

MP3 - Das Musikformat des Internet

Nikolaus Hammler

(Technische Universität Graz, Österreich
nobaq@sbox.tugraz.at)

Kurzfassung: Viele Menschen verwenden das MP3-Format, doch kaum jemand kennt die technischen Hintergründe zu dem Format. Die folgende Arbeit soll das beliebte Speicherformat vom technischen Standpunkt betrachten und dem Leser ermöglichen, die Technik hinter MP3 zu verstehen. Aufbauend auf den mathematischen Grundlagen wie der Fourier-Transformation bzw. der diskreten Fast Fourier-Transformation (FFT) wird erklärt, wie ein analoges Signal abgetastet und digitalisiert wird. Die Digitalisierung eines Signals beruht auf dem Abtasttheorem.

Ist ein Signal einmal digitalisiert, so müssen die großen Datenmengen komprimiert werden. Das geschieht einerseits über nicht verlustbehaftete Kompressionsalgorithmen wie z.B. der Huffman-Kodierung oder über verlustbehaftete Kompressionen basierend auf einem psychoakustischen Modell, bei dem unwichtige Daten entfernt werden.

Liegen die Daten fertig kodiert vor, so müssen diese als Datei gespeichert werden. Hierzu wird das MP3-Speicherformat definiert. Über einen optionalen ID3-Tag können zusätzliche Informationen zu einer MP3 Datei abgespeichert werden, wie z.B. Titel, Interpret oder Erscheinungsjahr.

Schlüsselwörter: MP3, MPEG1 Layer 3, Fourier-Transformation, Fourier-Analyse, Fast Fourier-Transformation, FFT, Abtasttheorem, Shannon, Huffman-Kodierung, psychoakustisches Modell, Hörschwellenmarkierung, Maskierung im Frequenzbereich, Maskierung im Zeitbereich, Joint-Stereo, ID3-Tag, MPEG Frame

Inhaltsverzeichnis

Abbildungsverzeichnis	3
Tabellenverzeichnis	4
1 Einleitung	5
2 Was macht MP3 so klein?	6
3 Mathematische Grundlagen	7
3.1 Die Fourier-Transformation	7
3.2 Die diskrete Fourier-Transformation	11
3.3 Die Fast Fourier-Transformation	11
4 Die Abtastung eines analogen Signals	12
4.1 Töne als Wellen	12
4.2 Die Abtastung	13
4.3 Das Abtasttheorem und Aliasing	15

4.4 Der Analog/Digital Umsetzer	16
5 Die Huffman-Kodierung	17
5.1 Ein präfixfreier Kode	17
5.2 Konstruktion des Huffman-Kodes	18
6 Joint-Stereo	19
7 Das psychoakustische Modell	20
7.1 Hörschwellenmaskierung	20
7.2 Maskierung im Frequenzbereich	20
7.3 Maskierung im Zeitbereich	21
8 Das Dateiformat MP3	21
8.1 Gliederung	22
8.2 Der ID3-Tag	22
9 Zusammenfassung - Ausblick	23
Literaturverzeichnis	25

Abbildungsverzeichnis

1	Rechteckschwingung	8
2	3 Oberschwingungen für Rechteckschwingung	8
3	3 Oberschwingungen für die Rechteckschwingung, addiert	9
4	Abgetastete Sinusschwingung [Tietze und Schenk 2002]	13
5	Eingangssignal mit Dirac-Kamm multipliziert [Tietze und Schenk 2002]	14
6	Spektrum vor und nach der Abtastung [Tietze und Schenk 2002]	14
7	Abtasttheorem verletzt [Tietze und Schenk 2002]	15
8	Sigma-Delta Modulator [Beis 2005]	17
9	Präfixfreier Kode als Binärbaum [Cormen et al. 2000]	17
10	Konstruktion Huffman (1) [Cormen et al. 2000]	18
11	Konstruktion Huffman (2) [Cormen et al. 2000]	18
12	Konstruktion Huffman (3) [Cormen et al. 2000]	19
13	Die Hörschwelle des Menschen [Leidinger 2006]	20
14	Frequenzmaskierung [Leidinger 2006]	21
15	Zeitmaskierung [Leidinger 2006]	22

Tabellenverzeichnis

1	MPEG Audio Frame Header [Supurovic 1999]	23
2	Der ID3 Tag [Supurovic 1999]	23

1 Einleitung

Sowie es immer das Bestreben der Menschheit war, Bilder und Momente festzuhalten, so wollte man auch immer Klänge, Töne und Musik speichern. Das erste Mal konnten Töne mittels des von Philipp Reis 1861 entwickelten Magnettelefons erfasst werden. Damit konnten Töne jedoch nicht gespeichert werden. 1877 erfand Thomas Alva Edison mit seinem Phonographen das erste Gerät, das Töne auch aufzeichnen, d.h. speichern und wiedergeben konnte. Dem Phonographen folgte die Schallplatte. Die Schallplatte schaffte es zwar, Töne in befriedigender Qualität aufzunehmen, jedoch war die Abspeicherung sehr aufwändig. Im Jahre 1963 kam der erste Kassettenrecorder auf den Markt. Der Kassettenrecorder bot erstmals die Möglichkeit, in einem handlichen Gerät gleichzeitig Töne aufzunehmen und sie wieder abzuspielen. [Paturi 1989]

Alle bisher entdeckten Verfahren hatten zwei große Nachteile: Die Daten wurden analog gespeichert und die Speichermedien wurden im Laufe der Zeit abgenutzt, sodass die Qualität der Aufnahme zusätzlich sank. Den großen Durchbruch schaffte im Jahre 1981 die Compact Disk, die ihre Daten erstmals digital abspeicherte und mit Hilfe eines Lasers abtastete. Durch die berührungslose Abtastung durch den Laser war die Compact Disk keinem Verschleiß beim Abspielen ausgesetzt und durch die digitale Speicherung der Daten konnten Musikstücke erstmals in zufrieden stellender Qualität abgespeichert werden. [Paturi 1989]

In den Neunziger Jahren wurde das Internet für die breite Masse zugänglich. Der in den Achtziger Jahren entwickelte Personal Computer verbreitete sich und wurde verbessert. Während die ersten Modelle des Personal Computers noch zu schwach waren, um Musik zu verarbeiten, wurde es durch Entwicklung der „Soundkarte“ möglich, auch auf dem Personalcomputer digital Musik aufzunehmen, abzuspeichern, zu verarbeiten und abzuspielen. Jedoch sah man sich mit einem großen Problem konfrontiert: Gerade in Zeiten, in denen die Massenspeicher noch sehr klein waren, waren die Datenmengen für Musik in guter Qualität so hoch, dass Archivierung und Verbreitung über den Computer unmöglich waren.

Den großen Durchbruch schaffte man mit der Entwicklung von MP3. MP3 - die Abkürzung für Moving Picture Experts Group 1 Audio Layer 3 - definiert ein Verfahren für die Audiokompression. Zwar fing die Entwicklung dieser verlustbehafteten Audiokompressionstechnik bereits 1982 an, ein großer Schritt vorwärts gelang jedoch erst Ende der Neunziger Jahre, da nun auch die Personalcomputer genügend Rechenleistung besaßen, um das rechenintensive Dekodieren einer MP3-Datei durchführen zu können. Einen zusätzlichen Aufschwung brachte auch das Internet, das zu dieser Zeit bereits weit verbreitet war. Nun war es wegen der kleinen Datenmengen möglich, Musik über den Computer und das weltweite Internet zu teilen, zu kopieren und zu verbreiten.

