
BACHELORARBEIT

Herr
Moritz Schiek

**Anforderung an den Master-
prozess im gegenwärtigen
Kontext**

2020

BACHELORARBEIT

Anforderung an den Master- prozess im gegenwärtigen Kontext

Autor/in:
Herr Moritz Schiek

Studiengang:
2016

Seminargruppe:
MG16wA-B

Erstprüfer:
Prof. Dr.-Ing. Michael Hösel

Zweitprüfer:
B. Eng Kai Hoffmann

BACHELOR THESIS

1

**Requirement for the master-
process in the current context**

author:

Mr. Moritz Schiek

course of studies:

2016

seminar group:

MG16wA-B

first examiner:

Prof. Dr.-Ing. Michael Hösel

second examiner:

B. Eng Kai Hoffmann

submission:

Mittweida, 01.07.2020

Bibliografische Angaben

Nachname, Vorname: Schiekel, Moritz

Thema der Bachelorarbeit: Aktuelle Bedeutung von Mastering und Anwendung

Requirement for the masterprocess in the current context

55 Seiten, Hochschule Mittweida, University of Applied Sciences,
Fakultät Medien, Bachelorarbeit, 2020

Abstract

Im Prozess des Masterings geht es darum, dass durch den Urheber erschaffene Musikstück für die Veröffentlichung vorzubereiten. Die Anforderungen an den Prozess ergeben sich aus den Wünschen des Kunden und den Anforderungen der Medien und Wiedergabesysteme, welche für das Wiedergeben der Musik genutzt werden. Durch den technischen Fortschritt haben sich die Anforderungen an das Mastering verändert. In der Vergangenheit war die Lautheitssteigerung einer der wichtigsten Teilaspekte des Masteringprozesses. Es hat sich gezeigt, dass der Trend zur Nutzung von Streamingdiensten geht, welche zum Großteil die Lautheit von Liedern anpassen. Es werden jedoch immer noch Medien verwendet, welche nicht die Lautheit von Liedern normalisieren, wodurch zwei unterschiedliche Ausrichtung des Masteringprozesses resultieren. Da die Lautheitsnormalisierung selten durch Anbieter der Medien dokumentiert ist, wurden durch eine experimentelle Messung die Eigenschaften der verschiedenen Medien ermittelt. Aus den Erkenntnissen der Recherchearbeit und den Messungen wurden Anforderungen an einen aktuellen Masteringprozess abgeleitet und konkrete Werte für die Ausrichtung des Masteringprozesses geschaffen. Um die bestmögliche Qualität und Wettbewerbsfähigkeit eines Musikstückes nach dem Masteringprozess zu gewährleisten, müssen unterschiedliche Versionen mit unterschiedlichen Anforderungen erstellt werden. Ist dies aus ökonomischen Gründen nicht möglich, kann ein Master angefertigt werden welches den gemeinsamen Anforderungen aller Medien nachkommt.

Inhaltsverzeichnis

Inhaltsverzeichnis	I-II
Abbildungsverzeichnis	III
Tabellenverzeichnis	IV
1 Einführung und Zielstellug	1
2 Mastering	3
2.1 Mastering Eindordnung in den Produktionsprozess.....	3
2.2 Bereiche des Masterings.....	3
2.2.1 Aussteuerung.....	3
2.2.2 Lautheit und Zieldynamik	4
2.2.3 Frequenzgang.....	6
2.2.4 Abtastrate und Bittiefe.....	7
2.2.5 Quantisierungsfehler und Korrektur.....	8
2.2.6 Datenreduktion.....	10
2.2.7 Pausen,Fades,Texte und Menüs	12
2.2.8 Daten Distribution	13
2.3 Effekte für das Mastering	13
2.3.1 Dynamikprozessoren	13
2.3.2 Filter und Entzerrer	18
3 Lautheitssteigerung im Mastering	22
3.1 Wahrnehmung von Lautheit	22
3.1.1 Menschliches Gehör	22
3.1.2 Psychoakustik.....	22
3.1.3 Auswirkung der Psychoakustik auf das Mastering.....	24
3.1.4 Kompetitive Steigerung der Lautheit im Mastering	26
3.1.5 Lautheitsnormalisierung.....	29
3.2 Messung der Lautheit.....	30
4 Ausgangspunkt für die aktuellen Anforderungen an den Masterprozess.....	33
4.1 Aktuelle Mediennutzung für den Musikkosnum	33
4.2 Aktuelle Endgerätnutzung zur Wiedergabe von Musik	36
4.3 Verarbeitung von Musik in den Medien	37
4.3.1 UKW	37
4.3.2 Streaming-Dienste	38
4.3.3 CD	47
4.4 Eigenschaften von Wiedergabesystemen	47

4.4.1	Kleine Lautsprecher	47
4.4.2	Kopfhörer	48
4.4.3	Lautsprecher im PKW	
4.4.4	Lautsprecherboxen	
5	Anforderung an den Masterprozess	49
5.1	Ausrichtung des Masterprozesses	49
5.2	Anforderungen an den Masterprozess	50
5.2.1	Eigenschaften der Dynamik und Lautheit	50
5.2.2	Eigenschaften des Frequenzganges	52
5.2.3	Samplerate und Bittiefe	54
5.2.4	Dithering	54
5.2.5	Daten Distribution	54
6	Schlussbetrachtung.....	56
	Literaturverzeichnis	V-IX
	Eigenständigkeitserklärung	X

Abbildungsverzeichnis

Abbildung 1: Durchschnittlicher Frequenzgang der gemessenen Lieder des Streaming-Dienstes Amazone Music Unlimited	54
---	----

Tabellenverzeichnis

Tabelle 1: Messung des Streaming-Dienstes Spotify.....	40
Tabelle 2 Messung des Streaming-Dienstes Apple Music	42
Tabelle 3: Messung des Streaming-Dienstes Amazone Music Unlimited.....	44
Tabelle 4: Messung des Streaming-Dienstes Soundcloud.....	45
Tabelle 5: Messung von Music-Videos des Video-Streaming-Dienstes YouTube	46
Tabelle 6: Messung des Streaming-Dienste YouTube Music.....	46

1 Einführung und Zielstellung

Mastering ist Teil einer Kette von Bearbeitungsschritten, die im Umgang mit Audiomaterial durchgeführt werden. Um das Audiomaterial auf das Massenmedium vorzubereiten und die Vorstellung des Künstlers zu erfüllen, setzt sich der Masteringprozess erfahrungsgemäß aus subjektiven, kreativen und technischen Aspekten zusammen.

Massenmedium und Wiedergabesysteme legen durch ihre Beschaffenheit die technischen Anforderungen an ein Mastering fest. Diese Anforderungen müssen bei der Erstellung eines Masters eingehalten werden, um so die Qualität trotz subjektiven Spielraums zu gewährleisten. Bob Katz schreibt dazu in seinem Buch *Mastering Audio* „Regardless of the form in which product is sold, our job as mastering engineers remains: we help music to be presented in the best possible way.“ (Vgl. Katz 2015, S. 28)

Durch die sich ständig ändernde Technik ist auch der Prozess des Masterings einer stetigen Änderung unterworfen. Besonders durch Einführung von technischen Errungenschaften wie Brickwall Limiter, die CD als Massenmedium und Loudness Normalization haben sich die Anforderungen an das Mastering stark verändert. Der Autor konnte während seines Pflichtpraktikums schwerlich aktuelle Abhandlungen finden, die den Zusammenhang zwischen den Anforderungen der Massenmedien, Wiedergabesysteme und dem Masteringprozess beschreiben. Massenmedien ist eine „Sammelbezeichnung [...] für Mittel (= Medien), mit denen Nachrichten und Unterhaltung in Schrift, Ton und Bild zu einem breiten Publikum kommen.“ (Bundeszentrale für politische Bildung 2011)

Das Bearbeiten von Audiomaterial ist ein sehr kompetitives Geschäft. Der Endnutzer kann mit seinem Konsumverhalten die Auftragslage der Mastering-Firmen bestimmen. Wenn bestimmte Bearbeitungsmaßnahmen das Konsumverhalten von Endnutzern steigern können, ist es wichtig, die Anforderungen an den Masterprozess im aktuellen Kontext zu ermitteln. Dadurch kann das Wahrnehmungsempfinden des Konsumenten, die Qualität der Medien und der Absatz der Firmen gesteigert werden.

Die Arbeit kann sich nur auf die aktuellen und historischen technischen Voraussetzungen beziehen. Viele Prozesse des Masteringprozesses sind sehr ähnlich geblieben, allerdings haben sich Tendenzen und konkrete Werte immer wieder geändert. Es ist nicht möglich vorherzusagen, wie sich die Technik in den kommenden Jahrzehnten ändern wird. Zum Beispiel ist es denkbar, dass durch die Einführung von 5G und die damit verbundene Steigerung der Datenraten auch höherwertige Datenformate für das Konsumieren von Musik und Videos verwendet werden. Es wäre zugleich denkbar, dass eine Form der Sättigung eintritt und eine weitere Steigerung der Qualität auf Grund der geringen Wahrnehmbarkeit auf der Seite des Endkonsumenten nicht lohnenswert ist.

Der Inhalt von Audiosignalen lässt sich in viele Bereiche unterteilen. Es gibt Audiomaterial, was nur oder vorwiegend aus Sprache besteht, wie zum Beispiel Hörbücher oder Podcasts. In Filmen besteht das Audiosignal meist aus einem Gemisch von Sprache, atmosphärischen Aufnahmen, Musik und Effektgeräuschen. Diese verschiedenen Audiosignale weisen alle verschiedene Charakteristiken in Dynamik und Frequenzzusammensetzung auf. Die Charakteristiken dieser beiden Eigenschaften können starken Einfluss darauf nehmen, wie das Audiomaterial bearbeitet werden muss oder kann. Um der Arbeit einen klaren inhaltlichen Rahmen zu bieten, möchte sich der Autor auf den Signaltyp „Musik“ beschränken.

Ein Teil des Aufgabenbereiches vom Mastering kann das Stem-Mastering sein. „Stems are a special kind of submix“. (Katz 2015, S. 133) Die Submixe entsprechen der Summe des Mixdowns. Im Mastering werden diese Submixe wieder zu einer Summe zusammengeführt. Die Summe kann dann anschließend gemastert werden. Der Vorteil besteht darin, dass auf eventuell auftauchende Probleme in einzelnen Instrumentengruppen, reagiert werden kann. Das Stem-Mastering ist deshalb auf die Verschleppung von Fehlern im Produktionsprozess zurückzuführen. Die eigentliche Aufgabe des Masterings ist es, die Summe aus ausbalancierten Instrumenten, in der bestmöglichen Form, für die Distribution vorzubereiten. Obgleich in der heutigen Zeit durch enge Ablaufpläne und Mangel an Produktionsgeldern ein Stem-Mastering zuweilen nicht aus dem Produktionsprozess ausschließbar ist, wird in dieser Arbeit von einem idealisierten Aufgabenbereich des Mastering-Prozesses ausgegangen. Infolgedessen geht der Autor in Bezug auf das Mastering nur von der Bearbeitung eines fertig summierten Signales aus.

In der Arbeit soll als Erstes das Mastering in den Produktionsprozess von Musik eingeordnet und seine Aufgabenbereiche erschlossen werden. Ein besonderer Schwerpunkt wird die Lautheit sein. Ursache für die Priorisierung ist die starke Gewichtung der Lautheitssteigerung innerhalb des Masteringprozesses.

Das Ziel der Arbeit ist es, zu untersuchen, wie aktuelle Massenmedien, Wiedergabesysteme und andere Faktoren den Masterprozess beeinflussen und welche Anforderungen sie stellen. Sofern möglich, sollen daraus konkrete Werte gebildet werden, welche in der Praxis verwendet werden können.

2 Mastering

2.1 Mastering Einordnung in den Produktionsprozess

Als Mastering beziehungsweise „technical mastering“ wird die Vorbereitung einer Pressvorlage im Presswerk bezeichnet. Die richtige Bezeichnung für die abschließende Bearbeitung des Audiomaterials und das Vorbereiten eines Datenträgers für das Presswerk wird als „Premastering“ bezeichnet. (vgl. Katz 2015, S. 10) Erfahrungsgemäß ist es in der Branche üblich, den Begriff Mastering für den Prozess des Premasterings zu verwenden. Zu dem Aufgabenbereich des Mastering-Ingenieurs gehört, sicherzustellen, dass die Qualität des Masters im Studio dieselbe ist, wie im Presswerk und bei dem Endkonsumenten. (Vgl. Katz 2015, S. 13)

Mastering wird nicht nur bei der Herstellung von CDs angewandt. Die mediengerechte Klang- und Format-Optimierung spielt bei jeder Art von Tonträger-Vervielfältigung oder Distribution eine Rolle“. (Bock 2014, S. 881) Die abschließende Bearbeitung von Audiomaterial kann in mehrere Bereiche unterteilt werden, die je nach Zustand des Ausgangsmaterials komplex sind. Die Bereiche können technisch oder gestalterisch beschaffen sein. (Vgl. Katz 2015, S. 10) Beide Ausprägungen können sich gegenseitig beeinflussen. Aus technischer Sicht wird dem Rauschen eine qualitätsmindernde Eigenschaft zugeschrieben. In den meisten Fällen wird während der Bearbeitung versucht, das Rauschen in dem Audiomaterial zu unterdrücken. In seltenen Fällen wird Rauschen als gestalterisches Mittel in dem Audiomaterial beibehalten und als Element der Musik angesehen. (Vgl. Katz 2015, S. 20) Vorwiegend bildet die technische Beschaffenheit der einzelnen Bereiche die Basis für die gestalterische Freiheit. Zum Beispiel ist der Frequenzgang des Audiomaterials durch den Frequenzgang des Wiedergabesystems des Nutzers eingeschränkt. Der Bearbeiter des Audiomaterials kann Veränderungen ausführen, die den Charakter des Audiomaterials beeinflussen.

2.2 Bereiche des Masterings

2.2.1 Aussteuerung

Eines der wichtigsten Aufgaben des Masterings ist die Anpassung des Pegels beziehungsweise dessen korrekte Aussteuerung. „Der digitale Pegel wird relativ zur maximalen Aussteuerbarkeit des Systems (also zur größten darstellbaren Zahl) bestimmt und in dBFS („full Scale“) angegeben.“ (Görne 2015, S. 35) Eine korrekte Aussteuerung bedeutet, dass der maximale Wert des signalsendenden Systems nicht größer ist als, der

maximal zulässige Wert des signalempfangenden Systems. Die korrekte Aussteuerung ist deshalb immer vom gewählten Massenmedium des Audiomaterials abhängig. Ein Aussteuern über die Systemgrenzen wird auch als Übersteuern bezeichnet. Bei Übersteuerung wird die Wellenform begrenzt, wodurch Störgeräusche entstehen.

Bei der digital-analog-Wandlung kann es zu leichten Pegelanhebungen im Vergleich zum Quellsignal kommen. Das kann dazu führen, dass ein Signal, was mit einem Samplepeak von 0 dBFS ausgesteuert wurde, beim Endkonsumenten bei der Wiedergabe übersteuert. Um eine Übersteuerung beim Kunden zu vermeiden, sollte beim Mastering der Pegel mit True Peak gemessen werden. „Mit Hilfe von mindestens 4-fachem Oversampling des anzuzeigenden Signals wird eine Vorhersage auch von sog. „Intersample-Peaks“ möglich, also von Spitzenpegeln, die zwischen 2 Samples liegen. [...] [D]iese „wahren“ Spitzenpegel [können] bis zu 3 dB und mehr über dem höchsten Samplewert liegen.“ (Spikofski und Camerer 2014, S. 1294–1295) True Peak und Samplepeak unterscheiden sich nur im Messverfahren. Beide werden in dBFS angegeben. Das True Peak Messverfahren ist durch die ITU-R BS.1770-3 definiert. (Vgl. Spikofski und Camerer 2014, S. 1304)

„Um eine möglichst hohe Klangqualität zu erreichen, sollte das Programm-Material in seinem Maximalpegel knapp die Aussteuerungsgrenze (Full Scale) des Tonträgers (entspr. ≤ 0 dBFS) erreichen.“ (Bock 2014, S. 882) Das verhindert zudem, dass nach der Requantisierung sehr leise Signale durch das Dithering verrauscht oder nicht mehr ausreichend gut abgebildet werden.

