüler/-innen werden sich fragen, was es bedeutet, dass die Kurve in manchen Bereichen unter 0 dB liegt. Hierzu beachte man, dass "Dezibel" eine logarithmische Maßeinheit ist. Wegen $\log 1 = 0$ bedeutet 0 dB gerade *nicht*, dass völlige Stille herrscht.

2. Die Hörschwelle trennt die Bereiche von Signalen voneinander, die für das menschliche Ohr wahrnehmbar bzw. nicht wahrnehmbar sind. Sie kennzeichnet die Lautstärke, die ein Ton einer bestimmten Frequenz haben muss, damit er vom menschlichen Ohr wahrgenommen werden kann. Unterhalb dieser *Hörschwelle* wird der Ton vom menschlichen Gehör nicht wahrgenommen. Dieses Bild zeigt zudem, dass diese Hörschwelle von der Frequenz des Tones abhängt.
   a) nicht hörbar b) hörbar
   c) nicht hörbar d) nicht hörbar

3. Im Bereich von 2 bis 5 kHz (2000 bis 5000 Hertz) reichen schon sehr niedrige Lautstärken aus, damit ein Ton hörbar ist; von etwa 0,5 bis etwa 10 kHz ist die erforderliche Lautstärke auch noch sehr niedrig.

4. In den Frequenzbereichen kleiner als etwa 50 Hz bzw. größer als etwa 15 kHz.

5. Steigt die Frequenz der Töne über etwa 11 bis 12 kHz, so wächst die erforderliche Lautstärke extrem stark an. Die zur Hörbarkeit erforderliche Lautstärke scheint unendlich groß zu werden, wenn der Bereich um 20 kHz erreicht wird.

Arbeitsblatt: Anhebung der Hörschwelle

Das menschliche Gehör kann nicht alle Töne wahrnehmen. Abb. 1 zeigt exemplarisch, wie die Hörschwelle durch die Anwesenheit von Tönen mit einer Frequenz von 1 kHz angehoben wird.

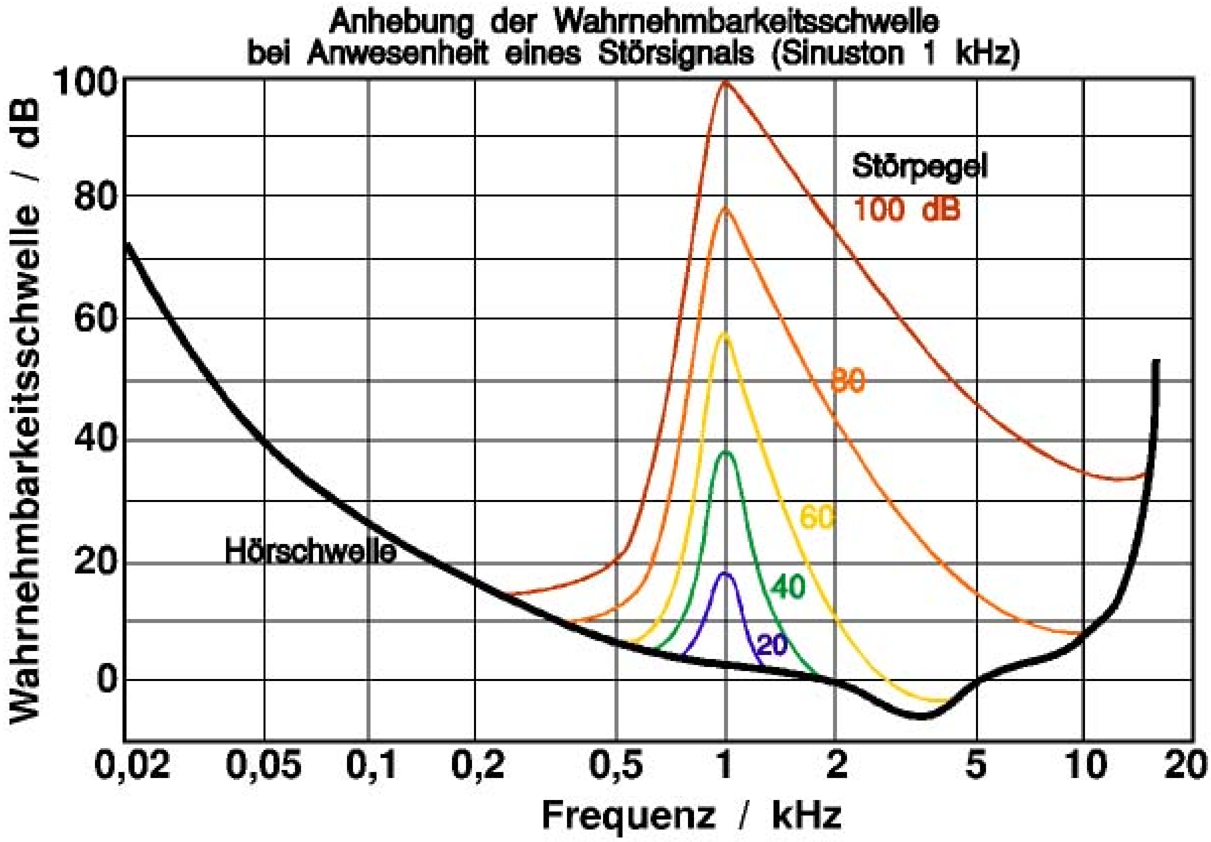


Abbildung 1: Grenzen des menschlichen Gehörs: Anhebung der Hörschwelle durch die Anwesenheit eines Tons von 1 kHz bei verschiedenen Lautstärken.