2 Was macht MP3 so klein?

Die MP3-Kodierung bedient sich mehrerer Methoden, um die Qualität der aufgenommenen akustischen Daten hoch zu halten und gleichzeitig die Datenmenge gering zu halten. Vergleicht man die Größe einer unkomprimierten, ein minütigen Musiksequenz in CD-Qualität (ca. 10 MB) mit der Größe einer ein minütigen, MP3 kodierten Musiksequenz in hoher Qualität (ca. 1 MB), so sieht man, dass mit MP3 eine Datenreduktion um den Faktor 10 erreicht werden kann. MP3 verwendet dafür sowohl verlustlose Komprimierung, als auch verlustbehaftete Komprimierung.

Bevor die Daten überhaupt verarbeitet werden können, müssen diese digitalisiert werden. Die mathematischen Grundlagen dazu werden in [Kapitel 3] besprochen. Es wird gezeigt, wie man aus den aufgenommenen Audiodaten das *Frequenzspektrum* berechnen und analysieren kann. Wie diese Signale tatsächlich digitalisiert (*abgetastet*) werden, wird in [Kapitel 4] erläutert.

Ist das analoge Signal erst einmal digitalisiert gespeichert, so stellt sich die Frage, wie man es verlustfrei komprimieren kann. MP3 wendet zwei Techniken zur verlustfreien Kompression an. Einerseits wird die *Huffman*-Kodierung verwendet, um die Redundanz zu minimieren. Das Verfahren wird genau in [Kapitel 5] beschrieben. Andererseits wird das *Middle/Side Stereo*-Verfahren verwendet, sofern das Signal ein Stereosignal ist. Das Verfahren basiert auf der Beobachtung, dass der linke und der rechte Stereokanal viel Redundanz besitzen. Das Verfahren wird ausführlicher in [Kapitel 6] besprochen.

Aber erst durch den zusätzlichen Einsatz des *psychoakustischen* Modells können Audiodateien spürbar verkleinert werden. Hierbei handelt es sich um eine verlustbehaftete Komprimierung, die auf der Tatsache beruht, dass der Mensch gewisse Töne gar nicht hört. Diese Töne werden aus den Daten einfach weggeschnitten. Die betreffenden Informationen können dadurch zwar nicht wiederhergestellt werden, im Regelfall handelt es sich dabei aber um Daten, die der Mensch sowieso nicht hören würde. Dieses verlustbehaftete Verfahren erklärt den minimalen Qualitätsunterschied zwischen einem normalen, unkomprimierten Musikstück auf einer Audio-CD und einem MP3. Eine Audio-CD nimmt den gesamten Hörbereich des Menschen von 0 Hz bis ca. 20 KHz verlustfrei auf. In diesem Spektrum befinden sich viele Zonen, die ein Mensch nicht wahrnimmt und so getrost weggeschnitten werden können. Das psychoakustische Modell wird in [Kapitel 7] besprochen.

[Kapitel 8] beschreibt wie die Audiodaten tatsächlich in einer Datei gespeichert werden.

3 Mathematische Grundlagen

Vorraussetzung für die Digitalisierung analoger Signale ist die Entwicklung der mathematischen Grundlagen, vor allem die von Jean Baptiste Joseph Fourier im Jahre 1822 entwickelte Fourier-Transformation. Die Fourier-Transformation bildet zusammen mit der Laplace-Transformation, der z -Transformation und einigen anderen die Gruppe der Integraltransformationen.

3.1 Die Fourier-Transformation

Die Fourier-Transformation liefert eine völlig neue Betrachtungsweise für Funktionen. Anstatt im so genannten *Zeitbereich* können Funktionen nun im *Frequenzbereich* untersucht werden. Sie lässt sich im Gegensatz zur Laplace-Transformation recht anschaulich erklären. Dadurch wird auch klar, wieso dieser Bereich Frequenzbereich heisst.

Die Idee ist ähnlich der Taylor-Polynome. Bei einem Taylor-Polynom wird versucht, eine Funktion durch Polynome in einem Punkt und dessen Umgebung möglichst gut anzunähern [Bartsch 2004], siehe [Gleichung 1].

$$T(x) = \sum_{k=0}^n \frac{1}{k!} f^{(k)}(x_0)(x - x_0)^k \quad (1)$$

Eine Funktion kann durch beliebig viele Polynome (und Ableitungen) beliebig gut angenähert werden.

Die Idee der Fourier-Transformation beruht auf der Fourier-Reihe. Mit ihr ist es möglich, eine beliebige, periodische Funktion durch eine Linearkombination von trigonometrischen Polynomen anzunähern. Verblüffend dabei ist, dass die Ursprungsfunktion ident mit der transformierten Funktion ist, sofern die Anzahl der trigonometrischen Polynome gegen unendlich geht.

Als einfaches Beispiel sei eine periodische Rechtecksfunktion gegeben. Man betrachte die Rechteckschwingung in [Abbildung 1]. Diese Rechteckschwingung hat die gleiche Phase wie ein Kosinus und soll nun durch mehrere trigonometrische Polynome angenähert werden. [Abbildung 2] zeigt die Grundschiwingung und 3 Oberschwingungen übereinander. Man erkennt, dass sich die Frequenz immer verdoppelt bzw. die Periodendauer halbiert. In [Abbildung 3] werden die Schwingungen addiert und man erkennt bereits die Gestalt der ursprünglichen Rechteckschwingung. Die Annahme ist nun, dass unendlich viele Oberschwingungen die originale Rechteckfunktion ergeben.

Ausgehend von obiger Annahme wird nun der Zusammenhang gebildet in [Gleichung 2] gebildet.

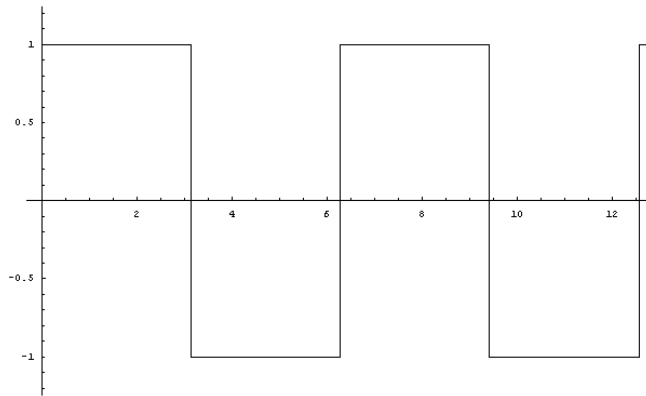


Abbildung 1: Rechteckschwingung

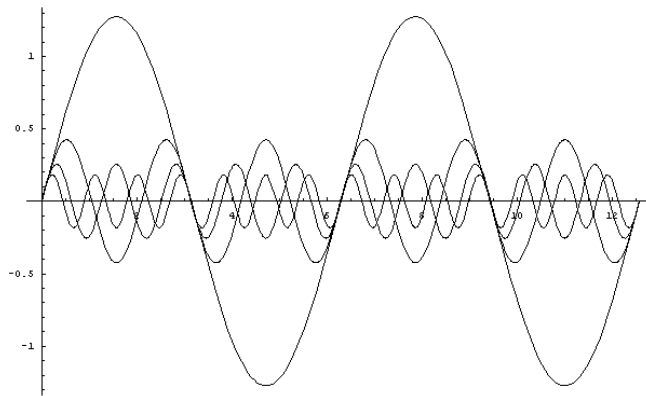


Abbildung 2: 3 Oberschwingungen für Rechteckschwingung

$$\begin{aligned}
 f(t) &= A_0 + A_1 \cos(\omega t + \phi_1) + A_2 \cos(2\omega t + \phi_2) + \dots + A_N \cos(N\omega t + \phi_N) \\
 &= \sum_{n=0}^N A_n \cos(n\omega t + \phi_n) \quad (2)
 \end{aligned}$$