2.2.2 Lautheit und Zieldynamik

Lautheit und Dynamik sind Eigenschaften, die helfen, das Audiomaterial zu beurteilen. Die Manipulation dieser Parameter kann sich auf das Hörerlebnis des Konsumenten positiv oder negativ auswirken. Beide Parameter können sich gegenseitig beeinflussen. Eine Änderung der Lautheit kann eine größere Dynamik bewirken. Gleichzeitig kann eine Änderung der Dynamik eine größere Lautheit erzeugen. Im Mastering wird erfahrungsgemäß immer die Dynamik verringert und die Lautheit gesteigert. Dafür gibt es zwei Gründe. Durch die geringere Dynamik wird während der Wiedergabe der Abstand von Nutzsignal zu den Umgebungsgeräuschen größer. Leise Passagen im Lied sind so bei angemessener Aussteuerung immer noch verständlich. Friesecke schreibt dazu: „Da der Endkonsument unter normalen Hörbedingungen die hohe Dynamik nicht hören kann, die die Instrumente bei einer Aufnahme erzeugen, muss man die Dynamik einschränken.“ (Friesecke 2007, S. 678) Dabei sollte es dem Hörer ermöglicht werden, „dass eine Produktion im Idealfall ohne Nachregeln des Pegels abgehört werden kann.“ (Bock 2014, S. 882) Auch Bob Katz ist derselben Auffassung: „Usually the recording medium and

intended listening environment simply cannot keep up with the full dynamic range of real life, so the mastering engineer is often called upon to raise the level.“ (Katz 2015, S. 75) Gleichzeitig bringt eine höhere Lautheit den psychoakustischen Effekt, dass Konsumenten ein lauterer Lied qualitativ höherwertig einschätzen, als ein vergleichbar leiseres. Im Mastering müssen deshalb Entscheidungen für Lautheit und Dynamik getroffen werden. Die Entscheidungen hängen dabei von den Anforderungen und der Beschaffenheit des Zielmediums, Wiedergabesystem und der beabsichtigten Klangästhetik ab. „Im Bereich der Unterhaltungsmusik ist die Dynamikbearbeitung ein wichtiges Stilmittel zur Erzeugung von „Druck“: Ein Rockmusiktitel soll kompakt und wuchtig klingen.“ (Bock 2014, S. 882)

Lautheit beeinflusst das subjektive Empfinden des Konsumenten. Sie hat deshalb einen großen Einflussfaktor auf die Qualität. Zum Beispiel wurde in der EBU R 128 Empfehlung eine einheitliche Richtlinie für die Programmlautheit geschaffen. Sie findet im öffentlichen Fernsehen und Rundfunk von Deutschland Verwendung. Lautheit beschreibt, wie laut der Mensch eine Quelle gegenüber einer anderen wahrnimmt. Bob Katz verdeutlicht, dass die Lautheit zur Beschreibung der Wahrnehmung des Zuhörers verwendet wird. (Vgl. Katz 2015, S. 218) Da es sich bei Lautheit um Eindrücke handelt, ist es schwer, diese in Zahlen zu fassen. Bob Katz begründet das damit: „Exposure time and context also affect our perception“. (Katz 2015, S. 218) Lautheit wird in LU gemessen. Wird die Einheit auf das digitale Vollaussteuerung bezogen, so erhält man die Einheit LUFS für full scale.

Dynamik beschreibt, wie groß oder klein der Abstand zwischen geringstem und höchstem Wert ist. Dabei kann zwischen der Dynamik des Pegels und der Lautheit unterschieden werden. Die Dynamik ist mit bestimmten Zeitintervallen korreliert. Diese Zeitintervalle werden als Macro und Micro bezeichnet. Bob Katz beschreibt, dass Macrodynamik der Unterschied in der Lautheit zwischen Teilen von Liedern oder zwischen Liedern selbst sein kann. (Vgl. Katz 2015, S. 74) Macrodynamik ist vergleichbar mit der Abstufung der Tonlautstärke in der Musik bis auf die Tatsache, dass sich bei einer Änderung der Lautheit im Mastering das Frequenzspektrum nicht zwangsweise ändert. Die Macrodynamik kann durch ein, in der EBU R 128 Empfehlung festgelegtes, Verfahren gemessen werden. Das Verfahren ermittelt dabei den LRA in dB.

Microdynamiken sind dagegen Lautstärkeänderungen in einem kurzen Intervall. Für Bob Katz umfassen Microdynamiken „transient quality, integrity or bounce, which involves the music's short-term peaks.“ (Katz 2015, S. 74) Zur Beurteilung der Microdynamik kann der Crestfaktor verwendet werden. „Der Crestfaktor (F_c), oder auch Scheitelfaktor, beschreibt das Verhältnis zwischen dem Spitzenwert (U_s) einer Wellenform und deren Effektivwert.“ (Friesecke 2007, S. 208)

2.2.3 Frequenzgang

Der Frequenzgang beschreibt den Zusammenhang zwischen Amplitude und Frequenz. Im Mastering wird der Frequenzgang aus technischen oder ästhetischen Gründen beeinflusst. Stefan Bock schreibt: „Ein Mastering-Ingenieur kann die Mischung unvoreingenommen beurteilen, um dann ein Maximum an Klangqualität und Wiedergabekompatibilität zu schaffen.“ (Bock 2014, S. 881)

Jedes Wiedergabesystem hat seinen eigenen Frequenzgang. Jedes Glied in der Signalkette kann eine Änderung hervorrufen. Die Signalkette bei dem Endkonsumenten besteht immer aus einem digital-analog-Wandler, Verstärker, Treiber, Gehäuse und Raum. Die unterschiedlichen Einflüsse der Komponenten addieren sich. Es entsteht eine Differenz zwischen dem originalen Frequenzgang und dem Wiedergabesystem des Konsumenten. Dieser Frequenzgang kann deutlich von dem im Mastering abweichen. Erfahrungsgemäß modifizieren Hersteller den Frequenzgang des Wiedergabesystems um sich einen Wettbewerbsvorteil zu erschaffen. Dabei wird der Frequenzgang so angepasst, dass er für den Kunden gefälliger wirkt oder einen bestimmten Klang aufweist, der mit dem Hersteller verbunden werden kann. Manche Wiedergabesysteme sind auch durch ihre Bauform begrenzt und können nur einen bestimmten Frequenzbereich effektiv abbilden. Bob Katz ist der Meinung, dass vor allem niedrigpreisige Lautsprecher einen eingeschränkten Wiedergabefrequenzgang haben. (Vgl. Katz 2015, S. 28) Die aufgezählten Beispiele führen dazu, dass das Hörerlebnis sich deutlich unterscheidet. Aufgabe im Masteringprozess ist es deshalb, das Hörerlebnis auf vielen Wiedergabesystemen annähernd anzugleichen.

Im Masteringprozess muss der Amplitudenfrequenzgang eines Liedes auch aus Sicht eines ästhetischen Aspektes geprüft und wenn nötig bearbeitet werden. Erfahrungsgemäß sollte dieser ausgewogen sein. Das bedeutet, dass tiefe, mittlere und hohe Frequenzen nicht über oder unterrepräsentiert werden. Es können keine einheitlichen Vorgaben geschaffen werden, weil durch unterschiedliche Musikstile und Instrumente unterschiedliche Amplitudenfrequenzgänge zustande kommen. Die Einschätzung, ob ein Stück den passenden klanglichen Charakter hat, hängt stark von der Erfahrung des Bearbeiters ab. Bob Katz schreibt: „As we become more sophisticated in our approach to listening, we develop a greater awareness of the subtleties of sonic and musical reproduction.“ (Katz 2015, S. 25)

Während des Masteringprozess werden störende Frequenzen häufiger abgeschwächt, als das gefällige Frequenzanteile gehoben werden. Eine Erklärung dafür ist „the ear hears boosts or resonances more easily than dips or absences.“ (Katz 2015, S. 26) Filterprozesse werden deshalb in subtraktive Filterung und additive Filterung unterschieden. Wird eine subtraktive Bearbeitung vorgenommen, werden die Amplituden der

ausgewählten Frequenzen mit einer positiven rationalen Zahl kleiner eins multipliziert. Wird additiv gearbeitet, werden die Amplituden mit einer rationalen Zahl größer eins multipliziert.

2.2.4 Abtastrate und Bittiefe

Für die Fertigstellung des Masters müssen Entscheidungen über Abtastrate und Bittiefe getroffen werden. Diese richten sich nach den Anforderungen des Zielmediums. Eine Audio-CD wird zum Beispiel mit einer Frequenz von 44100 Hz abgetastet und hat eine Bittiefe von 16 Bit. (Vgl. Görne 2015, S. 335) Bittiefe und Abtastrate sind Eigenschaften, die das Beschreiben von analogen Signalen in einem digitalen System ermöglichen.

Erfahrungsgemäß arbeiten Masteringstudios mit höheren Sampleraten, als die Sample-rate des Zielmediums. „Eine maximale Klangqualität beim Masteringprozess lässt sich dadurch erreichen, dass erst unmittelbar vor dem Erstellen des Masters die Wortbreite auf z.B. 16Bit und 44,1kHz Sampling-Frequenz (bei der Audio-CD) reduziert wird.“ (Bock 2014, S. 883) Häufig verwendete Sampleraten sind 48 kHz und 96 kHz. Dem zugrunde liegt, dass die Filtergüte von anti-aliasing-Filtern bei der Konvertierung von 44,1 kHz steiler konzipiert sein muss, um das Abtasttheorem zu erfüllen. Das Abtasttheorem definiert, dass für die Beschreibung eines Signals mit einer oberen Grenzfrequenz die Abtastfrequenz doppelt so groß sein muss, wie die Grenzfrequenz. (Vgl. Görne 2015, S. 159) Bruno Putzeys äußert zur Steilheit von anti-aliasing-Filtern: „I found that when you use fantastically steep anti-alias filters, you do get an audible signature [...]. I believe a 20 kHz LPF becomes transparent around the point when the transition band is 4 kHz or more“. (Katz 2015, S. 214) Ein Übertragungsband von 4 kHz machen Sampleraten von 48 kHz oder höher notwendig. Der gleichen Meinung ist auch Herr Görne: „Durch Vervielfachung der Abtastrate [...] kann [man] analoge Tiefpassfilter mit geringerer Flankensteilheit einsetzen.“ (Görne 2015, S. 164) Auch im Handbuch der Audiotechnik wird die Qualitätssteigerung bei hohen Abtastraten auf die Eigenschaften des Filters der AD-Wandler zurückgeführt. „Zwar gibt es keine Belege dafür, dass Schallanteile über 20 kHz gehört werden, aber es wird argumentiert, dass die Analog-Digital-Wandlung herkömmlicher Systeme Artefakte im hörbaren Bereich produziert (z.B. Abweichungen von einem „flachen“ Frequenzgang; Dunn 1998), die durch die hohe Abtastrate reduziert werden können.“ (Ellermeier W. 2008b, S. 55) Filter für hohe Sampleraten weisen weniger Phasenfehler in mittleren und hohen Frequenzen auf, als die steiflankigen Filter. (Vgl. Görne 2015, S. 164) Das führt zu einer getreueren Wiedergabe der Transienten und des Stereobildes.

Die Bittiefe des Audiomaterials wird gewöhnlich aus dem vorangegangenen Prozess, dem Mischen, übernommen. Bei der Umrechnung in das Zielformat ändert sich meistens die Bittiefe. Zum Beispiel hat eine Audio-CD eine Bittiefe von 16 Bit während in den

verschiedenen Produktionsphasen eine Bittiefe von 24 Bit zum Einsatz kommen sollte. Höhere Bittiefen haben den Vorteil, dass Berechnungen um ein Vielfaches genauer sind. Fehler in der Präzision sind im Audiomaterial wesentlich weniger hörbar. Präzisionsfehler entstehen zum Beispiel bei der Verstärkung des Pegels. „Multiplications in generally results in a longer wordlength [...] and the wordlength can increase, up to the precision of the DSP.“ (Katz 2015, S. 204) Durch die Verkleinerung der Anzahl an Quantisierungsstufen am Ende des Masteringprozesses treten Fehler auf, die als Rundungsfehler bezeichnet werden. Bob Katz schreibt, dass diese Fehler als „[...] buzz, grit, harshness, coldness and/ or loss of depth in the sound“ (Katz 2015, S. 200) wahrgenommen werden. Um während des Masteringprozesses die Quantisierungs- und Rundungsfehler im Zielmedium zu marginalisieren, wird Dithering am Schluss der Signalkette eingesetzt.

2.2.5 Quantisierungsfehler und Korrektur

Quantisierungsfehler entstehen bei analog-digital-Wandlung und bei der Änderung der Bittiefe. Das Signal muss während der analog-digital-Wandlung aus einem kontinuierlichen in einen Wert diskreten Bereich übertragen werden. Das bedeutet, dass die Funktion nicht mehr unendlich viele Werte, sondern nur noch eine bestimmte Anzahl an festgelegten Werten annehmen kann. Das Signal wird deshalb vor der Quantisierung gestuft, um es mit einem Referenzwert vergleichbar zu machen. Dieser Prozess wird als „Sample and Hold“ bezeichnet. „Die momentan gemessene Signalspannung wird für eine gewisse Zeit gehalten, damit der nachfolgende Quantisierer diese Spannung mit einer Referenzspannung vergleichen kann.“ (Görne 2015, S. 163) Die Signalspannung wird dann dem nächstgelegenen Referenzwert zugeordnet. Dadurch, dass es sehr unwahrscheinlich ist, dass die Signalspannung dem Referenzwert entspricht, entsteht ein Fehler, der sich in der Differenz zwischen Referenzwert und Signalspannung ausdrückt. „Die Differenz zwischen der Amplitude des ursprünglichen analogen Signals und der nächstgelegenen Stufe der Quantisierungs-Kennlinie ist der *Quantisierungsfehler*.“ (Görne 2015, S. 170) Im Handbuch der Tonstudioteknik wird beschrieben, dass bei Signalen mit großen Amplituden der Quantisierungsfehler vom Klang einem weißen Rauschen gleichen würde, wohingegen bei kleinen Amplituden der Fehler als ein in Abhängigkeit vom Eingangssignal stehendes, grobes Rauschen wahrgenommen würde. (vgl. Götz 2014, S. 670) Weißes Rauschen ist ein „fortdauerndes Schallsignal, das theoretisch aus unendlich vielen, in ihrer Frequenz unendlich dicht nebeneinander liegenden sinusförmigen Schwingungen besteht.“ (Dickreiter 2014b, S. 3) Die „wesentliche Eigenart des Rauschens ist die zufällige Folge von Amplituden, die [...] eine Gauß'sche Häufigkeitsverteilung haben.“ (Görne 2015, S. 174)

Ein weiterer Quantisierungsfehler entsteht, wenn die Bittiefe verringert wird. Das Verringern der Bittiefe ist vergleichbar mit dem Einkürzen der Kommastellen einer rationalen

Zahl. Wenn Kommastellen eingekürzt werden, wird die Präzision der Zahl ungenauer. Das Verringern der Bittiefe geschieht während des Exports, weil die meisten DAWs und DSPs intern mit einer höheren Bittiefe rechnen und das Zielmedium für die Distribution oft eine geringere Bittiefe aufweist. Zum Beispiel kann die Mastering-Software *Wavelab Pro* in der Version 9.5 „viele Formate laden, verarbeitet sie intern jedoch als 64-Bit-Float-Samples.“ (Steinberg Media Technologies GmbH 2018, S. 39) Beim Export muss die Bittiefe auf den ausgewählten Wert reduziert werden. Die überflüssigen Bits werden dabei weggelassen oder gerundet. (Vgl. Götz 2014, S. 681) Dabei entstehen Differenzen zwischen den neuen und den originalen Werten. Diese Fehler werden als Rundungsfehler bezeichnet. Sie „sind identisch mit Quantisierungsfehlern, nur entstehen sie im digitalen Signal. Man bezeichnet sie daher auch als Rundungsrauschen.“ (Görne 2015, S. 171) Da Quantisierungsfehler und Rundungsfehler sich in ihren Eigenschaften nicht unterscheiden, entsteht auch derselbe qualitätsmindernde Höreindruck.