Aufgaben:

1. Beschreibe, wie sich die Hörschwelle ändert, wenn ein Ton mit einer Frequenz von 1 kHz und 100 dB anwesend ist.
2. Interpretiere, welchen Auswirkungen ein Ton mit 1 kHz und 80 dB auf die vom menschlichen Ohr wahrnehmbaren Tonsignale hat.
   Gib in diesem Zusammenhang an, ob folgende Töne für das menschliche Ohr hörbar sind, wenn gleichzeitig ein Ton von 1 kHz und 80 dB erklingt:
   a) 0,2 kHz mit 10 dB b) 0,5 kHz mit 10 dB
   c) 2 kHz mit 20 dB d) 5 kHz mit 20 dB
3. Wie könnte man das Phänomen der Anhebung der Hörschwelle technisch ausnutzen, wenn man Musikdaten in komprimierter Form speichern will?

Arbeitsblatt: Anhebung der Hörschwelle

Lösungen der Aufgaben:

1. Bei Anwesenheit eines Tones mit 1kHz und 80 dB wird die Hörschwelle im Bereich zwischen etwa 0,25 kHz und 15 kHz angehoben. Diese Anhebung ist um so höher, je näher die Frequenz bei der des Tones mit 1 kHz liegt. Außerhalb dieses Frequenzbereichs ändert sich die Hörschwelle nicht.

2. Der Ton mit etwa 1 kHz und 80 dB schränkt die Menge der vom menschlichen Ohr wahrnehmbaren Tonsignale im Bereich zwischen 0,3 und 10 kHz ein. Nur Töne, deren Lautstärke die angehobene Hörschwelle übersteigen, sind wahrnehmbar.
   a) nicht hörbar b) nicht hörbar
   c) nicht hörbar d) hörbar

3. All das, was das menschliche Ohr nicht wahrnehmen kann, d.h. was unter der Hör- bzw. der angehobenen Hörschwelle (Mithörschwelle) liegt, braucht man auch nicht (genau) abzuspeichern, wenn man Musik in komprimierter Form abspeichern will. mp3 macht sich dieses Prinzip zunutze.

Zusatzinfos:

- Entsprechende Kurven für die Anhebung der Hörschwelle kann man auch für (dominante) Töne anderer Frequenz darstellen.

- Für mp3 werden die (musikalischen) Töne wieder in die Einzelfrequenzen zerlegt, aus denen sie zusammengesetzt sind. Die für das menschliche Ohr hörbaren Anteile müssen nur mit der nötigen Genauigkeit gespeichert werden. Unhörbare Anteile können ungenauer gespeichert werden.

Arbeitsblatt: “Multiskalenanalyse” einer Zahlenfolge durch fortgesetzte Mittelwertbildung

Wir betrachten als Beispiel folgende Zahlenfolge von Quadratzahlen: 0 1 4 9 16 25 36 49. Fassen wir die Zahlen in Paare zusammen und bilden die Mittelwerte dieser Paare, so erhalten wir die Folge $0,5\ \ 6,5\ \ 20,5\ \ 42,5$. Fassen wir diese Zahlen wieder zu Paaren zusammen und bilden die Mittelwerte der Paare, so erhalten wir die Folge $3,5\ \ 31,5$. Für dieses Zahlenpaar haben wir den Mittelwert 17,5. Damit haben wir die ursprüngliche Zahlenfolge in mehrere Skalen von Mittelwerten überführt:

Zahlenfolge	0	1	4	9	16	25	36	49
feine Skala von Mittelwerten	0,5		6,5		20,5		42,5	
mittlere Skala von Mittelwerten	3,5				31,5			
grobe Skala von Mittelwerten	17,5							

Um von einer Mittelwertskala wieder zur nächstfeineren zu gelangen, benötigen wir die Abweichungen der Mittelwerte von den zugehörigen Werten auf der vorherigen Skala:
$17,5 - 14 = 3,5$ bzw. $17,5 + 14 = 31,5$.

Entsprechend auf der nächstgröberen Skala:
$3,5 - 3 = 0,5 \quad 3,5 + 3 = 6,5 \quad 31,5 - 11 = 20,5 \quad 31,5 + 11 = 42,5$.

Ganz analog können wir auch von der feinen Skala von Mittelwerten zu unserer ursprünglichen Folge zurückkommen:

$$0,5 - 0,5 = 0 \qquad 0,5 + 0,5 = 1 \qquad 6,5 - 2,5 = 4 \qquad 6,5 + 2,5 = 9$$
$$20,5 - 4,5 = 16 \qquad 20,5 + 4,5 = 25 \qquad 42,5 - 6,5 = 36 \qquad 42,5 + 6,5 = 49$$

Die gröbste Skala von Mittelwerten und diese Abweichungen können wir uns wie in folgendem Schema merken. Wir haben zusätzlich die ursprüngliche Zahlenfolge nochmals dazu geschrieben:

Zahlenfolge	0	1	4	9	16	25	36	49
feinste Skala von Abweichungen	-0,5	+0,5	-2,5	+2,5	-4,5	+4,5	-6,5	+6,5
zweitfeinste Skala von Abweichungen	-3		+3		-11		+11	
zweitgröbste Skala	-14				+14			
gröbste Skala	17,5							

Zu den ursprünglichen Zahlen zurück kommen wir jetzt, indem wir den Mittelwert auf der gröbsten Skala und die entsprechenden gespeicherten Abweichungen auf allen feineren Skalen einfach addieren. Ein Beispiel:

Zahlenfolge	0	1	4	9	16	$\boxed{25}$	36	49
feinste Skala von Abweichungen	-0,5	+0,5	-2,5	+2,5	-4,5	$\boxed{+4,5}$	-6,5	+6,5
zweitfeinste Skala von Abweichungen	-3		+3		$\boxed{-11}$		+11	
zweitgröbste Skala	-14				$\boxed{+14}$			
gröbste Skala	$\boxed{17,5}$							

$17,5 + 14 - 11 + 4,5 = 25$.