Der Sinus ist lediglich ein phasenverschobener Cosinus, weswegen man [Gleichung 2] mit Hilfe der trigonometrischen Sumsätze umschreiben kann (mit $a_0 = A_0$, $a_n = A_n \cos \phi_n$ und $b_n = A_n \sin \phi_n$) [Wikipedia 2006c]:

$$f(t) = a_0 + \sum_{n=1}^N (a_n \cos(n\omega t) - b_n \sin(n\omega t)) \quad (3)$$

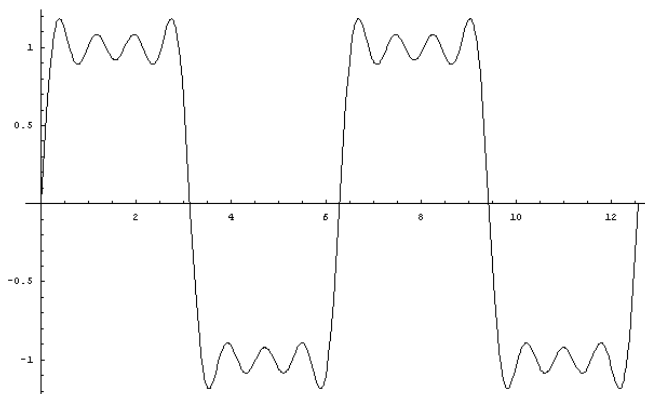


Abbildung 3: 3 Oberschwingungen für die Rechteckschwingung, addiert

a_0 , a_n und b_n in [Gleichung 3] nennt man Fourierkoeffizienten. Mit Hilfe der Formel von Euler und den daraus resultierenden Definitionen des Sinus und Kosinus kann daraus eine komplexwertige Reihe erstellt werden, die nur mehr einen komplexen Koeffizienten c_n umfasst [Wikipedia 2006c]:

$$f(t) = \sum_{n=-N}^N c_n e^{in\omega t} \quad (4)$$

Es stellt sich nun die Frage, wie der Fourierkoeffizient c_n (bzw. die Amplituden für den Sinus und den Kosinusanteil) gefunden werden kann. Dazu multipliziert man [Gleichung 4] mit $e^{-im\omega t}$ und integriert die Gleichung über eine Periode der Grundschwingung. Nach einigen Vereinfachungen bekommt man das Ergebnis aus [Gleichung 5] [Wikipedia 2006c].

$$c_m = \frac{1}{T} \int_0^T f(t) e^{-im\omega t} dt \quad (5)$$

Eingesetzt in [Gleichung 4] kommt man zum Ergebnis aus [Gleichung 6].

$$f(t) = \sum_{n=-\infty}^{\infty} e^{in\omega t} \frac{1}{T} \int_0^T f(t) e^{-in\omega t} dt \quad (6)$$

Es lässt sich zeigen, dass sich mit dieser Gleichung jede periodische Funktion, die einigen Randbedingungen genügt, durch trigonometrische Polynome annähern bzw. sogar ersetzen lässt, sofern $n \rightarrow \infty$. Es sei auch gesagt, dass manche „angenehme“ Funktionen nur eine endliche Anzahl von Fourierkoeffizienten benötigen, der Rest der Koeffizienten verschwindet dann ganz einfach. Funktionen, in denen nicht stetige Punkte vorkommen (wie z.B. die Ecken der Rechteckfunktion

aus [Abbildung 1]) werden immer unendlich viele Fourierkoeffizienten benötigen. Eine Fourierreihe konvergiert dann, wenn die *Dirichlet*-Bedingungen erfüllt sind [Spiegel 1990]:

1. Die Funktion ist definiert und eindeutig mit Ausnahme einer endlichen Zahl von Punkten
2. Die Funktion ist periodisch
3. Die Funktion ist stückweise stetig

Was ist aber mit Funktionen, die nicht periodisch sind? Auch diese können bearbeitet werden. Es wurde bereits erwähnt, dass die Frequenz der Oberschwingungen $\frac{n}{T}$ beträgt, der Abstand zwischen zwei Schwingungen beträgt also $\frac{1}{T}$ [Wikipedia 2006c]. Muss man nun eine nicht periodische Funktion bearbeiten, so kann man das Integrationsintervall bis in die Unendlichkeit ausdehnen und den gesamten Abbildungsbereich als Periode betrachten ($T \rightarrow \infty$). Daraus folgt auch, dass der Abstand zwischen zwei Oberschwingungen Null wird und statt einer Reihe für die Fourierkoeffizienten erhält man ein kontinuierliches Frequenzspektrum. Die differentiell dünnen Spektralanteile werden nun aufsummiert, was exakt der Definition des Riemann Integrals entspricht. Die Summe wird zum Integral. Man spricht nicht mehr von der Fourierreihe, sondern vom Fourierintegral bzw. der Fourier-Transformation. [Gleichung 7] zeigt diesen Schritt.

$$f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} a(\omega) e^{i\omega t} d\omega \quad (7)$$

In [Gleichung 7] ist die Funktion $a(\omega)$ die Funktion im *Frequenzbereich*, da diese von der Frequenz ω abhängt. Die große Bedeutung der Fourier-Transformation soll damit gezeigt sein: Es ist möglich, eine gegebene Funktion, die abhängig von einem Parameter ist (im Regelfall t), in den Frequenzbereich zu übersetzen. Das Frequenzspektrum entspricht nun den Amplituden für jeden enthaltenen Frequenzanteil. Die (direkte) Abhängigkeit von der Zeit verschwindet. [Gleichung 7] zeigt die sogenannte Fourier *Synthese*. Mit ihr ist es möglich, zu einem gegebenen Frequenzspektrum die zugehörige Funktion im Zeitbereich zu finden.

Im Allgemeinen wird es jedoch wichtig sein, das Frequenzspektrum für eine Funktion im Zeitbereich zu finden. Dieser Vorgang wird *Fourier-Analyse* genannt und entspricht der eigentlichen Fourier-Transformation ([Gleichung 7] entspricht der inversen Fourier-Transformation). Sie ist in [Gleichung 8] angegeben.

$$a(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt \quad (8)$$

3.2 Die diskrete Fourier-Transformation

Das Audiosignal entspricht nach der Abtastung einer Folge von diskreten Werten, wie genauer in [Kapitel 4] beschrieben wird. Darauf kann nicht einfach die oben besprochene Fourier-Transformation für kontinuierliche Signale angewandt werden. Einen ersten Schritt zur Anwendung der Fourier-Transformation auf das abgetastete Signal stellt die *Discret Time Fourier-Transform* (DTFT) dar. Dabei kann das Eingangssignal diskret sein, das Spektrum ist jedoch nach wie vor eine kontinuierliche Funktion. Zur vollständigen Beschreibung wird die *diskrete Fourier-Transformation* (DFT) benötigt.