Dithering ist ein Prozess, welcher verhindert, dass Quantisierungsfehler und Rundungsfehler zu einem merklich schlechteren Hörerlebnis beitragen. Während des Dithering wird dem Signal ein Rauschen hinzugefügt. Dabei sollte der Effektivwert des Rauschens dem Wert einer halben Quantisierungsstufe entsprechen. Werte die zwischen zwei Quantisierungsstufen liegen, werden durch das Rauschen der höheren oder niedrigeren Quantisierungsstufe abwechselnd zugewiesen. Der Durchschnitt des über zwei Quantisierungsstufen oszillierenden Signals entspricht dem Originalwert. „im zeitlichen Mittel ist also die geditherte Quantisierung genauer als das Quantisierungsraster selbst.“ (Weinzierl und Lerch 2008, S. 796) Dadurch kann auch bei sehr kleinen Signalen, die eine Amplitude unterhalb einer Quantisierungsstufe haben, die Information in das Rauschen kodiert werden.

Das Ditherrauschen sollte eine maximale Amplitude von einer halben Quantisierungsstufe der Ziel-Bittiefe haben. Dabei können die Amplituden verschiedene Werte zwischen $-q/2$ und $+q/2$ annehmen. Verschiedene Arten von Dithering können durch die Wahrscheinlichkeitsverteilung der möglichen Amplituden des Rauschens beschrieben werden. Zum Beispiel können die Wahrscheinlichkeiten für die Amplitudenwerte zwischen $-q/2$ und $+q/2$ gleich groß sein. Man spricht dann von einer „rechteckförmiger Amplitudendichteverteilung dRect (Rec-tangular Dither).“ (Weinzierl und Lerch 2008, S. 800) Nimmt die Wahrscheinlichkeit mit wachsendem Amplitudenwert linear ab, so ergibt sich eine dreieckige Amplitudendichteverteilung und es wird als Triangular Dither bezeichnet. Diese zwei Methoden sind bei Ditherprogrammen meist frei wählbar. Im Bereich der Musikproduktion wird meistens Triangular Dither verwendet. Es weist im Gegensatz zu Rectangular Dither bei kleinen Amplituden keine Rauschmodulation auf. Eine weitere Form der Amplitudenverteilung ist die Gauß'sche Normalverteilung. In dem Buch *Tontechnik* wird beschrieben, dass dieses Dithering vor allem für analog-digital-Wandler

sehr gut geeignet ist, da die Amplitudenverteilung der des thermischen Rauschens entspricht. (vgl. Görne 2015, S. 175)

Der Pegel des Dithering berechnet sich aus dem zwanzigfachen logarithmischen Verhältnis von einer Quantisierungsstufe zu der Gesamtanzahl der durch die Wortbreite möglichen Quantisierungsstufen. Dieser Wert wird dann mit 1,76 dB, für zum Beispiel Rectangular Dither, addiert. „1,76 dB bezieht sich auf die statistische Verteilung des Quantisierungsfehlers.“ (Friesecke 2007, S. 508) Wird Rectangular Dithering bei einem Signal mit einer Bittiefe von 16 Bit eingesetzt, so ergibt sich ein Pegel des Rauschsignals von -98 dBFS. Bei einem Signal mit einer Bittiefe von 24 Bit beträgt der Pegel des Rauschens nur noch -146 dBFS.

Das Dithering kann bei hohen Wiedergabepegeln hörbar werden. Bob Katz ist zudem der Meinung, dass 16 Bit ungefiltertes Dithering „adds a slight veil to the sound, narrows the imaging and reduces the depth.“ (Katz 2015, S. 206) „Immerhin ist es aber möglich, durch eine geeignete spektrale Gewichtung seine Hörbarkeit zu vermindern, indem man die Unempfindlichkeit des Ohrs bei sehr hohen Frequenzen ausnutzt [...].“ (Görne 2015, S. 175) Dieses Verfahren wird als Noise Shaping bezeichnet. Es gibt verschiedene Algorithmen, um das Rauschen zu formen.

Algorithmen, die für eine Abtastfrequenz von 44.1 kHz geeignet sind, verschieben das Rauschen in den oberen Frequenzbereich des menschlichen Hörvermögens. „Dies bedeutet, dass die hohen Frequenzen des Rauschens meist deutlich stärker vorhanden sind, als die mittleren oder tiefen Frequenzen.“ (Friesecke 2007, S. 547) In dem Buch *Audio Enzyklopädie* wird beschrieben, dass es Filter gibt, die nach den Kurven gleicher Lautstärke angepasst worden sind. (Vgl. Friesecke 2007, S. 547) Diese eignen sich sehr gut für eine Abtastfrequenz von 44.1 kHz, da durch sie eine optimale Anpassung im Hörbereich vorgenommen werden kann. Die besten Ergebnisse lassen sich erzielen, wenn höhere Abtastraten wie zum Beispiel 96 kHz verwendet werden. Bei einer Abtastfrequenz von 96 kHz würde die Grenzfrequenz bei 48 kHz liegen. Ein Teil des Rauschens kann nun in den Bereich zwischen der Grenzfrequenz des menschlichen Gehörs und der Grenzfrequenz der Abtastrate verschoben werden. Außerdem wird in der Audio Enzyklopädie beschrieben, dass sich der Pegel des Ditherings durch eine Höhere Abtastrate um 3 dB im Hörbereich absenkt. (Friesecke 2007, S. 506)

2.2.6 Datenreduktion

Datenreduktion kommt zum Einsatz, wenn im Datenträger die Speicherkapazität oder die Datenrate im Datenkanal begrenzt sind. Datenreduktion ist die Einschränkung des

Informationsgehaltes eines Signals. In dem Buch *Tontechnik* wird beschrieben, dass man jedes Signal in drei verschiedene Informationskategorien unterscheiden könne. Jedes unveränderte Quellsignal bestehe aus relevanten, irrelevanten und redundanten Informationen. (Görne 2015, S. 184) Der maximale Informationsgehalt eines Signals [...] lässt sich aus *Bandbreite B*, *Dynamik D* und *Signaldauer T* berechnen“ (Görne 2015, S. 186) Ausschlaggebend für die Menge an Informationsgehalt in einem Signal ist die Quellcodierung. Die Quellcodierung kann in verschiedene Kategorien unterteilt werden.

Wird das Quellsignal nicht verändert, so enthält das Signal den vollen Informationsgehalt. Bis zur Distribution des Audiomaterials wird erfahrungsgemäß während des Produktionsprozesses mit verlustfreier Quellcodierung gearbeitet. „Die heute überwiegend eingesetzte Quellcodierung verwendet Pulscodemodulation (PCM)“. (Heidrich und Maniak 2014, S. 902) Ein pulscodemoduliertes Signal, ist ein quantisiertes und wertkonkretes Signal, was durch ein gestuftes analoges Signal erzeugt wurde. Das im Mastering und der Audioproduktion am meisten verwendete Dateiformat wird als *wave* bezeichnet. (Vgl. Büsching und Goderbauer-Marchner 2014, S. 171) Beim Wave-Format (Datei-Suffix *.wav*) handelt es sich um einen Audio-Daten-Container“. (Heyna et al. 2004, S. 136) Die Datei ist dabei in verschiedene Bereiche unterteilt, in welchen verschiedene Informationen wie zum Beispiel Bittiefe und Samplerate neben den Daten des PCM Signals abgelegt werden. In dem Handbuch der Tonstudioteknik wird ein Container als ein „Vielfach (Multiplex) aus Essenz-, Meta- und Steuerdaten“ beschrieben. (Schwarzenberger 2014, S. 704)

Wenn im Quellsignal die Redundanz verringert wird, aber relevante und irrelevante Daten beibehalten werden, spricht man von einer verlustfreien Datenreduktion. (vgl. Görne 2015, S. 196) Nach der Übertagung des komprimierten Signals kann das Quellsignal durch Expansion kongruent wiederhergestellt werden. (Vgl. Götz 2014, S. 685) Bandbreite und Dynamik werden dadurch erhalten. Verlustfreie Datenreduktion kann im Prozess des Mastering für ein Auditorium genutzt werden. „Das System FLAC wird u.a. in Verbindung mit dem Prozess Digitale Bemusterung [...] beim breitbandigen Audio-Filetransfer zwischen den Rundfunkanstalten eingesetzt“. (Hoeg 2014b, S. 765) FLAC gehört zu der Kategorie der verlustfreien Quellcodierung. Die Datenmenge lässt sich bei FLAC, abhängig vom Signalinhalt, um nahezu die Hälfte gegenüber einem PCM Format reduzieren. Die Stärke der Kompression lässt sich meist beim Export in ganzzahligen Stufen von null bis acht regeln. Die nullte Stufe wird am schnellsten berechnet, senkt aber die Datenrate am wenigsten. Die höchste Stufe ist die rechenaufwändigste und zugleich effektivste in Bezug zur Datenrate.

Bei der verlustbehafteten Datenreduktion werden „nicht nur redundante, sondern auch irrelevante Informationen aus dem Signal [entfernt] und verringern damit die Datenmenge drastisch [...]. Damit sind sie hervorragend zur Speicherung und Distribution

beim Endverbraucher geeignet – für professionelle Anwendungen aber *nicht*.“ (Görne 2015, S. 197) Der Informationsgehalt kann durch Einschränkung der Bandbreite oder der Dynamik verringert werden. Durch psychoakustische Modelle können Bandbreite und Dynamik so moduliert werden, dass die Qualität des Höreindrucks bei gleichzeitiger Reduktion der Datenmenge möglichst erhalten bleibt. Im Buch *Tontechnik* wird beschrieben, wie durch den Verdeckungseffekt, abhängig von dem Verlauf der Dynamik des Signals, die Abstufung der Quantisierung verändert werden kann. (Görne 2015, S. 198) Die Distribution durch ein verlustbehaftetes Format kann den Höreindruck verändern. Bei zu starker Reduzierung der Datenrate entstehen hörbare Artefakte, die als qualitätsmindernd wahrgenommen werden. Im Masterprozess gibt es die Möglichkeit mit verschiedenen Programmerweiterungen eine verlustbehaftete Codierung vorzuhören und so Anpassungen zu machen, die eine bessere Qualität auf dem Zielmedium ermöglichen.

2.2.7 Pausen, Fades, Texte und Menüs

Pausenzeiten, Fades, Texte und Menüs sind vor allem für die Albumproduktion sehr relevant. Gerade bei Alben, die nicht nur aus einer Auswahl von Liedern bestehen, sondern ein komplexes Werk verkörpern, sind die genannten Eigenschaften ein wichtiger Bestandteil, um den Konsum als ein Event zu etablieren. Gleichzeitig müssen auch alle Angaben die später im Zielmedium ausgelesen werden sollen, im Masteringprozess eingebettet werden. Zum Beispiel können in dem CD-Text Informationen über „song title, artist, album name, and genre“ gespeichert werden. (Katz 2015, S. 18)

Mit dem Festlegen der Pausenzeit kann der Mastering-Ingenieur das Hörerlebnis bei dem kontinuierlichen Hören eines ganzen Albums beeinflussen. „Spacing and leveling contribute greatly to the listener’s emotional response and overall enjoyment of an album.“ (Katz 2015, S. 46) Pausenzeiten in einem Album sind vergleichbar mit den Pausenzeiten in der Musik. Sie sind entscheidend für den Rhythmus. Pausenzeiten können antizipatorische Wirkung haben, die die Emotionalität eines Musikstückes verstärken. Dadurch beeinflussen sie auch den dramaturgischen Verlauf des Albums. Auf Grund der unterschiedlichen subjektiven Auffassung eines jeden Menschen werden Pausenzeiten immer unterschiedlich wahrgenommen. Bob Katz beschreibt, dass die Wahl der Länge auch von der Tageszeit, Verfassung und dem Herzschlag abhängig ist. (Vgl. Katz 2015, S. 46)

Konzeptalben sind Alben, bei denen sich eine Erzählung über die Länge des ganzen Albums erstreckt. Einzelne Lieder tragen zur Haupterzählung bei. Bei solchen Alben kann es Sinn ergeben, Lieder übergangslos oder mit einem Zwischenspiel zu gestalten. Das daraus resultierende Problem ist, dass bei Zufallswiedergabe das Lied mit dem Intermezzo beginnt. Bei einer CD lässt sich dieses Problem leicht lösen, indem das Intermezzo als separater Track mit dem Index *null* abgespeichert wird. Dieser Index wird

beim Springen in der Tracklist ausgeblendet, aber beim Linearen Abspielen des Albums mit wiedergegeben. Für Bob Katz ergibt sich daraus die Notwendigkeit, dass unterschiedliche Strukturierungen für unterschiedliche Zielmedien angefertigt werden müssen. (vgl. Katz 2015, S. 46)

Fades sind ein wichtiger Bestandteil des Masteringprozesses. Fades können in zwei Kategorien unterschieden werden. Technische Fades verhindern, dass die Membran eines Lautsprechers aus der Ruhelage schlagartig ausgelenkt wird und ein hörbares Knacksen erzeugt. Diese Fades können daher sehr kurz sein. Ästhetische Fades ermöglichen das gezielte Ein- und Ausführen des Konsumenten in das Hörerlebnis eines Liedes. Sie sind deshalb meist bewusst hörbar gemacht. Bob Katz empfiehlt „that mixing engineers send unfaded material so it can be refined in the mastering.“ (Katz 2015, S. 44) Dort können dann gezielte Fades verwendet werden, die auch in einem Album Übergänge ermöglichen, die nicht negativ auffallen.