Aufgabe: Führe analog Multiskalenanalysen durch für die Folgen
a) 0, 1, 2, 3, 4, 5, 6, 7
b) 0, 1, 8, 27, 64, 125, 216, 343

Arbeitsblatt: “Multiskalenanalyse” einer Zahlenfolge durch fortgesetzte Mittelwertbildung

Infos für Lehrkräfte:

Mit Hilfe dieses Arbeitsblatts sollen die Schüler/-innen feststellen, dass sich Zahlenfolgen nicht nur durch Hintereinanderreihung aufschreiben lassen, sondern auch als eine hierarchische Darstellung wie in dem Arbeitsblatt beschrieben.

Löst man diese Aufgabe zeichnerisch, so entspricht dies gerade einer Zerlegung einer Funktion, die durch ausgewählte Funktionswerte approximiert wird, in Rechteckschwingungen. Dies ist der Inhalt des nächsten Arbeitsblattes, das aber auch direkt (d.h. ohne das vorliegende Arbeitsblatt) bearbeitet werden kann.

Lösung der Aufgabe:

a)

Zahlenfolge	0	1	2	3	4	5	6	7
feine Skala von Mittelwerten	0,5		2,5		4,5		6,5	
mittlere Skala von Mittelwerten	1,5				5,5			
grobe Skala von Mittelwerten	3,5							

Zahlenfolge	0	1	2	3	4	5	6	7
feinste Skala von Abweichungen	-0,5	+0,5	-0,5	+0,5	-0,5	+0,5	-0,5	+0,5
zweitfeinste Skala von Abweichungen	-1		+1		-1		+1	
zweitgröbste Skala	-2				+2			
gröbste Skala	3,5							

b)

Zahlenfolge	0	1	8	27	64	125	216	343
feine Skala von Mittelwerten	0,5		17,5		94,5		279,5	
mittlere Skala von Mittelwerten	9				187			
grobe Skala von Mittelwerten	98							

Zahlenfolge	0	1	8	27	64	125	216	343
feinste Skala von Abweichungen	-0,5	+0,5	-9,5	+9,5	-30,5	+30,5	-63,5	+63,5
zweitfeinste Skala von Abweichungen	-8,5		+8,5		-92,5		+92,5	
zweitgröbste Skala	-89				+89			
gröbste Skala	98							

Arbeitsblatt: Zerlegen von Tonsignalen in Rechteckschwingungen unterschiedlicher Frequenzen

Um Tonsignale oder allgemeiner, eine gegebene Funktion $f(x)$, in Rechteckschwingungen unterschiedlicher Frequenzen zu zerlegen, betrachten wir die Funktion (die Tonsignale) nur an ausgewählten Werten. Diese Werte werden in Abb. 1 durch die gefärbten Balken wiedergegeben. Rechts in dieser Abbildung ist der umkreiste Ausschnitt der Funktion vergrößert dargestellt. Wir erläutern das Prinzip unseres Vorgehens im Folgenden anhand dieses Ausschnitts.

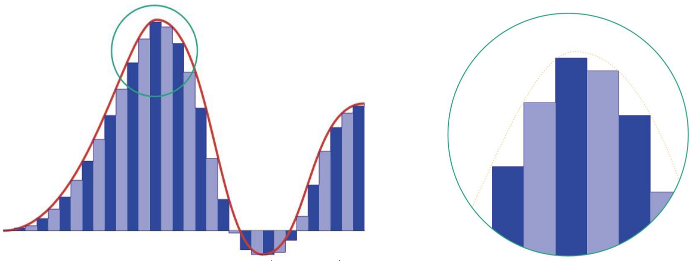


Abbildung 1: Funktion im Ganzen (links) und ein vergrößerter Ausschnitt (rechts)

Die linke Skizze in Abb. 2 zeigt, dass wir Mittelwerte der Funktionswerte bilden, um auf eine erste gröbere Skala zu kommen. Zurück zur feinen Skala können wir wieder kommen, indem wir die Abweichungen der ursprünglichen Funktionswerte vom zugehörigen Mittelwert wieder zum Mittelwert addieren.
Dieses Vorgehen können wir rekursiv fortsetzen, um auf gröbere Skalen zu kommen. Die mittlere Skizze in Abb. 2 zeigt den Mittelwert auf der entsprechenden nächstgröberen Skala.

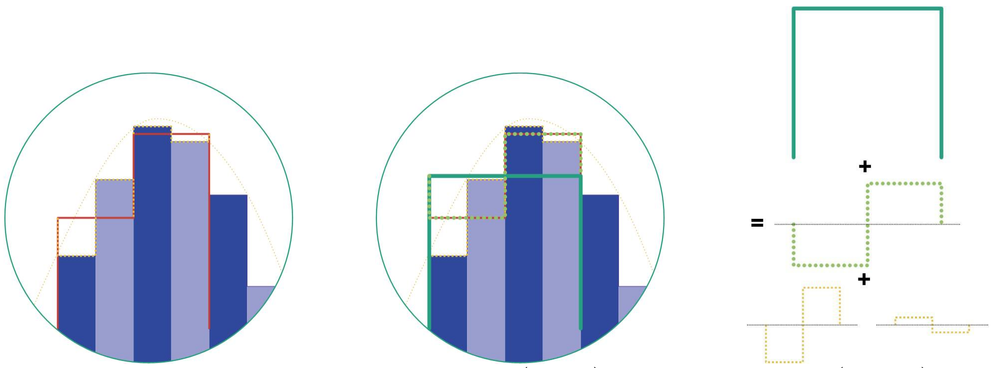
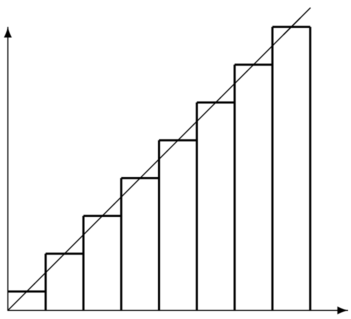