Die DFT ist eher eine Folge als eine Funktion einer kontinuierlichen Variable. Sie entspricht den Abtastwerten der Fourier-Transformation zu äquidistanten Abständen. Da der Computer nur mit endlichen Folgen arbeiten kann, setzt man die Eingangsauidiodaten zu einer periodischen Folge hintereinander [Oppenheim et al. 2004]:

$$x[n] = x[(n \bmod N)] \quad (9)$$

Arbeitet man im diskreten Frequenzbereich, so gibt es nur endlich viele Frequenzen. Die Frequenzen wiederholen sich dabei am Einheitskreis, da es nur ganzzahlige Exponenten gibt. Deswegen entspricht der Einheitskreis bei der diskreten Fourier-Transformation der imaginären Achse bei der kontinuierlichen. Ansonsten arbeitet die diskrete Fourier-Transformation gleich, wie bereits für den kontinuierlichen Fall abgeleitet [Oppenheim et al. 2004]:

$$X[k] = \sum_{n=0}^{N-1} x[n] \cdot e^{-j\frac{2\pi kn}{N}} \quad (10)$$

$$x[n] = \frac{1}{N} \sum_{k=0}^{N-1} X[k] \cdot e^{j\frac{2\pi kn}{N}} \quad (11)$$

[Gleichung 10] stellt dabei die Analysegleichung, [Gleichung 11] die Synthesgleichung dar. Der große Vorteil der diskreten Variante ist, dass es sehr effiziente Algorithmen zur Berechnung der Fourier-Transformation gibt, z.B. die *Fast Fourier-Transformation* (FFT) [Oppenheim et al. 2004].

3.3 Die Fast Fourier-Transformation

Würde man die diskrete Fourier-Transformation direkt mit [Gleichung 10] und [Gleichung 11] berechnen, so würde das eine Komplexität von $O(n^2)$ verursachen.

Zwischen den reellen Fourierkoeffizienten a_n bzw. b_n und dem komplexen Koeffizienten besteht der Zusammenhang in [Gleichung 12] [Berglez 2004].

$$a_n + jb_n = \frac{1}{N} c_k (-1)^k \quad (12)$$

Durch direkte Bestimmung der Koeffizienten lt. [Gleichung 12] ergeben sich viele unnötige Multiplikationen mit 1 und -1 , die hohen Rechenaufwand verursachen (für die Multiplikation mit -1 müsste lediglich ein Bit invertiert werden).

Viele Daten sind also redundant vorhanden [Berglez 2004].

Mittels einer *Divide & Conquer Strategie* (Aufteilen des Problems in Teilprobleme) kann man den Rechenaufwand auf $O(n \cdot \log n)$ reduzieren. Die FFT eignet sich also perfekt für die Berechnung des Frequenzspektrums eines diskretisierten Audiosignals.

4 Die Abtastung eines analogen Signals

4.1 Töne als Wellen

Das menschliche Ohr als Sinnesorgan nimmt, ähnlich wie ein Mikrofon, Luftdruckunterschiede als Töne, Sprache oder Geräusche wahr. Der Schall entspricht einer Änderung des Luftdrucks. Es gibt unterschiedliche Arten von Schall, z.B. Töne, Geräusche oder Knall. Allen ist gemein, dass sie durch eine Druck- oder Dichteänderung eines Mediums, in der Regel Luft, übertragen werden [Roederer 2000].

Ein einzelner Ton kann als eine Sinusschwingung mit bestimmter Frequenz aufgefasst werden. Musik entspricht einer Überlagerung aller Einzeltöne. Geräusche bzw. ein Knall entsprechen einer plötzlichen Luftdruckänderung und vieler damit einhergehenden, stochastischen Oberschwingungen. Jeglicher Schall kann so als Frequenzgemisch von unterschiedlichsten Basisfrequenzen (mit jeweils eigener Amplitude und Phasenverschiebung) aufgefasst werden.

Die Wellenlänge eines Tons ist gegeben durch [Gleichung 13] [Roederer 2000].

$$\lambda = \frac{c}{f} \quad (13)$$

c in [Gleichung 13] ist die Ausbreitungsgeschwindigkeit einer Welle in einem Medium, entspricht also der Schallgeschwindigkeit. Es ist bemerkenswert, dass sich durch die Trägheit des Übertragungsmediums der Schall nicht unendlich schnell ausbreiten kann. Dementsprechend hat der Schall immer eine höchste Frequenz und damit ein endliches Spektrum.

Der Schall kann durch ein Mikrofon aufgenommen werden, das ähnlich wie das menschliche Ohr funktioniert: Über eine Membran oder Kohlepulver (Kohlemikrofon) werden die minimalen Druckunterschiede aus der Luft aufgenommen

und in eine elektrische Spannung proportional zum aufgenommenen Schall umgesetzt. Die so entstandene elektrische Welle, die (abgesehen von zusätzlichem Rauschen) band begrenzt ist, lässt sich nun hervorragend elektronisch verarbeiten.

4.2 Die Abtastung

Das aufgenommene Signal soll nun digitalisiert werden. Dazu müssen endlich viele Proben aus dem ursprünglichen, analogen Signal entnommen und gespeichert werden. In der Praxis wandelt man ein kontinuierliches Signal in ein diskretes um, indem man mittels eines Abtast-Halte-Gliedes in äquidistanten Zeitpunkten $t_\mu = \mu T_a$ Proben entnimmt [Brasseur 2006].

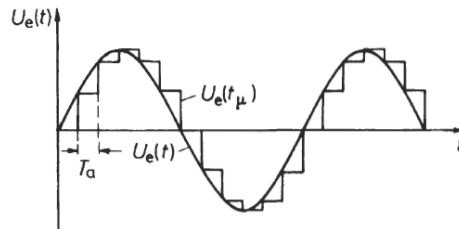


Abbildung 4: Abgetastete Sinusschwingung [Tietze und Schenk 2002]

Gegeben sei nun die Eingangsfunktion in [Abbildung 4]. Es ist nahe liegend, dass sich die abgetastete Treppenfunktion immer mehr der ursprünglichen Sinusfunktion nähert, je mehr Proben entnommen werden bzw. je kürzer der Abstand zwischen zwei Abtastungen wird.

Der schaltungstechnische Aufwand wird jedoch immer größer, je mehr Proben entnommen werden, d.h. desto höher die *Abtastfrequenz* ist. Die große Frage ist nun: Was ist die niedrigste Abtastfrequenz, bei der sich das Originalsignal noch fehlerfrei rekonstruieren lässt? [Tietze und Schenk 2002].