2.2.8 Daten Distribution

Wie bereits dargelegt, ist es die Aufgabe des Masteringstudios sicher zu stellen, dass alle Daten, die von dem Presswerk oder dem Distributor empfangen werden, kongruent sind mit den Daten, die vom Masteringstudio erzeugt wurden. (Vgl. Katz 2015, S. 13) Dabei kann man den Träger für die Distribution in ein physikalisches oder virtuelles Medium unterscheiden. „For CDs, the master can be CDR, which is physically delivered to the plant, or a DDP file, which can be delivered electronically.“ (Katz 2015, S. 10) Um die Integrität der Daten zu gewährleisten, werden verschiedene Mechanismen während der Distribution eingesetzt. Eine sehr effiziente Möglichkeit sind Prüfsummen, die bei der Distribution leicht einzusetzen sind. „It is customary to include a checksum along with the file set to test for possible corruption during hard disk or interned copying.“ (Katz 2015, S. 18)

2.3 Effekte für das Mastering

2.3.1 Dynamikprozessoren

Regelverstärker sind wichtige Werkzeuge während des Masteringprozesses. Sie ermöglichen die Modifikation von Macrodynamiken und Mikrodynamiken sowie das korrekte Aussteuern des Audiomaterials. Regelverstärker arbeiten dabei immer nach dem gleichen Prinzip. Durch ein Steuersignal verstärkt oder reduziert ein Dynamikprozessor nach festgelegten Prinzipien das Nutzsignal. Diese Prinzipien lassen sich in dynamisches und statisches Verhalten unterteilen. Das statische Verhalten von

Dynamikprozessoren lässt sich durch Kennlinien darstellen. Die Kennlinie beschreibt, wie der Eingang im Verhältnis zum Ausgang übertragen wird. Je nach Form der Kennlinie lassen sich Dynamikprozessoren in Kompressor, Limiter, Expander und Gate unterteilen. An der Kennlinie kann unterschieden werden, ob es sich bei Kompressor und Expander um einen up oder downward Dynamikprozessor handelt. Jeder Dynamikprozessor lässt sich auch durch sein dynamisches Verhalten beschreiben. Das dynamische Verhalten wird durch die drei Parameter Einregelzeit, Ausregelzeit und Haltezeit bestimmt. Die Einregelzeit beschreibt, wie schnell der Dynamikprozessor bei einem Über- oder Unterschreiten des Schwellwertes eine Modifikation des Nutzsignales vornimmt. Die Ausregelzeit beschreibt, wie schnell der Dynamikprozessor nach über- oder unterschreiten des Schwellwertes in einen linearen Bereich zurückkehrt. Die Haltezeit eines Dynamikprozessors beschreibt ein festgelegtes Mindestzeitintervall, welches zwischen Einregel- und Ausregelvorgang eingehalten werden muss.

Kompressor

Kompressoren lassen sich anhand der Kennlinie und Funktionsweise in upward und downward Kompressoren unterteilen. Bei downward-kompressoren werden alle Amplituden über dem Schwellwert um ein bestimmtes Verhältnis reduziert. Upward-Kompressoren verstärken alle Amplituden, die unter einen bestimmten Schwellwert fallen.

Der Parameter Schwellwert bestimmt, wie stark der Kompressor arbeitet und wie er eingestellt wird. In der Kennlinie ist der Schwellwert der Punkt, von dem der Dynamikprozessor aus einem neutralen in einen abgeflachten Bereich wechselt. Im neutralen Bereich der Kennlinie wird das Eingangssignal im Verhältnis eins zu eins auf den Ausgang übertragen. Hat das Steuersignal einen größeren Pegel als der Schwellwert, wird durch den Verstärker das Nutzsignal heruntergeregelt. Der Schwellwert in einem Kompressor kann dynamisch oder statisch realisiert werden. Wird er dynamisch angelegt, kann der Schwellwert mit einem Wert frei eingestellt werden. Ist der Schwellwert statisch angelegt, wird das Signal so lange verstärkt, bis es den fixen Schwellwert überschreitet und das erforderliche Maß an Pegelreduktion aufweist. Kompressoren mit fixem Schwellwert haben eine Eingangsverstärkung, um das Signal entsprechend zu verstärken.

Das Kompressionsverhältnis ist als Verhältnis zwischen der „Eingangspegeldifferenz ΔL_{in} und Ausgangspegeldifferenz ΔL_{out} oberhalb des Schwellwertes definiert“. (H.-J. Maempel, S. Weinzierl, P. Kaminski 2008b, S. 732) Upward- und Downward-Kompressoren können durch eine Klammerschreibweise des Kompressionsverhältnisses unterschieden werden. Ein Downward-Kompressor reduziert alle Amplituden, die oberhalb des Schwellwertes liegen um das vorgegebene Kompressionsverhältnis. „Der Upward-Kompressor arbeitet dagegen umgekehrt: Die Dynamik-Reduktion erfolgt unterhalb des Schwellwertes.“ (Maempel 2014b, S. 380) Bei einem Kompressionsverhältnis von 2,5:1

wird jede Amplitude die den Schwellwert passiert mit dem Faktor 0,4 reduziert. Der Faktor lässt sich für Downward-Kompressoren aus dem Reziproken des Kompressionsverhältnisses errechnen. Das Kompressionsverhältnis kann im Arbeitspunkt modifiziert werden.

Diese Modifikation wird als *Soft Knee* bezeichnet. Dabei steigert sich die Ratio von einem kleineren Wert über den Arbeitspunkt zum vollen, eingestellten Wert. „Bei der sog. *Soft Knee*- oder *Overeasy*-Schaltung hat die Kennlinie keinen scharfen Knick am Schwellwert, sondern zeigt einen allmählichen Anstieg des Kompressionsverhältnisses.“ (Görne 2015, S. 357)

Jeder Kompressor hat neben dem statischen Verhalten, also dem Kompressionsverhältnis und dem Schwellwert, auch noch ein dynamisches Verhalten. „Ausregelvorgänge nach Über- oder Unterschreiten des Schwellwerts bezeichnet man zusammenfassend als dynamisches Verhalten.“ (Maempel 2014b, S. 381) Erfahrungsgemäß sind Ein- und Ausregelzeit die typischsten Parameter, die bei Kompressoren verwendet werden. Sie sind fest eingestellt oder durch den Benutzer wählbar. Die Haltezeit ist selten Bestandteil eines Kompressor-Aufbaus und wird auch nur selten als wählbarer Parameter festgelegt. Die Einregelzeit gibt an, wann der Kompressor nach der Überschreitung des Schwellwertes eine Pegelreduzierung vornimmt. Die Ausregelzeit beschreibt, wann der Kompressor die Pegelreduktion nach Unterschreiten des Schwellwertes einstellt. Wie lang die Pegelreduktion bis zum Abklingvorgang andauert, wird von der Haltezeit bestimmt. „Da die Regelvorgänge exponentiell verlaufen und theoretisch unendlich lange andauern, sind die Zeitparameter als die Zeit definiert, die der Kompressor zum Ausregeln von 63% ($1 - 1/e$) der Spannungsdifferenz zwischen Ist- und Sollpegel benötigt.“ (Maempel 2014b, S. 381) Bei Kompressoren muss beachtet werden, dass bei Einregelzeiten, die sehr klein sind, eine nichtlineare Verzerrung stattfindet. (Vgl. Görne 2015, S. 357) Dieses Problem entsteht häufig bei tiefen Frequenzen. Liegt die Einregelzeit unterhalb von 50 ms, so werden Frequenzen oberhalb von 20 Hz nichtlinear verzerrt.

Downward-Kompressoren haben eine Ausgangsverstärkung. Durch die Pegelreduktion während des Verstärkens entsteht ein Signal mit einem geringeren Crestfaktor und einer geringeren Differenz zum maximal abbildbaren Pegel. „Durch Absenken der Schwelle [...] erhält man einen ungenutzten oberen Dynamikbereich, den sog. Kompressionshub.“ (Maempel 2014b, S. 380) Durch die Ausgangsverstärkung kann der nutzbar gemachte Aussteuerungsbereich aufgeholt werden.

Das Steuersignal ist ein sehr wichtiger Bestandteil eines Kompressors. Für die Nutzung des Steuersignals gibt es zwei mögliche Bauarten. Beim Feed-forward Kompressor wird das Nutzsignal dupliziert, gleichgerichtet und dann logarithmiert. (Vgl. Friesecke 2007, S. 677) Es wird dann zur Steuerung des Verstärkers verwendet. Bei dem Feed-

backward- Kompressor wird das Ausgangssignal des Verstärkers zurückgeführt, um als Steuersignal zu dienen. (Vgl. Friesecke 2007, S. 677) Das Steuersignal kann auch durch einen externen Eingang gespeist werden, um das Nutzsinal unabhängig zu steuern. Das Steuersignal eines Kompressors kann zusätzlich auch noch gefiltert werden.

Durch die Filterung des Steuersignals mit einem Bandpass kann auch ein De-Esser realisiert werden. Ein De-Esser kann im Masteringprozess zur nachträglichen Bearbeitung der Zischlaute eingesetzt werden, da De-Esser meist sehr transparent eingestellt werden können. Sie können aber auch zur Reduktion nicht statischer Resonanzen genutzt werden.

Limitier

Ein Limitier ist ein Kompressor mit einer sehr hohen Ratio. Wird ein Kompressor mit einer Ratio von über 10:1 verwendet, wird er als Begrenzer bezeichnet. (Vgl. Görne 2015, S. 359) Auch Bob Katz schreibt: „Most authorities call a compressor with ratio of 10:1 or greater a *limiter*“. (Katz 2015, S. 84) Im Handbuch der Tonstudioteknik ist der Limiter anders definiert. „Für die Arbeitsweise als Limiter oder Begrenzer ist ein hohes Kompressionsverhältnis von 20:1 bis ∞ :1 erforderlich, in der Regel kombiniert mit einem hohen Schwellwert“. (Maempel 2014b, S. 380) Allgemein kann der Limiter als ein Kompressor mit wesentlich höherem Kompressionsverhältniss betrachtet werden.

Der Limitier zeichnet sich durch eine Kennlinie aus, die von einem neutralen Bereich in eine Waagrechte übergeht. Das bedeutet, dass alle Werte, welche den Schwellwert übersteigen, auf einen Ausgangswert reduziert werden. Limiter mit sehr hohem Kompressionsverhältnis, wie zum Beispiel 1000:1 können deshalb für die korrekte Aussteuerung verwendet werden. Um bei schnellen Spitzenpegeln eine Überschreitung des Schwellwertes zu vermeiden, müssen Limiter schnelle Ausregelzeiten besitzen. Einige Limiter sowie Kompressoren besitzen eine „Look Ahead“ Funktion. „Wird das Signal verzögert, so kann die Pegelanalyse zeitlich vor der Regelung erfolgen.“ (Görne 2015, S. 358) Das bedeutet, dass der Limiter schon ausregeln kann, bevor die Signalspitze erreicht ist. Bob Katz schreibt, dass Limiter mit kurzen Einregelzeiten weniger auffällig sind, da sie kurzzeitig eingreifen und sie nur die Transienten beeinflussen. (Vgl. Katz 2015, S. 85) Zu kurze Einregelzeiten führen zu einer stark begrenzten Wellenform. Je nach Länge der Limitierung kann der Konsument die Begrenzung als Verzerrung wahrnehmen. Um dieses Problem zu reduzieren, können die Einregelzeiten programmabhängig gestaltet werden. Bob Katz schreibt dazu: „The most successful digital limiters have auto-release control, which slows down the release time if the duration of the limiting is greater than a few milliseconds.“ (Katz 2015, S. 85)

Als True-Peak-Limiter werden Limiter bezeichnet, die Intersample-Peaks limitieren können. Intersample-Peaks sind Spitzenwerte, welche zwischen zwei Abtastwerten liegen. Diese können von Geräten, die auf die Abtastrate des Systems geeicht sind, nicht ausgelesen werden. Wird das Material nach diesen Messwerten auf eine Vollaussteuerung von 0 dBFS normalisiert, kommt es zu einer Übersteuerung bei der Wiedergabe. True-Peak-Limiter können dies vermeiden, indem sie mit dem Verfahren des Oversampling arbeiten. Oversampling bedeutet, dass das Audiomaterial in eine höhere Abtastrate konvertiert wird. Intersample-Peaks können nach der Sampleraten-Konvertierung ausgelesen und begrenzt werden. Nach der Bearbeitung wird das Audiomaterial wieder in die ursprüngliche Abtastfrequenz zurückgewandelt.

Einige Limiter bieten die Möglichkeit eine Obergrenze einzustellen. Die Obergrenze wird im Englischen als „ceiling“ bezeichnet. Das Audiomaterial wird nach der Limitierung soweit aufgeholt, bis die höchsten Amplituden den zuvor eingestellten Wert erreichen.

Expander

Ein Expander kann Micro- und Macrodynamiken erweitern. Er kann zur Verbesserung des Rauschabstandes oder zur Restaurierung von zu stark komprimierten Mixen verwendet werden, dabei jedoch nur vorhandene Dynamik erweitern. Ist keine Dynamik vorhanden, können keine Amplituden verstärkt werden. Die Kennlinie des Expanders zeichnet sich durch eine steile Linie und eine normale Linie aus. Je nach Form der Kennlinie kann hier auch wieder in Upward- und Downward-Expander unterschieden werden. „Besteht die Kennlinie nur aus zwei Abschnitten, unterscheidet man zwischen einem Downward-Expander – der steile Abschnitt ist unten und [...] Upward Expander – der steile Abschnitt ist oben.“ (Maempel 2014a, S. 384) Upward-Expander erweitern alle Amplituden die über dem Schwellwert liegen um den Faktor des Expandierverhältnisses. Wird ein Expander mit einem Expandierverhältnis von 0,5: 1 verwendet, so werden alle Amplituden oberhalb des Schwellwertes mit dem Faktor 2 multipliziert.

Gate

Ein Gate wird beim Mastering fast nie verwendet. Wenn das Hören von Audiomaterial unterdrückt werden soll, wird dies erfahrungsgemäß manuell, mit einer Automation der Verstärkung gemacht. Das Gate ist wie der Limiter eine extreme Form des Expanders. „Ein Expander mit extremer Parameter-Einstellung wird zum Gate (Tor).“ (Görne 2015, S. 359) Ein Gate ist ein Dynamikprozessor mit einer Kennlinie, die vor dem Arbeitspunkt senkrecht und nach dem Arbeitspunkt neutral verläuft. Liegt das Steuersignal unter dem Arbeitspunkt, wird jeder Eingangsamplitude der Ausgangswert 0 zugeordnet. Alle Amplituden über dem Arbeitspunkt werden ohne Modifikation übertragen.

AGC

Die Abkürzung AGC steht für Automatic Gain Control und. AGC Regelverstärker werden verwendet, um die Macrodynamik zu kontrollieren. „AGC (Automatic Gain Control) [...] bewirkt ein langsames Nachfahren des Gesamtpegels und schafft es somit, verschiedenlaute Titel gut einander anzugleichen.“ (Friesecke 2007, S. 787) Ein AGC kann im Masteringprozess zum Einsatz kommen. Es ist jedoch der Regelfall, dass die Automation der Verstärkung von dem Bearbeiter manuell eingetragen werden.

Multiband Dynamikprozessoren

Multiband Dynamikprozessoren können Kompressor, Limiter, Expander oder Gate sein. Diese Prozessoren besitzen ein oder mehrere Frequenzbänder, die unabhängig voneinander bearbeitet werden können. Die Frequenzbänder werden durch Bandpassfilter erzeugt. Je nach Bauart können die Grenzfrequenzen der Bandpassfilter starr oder verschiebbar konstruiert sein.

2.3.2 Filter und Entzerrer

Filter und Entzerrer werden im Mastering verwendet, um das Spektrum beziehungsweise den Frequenzgang zu beeinflussen. Diese Definition ist auch im Handbuch der Tonstudioteknik zu finden. „Die wichtigste Bestimmungsgröße ist das Frequenzspektrum des Audiosignals. Seine Veränderung erfolgt im elektrischen Signalweg durch Filter.“ (Maempel 2014c, S. 387) In dem Kapitel zur Klanggestaltung wird auch aufgezeigt, dass Filter meistens unterteilt werden in einfache Filter mit einfachen Aufbau, welche die Übertragungsqualität steigern können und komplexe Filter, die hauptsächlich zur Klanggestaltung genutzt werden. (Vgl. Maempel 2014c, S. 387) Einfache Filter können zum Beispiel Hoch- und Tiefpassfilter sein. Komplexe Filter sind dagegen zum Beispiel parametrische Filter. Sie werden auch als Entzerrer bezeichnet. Filter lassen sich je nach Bauform mit verschiedenen Parametern beschreiben. Diese Parameter sind Verstärkung, Gütefaktor beziehungsweise Filtergüte, Mittenfrequenz, Grenzfrequenz und Flankensteilheit. Dynamische Filter besitzen noch die Parameter Schwellwert, Einregelzeit und Ausregelzeit.