Abbildung 2: Multiskalenanalyse: feine Skalen (links) und grobe Skala (Mitte); Darstellung der Funktion durch Rechtecksschwingungen (rechts)

Die rechte Skizze in Abb. 2 zeigt, wie wir durch Addition des Mittelwerts auf der gröbsten Skala und den Abweichungen der Mittelwerte auf den feineren Skalen wieder zu unseren ursprünglichen Funktionswerten zurückkommen. Betrachten wir diese Skizze genauer, so stellen wir fest, dass wir tatsächlich unseren Funktionsausschnitt in eine Folge von Rechteckschwingungen zerlegt haben.

Aufgabe: Zerlege die Funktion $f(x) = x$ in Rechteckschwingungen unterschiedlicher Frequenzen. Wir betrachten nur den Bereich $0 \leq x \leq 8$ und nehmen die x-Werte 0,5; 1,5; 2,5; 3,5; 4,5; 5,5; 6,5 und 7,5, um die rechts skizzierte Balkendarstellung zu erhalten. Gehe dann so vor wie oben beschrieben, bis du auf einer gröbsten Skala angekommen bist. Skizziere die entstandenen Rechteckschwingungen.



Arbeitsblatt: Zerlegen von Tonsignalen in Rechteckschwingungen unterschiedlicher Frequenzen

Infos für Lehrkräfte:

Die Schüler/-innen sollten verstehen, dass man Funktionen näherungsweise in Rechteckschwingungen zerlegen kann und dieses Verständnis durch die (selbst angefertigte) Skizze vertiefen sowie die Einfachheit dieses Vorgehens erkennen.

Sollte im Unterricht auch das Arbeitsblatt *"Multiskalenanalyse" einer Zahlenfolge durch fortgesetzte Mittelwertbildung* behandelt worden sein, so sollte die Lehrkraft bei der Besprechung des vorliegenden Arbeitsblatts explizit darauf hinweisen, dass das Vorgehen in beiden Fällen analog ist: Die Zahlenfolge können wir auch auffassen als Werte der Funktion $f(x) = x^2$. Daher haben wir bereits in dem Arbeitsblatt *"Multiskalenanalyse" einer Zahlenfolge durch fortgesetzte Mittelwertbildung* eine Zerlegung dieser Funktion in Rechteckschwingungen durchgeführt. Dies wird deutlich, wenn wir die Abweichungen auf den einzelnen Skalen nochmals genauer betrachten. Wir stellen dabei fest, dass je zwei dieser Abweichungen den gleichen Betrag haben, sich aber im Vorzeichen unterscheiden; so können z.B. die Werte -3 und $+3$ (bzw. die Werte -11 und $+11$) auf der zweitfeinsten Skala von Abweichungen als eine Rechteckschwingung (der Höhe 3 bzw. 11) aufgefasst werden.

Lösung der Aufgabe:

Man kann zur Lösung der Aufgabe rein zeichnerisch vorgehen oder zuerst die Mittelwerte und die Abweichungen von den Mittelwerten berechnen und erst dann die Zeichnung anfertigen.

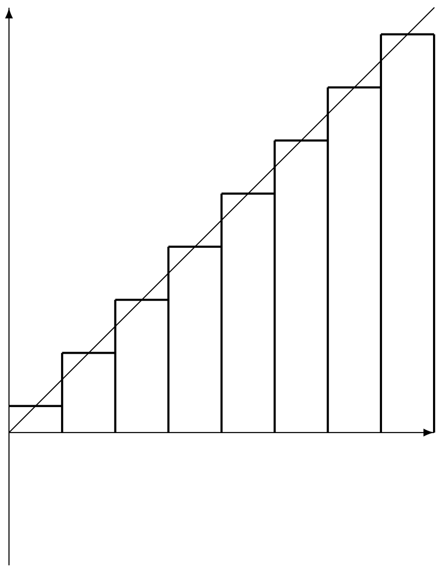
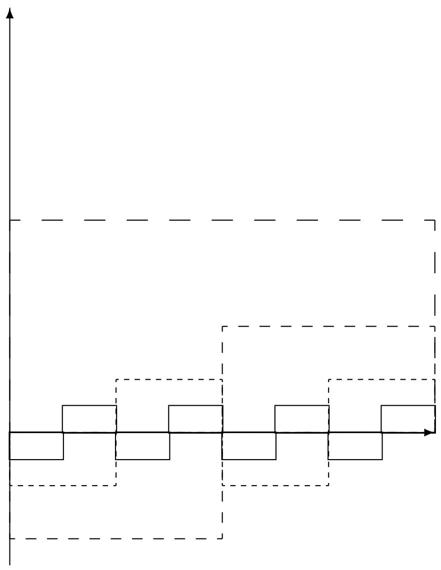
Für die Mittelwerte auf en einzelnen Skalen finden wir:

Funktionswerte	0,5	1,5	2,5	3,5	4,5	5,5	6,5	7,5
feine Skala von Mittelwerten	1		3		5		7	
mittlere Skala von Mittelwerten	2				6			
grobe Skala von Mittelwerten	4							