Eine einfache Sinusschwingung besitzt keinerlei Information, da sie in jeder Periode gleich ist. Durch drei Stichproben kann ein Sinussignal bereits für den kompletten Abbildungsbereich rekonstruiert werden (es müssen Informationen über Amplitude, Phasenverschiebung und Frequenz gespeichert werden). Wie sieht es nun mit einem beliebigen Signal aus? Wie bereits im vorigen Kapitel erwähnt, setzt sich das Eingangssignal aus unterschiedlichen Frequenzen zusammen. Wie oft muss ein beliebiges Signal nun abgetastet werden?

Mathematisch beschreibt man den Abtastvorgang am besten mittels eines Dirac-Kamms. Der Dirac-Kamm ist in [Gleichung 14] angegeben [Wikipedia

2006b].

$$\Delta_T(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT) \quad (14)$$

Ein Dirac-Kamm entspricht einer Folge von unendlich dünnen, unendlich hohen Impulsen mit der Fläche 1, genannt Dirac-Impuls. Der Dirac-Impuls ist keine Funktion im herkömmlichen Sinne, sondern eine Distribution, die nur über ihr Integral definiert ist. Integriert man über einen Dirac-Impuls, so erhält man die Fläche 1. Wird das analoge Signal abgetastet, so entspricht das mathematisch der Multiplikation des Dirac-Kamms mit der Eingangsfunktion. Im Abstand T werden die Impulse mit dem Wert der Eingangsfunktion *gewichtet*, zwischen zwei Abtastwerten verschwindet die Funktion. [Abbildung 5] zeigt die Funktion graphisch.

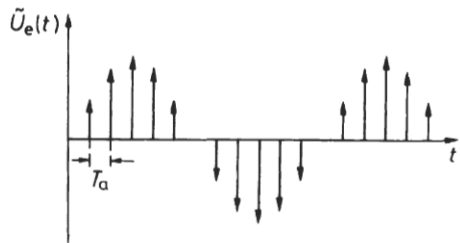


Abbildung 5: Eingangssignal mit Dirac-Kamm multipliziert [Tietze und Schenk 2002]

Es sei angemerkt, dass die Gewichtung der Dirac-Impulse durch die Höhe der Pfeile graphisch veranschaulicht ist. Das Spektrum der Eingangsfunktion ist ein kontinuierliches Spektrum zwischen den zwei Grenzfrequenzen $-f_{max}$ und f_{max} (die negativen Frequenzen entsprechen den komplexen Spiegelfrequenzen), da das Eingangssignal ja band begrenzt ist (siehe [Abbildung 6] oben).

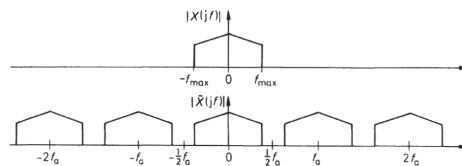


Abbildung 6: Spektrum vor und nach der Abtastung [Tietze und Schenk 2002]

Wird das Signal abgetastet, so vervielfacht sich das komplette Spektrum im Abstand zur Abtastfrequenz. Das liegt an der Linearität der Fourier-Transformation, da das Eingangssignal mit dem Dirac-Kamm multipliziert wurde. [Abbildung 6] unten verdeutlicht dies. Soll das ursprüngliche Signal wieder hergestellt werden, so reicht ein Tiefpass mit der Grenzfrequenz f_{max} ; es bleibt das ursprüngliche Spektrum übrig (der Rest wird durch den Tiefpass abgeschnitten) und das ursprüngliche Signal ist rekonstruiert.

Die einzelnen Abtastwerte können nun in einem Array gespeichert und digital weiterverarbeitet werden.

4.3 Das Abtasttheorem und Aliasing

Nun wird die Abtastfrequenz verringert. Dadurch verkleinert sich der Abstand zwischen den benachbarten Spektren in [Abbildung 6] unten. Genau ab dem Zeitpunkt, ab dem die Abtastfrequenz kleiner als doppelt so groß ist wie die Grenzfrequenz, fangen sich die Spektren an zu überlappen. Ab diesem Zeitpunkt ist es nicht mehr möglich, die Originalfunktion zu rekonstruieren, da das Spektrum nicht mehr eindeutig ist. [Abbildung 7] verdeutlicht die Überlappung der Spektren.

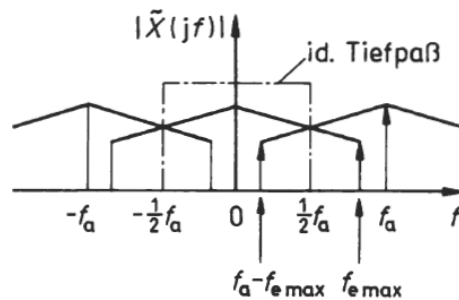


Abbildung 7: Abtasttheorem verletzt [Tietze und Schenk 2002]

Daraus lässt sich nun die Bedingung für die Frage nach der niedrigsten Abtastfrequenz beantworten:

„Das Originalspektrum erscheint nur dann unverändert, wenn die Abtastfrequenz mindestens doppelt so hoch gewählt wird, dass sich die periodisch wiederkehrenden Spektren nicht überlappen. Das ist nach [Abbildung 6] für

$$f_a > 2 \cdot f_{max}$$

der Fall. Diese Bedingung wird als Abtasttheorem bezeichnet.“ [Tietze und Schenk 2002]

Ist diese Bedingung nicht erfüllt, so entsteht *Aliasing*. Das bedeutet, dass sich zwei Spektren überlappen. Im Zeitbereich wird in diesem Fall eine niedrigere Frequenz als die tatsächliche Frequenz rekonstruiert. Ein Beispiel für dieses Phänomen sind die Wagenräder bei alten Filmen, die sich nur langsam (oder sogar rückwärts) zu drehen scheinen, da die Kamera nicht genügend Bilder anfertigen konnte.

Um Aliasing zu vermeiden, schaltet man vor der Abtastung *immer* einen Tiefpass mit der Grenzfrequenz f_{max} , um sicherzugehen, dass im Signal keine höhere Frequenz enthalten ist. Aufgrund der Tatsache, dass in der Praxis keine kausalen¹ Systeme hergestellt werden können, ist es nicht möglich, einen Tiefpass mit unendlicher Steilheit zu bauen. Aus diesem Grund wählt man in der Praxis den Faktor 2,2 statt 2: Bei einer Audio-CD werden Frequenzen bis 20 kHz übertragen und damit mit 44,1 kHz abgetastet [Wikipedia 2006a].

4.4 Der Analog/Digital Umsetzer

In der Praxis werden für die Umsetzung analoger Signale Analog-Digital Umsetzer, kurz *ADU* verwendet. Diese beinhalten, falls nötig, bereits Abtast-Halteglieder. Ein ADU misst das Verhältnis einer analogen Eingangsgröße a zu einer Referenzgröße a_r und gibt diese in Form eines digitalen Wertes an [Leopold 2005].

Der einfachste Vertreter dieser Gattung ist der Flash Umsetzer, der für jeden möglichen binären Wert einen Komparator besitzt. Der Vorteil ist die hohe Geschwindigkeit; der umfangreiche Schaltungsaufwand ist der große Nachteil. Neben anderen Vertretern gibt es den Sigma-Delta Umsetzer, der aufgrund seiner Genauigkeit vorwiegend im Audiobereich Verwendung findet.