„Die Grenzfrequenz eines Filters liegt definitionsgemäß an der Frequenz, bei der das Filter einen Pegelverlust von -3dB aufweist.“ (Friesecke 2007, S. 308) Praktisch ist die Grenzfrequenz die Frequenz, ab welcher das Filter beginnt zu wirken. Sie wird verwendet, um einfach Filter zu beschreiben.

Der Gütefaktor beschreibt, in welchem Frequenzbereich der Filter wirksam ist. Er wird aus dem Quotienten der Mittenfrequenz und der Bandbreite gebildet (Vgl. Maempel 2014c, S. 393)

Entzerrer arbeiten mit einer frequenzabhängigen Verstärkung. Sie werden auch als Equalizer bezeichnet. Die Frequenz, die verstärkt oder bedämpft werden soll, wird durch eine Bandpass oder Bandsperre selektiert und dann mit dem Quellsignal gemischt. „Equalizer sind heute entweder analog als aktive RC-Netzwerke mit Operationsverstärkern aufgebaut oder digital realisiert, zum Beispiel durch Allpass-Filter, deren Ausgangssignale den Originalsignalen abgeschwächt zugemischt werden.“ (Maempel 2014c, S. 393)

Die Flankensteilheit ist ein Maß, welches für die Beschreibung von einfachen Filtern verwendet wird. Sie gibt an, wie stark das Filter unter- oder oberhalb der Grenzfrequenz mit steigender oder fallender Frequenz bedämpft. Die Flankensteilheit kann in dB pro Oktave oder dB pro Dekade angegeben werden. Erfahrungsgemäß findet die Angabe in dB pro Oktave am meisten Verwendung. Eine Oktave bedeutet eine Frequenz-Verdoppelung oder -Halbierung. Wird ein Hochpassfilter mit 6 dB pro Oktave verwendet, dann nimmt bei einer Frequenzhalbierung die Dämpfung um 6 dB zu.

Die Mittenfrequenz ist ein Parameter, der für die Beschreibung von einem komplexem Filter verwendet wird. Sie gibt an, bei welcher Frequenz die höchste Dämpfung oder Verstärkung zu verzeichnen ist. Die Mittenfrequenz eines Bandpassfilters kann durch die Wurzel aus dem Produkt der oberen und der unteren Grenzfrequenz berechnet werden. (Vgl. Maempel 2014c, S. 745)

Einfach Filter

Einfache Filter werden auch als klassische Filter bezeichnet. „Bei klassischen Filtern ist in der Regel nur die Grenzfrequenz als Parameter einstellbar.“ (Maempel 2014c, S. 388) Ein Hochpassfilter ist ein Filter, der unterhalb seiner Grenzfrequenz die Frequenzen kontinuierlich abdämpft. Ein Tiefpassfilter ist ein Filter, welcher Frequenzen über seiner Grenzfrequenz kontinuierlich abdämpft. Ein Bandpassfilter ist ein Filter der, aus einem Hochpassfilter und einem Tiefpassfilter zusammengesetzt ist.

Hochpassfilter werden im Mastering erfahrungsgemäß oft eingesetzt. Filter mit einer hohen Flankensteilheit werden dazu verwendet, den Bassbereich zu begrenzen. Filter mit einer niedrigen Flankensteilheit werden verwendet, um den Bassbereich zu attenuieren.

Tiefpassfilter werden im Mastering erfahrungsgemäß weniger bis selten eingesetzt. „Außerdem werden sie gestalterisch zur Verminderung der Helligkeit des Nutzsignals verwendet“. (Maempel 2014c, S. 389)

Komplexe Filter

„Equalizer oder Entzerrer sind Filter, die eine Anhebung oder Absenkung des Pegels innerhalb definierter Frequenzbereiche ermöglichen.“ (Maempel 2014c, S. 393) „Im Unterschied zum einfachen Filter kann neben der Grenzfrequenz daher auch der Verstärkungspegel im entsprechenden Frequenzbereich als Parameter verändert werden[...]“. (H.-J.Maempel, S.Weinzierl, P.Kaminski 2008a, S. 745) Komplexe Filter werden im Mastering deshalb zur Verbesserung und Gestaltung des Klangs eingesetzt. Die komplexen Filter können durch ihre unterschiedlichen Amplitudenfrequenzgänge unterschieden werden.

Der Kerbfilter ist der steiflankigste Filter. Er wird meistens zur Reduktion von Resonanzen eingesetzt. Der Glockenfilter sieht dem Kerbfilter ähnlich. Er weist jedoch weniger steileren Flanken auf. Glockenfilter können je nach Filtergüte für ein breites Aufgabespektrum eingesetzt werden. Ein Kuhschwanzfilter bewirkt eine Anhebung oder Dämpfung der Frequenzen oberhalb oder unterhalb einer Grenzfrequenz. Bei diesem kann die Grenzfrequenz und die Flankensteilheit gewählt werden. „Die Steilheit analoger Geräte ist allerdings selten einstellbar. Digitale sog. continuous slope-Filter ermöglichen eine kontinuierliche Veränderung der Flankensteilheit von Schelf-Filtern und Pässen [Algorithmix, 2007].“ (Maempel 2014c, S. 395) Eine besondere Form des Kuhschwanzfilters ist der Niveau-Filter. Der Niveau-Filter besteht aus zwei, durch einen Rotationspunkt gekoppelte Kuhschwanzfilter, die entgegengesetzt arbeiten. Dieser Filter ist besonders gut für Mastering geeignet, damit die Balance zwischen Tief- und Hochtonbereich leicht verändert werden kann. In den meisten komplexen Filtern-Anwendungen lassen sich auch einfache Filterkurven, wie Hoch- und Tiefpass mit wählbarer Flankensteilheit, einstellen.

Graphische Equalizer sind Entzerrer, die aus einer Vielzahl an Filtern bestehen, die über den Frequenzbereich verteilt sind. „[Sie] bestehen in der Regel aus Bänken von meist 6 bis 30 parallel geschalteten Glockenfiltern“. (Maempel 2014c, S. 394) Die Filterbreite wird meist im Terz- oder Oktavabstand realisiert. Diese Filter haben alle eine fest eingestellte Güte. In einem festgelegten Bereich kann die Verstärkung des Filters frei gewählt werden. Der Graphische Equalizer kommt selten während des Masterings zum Einsatz, da der voll parametrische Equalizer flexibler ist und so immer auf Problemstellungen angepasst werden kann.

Parametrische Equalizer sind Entzerrer, welche aus mehreren Filtern bestehen, die unterschiedlich einstellbare Parameter haben. Die Filter können verschiedene Formen wie

Bandsperre, Glocke, Kerbe, Kuhschwanz, Niveau oder Baxandal haben. Einige Geräte können zusätzlich in halb- und vollparametrisch unterteilt werden. „Sog. Halb-, semi- oder teilparametrische Equalizerbänder lassen die Einstellung nicht nur des Verstärkungspegels [...] sondern auch der Mittenfrequenz zu [...], vollparametrische Equalizer ermöglichen zusätzlich die Veränderung der Güte.“ (Maempel 2014c, S. 395)

Dynamische Equalizer sind Filter, deren Verstärkung abhängig von einem Steuersignal ist. Wird der Schwellwert stärker überschritten, wird das Nutzsignal stärker gefiltert. Der Parameter Verstärkung gibt die maximale Dämpfung beziehungsweise Verstärkung des Filters an, welche abhängig von der Überschreitung des Schwellwertes durchgeführt wird. Bei dynamischen Entzerrern kann dabei oft zwischen Expandierung und Kompression gewählt werden.

3 Lautheitssteigerung im Mastering

3.1 Wahrnehmung von Lautheit

3.1.1 Menschliches Gehör

Schall wird über das Außenohr aufgenommen. Er gelangt über den Gehörgang an das Trommelfell. Hinter dem Trommelfell sitzen Ambos, Hammer und Steigbügel. Sie gehören zum Mittelohr. Über die Gehörknöchelchen wird der Schall als mechanische Schwingung übertragen. Der Steigbügel ist das letzte Gehörknöchelchen und gibt die Schwingung an das ovale Fenster weiter. Die mechanische Welle breitet sich in der Gehörschnecke aus. Die Gehörschnecke unterteilt sich in drei Kammern: Vorhoftrappe, Schneckengang und Paukentreppe. Diese Gänge sind mit Flüssigkeit gefüllt. „Druckbewegungen am ovalen Fenster des mit Flüssigkeit gefüllten Innenohrs bewirken Wellen, die entlang der Basilarmembran laufen“. (Ellermeier W. 2008a, S. 47) Die Basilarmembran ist elastisch gebaut und weist deshalb Dämpfungseigenschaften auf. Im Handbuch der Audiotechnik wird beschrieben, dass je nach Frequenz der Welle die Basilarmembran an unterschiedlichen Stellen ausgelenkt wird. (Vgl. Ellermeier W. 2008a, S. 48) „Das Schallsignal erfährt auf der Basilarmembran eine Frequenz-Orts-Transformation.“ (Görne 2015, S. 114) Die Basilarmembran berührt bei der Auslenkung die auf den äußeren Haarzellen sitzenden Stereozilien. „Die inneren Haarzellen wandeln diese Erregung in elektrische Impulse auf dem Hörnerv.“ (Görne 2015, S. 115)

3.1.2 Psychoakustik

„Die wissenschaftliche Disziplin der Psychoakustik untersucht, wie akustische Reize wahrgenommen werden.“ (Ellermeier W. 2008b, S. 52) Viele Phänomene, wie zum Beispiel die Unempfindlichkeit gegenüber tiefen Frequenzen kann durch den Aufbau des Gehörs begründet werden. Das Ohr kann dabei als Überträger in einer Signalkette verstanden werden. Es ist jedoch weit komplexer, als alle anderen Glieder in der Signalkette. Zum Beispiel besitzt es, wie jede andere Komponente auch, einen Übertragungsfrequenzgang. Der Unterschied zu den anderen Gliedern ist, dass der Frequenzgang des menschlichen Ohres vom Schalldruck abhängig ist. Auf Grundlage der Eigenschaften des menschlichen Gehörs können Eigenschaften des Audiomaterials gezielter manipuliert werden.

Die Bandbreite des menschlichen Gehörs wird in der Audiotechnik von 20 Hz bis 20 kHz beschrieben. Diese Bandbreite ist sehr idealisiert. Die obere Grenzfrequenz wird durch

die Beweglichkeit der Haarzellen bestimmt. Im Handbuch der Tonstudientechnik wird beschrieben, dass die äußeren Haarzellen sich bei Schwingungen mit bis zu 20 kHz mitbewegen können. (Vgl. Goeres-Petri und Dickreiter 2014, S. 117) Die tiefste Frequenz, die wahrgenommen werden kann, beträgt 16 Hz. Im Buch Tontechnik wird beschrieben, dass der Grund dafür die Datenverarbeitung im Gehirn ist. Unterhalb von 16 Hz könnten die Druckänderungen als einzelne Hörereignisse wahrgenommen werden. (Vgl. Görne 2015, S. 121) Im Handbuch der Audiotechnik wird zusätzlich beschrieben, dass für die Wahrnehmung von solch tiefen Frequenzen, Drücke von mehr als 100 dB SPL notwendig sind. (Vgl. Ellermeier W. 2008b, S. 54) Die Obere Grenzfrequenz des Gehörs beträgt im Idealfall 20 kHz. Die meisten Menschen hören effektiv bis 16 kHz. Die Bandbreite wird bei Anstieg des Alters stark eingeschränkt. Außerdem können durch Dauerbelastung Hörschäden entstehen die ebenfalls die Bandbreite begrenzen. In der Regel wird der obere Frequenzbereich zuerst geschädigt. „Die allgemeine Hörfähigkeit wird durch die Altersschwerhörigkeit aber relativ geringfügig beeinträchtigt, da der wichtigste Hörbereich unter 4 kHz liegt“. (Dickreiter und Goeres-Petri 2014, S. 120)

„Die Amplitude der eintreffenden Frequenz wird durch die Menge der Nervenimpulse übertragen.“ (Friesecke 2007, S. 115) Je größer die Amplitude des eintreffenden Schalleignisses ist, desto mehr Reize werden von den inneren Haarzellen an das Gehirn gesendet. Fletscher und Mundson schreiben in einem Zeitschriftenaufsatz aus dem Jahr 1933: „The magnitude of an auditory sensation, that is, the loudness of the sound, is probably dependent upon the total number of nerve impulses that reach the brain per second along the auditory tract.“ (Fletcher und Munson 1933, S. 82) Das Lautheitsempfinden der Menschen ist extrem frequenzabhängig. Werden zwei verschiedene Frequenzen mit einem Abstand größer einer Oktave und gleichem Schalldruck wiedergegeben, so werden die Töne unterschiedlich laut wahrgenommen. Wenn man eine Frequenz als Referenz verwendet und alle anderen Frequenzen angleicht, bis sie gleich laut erklingen, so erhält man die Kurven gleicher Lautstärkepegel. Diese Kurven, die zuerst von Fletscher und Mundson entdeckt wurden und später von Eberhard Zwicker korrigiert wurden, bilden indirekt den Frequenzgang des Ohres bei verschiedenen Pegeln ab.

Breitbandige Schalleignisse werden als lauter empfunden, da sie mehrere Bereiche an unterschiedlichen Stellen auf der Basilarmembran anregen. Das führt zu einer erhöhten Reizauslösung. Im Handbuch der Audiotechnik wird von der Addition der Schallintensitäten innerhalb der Frequenzgruppenfilter gesprochen. (Vgl. Ellermeier W. 2008b, S. 61) Als Frequenzgruppenfilter bezeichnet man den Bereich auf der Basilarmembran, welche unterschiedliche Frequenzbereiche wahrnehmen können. (Vgl. Görne 2015, S. 123)

Menschen sind besonders empfindlich für mittlere Frequenzen. Die Kurven gleicher Lautstärkepegel zeigen, dass für tiefe und hohe Frequenzen deutlich mehr Schalldruck

benötigt wird, um eine vergleichbare Laustärke wie bei einem Schallereignis im Bereich zwischen 500 Hz und 5 kHz zu erzeugen. Das bestätigt auch das Handbuch der Audio-technik: „Die größte Empfindlichkeit besteht für Frequenzen zwischen 2 und 5 kHz [...]“ (Ellermeier W. 2008a, S. 46) Der Grund dafür ist die Notwendigkeit in der Evolutionsgeschichte des Menschen. In dem Frequenzbereich zwischen 1000 und 5000 Hz liegen viele der für das Sprachverstehen wichtigen Konsonanten und auch viele Umweltgeräusche, die für das Überleben des prähistorischen Menschen wahrscheinlich bedeutsame Warnsignale gewesen sein dürften, wie zum Beispiel brechende Zweige, Knistern im Gras usw..“ (Ellermeier W. 2008a, S. 47) Die trichterähnliche Formung des Außen- und Mittelohrs fördert vor allem die Empfindlichkeit für Frequenzen zwischen 1 und 5 kHz.

Das Ohr besitzt auch ein zeitliches Verhalten. Zum Beispiel ist es in der Lage, Schallergebnisse durch eine künstliche Schwingungsverstärkung zu verlängern. „Auch scheinen die Stereozilien durch Verformung in der Lage zu sein, die Resonanz künstlich weiter zu verstärken“. (Friesecke 2007, S. 114) Auch besitzt das Ohr einen Schutz vor Tieffrequenten Signalen. Bei einem Schalldruck von über 80 dB SPL kann innerhalb von 200 ms ein Muskel im Ohr kontrahiert werden, was dafür sorgt, dass Frequenzen unter 1000 Hz abgedämpft werden. (Vgl. Ellermeier W. 2008a, S. 47) Wie schon von Bob Katz beschrieben, ist das Lautheitsempfinden auch von Signaleigenschaften und der Exposition abhängig. „In der Regel nimmt die Lautstärke bis zu einer Schalldauer von etwa 200 ms zu, man spricht von zeitlicher Integration.“ (Ellermeier W. 2008b, S. 60) Liegen zwei Töne dicht beieinander und werden schnell genug hintereinander abgespielt, so wird der Folge-Ton durch den Ersten verdeckt. Dieser Effekt wird zum Beispiel bei der Reduzierung des Informationsgehaltes während der Kodierung von MP3 Dateien verwendet.