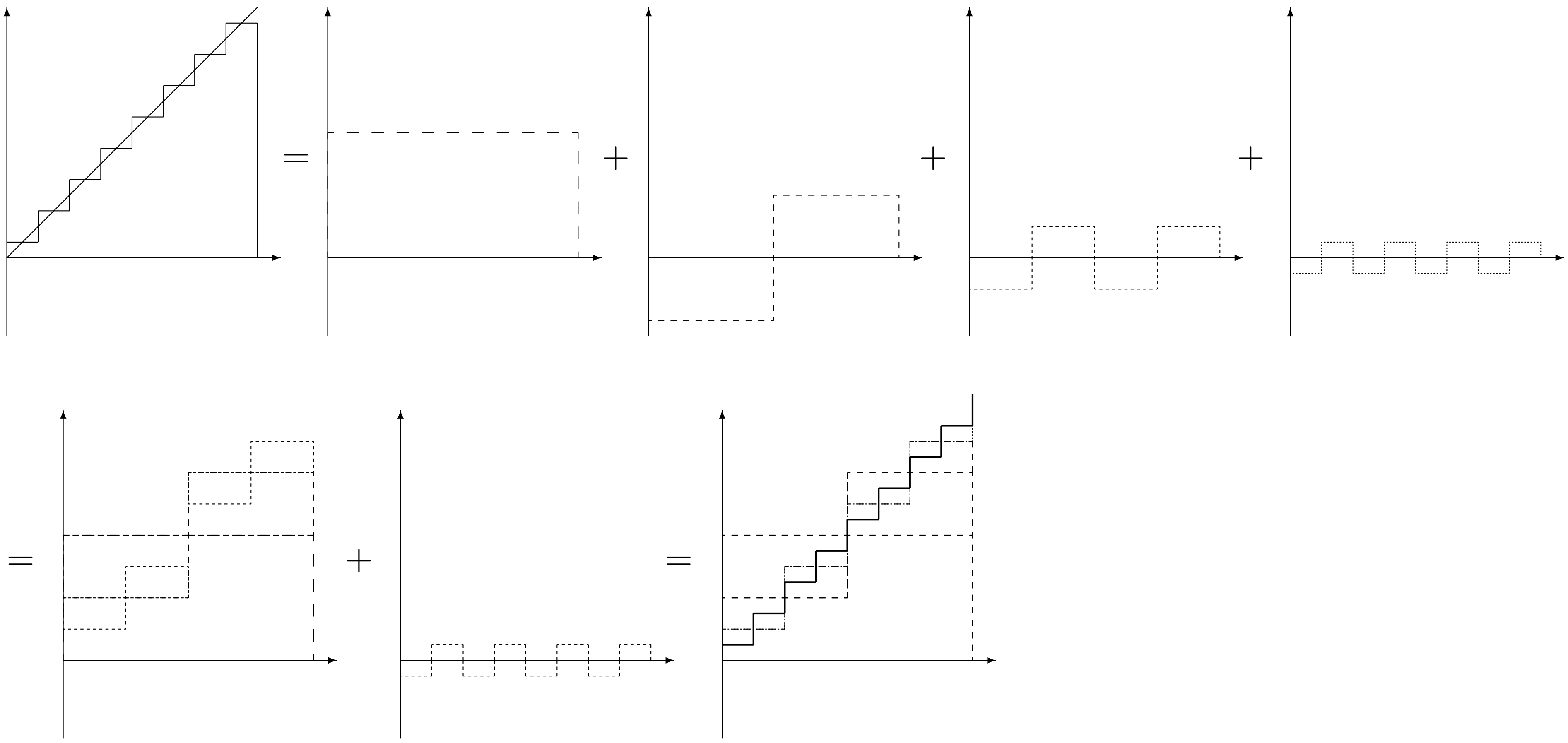
Die Abweichungen von den Mittelwerten lauten auf den einzelnen Skalen:

Funktionswerte	0	1	2	3	4	5	6	7
feinste Skala von Abweichungen	-0,5	+0,5	-0,5	+0,5	-0,5	+0,5	-0,5	+0,5
zweitfeinste Skala von Abweichungen	-1		+1		-1		+1	
zweitgröbste Skala	-2				+2			
gröbste Skala	4							

Skizzen: Die Funktion $f(x)$ und die Darstellung durch die Funktionswerte für $x = 0{,}5$; 1,5; 2,5; 3,5; 4,5; 5,5; 6,5; 7,5; ist links dargestellt. Rechts sieht man die einzelnen Rechteckschwingungen (vier Skalen):

Deutlicher wird die Überlagerung durch folgende Skizzen:



Arbeitsblatt: Datenkompression

In vielen Anwendungsbereichen müssen sehr große Datenmengen gespeichert werden. Hier ist es dann wichtig, möglichst wenig Speicherplatz zu verbrauchen, d.h. die Daten möglichst komprimiert zu speichern.

Aufgabe:

Versuche die folgenden Daten so zu schreiben, dass du mit möglichst wenig Ziffern/Buchstaben auskommst.