Der Sigma-Delta Umsetzer besteht aus einem Modulator und einem Filter. Der Modulator erzeugt ein pulsverhältnismoduliertes Signal (PPM), das immer dem Mittelwert der Eingangsgröße entspricht. Dazu ist der Modulator als geschlossener Regelkreis aufgebaut. Wie bei jedem negativ rückgekoppelten Regelkreis soll die Ausgangsgröße möglichst genau der Eingangsgröße folgen [Horn und Dourdoumas 2004]. Das Ausgangssignal wird dabei vom Eingangssignal subtrahiert. Dieses Signal wird integriert und mittels eines 1 Bit Flash Umsetzers (Komparator) in einen digitalen Wert umgesetzt [Beis 2005]. [Abbildung 8] verdeutlicht den Vorgang.

Der abschließende digitale Tiefpass gewinnt durch den Mittelwert des PPM Signals den digitalen Wert.

¹ Kausale Systeme sind Systeme, die auch für Zeitpunkte $t < 0$ definiert sind. Aus diesem Grund lassen sich solche Systeme nicht bauen

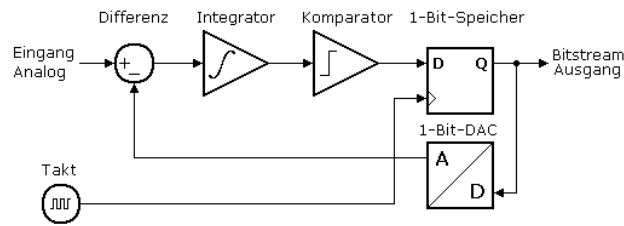


Abbildung 8: Sigma-Delta Modulator [Beis 2005]

5 Die Huffman-Kodierung

5.1 Ein präfixfreier Kode

Die Huffman-Kodierung ist eine sehr effektive Kodierung, die auch im bekannten und weit verbreiteten ZIP-Format Verwendung findet. Kompressionsraten von 20% bis 90% sind dabei nicht untypisch [Cormen et al. 2000].

Bei der Huffman-Kodierung handelt es sich um eine Entropie-Kodierung. Das bedeutet, alle vorkommenden Zeichen werden nach der Wahrscheinlichkeit des Auftretens gewichtet und mit unterschiedlich langen Binärkodenfolgen kodiert.

Arbeitet man jedoch mit unterschiedlich langen Codes, so muss man eine wichtige Tatsache beachten: Keine Binärkodenfolge darf sich am Beginn einer anderen Binärkodenfolge wieder finden. Diese Eigenschaft nennt man *präfixfreien* Kode. Um sicher zu gehen, dass ein Kode präfixfrei ist, kann ein Binärbaum verwendet werden.

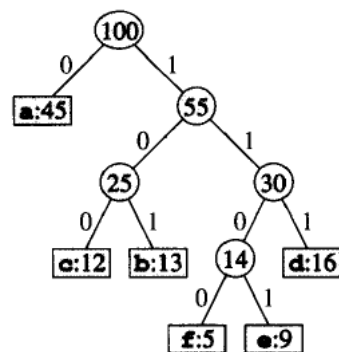


Abbildung 9: Präfixfreier Kode als Binärbaum [Cormen et al. 2000]

[Abbildung 9] verdeutlicht dies. Die Blätter bilden die Zeichen mit der Auftrittswahrscheinlichkeit, die internen Knoten enthalten die Summe der Blätter im Teilbaum und der Pfad enthält den binären Code. Zum Beispiel tritt das „a“ am häufigsten auf, und erhält daher den kurzen Code „0“. Damit der Code präfixfrei bleibt, darf aber keine andere Binärkodenfolge mit „0“ beginnen (rechter Teilbaum). Das „f“ tritt hingegen sehr selten auf und erhält daher die lange Binärkodenfolge „1100“.

5.2 Konstruktion des Huffman-Kodes

Wie konstruiert man nun anhand einer Häufigkeitstabelle einen optimalen präfixfreien Kode? Huffman hat einen Greedy² Algorithmus erfunden, der diesen konstruiert [Cormen et al. 2000].

Am besten veranschaulicht man sich den Algorithmus graphisch. Zuerst werden alle benötigten Zeichen mit der Anzahl der Auftritte angeschrieben, siehe [Abbildung 10]. Nun werden die Zeichen mit der kleinsten Auftrittswahrscheinlichkeit ausgewählt. Aus ihnen wird ein Teilbaum gebildet. Die Wurzel enthält die Summe der beiden Blätter, siehe [Abbildung 11]. Dieser Vorhang wird nun solange wiederholt, bis aus allen Elementen ein Baum gebildet ist. [Abbildung 12] zeigt einen weiteren Schritt. Zum Schluss ist der fertige Baum wie in [Abbildung 9] gebildet.

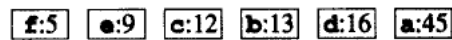


Abbildung 10: Konstruktion Huffman (1) [Cormen et al. 2000]

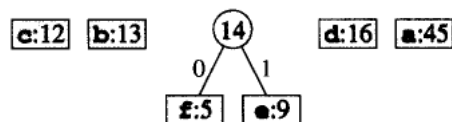


Abbildung 11: Konstruktion Huffman (2) [Cormen et al. 2000]

Die Bildung des Huffman Codes benötigt $O(n \cdot \log n)$.

² Ein Greedy-Algorithmus wählt bei jedem Schritt immer den Weg, der zum aktuellen Zeitpunkt den höchsten Gewinn erzielt

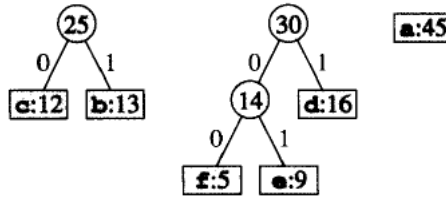


Abbildung 12: Konstruktion Huffman (3) [Cormen et al. 2000]

6 Joint-Stereo

Unter *Joint-Stereo* versteht man die Verschmelzung der beiden Stereokanäle zum Zwecke der Datenreduktion. Joint-Stereo bezeichnet eigentlich zwei Verfahren, nämlich *Intensity-Stereo* und *Middle/Side-Stereo* [Wikipedia 2006d]. Bei MP3 wird das Middle/Side-Verfahren eingesetzt.

Das Middle/Side-Verfahren basiert auf der Beobachtung, dass bei den beiden Stereokanälen (links und rechts) eine hohe Korrelation vorliegt. Viele Daten in den Kanälen sind also redundant. Aus diesem Grund speichert man den Mittelwert beider Kanäle in dem so genannten *Middle*-Kanal und die Differenz zu den beiden Kanälen im *Side*-Kanal. Dadurch ist das Verfahren verlustfrei, wie [Gleichung 17] und [Gleichung 18] zeigen [Wikipedia 2006d].

$$K_m = \frac{K_l + K_r}{2} \quad (15)$$

$$K_s = \frac{K_l - K_r}{2} \quad (16)$$

Die beiden Kanäle berechnen sich laut [Gleichung 15] und [Gleichung 16]. Die beiden ursprünglichen Kanäle können durch die [Gleichung 17] und [Gleichung 18] rekonstruiert werden [Wikipedia 2006d]:

$$K_l = K_m + K_s \quad (17)$$

$$K_r = K_m - K_s \quad (18)$$

Wie bereits gezeigt, enthalten die beiden Kanäle die Fourierkoeffizienten für das zeitdiskrete Signal, also die Amplituden für jede vorkommende Frequenz. Addiert man beide Amplituden, ändert sich in Bezug auf den Speicherbedarf wenig. Werden die Werte jedoch subtrahiert, so bekommt man aufgrund der hohen Korrelation sehr geringe Werte, in manchen Fällen sogar den Wert Null. Je kleiner die Zahl ist, desto weniger Bits werden auch für die Speicherung benötigt. So kann der Differenzkanal mit Hilfe der Huffman-Kodierung sehr gut komprimiert werden.