3.1.3 Auswirkung der Psychoakustik auf das Mastering

Die Beschaffenheit des menschlichen Gehörs und die damit verbundenen psychoakustischen Effekte werden im Mastering zur gezielten Manipulation des Audiomaterials genutzt. Dabei ist zu vermerken, dass die Effekte meist instinktiv ausgenutzt wurden. Während zuvor die Lautheit noch von den Massenmedien begrenzt war, entdeckte man, dass sich laute Musik besser anhörte. Dieser psychoakustische Effekt wurde ausgenutzt, um die Absatzzahlen von Tonträgern und die Einschaltquoten der Rundfunksender zu erhöhen. Andere psychoakustische Effekte wurden ausgenutzt, um die Lautheit weiter steigern zu können.

Der Effekt des scheinbaren Zuwachses an Qualität durch Lautheitssteigerung ist bedingt durch das schalldruckabhängige Wahrnehmen von Frequenzen. Wird Musik mit einem höheren Schalldruck wiedergegeben, linearisiert sich der Frequenzgang des Ohres zunehmend. Die Folge ist, dass wir hohe und tiefe Frequenzen deutlich besser wahrnehmen. Die Musik klingt dadurch kräftiger und voller und wird somit als qualitativ

hochwertiger wahrgenommen. Wird die Lautstärke angehoben, tritt gleichzeitig noch ein anderer Effekt auf. Wie zuvor beschrieben, lösen größere Amplituden mehrere Reize aus. Diese können gleichzeitig dafür sorgen, dass Glückshormone freigesetzt werden. Eine größere Anzahl an Reizen bedeutet auch eine größere Ausschüttung von Glückshormonen. Da die Ausschüttung von Glückshormonen süchtig macht, kann eine positive Neigung zu lauter Musik entstehen.

Die Bandbreite des Signals kann im Mastering fast immer begrenzt werden. Die Abtastfrequenz von 44,1 kHz hat sich zum Beispiel aus der Grenzfrequenz des Ohres ergeben. Bei einer Abtastung mit 44,1 kHz kann maximal eine Frequenz von 22,05 kHz abgebildet werden. Werden die Stopbänder, welche für die Wandlungen notwendig sind, subtrahiert, kommt man auf eine Grenzfrequenz, welche dem menschlichen Gehör entspricht. Dass sehr viele Menschen nicht mehr als 16 kHz hören, findet zum Beispiel im MP3 Format Verwendung. Die Entnahme der hohen Frequenzen führt jedoch nicht zu einem Lautheitsverlust. Der Lautheitsverlust ist nur auffällig, wenn Frequenzen aus den empfindlichen Frequenzgruppen gefiltert werden. Auch kann eine untere Grenzfrequenz festgelegt werden. Diese Grenzfrequenz wird dabei indirekt durch einen Hochpassfilter eingerichtet, welches durch den Mastering-Ingenieur festgelegt wird. Die Bandbreite des Formates bleibt bestehen. Die Begrenzung des Bassbereiches bietet mehrere Vorteile. Die meisten Wiedergabesysteme sind nur auf einen Übertragungsbereich von 20 Hz bis 20 kHz ausgelegt. Das bedeutet, dass alle Frequenzen unterhalb 20 Hz in den meisten Fällen nicht wiedergegeben werden können. Außerdem können Fehler durch zu große Auslenkung der Membran, wegen zu tiefer Frequenzen entstehen. Diese Fehler entstehen vor allem bei Einweg-Systemen. „Durch den Dopplereffekt wird die höhere Schwingung frequenzmoduliert, der höhere Ton wirkt rau, es entstehen Intermodulationsverzerrungen, die sehr störend wirken können, da Summen und Differenzöne der Frequenzkomponenten entstehen.“ (Dickreiter 2014a, S. 210) Ein weiterer Vorteil ist, dass man durch Reduktion der Bassfrequenzen einen höheren Aussteuerungsbereich erzeugen kann. Auf Grund der geringen Empfindlichkeit des Gehörs bei tiefen Frequenzen müssen diese deutlich größere Amplituden aufweisen, als höhere Frequenzen, um wahrnehmbar zu sein. Bob Katz Aussage unterstützt diese Tatsache: „A low frequency boost introduces so much energy that it can reduce the highest clean program level we can give to a song (in cases where the client is demanding a 'loud' master).“ (Katz 2015, S. 63)

Die Empfindlichkeit des Menschen hin zu Frequenzen zwischen 1 und 5 kHz ist besonders wichtig für den Masteringprozess. Durch die erhöhte Aufmerksamkeit in diesem Frequenzbereich fallen kleinste Fehler auf. Störgeräusche, sich gegenseitig maskierende Instrumente oder unausgeglichene Frequenzgänge fallen dem Konsumenten am wahrscheinlichsten auf. Es ist deshalb sicher zu stellen, dass im Masteringprozess der mittlere Frequenzbereich besonders gut repräsentiert wird. Die Stimme sollte deshalb

nie die Konsequenzen einer Entscheidung im Masteringprozess tragen. In einem Artikel über starke Emotionen und Gänsehaut beim Musikhören in der Zeitschrift für Audiologie wird beschrieben, dass sich gezeigt hat, dass die so genannte Chill-Reaktion öfter bei Liedern eintrat, welche eine höhere Brillanz zwischen 920 und 4400 Hz aufwiesen. (Vgl. Altenmüller und Kopiez 2012, S. 55) Die Chill-Reaktion „Derartige »Chills«, »Thrills«, oder »Gänsehauterlebnisse« sind mit dem Gefühl eines Fröstelns und mit Schauern, die den Rücken hinunterlaufen, verbunden. [...] Chills [gehen] mit einer dopaminergen Aktivierung im Bereich der Belohnungszentren des Striatums und des Accumbens-Kerns einher.“ (Altenmüller und Kopiez 2012, S. 56) Durch den Prozess der Lautheitssteigerung im Mastering werden mit Begrenzern die größten Amplituden reduziert. Durch die Begrenzung tiefer Frequenzen werden durch die Kompensation des Kompressionshubes alle Frequenzen oberhalb der limitierten Frequenzen angehoben. Das kann im Gegensatz die Möglichkeit einer Chill-Reaktion vergrößern.

Die zeitlichen Eigenschaften des Ohres wirken sich indirekt auf das Mastering aus. Es lässt sich zum Beispiel daraus schließen, dass das Ohr eine Mittelung des Schalls durchführt. Schallereignisse von kurzer Dauer und hoher Intensität können genauso laut wahrgenommen werden wie lange Schallereignisse mit mittlerer Intensität. Dieser Effekt ist nicht nur bedeutsam für den Höreindruck, der durch Kompression entsteht, sondern auch für die Beurteilung des Lautheitsempfinden in einem metrischen System.

3.1.4 Kompetitive Steigerung der Lautheit im Mastering

Barrierefreiheit für die Lautheitssteigerung

Vor der Einführung der CD im Jahr 1982 dienten Schallplatte und das Magnettonband als Massenmedium für Audiomaterial. Das Audiomaterial musste für die Vinylplatte auch korrekt ausgesteuert werden. Nicht korrektes Aussteuern und Übersteigern der Lautheit hatte zur Folge, dass die Vinylplatte nicht mehr fehlerfrei abgespielt werden konnte. Bob Katz beschreibt die Fehler: „The LP would either skip or cause extreme sibilance if the average or the peak level was made too high, so its loudness was self-limiting, or worse, the expensive cutterhead would burn up.“ (Katz 2015, S. 243) Durch Publikation des digitalen Mediums CD, wurde ein Massenmedium eingeführt, welches im Gegensatz zur Schallplatte wenig physikalische Limitierung aufwies. Die CD hat einen größeren Dynamikraum als die Schallplatte. Abgesehen von Intersamplepeaks kommt es bei digitaler Vollausteuern nicht zu Übertragungsfehlern zwischen Massenmedium und Wiedergabesystem. Das führte nicht nur dazu, dass sich die Aussteuerungsgrenze im Laufe der Zeit in Richtung 0 dBFS verschob, sondern auch Veröffentlichungen immer lauter wurden. Wird gegen 0 dBFS ausgesteuert, wird von Peak-Normalisation gesprochen. Das

Bedeutet, dass die größten Amplituden einem bestimmten Wert, in diesem Fall dem höchsten digitalen Wert, entsprechen.

Wettbewerbsartige Steigerung der Lautheit

Die Entwicklung der CD beschleunigte die wettbewerbsartige Steigerung der Lautheit extrem. Peak-Normalisation wurde durch die Entwicklung von digitalen Aufnahmen- und Bearbeitungsverfahren ermöglicht. (Vgl. Katz 2015, S. 243) Mastering-Ingenieure begannen die Peak-Werte von Audiomaterial zu normalisieren. Durch die Anhebung der Spitzenwerte bis zur Vollaussteuerung konnte die Lautheit erhöht werden. Im Gegensatz zur Schallplatte führt dies nicht zu Problemen während der Wiedergabe. Man begann durch Kompressoren den Dynamikbereich einzuschränken und danach die Spitzenwerte wieder zu normalisieren. Dadurch entstanden Lieder, welche wesentlich lauter waren als die Lieder welche nur normalisiert waren. Ein weiterer großer Einfluss auf die Steigerung der Lautheit war die Entwicklung des Brickwall-Limiters. Durch sie gelang es, die Dynamik noch mehr einzuschränken und die maximalen Amplituden gegen digital Null auszusteuern. Mastering-Ingenieure versuchten immer neue Methoden zu finden und zu nutzen, um die Lautheit noch mehr zu steigern und so ihre Marktrelevanz zu erhalten. Schlussendlich wurde sogar ein Clipping der digital-analog-Wandler künstlich herbeigeführt, um durch den schnellen Anstieg der Obertöne eine höhere Lautheit zu erzeugen. Dieser Vorgang wird als Shredding bezeichnet. (Vgl. Katz 2015, S. 247)

Ein weiterer Faktor, der die Lautheitssteigerung über die Zeit extrem beschleunigt hat, ist die unterschiedliche Voraussetzung für Lautheit von verschiedenen Programmen. Bob Katz beschreibt, dass Rockproduzenten sehr enttäuscht waren, weil ihre Produktionen im Gegensatz zu akustischen Produktionen sehr viel leiser waren. (Vgl. Katz 2015, S. 244) Dieser Unterschied entsteht hauptsächlich aus einem unterschiedlichen Crestfaktor. Der Crestfaktor der Stereosumme wird vor allem durch perkussive Instrumente im Mix deutlich gehoben. Rockproduktionen können durch kräftig gespielte Schlagzeuge extrem hohe Spitzenwerte aufweisen. Durch einen geringeren Crestfaktor kann das Audiomaterial bei einer Peak-Normalisierung deutlich in der Lautheit gesteigert werden, ohne hörbare Artefakte durch Dynamikprozessoren zu besitzen. Soll das Audiomaterial mit hohem Crestfaktor eine ähnliche Lautheit aufweisen wie das Audiomaterial mit geringem Crestfaktor, so muss bei der Dynamikbearbeitung mit einer hohen Verminderung der klanglichen Qualität gerechnet werden.

In den 90iger Jahren fingen Mixing-Ingenieure an, die Masterkanalzüge der Mixe zusätzlich mit Dynamikprozessoren zu bearbeiten. „The mix engineers [...] started to imitate the level and sound of already mastered music.“ (Katz 2015, S. 247) Das führte dazu, dass Mastering-Ingenieure die Lautheit im Masteringprozess weiter anhoben, um den Kunden einen weiteren scheinbaren Zuwachs an Qualität zu bieten. Schlussendlich fand

in jedem Schritt von der Aufnahme bis zur Ausspielung eine Einschränkung der Dynamik durch Dynamikprozessoren statt.

Nicht nur im Mastering entstand ein Rennen um die höchste Lautheit. Radio- und Fernsehsender versuchten sich gegenseitig zu übertreffen. Wenn Konsumenten das Radioprogramm wechselten, sollte die Aufmerksamkeit durch eine höhere Lautstärke gewonnen werden. Der Ansatz der Gewinnung der Aufmerksamkeit durch Lautheit wurde auch in TV-Sendungen verwendet und sogar innerhalb der Sender, zum Beispiel wurden bestimmte Programmteile, wie Werbung, besonders laut übertragen. Musiktitel, welche durch den Masteringprozess sehr laut geworden sind, wurden durch die Dynamikprozessoren der Funkhäuser noch mehr eingeschränkt. Bob Katz beschreibt, dass das Audiomaterial, welches während des Mastering bereits übersteuert wurde, in den Funkhäusern zu weiterer Übersteuerung führte. (vgl. Katz 2015, S. 247)

Dass die Lautheitsanhebung teilweise nur einen temporären Vorteil erschaffen konnte wurde von Verantwortlichen übersehen. Der Endkonsument hat die volle Kontrolle über die Lautheit. Durch die Verstärkungseinstellung an seinem Wiedergabesystem kann er laute Medien leiser und leise Medien lauter stellen. Es gibt Ausnahmefälle, in denen der Konsument keine Kontrolle über die Wiedergabelautstärke hat. Solche Situationen treten im Livebereich bei Konzerten, im Club oder bei einem Wechsel des Programminhaltes auf.

Lautheit als Verkaufsgrund

Kunden, die Musik veröffentlichen wollen, möchten die bestmögliche Wettbewerbsfähigkeit ihres Werkes erreichen. Der Grund dafür findet sich in dem Willen zur Selbstverwirklichung und der Notwendigkeit der Profitabilität. Eine perfekte Fertigung kann nur erlangt werden, wenn alle Bereiche der Produktion einen Mehrwert erzielen. Wird ein Mastering des Werkes in Anspruch genommen, erwartet der Auftraggeber auch dort einen Mehrwert für die Produktion. Soll ein Mastering für ein nicht normalisiertes Massenmedium durchgeführt werden, so wird der Auftraggeber die Forderung stellen, dass das Werk gleich oder lauter ist als vergleichbare Werke der Konkurrenz. Der Kunde beschränkt sich dabei auf den Eindruck der Überlegenheit, im Sinne der Lautheit und akzeptiert die Degradation der Qualität. Bob Katz beschreibt, wie Kunden ihre Meinung änderten: „If a rocker liked a more open, clear sound, and discovered his album was 2 dB softer than the competition, he may have decided to compromise his standards in the name of competition.“ (Katz 2015, S. 244) Um den Forderungen der Kunden nachzugehen und somit auch eine Konkurrenzfähigkeit zu erhalten, müssen Masteringstudios sich diesen Anforderungen stellen. Das führt rückschlüssig zu einer starken Priorisierung innerhalb des Masteringprozesses. Durch die hohe Nachfrage lässt sich leichter eine Relevanz am Markt erschaffen, als durch einen wiedererkennbaren Klang oder einen

ausgewogeneren Frequenzgang. Das Ergebnis ist ein ständiger Wettbewerb innerhalb der nicht normalisierten Massenmedien

3.1.5 Lautheitsnormalisierung

Die Lautheitssteigerung wird wegen ihres psychoakustischen Effekts instrumentalisiert. Allgemein findet ein Qualitätsverlust durch die übermäßige Steigerung der Lautheit statt. Der Lautstärkeunterschied zwischen verschiedenen Produktionen sorgte zudem für eine Verschlechterung des Hörerlebnisses. Um ein gleichmäßigeres und damit besseres Hörerlebnis über ein gesamtes Programm zu gewährleisten, wurde die Lautheitsnormalisierung eingeführt. Gleichzeitig wurden damit indirekt Grenzwerte für die Lautheitssteigerung geschaffen.