1. 0 0 0 0 0 1 0 0 0 0 1 1 1 1 1 0 1 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 1 1 1 0 0 1
2. 0 1 1 0 0 1 0 0 0 1 0 1 1 0 0 1 0 0 1 0 1 0 0 1 1 0 1 1 1 1 1 1 1 0 1 1 1 1 0
3. c c c a b b b c c a c c b c a a b c c b a a b b c a a b b b b b c b a b c a c
4. b a a a a a c a c c a c a a a c b b b b b a a b a a c a b a a c b b b c a b a
5. 999993 1000005 1000007 999993 999992 1000010 999996 999991 999990 999991 1000007 1000003 999995 1000008 999992 999998 999990 999996 1000010 1000003
6. 1000004 1000006 1000004 1000008 1000005 1000011 1000012 1000011 1000009 1000010 1000020 1000013 1000019 1000021 1000028 1000018 1000025 1000029

Arbeitsblatt: Datenkompression

Einige mögliche Lösungen:

1. Eine Lösungsansatz könnte darin bestehen, Ziffern, die mehr als zweimal hintereinander auftauchen, nicht einfach aufzuzählen, sondern mit einem Faktor aufzuführen, der angibt, wie oft die Ziffer jetzt kommt. 00000 würde dann übergehen in 50 (zu lesen als 5×0). Hiermit ergibt sich:
   5014051019001001003100 1 (man beachte hierbei, dass 10×0 nicht als 100 geschrieben werden darf, weil dies nicht mehr eindeutig wäre; daher wurde hier für diese Teilfolge 900 gewählt; 5050 wäre aber ebenso möglich wie 6040)
   Eine andere Lösung könnte darauf basieren, dass man je drei Elemente der Folge zusammenfasst und als (dreistellige) Zahl im Zweiersystem interpretiert:
   000 001 000 011 111 010 000 000 000 100 100 111 001. Dieser Ansatz würde zu folgender Darstellung führen:
   0103720004471

2. Der oben beschriebene erste Lösungsansatz (00000 $\rightarrow$ 50) führt hier zu
   011001301011001001010011071041 0, der zweite (Dualsystem) zu
   3105445157736.

3. Mehr als zweimal nacheinander auftauchende Buchstaben könnte man wieder mit einem Faktor aufführen, also 5a statt aaaaa. Dies führt zu
   3ca3bccaccbcaabccbaabbcaa5bcbabcac . Der Gewinn ist hier allerdings relativ klein. Er wird größer, wenn man zusätzlich doppelt vorkommende Buchstaben durch einen Großbuchstaben ersetzt (z.B. aa $\rightarrow$ A). Dann erhält man in diesem Beispiel
   3ca3bCaCbcAbCbABcA5bcbabcac .

4. Mit dem gleichen Ansatz wie in der vorigen Aufgabe erhält man hier
   b5acaccac3ac5baabaacabaac3bcaba . Auch hier hält sich der Gewinn in Grenzen. Ersetzt man wieder zusätzlich doppelt vorkommende Buchstaben durch entsprechende Großbuchstaben, so erhält man
   b5acaCac3ac5bAbAcabAc3bcaba .

5. Die Schüler/-innen sollten hier zunächst feststellen, dass alle Zahlen der Folge in der Nähe von 1000000 liegen. Ein Ansatz könnte dann darin bestehen, zunächst diese Zahl und dann nur die Abweichungen von 1000000 zu speichern (analog könnte man auch vom Mittelwert der Zahlenfolge als Basis ausgehen). Also etwa:
   1000000 -7 5 7 -7 -8 10 -4 -9 -10 -9 7 3 -5 8 -8 -2 -10 -4 10 3

6. Geht man wie in der Lösung der vorigen Aufgabe vor, so erhält man:
   1000000 4 6 4 8 5 11 12 11 9 10 20 13 19 21 28 18 25 29 oder, wenn man von 1000011 ausgeht:
   1000011 -7 -5 -7 -3 -6 0 1 0 -2 -1 9 2 8 10 17 7 14 18

Arbeitsblatt: Huffman-Codierung

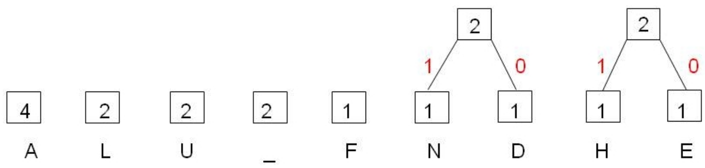
Um sehr große Datenmengen so zu speichern, dass sie möglichst wenig Speicherplatz verbrauchen, kann man eine sogenannte Huffman-Codierung verwenden. Anhand des Textbeispiels “ALAAF UND HELAU” erläutern wir, wie solch eine Huffman-Codierung funktioniert. Dabei ist die Idee der Huffman-Codierung, einen nur sehr kurzen Code für häufig vorkommende Buchstaben, längeren Code hingegen für Buchstaben, die nur selten vorkommen, zu verwenden. Dies minimiert die Länge des gesamten Codes für den ursprünglichen Text. Gleichzeitig ist aus einer Huffman-Codierung die ursprüngliche Information schnell, eindeutig und exakt reproduzierbar.

Ausgangspunkt für die Huffman-Codierung ist eine Häufigkeitsanalyse: Die folgende Tabelle beschreibt, wie oft welcher Buchstabe in unserem Beispieltext vorkommt. Dabei steht _ für das Leerzeichen zwischen zwei Worten:

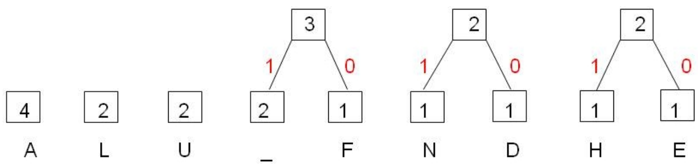
A	L	U	_	F	N	D	H	E
4	2	2	2	1	1	1	1	1

Um einen Huffman-Code zu erzeugen, schreiben wir die im Text vorkommenden Buchstaben nebeneinander und zeichnen darüber je ein Kästchen, in dem steht, wie oft die Buchstaben in dem Text jeweils vorkommen. Dann werden jeweils die beiden Kästchen mit den kleinsten vorkommenden Zahlen mit einem neuen Kästchen verbunden, in das die Summme der Zahlen aus den beiden Ursprungskästchen eingetragen wird. Die Linien, die die neuen Kästchen mit seinen Ursprungskästchen verbinden, werden jeweils abwechselnd mit “1” und “0” markiert. Dieses Vorgehen wird solange wiederholt, bis nur noch ein Kästchen übrig ist.