7 Das psychoakustische Modell

Die Psychoakustik untersucht die Wahrnehmung von Geräuschen. Um die Datenmenge zu reduzieren, werden alle Informationen weggelassen, die der Mensch sowieso nicht hören kann. Die Erkenntnisse sind auf Dr. Karlheinz Brandenburg zurückzuführen, der am Fraunhofer Institut bei Professor Seitzer promovierte. Dr. Brandenburg beschrieb erstmals, wie das menschliche Gehör arbeitet und welche Signale es wahrnimmt. Er kam dabei zu erstaunlichen Resultaten. [Schriber 2000]

7.1 Hörschwellenmaskierung

Der Hörbereich des Menschen reicht von ca. 16 Hz bis ca. 20 kHz. Doch auch innerhalb dieses Frequenzbandes lösen nur bestimmte Signale einen Hörreiz aus. Zum Beispiel ist das Ohr zwischen 2 kHz und 4 kHz am empfindlichsten. Dabei kommt es abgesehen von der Frequenz noch auf die Amplitude an, ob ein Mensch einen Ton wahrnimmt oder nicht. [Schriber 2000]

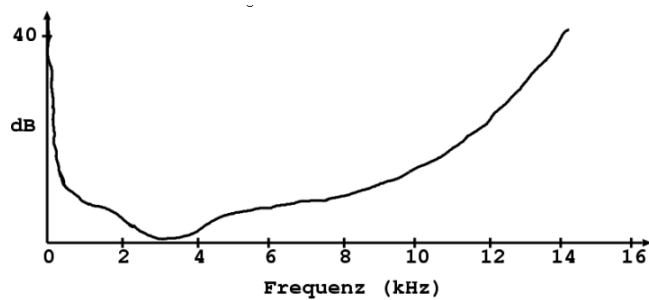


Abbildung 13: Die Hörschwelle des Menschen [Leidinger 2006]

[Abbildung 13] zeigt die Hörschwelle des Menschen. Alle Signale, die sich unter der eingezeichneten Schwelle befinden, können weggelassen werden. Das Entfernen der Daten geschieht im Frequenzbereich. Die Fourierkoeffizienten, die für eine bestimmte Frequenz einen zu niedrigen Wert haben, werden einfach aus dem Datenmaterial entfernt.

7.2 Maskierung im Frequenzbereich

Nach der Hörschwellenmaskierung werden Töne entfernt, die von anderen Tönen übertönt werden [Leidinger 2006]. Dies geschieht ebenfalls durch Entfernung der betreffenden Fourierkoeffizienten aus dem Spektrum.

Wenn ein Ton über der Hörschwelle liegt, wird die Hörbereichskurve aus [Abbildung 13] drastisch verändert. Ein Maskierungston von 2 kHz und einer Lautstärke von 40 dB überdeckt z.B. einen Großteil des menschlichen Hörbereichs, sodass andere Töne sehr laut sein müssen, um wahrgenommen zu werden. Besonders stark äussert sich dieses Verhalten wenn zwei Frequenzen nebeneinander liegen [Schriber 2000]. [Abbildung 14] zeigt die veränderte Kurve.

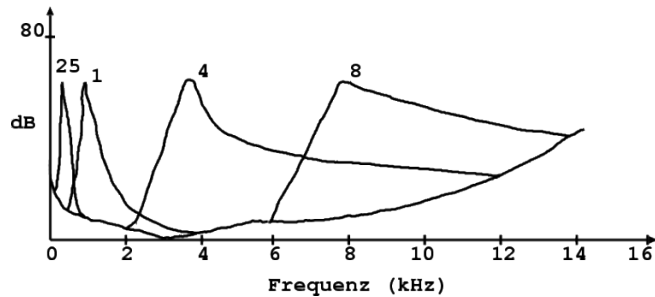


Abbildung 14: Frequenzmaskierung [Leidinger 2006]

7.3 Maskierung im Zeitbereich

Zum Schluss wird noch eine Maskierung im Zeitbereich durchgeführt. Im Frequenzbereich können ja nur Frequenzen und Amplituden geändert werden, die Zeit kommt nur mehr indirekt vor.

Psychoakustiker haben herausgefunden, dass es neben der simultanen Maskierung im Frequenzbereich auch eine zeitliche gibt. Das Ohr benötigt nämlich nach der Registrierung eines lauten Geräusches eine Erholungszeit, bis der nächste Ton wahrgenommen werden kann. Das Ohr nimmt sowohl kurz vor einem lauten Geräusch, als auch danach keine leiseren Töne mehr wahr. Dabei kann eine Verdeckungszeit von 20 ms beobachtet werden. [Schriber 2000].

[Abbildung 15] veranschaulicht die Zeitmaskierung graphisch. Im Beispiel wird ein 1,1 kHz Ton maskiert. Man sieht, dass man bis zu 5 ms vor dem eigentlichen Ton keinen anderen wahrnehmen kann bzw. der Ton über 20 ms benötigt, bis er abgeklungen ist.

Die Maskierung im Zeitbereich ist der aufwendigste Teil der MP3-Kodierung.

8 Das Dateiformat MP3

Liegen die MP3-Audiodaten fertig vor, so müssen diese strukturiert gespeichert werden. Das geschieht über die Definition eines Dateiformates.

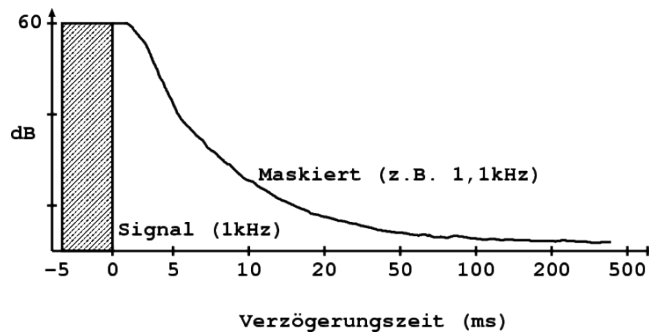


Abbildung 15: Zeitmaskierung [Leidinger 2006]

8.1 Gliederung

Eine MP3-Datei ist in unabhängige kleinere Teile, genannt *Frames*, geteilt [Supurovic 1999]. Jedes Frame ist unabhängig von den anderen mit den besprochenen Kompressionsmethoden komprimiert und hat daher seinen eigenen Header. Es ist nicht zwingend, dass alle Header in einer Datei die gleichen Parameter enthalten, im Regelfall sollte dies jedoch der Fall sein.

[Tabelle 1] zeigt einen Kurzüberblick, welche Daten in einem Frame-Header enthalten sind. Der Header ist 32 Bit lang. Für eine umfangreiche Beschreibung wird auf [Supurovic 1999] verwiesen.