Lautstärkennormalisierung wird durch Verfahren ermöglicht, welche die empfundene Lautheit numerisch ausdrücken können. Kann die empfundene Lautheit einzelner Programminhalte ermittelt werden, so kann diese an einen Referenzwert durch Verstärkung und Dämpfung angeglichen werden. Das Resultat sind geringere bis keine Lautheitssprünge zwischen verschiedenen Programmabschnitten. Die Einführung von festen Lautheitswerten führt auch zu einer Steigerung der Mikrodynamiken. Diese wurden durch das exzessive Limitieren und Komprimieren, im Wettbewerb um die höchste Lautheit, stark beschränkt.

Durch feste Grenzwerte gibt es keine Notwendigkeit, den Lautheitsgrad weiter zu steigern, da er später im Zielmedium angepasst wird. Die deutliche Steigerung der Lautheit über den vorgegebenen Wert des Zielmediums kann womöglich auch zu einem Wettbewerbsnachteil führen.

Durch die Lautheitssteigerung werden, ab einem gewissen Grad, immer sowohl Makro- also auch Mikrodynamiken reduziert. Der positive Effekt der Lautheitssteigerung überwiegt den Nachteil der reduzierten Makro-Mikrodynamik in einem nicht lautheitsnormalisierten Massenmedium. Wird zum Beispiel ein übermäßig lautes Mastering für ein lautheitsnormalisiertes Massenmedium angefertigt, so wird es später im Zielmedium attenuiert werden. Die Reduktion der Makro- und Mikrodynamiken bleibt bei der Attenuierung erhalten und wird jetzt nicht mehr von dem Effekt der Lautheitssteigerung überwogen. Die Qualität ändert sich damit deutlich für den Konsumenten. Bob Katz schreibt zu dem Vorteil der Lautheitsnormalisierung: „The concep of loudness normalization actually encourages more dynamic mixing and mastering, because producers will no longer be motivated to try to make a ‘loud’ Album.” (Katz 2015, S. 220) Der Qualitätszuwachs kann nur gewährleistet werden, wenn der Anbieter des lautheitsnormalisierten Massenmediums eine adäquaten Ziellautheit wählt.

Gleichzeitig muss beachtet werden, dass einige Massenmedien einen gesetzten Zielwert erreichen müssen. Muss ein dynamisches Programm auf Grund der Lautheitsnormalisierung angehoben werden, so ändern sich die Spitzenpegel. Das Signal könnte nach der Anhebung zur Erreichung der Ziellautheit Spitzenwerte über 0 dBFS aufweisen. Diese würden durch die digitale Grenze abgeschnitten werden und das Audiomaterial würde übersteuern. Um diese Situationen zu vermeiden, setzen manche Massenmedien Begrenzer ein. Ein Problem ergibt sich aus den festgelegten Werten der Begrenzer. Diese Werte können eine Begrenzung verursachen, welche hörbare Artefakte erzeugt oder den musikalischen Verlauf beeinträchtigt. Es entsteht eine deutlich schlechtere Qualität, als bei einem, durch einen Mastering Ingenieur, korrekt eingestellten Begrenzer.

Es ist zu schlussfolgern, dass sowohl die Überschreitung, als auch Unterschreitung der Lautheitsvorgaben der lautheitsnormalisierten Massenmedien zu Qualitätsverlust führen kann. Aktuell sind sowohl Massenmedien mit und ohne Lautheitsnormalisierung gebräuchlich. Abhängig von dem Anspruch des Kunden kann eine unverhältnismäßige Lautheitssteigerung auch aktuell noch notwendig und eine Anforderung an den Masteringprozess sein. Die Mehrheit der heute verwendeten Massenmedien verwenden Systeme zur Lautheitsnormalisierung, weshalb Lautheitsmessung während des Masterings für die Sicherstellung der Qualität im Massenmedium notwendig ist.

3.2 Messung der Lautheit

Die Schaffung einer standardisierten Lautheitsmessung hat ihren Ursprung im Öffentlichen Rundfunk. „Insbesondere beim Rundfunk (Hörfunk und Fernsehen) ist die Lautheitsanpassung ein vordringliches Problem, da Lautheitssprünge zwischen verschiedenen Sendern, aber auch [...] verschiedenem Programmmaterial (Wort/Musik), [...], von den Hörern als besonders störend empfunden werden.“ (Weinzierl 2008a, S. 562) Um das Problem zu beheben, musste nach einer Möglichkeit gesucht werden um die Lautheitsunterschiede im Programm zu messen.

In einer Arbeitsgruppe der ITU wurde ermittelt, dass eine gewichtete Leq-Messung die größte Übereinstimmung mit der menschlichen Lautheitsempfindung bringt. „Hierbei ergab sich eine maximale Übereinstimmung zwischen perzeptiv und technisch bestimmter Lautheit für eine Leq-Messung mit einer Frequenzbewertung“. (Weinzierl 2008a, S. 563) Weinzierl beschreibt in dem Kapitel Leq-Messung, dass bei einer Leq-Messung über ein bestimmte Zeitfenster der energieäquivalente Mittelwert zu einem Referenzwert berechnet wird. (Vgl. Weinzierl 2008b, S. 560) Bevor das Signal über einen Zeitraum gemittelt wird, wird es gefiltert. Dieser Filterung wird als K-weighting bezeichnet und besteht aus zwei kombinierten Filtern. „The K-weighting filter is composed of two stages of

filtering; a first stage shelving filter and a second stage high-pass filter.” (ITU 2017, S. 3) Zum Einsatz kommt ein Vorfilter, welcher die akustischen Eigenschaften des Kopfes kompensiert. Dieser Filter ist ein Kuhschwanzfilter, der eine Grenzfrequenz von 997 Hz und eine Anhebung um 4 dB besitzt. Der zweite Filter der Vorfilterung ist ein Hochpassfilter 2. Ordnung, der die Unempfindlichkeit des Gehörs in Bezug auf tiefe Frequenzen berücksichtigt. Bevor der Mittelwert berechnet wird, muss jeder Kanal mit einem Faktor beaufschlagt werden, um die Empfindlichkeit des menschlichen Gehörs gegenüber von hinten eintreffenden Signalen zu kompensieren. Für Stereosummen ist dieser Faktor nicht relevant, weil er für den linken und rechten Kanal eins beträgt.

Als Referenzwert wird der maximaler Pegel 0 dBFS eines digitalen Systems verwendet. Der Wert, der berechnet wird, hat die Einheit *LKFS* nach ITU 1770-1. „The numerical result for the value of loudness that is calculated in equation (2) should be followed by the designation LKFS.“ (ITU 2007, S. 5) Die Einheit wurde durch die EBU umdeklariert in *LUFS*. „‘LUFS’ is equivalent to ‘LKFS’ (which is used in ITU-R BS.1770). The EBU uses ‘LUFS’ which is compliant with international naming conventions.“ (EBU 2014, S. 3)

Die Bildung eines relativen Wertes hat die Einheit *LK* beziehungsweise *LU*. Werden zwei Werte miteinander verglichen kann der Lautstärkeunterschied in LU angegeben werden. 1 LU entspricht dabei 1 dB.

Zur Messung der Lautstärke werden drei verschiedene Zeitfenster verwendet. Die sogenannte Programmlautstärke wird über die Dauer des gesamten Programmes ermittelt. Es muss deshalb vom Anfang der Messung bis zum Ende integriert werden. Sie wird deshalb als „integrated loudness“ bezeichnet. Zusätzlich wird für die Messung über die Gesamtlänge eines Programms ein Gate verwendet, um das Verfälschen des Durchschnittes durch zu leise Abschnitte zu verhindern. Das Gate besteht aus zwei Gates. Das erste Gate ist ein absolutes Gate, welches alle Werte unterhalb von -70 LUFS ausschließt. Das Relative-Gate schließt alle Werte, welche 10 LU unterhalb des integrierten Lautheitswertes liegen, welcher sich aus der Berechnung mit dem Absoluten-Gate ergeben hat. (Vgl. ITU 2017, S. 6) Weitere Maße zur Beurteilung von Lautheit sind „short-term loudness“ und „momentary loudness“. Beide Messungen werden ohne Gate durchgeführt. Sie unterscheiden sich durch das jeweilige verwendete Messintervall. „The momentary loudness uses a sliding rectangular time window of length 0.4 s. [...] The short-term loudness uses a sliding rectangular time window of length 3 s.“ (EBU 2016a, S. 6)

Die statistische Verteilung des Umfangs der Lautstärke wird mit der Einheit LRA angegeben. „The measure ‘Loudness Range’ quantifies the variation in a time-varying loudness measurement“ (EBU 2016a, S. 7) Für die Messung wird ein gleitendes 3 s Fenster

verwendet. Es wird fast dasselbe Gate benutzt, wie für die Programmlautstärkemessung. Der Unterschied besteht darin, dass das relative Gate erst bei Werten, die -20 LU unterhalb des Wertes liegen, welche sich aus der Berechnung mit dem Absoluten-Gate ergeben hat, schließt. Von der Berechnung werden alle Werte unterhalb von 10 und oberhalb von 95 Prozent der gesamt gemessenen Werte ausgeschlossen. (vgl. EBU 2016b, S. 6)

Der LRA Wert kann als Richtlinie verwendet werden, da er Information über den Lautstärkeumfang gibt. „Individual maximum values for LRA can provide a good dynamic framework for different formats (for example, 5.1 vs. 2.0), genres, distribution platforms as well as different replay environments.“ (EBU 2016c, S. 18) Der Dynamikumfang kann nur ein hinreichendes, jedoch kein entscheidendes Kriterium zur Beurteilung eines Musikstückes sein. Der Dynamikumfang eines Stückes ist Bestandteil der künstlerischen Gestaltungsmittel. Zum Beispiel haben Orchesteraufnahmen schon durch die Komposition einen höheren LRA als eine moderne Rockproduktion. Bob Katz beschreibt das typische LRA-Werte bei klassischer Musik zwischen 8-23 LU und bei 80's pop/rock zwischen 4-13 LU liegen. (Vgl. Katz 2015, S. 221)

Eine weitere Methode zur Messung der Lautstärke ist in dem System Replaygain verankert. Replaygain wird genutzt, um die Lautheit von Audiomaterial zu analysieren und dann durch Verstärkung oder Dämpfung, Lautheitsunterschiede zwischen verschiedenen Programminhalten zu minimieren. Replaygain nutzt, ähnlich wie zur Analyse der Lautheit in LUFS, einen Vorfilter. Dieser Filter besteht aus einem Hochpassfilter mit einer Flankensteilheit von 12 dB pro Oktave. Die Grenzfrequenz liegt bei 150 Hz. Der zweite Filter, welcher für hohe Frequenzen zum Einsatz kommt, wird als „yulewalk“ Filter bezeichnet. Nach der Filterung wird das Signal in 50 ms lange Segmente unterteilt. Von jedem Segment wird der RMS-Wert berechnet. Nach der Berechnung der RMS-Werte, werden diese in dB umgewandelt. Die dB Werte werden nach Größe sortiert und der Wert welcher 95% zwischen höchstem und niedrigsten RMS-Wert entspricht der gemessenen Lautheit. Dieser Wert wird dann im Datenformat gespeichert. Im Wiedergabesystem kann dieser Wert ausgelesen und die Anpassung an einen Zielwert berechnet werden. (Vgl. Robinson 2002, S. 285–290)

4 Ausgangspunkt für die aktuellen Anforderungen an den Masterprozess

4.1 Aktuelle Mediennutzung für den Musikkonsum

Entwicklung des Marktes

Aktuell werden verschiedenste Formen von Massenmedien genutzt. Allgemein gilt, dass Massenmedien mit älterer Technologie von Massenmedien mit neuerer Technologie am Markt verdrängt werden. Als Aspekt für eine erfolgreiche Einführung am Markt haben sich Massenmedien erwiesen, welche durch technischen Fortschritt einen höheren Komfort für den Kunden etablierten. Zum Beispiel löste die CD 1982 die Schallplatte ab. Sie offerierte dem Kunden eine geringere Anfälligkeit gegenüber physischen Einwirkungen und man konnte frei zwischen verschiedenen Liedern eines Albums durchschalten. Tim Renner, ehemaliger Geschäftsführer von Universal Music Deutschland erklärt: „Gerade weil sich in der einfachen äußeren Benutzbarkeit des CD-Systems die zugrunde liegende technische Komplexität nicht spiegelte, gelang die extrem schnelle Markteinführung mit gewaltigen Verkaufserfolgen.“ (Renner 2004, S. 22)

Online-Händler, wie Amazone oder iTunes bieten Musik als bezahlten Download an. Der Download von Musik wird durch Streamingdienste abgelöst werden. Streamingdienste bieten den Vorteil, auf der Basis eines monatlichen Abonnements, frei aus einem Archiv von Musik wählen zu können. Zusätzlich bieten Streamingdienste Angebote wie Playlists an. Diese sind einer Klassifizierung entsprechend gewählte Zusammenstellungen von verschiedenen Musiktiteln. Zum Beispiel gibt es Playlists, welche nur Titel beinhalten, die einer bestimmten Musikrichtung entsprechen. Es gibt auch Playlists deren Titelauswahl nur eine konkrete Gefühlslage bedienen. Playlists besitzen aus diesem Grund den Vorteil, dass sie, im Vergleich zu Rundfunkprogrammen, sehr spezifisch sein können. Aus Nutzungszahlen geht hervor, dass sich die Streamingdienste, welche einen deutlich höheren Komfort gegenüber physischen Datenträgern und Downloadangeboten bieten, durchsetzen werden.

Entgegen der Erwartungshaltung, dass durch Erscheinen neuer Massenmedien die Verkaufszahlen der älteren Generation zurückgehen müssen, sind steigende Abverkaufszahlen von Schallplatten zu beobachten. Dieser Trend wird geschaffen durch verschiedene Faktoren. Besonders in nichtpopulären Musikrichtungen ist es typisch, Musiker durch den Kauf von physischen Datenträgern zu unterstützen. CDs und Schallplatten werden häufiger in Musikrichtungen wie Jazz und Klassik verkauft. Der Trend begründet sich teilweise auch aus einer Gegenbewegung zum immer schneller

wachsenden Fortschritt. Ein weiterer Grund ist, dass in diesem Segment Musik nicht als Gebrauchsprodukt, sondern als Kunst verstanden wird. Das führt dazu, dass Konsumenten einen bewussten Konsum, in Form eines Events, wertschätzen. Das Auflegen und bewusste Durchhören eines physischen Datenträgers wird von den Konsumenten mehr geschätzt, als das Anwählen eines digitalen Streams. Dieser nostalgische Trend beeinflusst den allgemeinen Trend jedoch nur wenig, ist aber trotzdem in den Zahlen erkennbar und zu berücksichtigen.

Betrachtet man die Verteilung der gesamt gehörten Zeit, auf unterschiedliche Medien, so kann man feststellen, dass das Radio immer noch eines der meist genutzten Massenmedien ist. Das ist aber nur der Fall, wenn nicht nach Alter differenziert wird. In den Altersgruppen unter 30 hat das Radio extrem an Bedeutung verloren. Mit dem Wechsel der Generationen wird sich auch die Radionutzung immer mehr verringern. Die verlorenen Nutzerzahlen werden sich dann in denen der Streaming-Diensten wiederfinden. Das Radio hat vor allem wegen seiner einfachen Handhabung und seiner einfachen Zugänglichkeit eine hohe Nutzungszeit. Es verliert jedoch durch fehlende Individualisierbarkeit an Bedeutung.