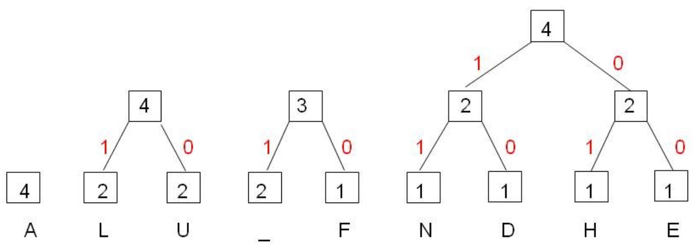
Wenden wir dieses Vorgehen auf unser Beispiel “ALAAF UND HELAU” an, so verbinden wir zunächst je zwei Buchstaben, die nur einmal vorkommen, also z.B. N und D sowie H und E zu einem neuen Kästchen, tragen in diese Kästchen die Summen aus den Ursprungskästchen (also 2) ein und markieren die Wege zu den neuen Kästchen jeweils abwechselnd mit 1 und 0 und erhalten so



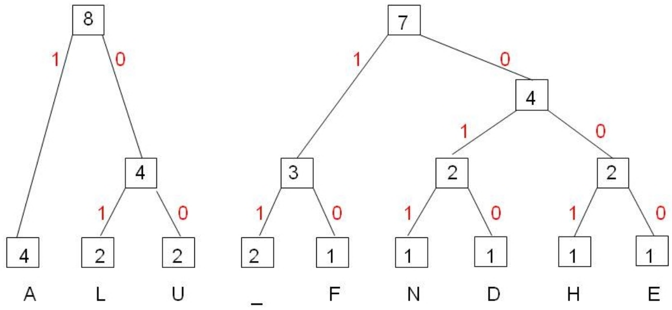
Damit ist nur noch ein Kästchen mit einer 1 übrig geblieben, das wir jetzt mit einem anderen Kästchen, in dem eine möglichst kleine Zahl steht, verbinden müssen. Hierfür könnten wir z.B. das Kästchen über dem Leerzeichen _ wählen:



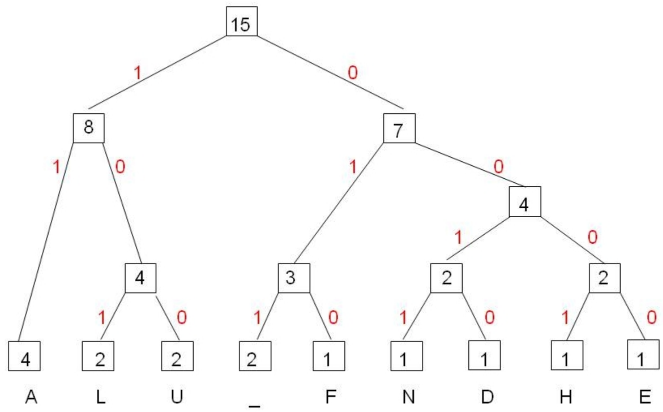
Als nächstes verbinden wir wieder jeweils zwei Kästchen mit den niedrigsten vorkommenden Zahlen, d.h. die Kästchen, in denen eine 2 steht, zu jeweils einem neuen Kästchen, in das wir die Summe (also 4) schreiben, und markieren wieder die Wege abwechselnd mit 1 und 0:



Verbinden wir jetzt das Kästchen mit einer 3 mit einem Kästchen, das eine 4 enthält, sowie die verbleibenden Kästchen, die eine 4 enthalten, so erhalten wir das nebenstehende Schema.



Jetzt können wir nur noch die Kästchen mit der 7 und der 8 verbinden und haben damit folgenden Baum erhalten, in dessen Wurzel die Anzahl der zu codierenden Zeichen steht (in unserem Fall die Zahl 15):



Mit Hilfe dieses Baumes können wir jetzt jeden Buchstaben, der in unserem Text vorkommt, codieren. Dazu starten wir bei der Wurzel oben im Baum und wandern von hier aus entlang der verbundenen Kästchen zu dem Buchstaben, den wir codieren möchten. Die Ziffern, die wir an den Verbindungswegen finden, stellen den Code für die Buchstaben dar. Der Buchstabe A wird also durch 11 codiert, der Buchstabe U durch 100, der Buchstabe H durch 0001 und ein Leerzeichen durch 011. "ALAAF UND HELAU" wird damit codiert durch
111011111010011100001100100110001000010111100 .
Man sieht sofort, dass dies deutlich kürzer ist als der entsprechende binäre ASCII-Code, denn dieser lautet:
0100000101101100011000010110000101100110001000000111010101101110011001000000100
00001001000011001010110110001100001011101 01 .

Die Decodierung verläuft entsprechend. Wir starten an der Wurzel des Baumes und folgen gemäß dem Code dem Baum, bis wir an einem Buchstaben ankommen, und wiederholen dieses Vorgehen, bis unsere Ziffernfolge abgearbeitet ist.

Dabei ist wichtig, dass wir sowohl bei der Codierung als auch bei der Decodierung oben im Baum (bei der Wurzel) starten und zu den Buchstaben wandern und nicht umgekehrt, weil sonst die Decodierung nicht mehr eindeutig möglich wäre. Würden wir etwa jeweils von den Buchstaben aus nach oben wandern, so könnte etwa ein Code 11001111... sowohl mit AU als auch mit NA beginnen können.