8.2 Der ID3-Tag

Der *ID3-Tag* ist aus dem Bedürfnis der Benutzer entstanden, zusätzliche Daten wie Titel oder Interpret zu einem MP3 zu speichern. Er ist nicht offizieller Bestandteil des MPEG Standards. Der ID3-Tag entspricht einem 128 Byte langem Tag am Ende der Datei, der Informationen zum Musikstück speichern kann. [Tabelle 2] zeigt den Aufbau dieses Headers.

Neben dem in [Tabelle 2] an die Datei angehängten ID3-Tag gibt es noch den *ID3v2* Tag, die zweite Version des ID3-Tags. Dieser ist viel komplizierter aufgebaut und wird am Anfang der MP3-Datei gespeichert. Vorteil ist, dass die Größe der Datenfelder dynamisch gespeichert sind. Nachteil ist jedoch, dass bei jeder Änderung die ganze Datei neu geschrieben werden muss, da sich die Informationen am Anfang der Datei befinden. Die Änderung kann daher zu längeren Aussetzern während des Abspielens führen.

Länge (Bits)	Position (Bits)	Inhalt
11	31-21	Frame sync (alle Bits gesetzt)
2	19-20	MPEG Audio Version
2	17-18	Layer, für MP3 01
1	16	CRC nach dem Header. Im Regelfall nicht verwendet
4	12-15	Bitrate, kodiert.
2	10-11	Abtastrate. Meistens 00 für 44.1kHz
1	9	Frame padding. Wenn gesetzt, folgen Padding-Daten
1	8	Frei verfügbar
2	6-7	Channel Modus. Meistens 01 für Joint-Stereo
2	4-5	Zusätzliche Einstellungen für Joint-Stereo
1	3	Audio ist geschützt (Copyright)
1	2	Stück ist ein Original
2	0-1	Emphase (Anhebung der Amplitude)

Tabelle 1: MPEG Audio Frame Header [Supurovic 1999]

Länge (Bytes)	Position (Bytes)	Inhalt
3	0-2	Magisches Wort TAG
30	3-32	Titel
30	33-62	Interpret
30	63-92	Album
4	93-96	Jahr
30	97-126	Kommentar
1	127	Genre

Tabelle 2: Der ID3 Tag [Supurovic 1999]

9 Zusammenfassung - Ausblick

In dieser Arbeit wurde Technik hinter MP3 eingehend beleuchtet: Wie die aufgenommenen Audiodaten in den Computer gelangen, wie die Kompression funktioniert und welche Teile entfernt werden können. Durch die ausgeklügelten Kompressionsmechanismen ist das Maximum an Kompression für ein Audioformat scheinbar erreicht.

Jedoch scharrt bereits der Nachfolger von MP3 in den Startlöchern: Advanced Audio Coding (AAC) basiert auf dem Nachfolger von MPEG-1, MPEG-2. Neben der verbesserten Kodierung, die Unzulänglichkeiten von MP3 ausbessern soll, erhält AAC ein neues Dateiformat: MP4. Dieses fügt neue Features wie die Kodierung von Metadaten oder das in anderen Formaten bereits verfügbare

und stark von der Industrie geforderte Digital Rights Management (DRM) hinzu. Mit DRM ist durch kryptographische Schutzmechanismen (die u.a. einen Kopierschutz umfassen) der kontrollierte Verkauf von digitaler Musik über das Internet möglich.

Literaturverzeichnis

Bücher

- [Bartsch 2004] BARTSCH, Hans-Jochen: *Taschenbuch Mathematischer Formeln*. 20. Auflage. Fachbuchverlag Leipzig, 2004
- [Cormen et al. 2000] CORMEN, Thomas H. ; LEISERSON, Charles E. ; RIVEST, Ronald L.: *Introduction to Algorithms*. 25. Auflage. McGraw-Hill Book Company, 2000
- [Horn und Dourdoumas 2004] HORN, Martin ; DOURDOUMAS, Nicolaos: *Regelungstechnik*. 2. Auflage. Pearson Studium, 2004
- [Oppenheim et al. 2004] OPPENHEIM, Alan V. ; SCHAFER, Roland W. ; BUCK, John R.: *Zeitdiskrete Signalverarbeitung*. 2., überarbeitete Auflage. Pearson Studium, 2004
- [Paturi 1989] PATURI, Felix R.: *Chronik der Technik*. 3. Auflage. Chronik Verlag, 1989
- [Roederer 2000] ROEDERER, Juan G.: *Physikalische und psychoakustische Grundlagen der Musik*. 3. Auflage. Springer, 2000
- [Spiegel 1990] SPIEGEL, Murray R.: *Fourier Analysis, Theorie und Anwendung*. McGraw-Hill Book Company GmbH, 1990
- [Tietze und Schenk 2002] TIETZE, Ulrich ; SCHENK, Christian: *Halbleiter Schaltungstechnik*. 12. Auflage. Springer, 2002

Diplom- und Doktorarbeiten

- [Schriber 2000] SCHRIBER, Pascal: *Das neue Audiospeicherformat MP3*, Institut für Informatik der Universität Zürich, Diplomarbeit, 5 2000. – <http://www.ifi.unizh.ch/mml/publications/diplomarbeiten/schriber.pdf>

Skripten

- [Berglez 2004] BERGLEZ, Peter: *Numerisches Rechnen und Lineare Algebra* / Institut für Mathematik A der Technischen Universität Graz. 2004. – Vorlesungsskriptum
- [Brasseur 2006] BRASSEUR, Georg: *Elektrische Messtechnik Labor* / Institut für Mathematik A der Technischen Universität Graz. 2006. – Vorlesungsskriptum
- [Leopold 2005] LEOPOLD, Hans: *Elektronische Schaltungstechnik 2* / Institut für Mathematik A der Technischen Universität Graz. 2005. – Vorlesungsskriptum

Internetressourcen

- [Beis 2005] BEIS, Uwe: *Eine Einführung in Sigma-Delta Wandler*. <http://www.beis.de/Elektronik/DeltaSigma/DeltaSigma.html>. 2005. – (18.12.2006)
- [Leidinger 2006] LEIDINGER, Alexander: *Audiokompression (Datenreduktion)*. <http://www.leidinger.net/publications/Audiokompression>. 2006. – (20.05.2006)
- [Supurovic 1999] SUPUROVIC, Predrag: *MPEG Audio Frame Header*. <http://www.dv.co.yu/mpgscript/mpeghdr.htm>. 1999. – (15.12.2006)
- [Wikipedia 2006a] WIKIPEDIA: *Abtasttheorem*. <http://de.wikipedia.org/wiki/Nyquist-Shannon-Abtasttheorem>. 2006a. – (18.12.2006)

- [Wikipedia 2006b] WIKIPEDIA: *Dirac-Kamm*.
<http://de.wikipedia.org/wiki/Dirac-Kamm>. 2006b. – (18.12.2006)
- [Wikipedia 2006c] WIKIPEDIA: *Fourier-Transformation*.
<http://de.wikipedia.org/wiki/Fourier-Transformation>. 2006c. – (18.12.2006)
- [Wikipedia 2006d] WIKIPEDIA: *Joint-Stereo*.
<http://de.wikipedia.org/wiki/Joint-Stereo>. 2006d. – (18.12.2006)