Bemerkung: Entwickelt man wie oben beschrieben einen Huffman-Code, so ist die entstehende Codierung nicht eindeutig, weil nicht vorgeschrieben ist, welche Kästchen miteinander verbunden werden. Im obigen Beispiel haben wir im ersten Schritt N mit D und H mit E verbunden. Wir hätten aber auch N mit H und D mit E verbinden und dann weitergehen können. Daher muss für die Decodierung auch der entstandene Codierungsbaum bekannt sein.

Aufgaben:

(1) Decodiere mit dem obigen Huffman-Baum den Code 0001110011001000001010011.
(2) Codiere mit dem obigen Huffman-Baum den Text "AUF HAENDEN LAUFEN".
(3) Erstelle für den Text "AUF HAENDEN LAUFEN" einen eigenen Huffman-Baum und nutze ihn, um diesen Text zu codieren. Vergleiche die Länge des Codes mit der in Aufgabe (2).

Arbeitsblatt: Huffman-Codierung

Lösung der Aufgaben:

(1) Der Code 0001110011001000001010011 bedeutet HANDELN.

(2) Der mit dem Huffman-Code, der auf dem Arbeitsblatt angegeben ist, entwickelte Code für den Text “AUF HAENDEN LAUFEN” lautet:

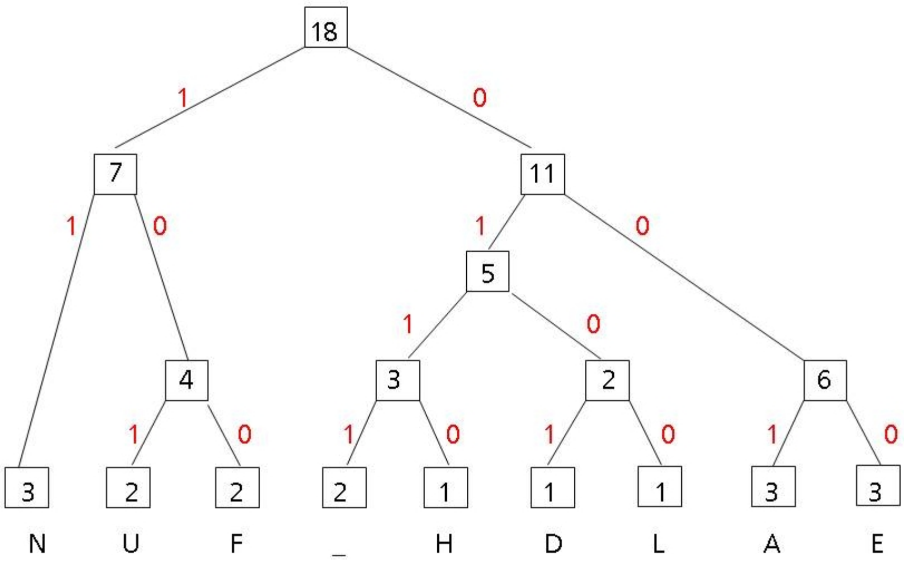
A	U	F		H	A	E	N	D	E	N		L	A	U	F	E	N
11	100	010	011	0001	11	0000	0011	0010	0000	0011	011	101	11	100	010	0000	0011

also:
11100010011000111000000110010000000110111011110001000000011

(3) Eine Häufigkeitsanalyse des Textes “AUF HAENDEN LAUFEN” ergibt

A	E	N	U	F	␣	H	D	L
3	3	3	2	2	2	1	1	1

Der entstehende Huffman-Baum ist nicht eindeutig. Es kann z.B. folgender Baum entstehen:



Damit ergibt sich die Codierung des Textes zu:

A	U	F		H	A	E	N	D	E	N		L	A	U	F	E	N
001	101	100	0111	0110	001	000	11	0101	000	11	0111	0100	001	101	100	000	11

also:
00110110001110110001000110101000110111010000110110000011

Dies Zeichenkette besteht aus 56 Ziffern. Die in Aufgabe (2) entstandene Zeichenkette hat eine Länge von 59 Ziffern. Wie zu erwarten, ist die für diesen Text optimierte Huffman-Codierung effizienter.

Literatur

[1] Helmut Neunzert: Einführungsvortrag auf dem Kongress *Mathematik in der Praxis*, Berlin, März 2009.

[2] http://www.iis.fraunhofer.de/bf/amm/mp3geschichte/erfolgsgeschichte/

[3] http://www.iis.fraunhofer.de/bf/amm/mp3geschichte/wasistmp3/

[4] http://www.mp3-geschichte.de

[5] http://de.wikipedia.org/wiki/Hörschnecke

[6] http://www.zeitnitz.de/Christian/scope_de

[7] http://de.wikipedia.org/wiki/MP3

[8] http://www.mp3-tech.org/tech.html

[9] http://www.mp3-tech.org/programmer/docs/trev_283-popp.pdf

[10] http://www.mp3-tech.org/programmer/docs/jacaba_main.pdf

[11] http://w3-o.cs.hm.edu/~ruckert/wiemp3/index.html

[12] http://www.movie-college.de/filmschule/ton/tonphysik.htm

[13] http://www.movie-college.de/filmschule/ton/psychoakustik.htm

[14] http://de.wikipedia.org/wiki/Welle_%28Physik%29

[15] http://www.ziegenbalg.ph-karlsruhe.de/materialien-homepage-jzbg/cc-interaktiv/huffman/index.htm

[16] http://www.iti.fh-flensburg.de/lang/algorithmen/code/huffman/huffman.htm

[17] http://www.flutepage.de/deutsch/goodies/physik.php