Der Markt in Zahlen

Die Zahlen bestätigen die Vermutung, dass sich Streamingdienste als Massenmedium gegenüber anderen Formaten, am Markt durchsetzen werden. Vergleicht man die Berichte des Bundesverbandes Musikindustrie e.V. aus den Jahren 2017 und 2019 dann wird ein deutliches Wachstum ersichtlich. Während bei einem Gesamtumsatz von 1,59 Milliarden Euro der Anteil an Streamingdiensten 2017 noch 34,6 % betrug, wuchs er 2019 auf 55,1 % bei einem Gesamtumsatz von 1,62 Milliarden Euro an. (Vgl. Bundesverband Musikindustrie e.V. 2020, S. 5) Ein weiterer Indikator für das Etablieren der Streamingdienste ist die Gesamtanzahl von Streams pro Jahr. 2013 fanden 5,9 Milliarden Aufrufe statt. Bis zum Jahr 2019 stieg die Zahl auf 107,4 Milliarden Streams. (vgl. Bundesverband Musikindustrie e.V. 2020, S. 17). Wie vermutet geht der Download als Massenmedium mit mehr Nachteilen stetig zurück. Seit dem Rekordhoch im 2012 mit 114,6 Millionen Downloads fielen die Absatzzahlen kontinuierlich. Im Jahr 2019 fanden nur noch 39,3 Millionen Downloads statt. (Vgl. Bundesverband Musikindustrie e.V. 2020, S. 16) Auch der Umsatz durch Downloads geht zurück. Seit dem 2013 mit Downloads 258 Millionen Euro erwirtschaftet wurden, sind die Einnahmen stetig gefallen. Im Jahr 2019 betragen die Einnahmen durch Downloads nur noch 100 Millionen Euro. (Vgl. Bundesverband Musikindustrie e.V. 2020, S. 11)

Die Aufstellung des deutschen Marktes von Streaming-Diensteanbietern kann durch die ARD/ZDF Onlinestudie aufgezeigt werden. Bei den Streamingdiensten, die nur Musik anbieten, ist Spotify mit Abstand der Anbieter mit den größten Nutzerzahlen. Im Jahr

2019 nutzten laut der ARD/ZDF Onlinestudie 22 % Spotify. Der zweithäufigst benutzte Streamingdienst ist Amazon Music mit 14 %, weit dahinter sind Apple Music mit 7 %, Google Play Music mit 6 % und Soundcloud mit 5 %. (Vgl. Mai et al. 2019, S. 417)

Gleichzeitig ist ein extremer Rückgang der Verkaufszahlen von CDs zu beobachten. 2017 wurden durch den Verkauf von CD-Alben 45,4 % des Gesamtumsatzes des Jahres erwirtschaftet. (Vgl. Bundesverband Musikindustrie e.V. 2018, S. 7) Im Gegensatz dazu betrug der Anteil am Gesamtumsatz 2019 nur noch 29,0 %. (vgl. Bundesverband Musikindustrie e.V. 2020, S. 6) Der Gesamtumsatz bezieht sich dabei auf den Verkauf von Musik. Der Absatz von Vinylplatten steigt seit 2006 wieder an. Wurden 2010 noch ca. 0,6 Millionen Vinylplatten verkauft, stieg die Zahl fast konstant bis auf 3,4 Millionen im Jahr 2019. (Vgl. Bundesverband Musikindustrie e.V. 2020, S. 15)

In den kommenden Jahren werden Streamingdienste zunehmend an Bedeutung gewinnen. So werden sich die Streamingdienste mit den besten Angeboten und mit der breitesten Auswahl durchsetzen. Die Streaming-Dienste lassen sich in Video-Streaming-Dienste und Audio-Streaming-Dienste unterscheiden. Dabei wird bei den Video-Streaming-Diensten der Video-Aspekt vernachlässigt und der Konsum der Musik priorisiert.

In der Onlinestudie von ARD und ZDF werden von 186 Minuten, die täglich gehört werden, 149 über Radio konsumiert. Das entspricht 79,7%. Betrachtet man die Altersgruppe zwischen 14 und 29 Jahren, so werden von 175 täglich gehörten Minuten nur 78 (44,6%) über das Radio konsumiert und 73 (41,7%) über Streaming-Anbieter. (Vgl. Mai et al. 2019, S. 408). In der Publikation des Bundesverband Musikindustrie e.V. aus dem Jahr 2019 wird dies noch drastischer dargestellt. In einer Altersgruppe 16 bis 24 Jahren hat Radio nur einen Anteil von 17 % an der wöchentlich gehörten Zeit. Im Gegensatz dazu werden in der gleichen Altersgruppe Streaming Dienste zu 45 % der Zeit für den wöchentlichen Musikkonsum verwendet. (vgl. Bundesverband Musikindustrie e.V. 2020, S. 25). Es wird ersichtlich, dass das Radio weiterhin eines der wichtigsten Massenmedien zum Konsum von Musik ist, aber immer mehr durch Streamingdienste abgelöst wird. Das Radio als Massenmedium kann nochmals unterteilt werden in seine Übertragungsweise. Laut der ARD/ZDF Onlinestudie werden von täglichen 149 Minuten Live Radio 123 Minuten über UKW empfangen und 17 über DAB+.

Ein wichtiger Nutzungsgrund bei Streaming-Diensten sind kostenlose Angebote und Bequemlichkeit. Besonders gerne wird der kostenlose Video-Streamingdienst YouTube verwendet. Bei ihm können kostenlos Videos mit Musikinhalt gestreamt werden. YouTube ist dabei auf den meisten Endgeräten frei zugänglich, vorinstalliert oder über einen Browser aufrufbar. 2019 lösten kostenpflichtige Musik-Streamingdienste mit 52,1 % YouTube mit 36,7 % ab. (Vgl. Bundesverband Musikindustrie e.V. 2020, S. 26)

Betrachtet man den internationalen Markt, wird Videostreaming zum Beispiel über YouTube immer noch am meisten genutzt. „Über die 19 in der Nutzungsbefragung von IFPI untersuchten Märkte hinweg spielte bei der On-Demand-Nutzung Video-Streaming mit durchschnittlich 47 Prozent eine deutlich größere Rolle als Premium-Audio-Streaming“. (Bundesverband Musikindustrie e.V. 2020, S. 27) Daraus lässt sich schließen, dass YouTube weiterhin eine wichtige Rolle für das Streaming von Audio spielen wird. Das Gleiche bestätigt auch die Onlinestudie von ZDF und ARD.“ Wie im Vorjahr hört fast die Hälfte der deutschsprachigen Bevölkerung ab 14 Jahren (48 %) Musik über YouTube. (Mai et al. 2019, S. 415)

4.2 Aktuelle Endgerätenutzung zur Wiedergabe von Musik

Die verwendeten Wiedergabegeräte hängen stark von dem verwendeten Massenmedium ab. Zum Beispiel ist die Wahrscheinlichkeit, dass das Massenmedium Schallplatte auf einem Autolautsprecher wiedergegeben wird sehr gering. Die Endgerätenutzung ist, wie die Nutzung der Massenmedien, stark abhängig von der komfortablen Bedienung. Je unproblematischer und leichter zugänglich ein Gerät ist, umso wahrscheinlicher ist seine Nutzung.

Wie zuvor festgestellt, ist eines der meist genutzten Massenmedien der Rundfunk. Rundfunk kann über viele Wiedergabesysteme abgespielt werden. Einige der Systeme werden durch ihren Einsatzort öfter verwendet als andere. So wird der Rundfunk am meisten über Autolautsprecher konsumiert, da Rundfunkempfangssysteme in fast jedem PKW verbaut sind. „Innerhalb der letzten 14 Tage betrachtet, ist für gut drei Viertel der Jugendlichen die Radionutzung über ein Autoradio (77 %) am häufigsten“. (Feierabend et al. 2019, S. 20)

Das am meist vorhandene Endgerät zur Musikwiedergabe in deutschen Haushalten, ist die Stereoanlage. Das geht aus einem Bericht von VuMA Touchpoints hervor. Dabei wurden 23068 persönliche Interviews mit deutschsprachigen Personen ab 14 Jahren durchgeführt. 76,3 % gaben dabei an, Stereoanlagen in ihrem Haushalt zu führen. Mit 65,9 % liegt das Radio dicht dahinter. 61,5 % der Haushalte verfügen über Laptops und 44,4 % PCs. In 39,6 % der Haushalte sind Bluetooth Lautsprecher vorhanden. (Vgl. VuMA Touchpoints 2019, S. 73)

Musik Streaming-Dienste werden am meisten auf dem Smartphone und an Laptops genutzt. Das geht aus einem Dokument der Verbraucherzentrale Rheinland-Pfalz hervor. Die Untersuchung wurde dabei an zwei Gruppen durchgeführt. Eine Gruppe bestehend aus Nutzer von kostenpflichtigen Streaming-Diensten, die andere von kostenlosen.

Beide Gruppen nutzen dabei das Smartphone als hauptsächliches Wiedergabegerät. (Vgl. Schmidt und Zaborowski 2017, S. 23)

Die Musikwiedergabe über Smartphones kann über verschiedene Wege erfolgen. Die Musik kann über die internen Lautsprecher wiedergegeben, über den 3,5 TRS Ausgang ausgespielt oder drahtlos über Bluetooth, WLAN oder NFC übertragen werden. Dabei kann aus einer Umfrage der Warner Music Group Germany aus dem Jahr 2016 schätzungsweise abgeleitet werden, dass mobiler Musikkonsum mit dem Smartphone am häufigsten durch In-Ears gewährleistet wird. Der integrierte Lautsprecher des Wiedergabesystem liegt in der Nutzung hinter den In-Ears. Am dritt häufigsten werden Kopfhörer, welche die Ohren halb oder ganz bedecken, verwendet. Im Vergleich kommen mobile Bluetooth Lautsprecher nur selten zum Einsatz. (Vgl. Weidenbach 2017)

Werden Kopfhörer für die Wiedergabe verwendet, so können aus einer Befragung aus dem Jahr 2019 von Statista die beliebtesten Kopfhörer Marken entnommen werden. An erster Stelle steht Sony mit 16% der befragten Kopfhörerbesitzer. Mit 12 % folgen Apple und Samsung. (Vgl. Kunst 2019)

4.3 Verarbeitung von Musik in den Medien

4.3.1 UKW-Rundfunk

Vor der Ausstrahlung des Radio-Programmes findet häufig eine Bearbeitung der Stereosumme statt. Die Stereosumme muss dabei angepasst werden, um ordnungsgemäß Übertragen zu werden. Radiosender normalisieren ihre Programme nach der EBU R128 Empfehlung. 2015 wurde das Programm SWRinfo zum ersten Mal nach der Lautheitsregulierung der EBU R128 durchgeführt. (Vgl. Zehentmeier 2015)

Radiosender setzen Begrenzer ein, um eine bestimmte maximale Amplitude zu erreichen. Nach der Frequenzmodulation entspricht eine höhere Amplitude einem breitbandigerem Frequenzhub im übertragenen Multiplex. Da in Deutschland der Frequenzhub begrenzt ist, muss die Amplitude zuvor begrenzt werden.

Die Frequenzbandbreite des Hörfunks muss vor der Übertragung begrenzt werden. „FM überträgt praktisch den vollen hörbaren NF-Frequenzbereich zwischen 30 Hz und 15 kHz mit einem relativ hohen Störabstand von etwa 40 dB.“ (Hoeg 2014a, S. 1108)

Zusätzlich kann durch Effekte wie Dynamikprozessoren und Entzerrer der Klang so verändert werden, dass ein Wiedererkennungswert geschaffen wird. Im Rundfunk kommen oft Multieffektgeräte mit dem Namen Optimod des Herstellers Orban zum Einsatz.

Besonders der Optimod 8700 bietet zum Beispiel verschiedene Möglichkeiten der Bearbeitung. Es können verschiedene Multibandeffekte eingesetzt werden um transparente oder konsistente Klänge zu erzeugen. Subharmonic Synthesizer können dazu eingesetzt werden um Frequenzen künstlich nach unten zu doppeln um einen kräftigeren Bassbereich zu erzeugen. Exponential Loudness soll die Dynamik in stark komprimierten Audiomaterial wiederherstellen und so Gehörermüdung vermeiden. (vgl. Lauterbach 2019) Die Einstellungen des Optimods können frei gewählt werden, wodurch eine optimale Anpassung für die Rundfunkbetreiber gewährleistet wird. Gleichzeitig ist die Verarbeitung während der Distribution über das Massenmedium Rundfunk nicht ausreichend einsehbar.

4.3.2 Streaming-Dienste

Einige Streamingdienste wie Apple Music oder Spotify treffen klare Aussagen, wie sie Music nach der Distribution verarbeiten beziehungsweise welche Eigenschaften eine möglichst artefaktfreie Übermittlung zum Konsumenten ermöglichen. Andere Streamingdienste wie Amazon Music Unlimited veröffentlichen keine Angaben für Anforderung an das Audiomaterial während der Distribution.

Wie zuvor festgestellt, ist Lautheit ein entscheidendes Kriterium. Die Lautheitsregulierung der Streaming-Dienste kann aus Sicht des Konsumenten untersucht werden. Aus dem Vergleich der Messwerte können Unterschiede in der Regulierung und den Anforderungen für den Masterprozess abgeleitet werden.

Bei dem Versuch wurde die Wiedergabe von fünf der am meist genutzten Streaming-Dienste gemessen. Bis auf YouTube Music und Soundcloud wurde für alle Streaming-Dienste die Desktop Anwendung verwendet. Alle Streaming-Dienste wurden vollausgesteuert. Ebenfalls wurden alle auf die höchste Streaming Qualität eingestellt. Bei Anwendungen mit wählbarer Lautheitsnormalisierung wurden die Einstellungen überprüft und gegebenenfalls auf Lautheitsnormalisierung aktiviert. Mit dem Programm Virtual Audio Stream von der Firma DDMF wurde der interne Audiostream von Windows durch das Plugin Insight von der Firma Izotope geleitet. Das Plugin misst Lautheit nach ITU BS 1770. Bei der Messung wurde das Gate nach ITU BS 1770 Teil 2 verwendet. Die Messung beginnt bei Anklingen des Liedes und endet mit Wechsel zu dem nächsten Titel. Bei der Messung werden die Integrierte Lautheit in dB LUFS und der Truepeak Pegel in dBTP gemessen. In einer Tabelle wurden alle Daten gesammelt. Die Titel wurden aus verschiedenen Musikrichtungen ausgewählt, um einen möglichst hohen Umfang an Dynamik abzudecken. Dabei können die Lieder in die Bereiche Klassik, Jazz, Rock, Metal, Rap und Elektro zugeordnet werden.

Besonders Auffallend sind die Titel „Die Moldau“ von Uwe Mund, „Hey Laura“ von Gregory Porter und „Dich zu Lieben“ von Roland Kaiser. Sie fallen durch eine sehr niedrige, integrierte Lautheit in verschiedenen Streaming-Diensten auf. „Hey Laura“ ist ein Jazzstück, während Die Moldau ein klassisches Stück ist. Wie in Kapitle 3.1.4 beschrieben, weisen Musikstücke dieser Gattungen schon bei der Aufnahme einen hohen Dynamikumfang auf. Es ist zu vermuten, dass die Abweichung durch Ausnahmeregelungen für die Normalisierung bestimmter Musikstile zustande kommt.

Spotify

Spotify liefert auf der Internetseite Spotify for Artists Informationen darüber wie Audio-material eingelesen und wiedergegeben wird. Spotify fordert eine Distribution im Wave-Format oder in FLAC. Die empfangene Datei wird auf Korrumpierung überprüft. Anschließend wird die Audiodatei in das Wave-Format mit einer Abtastrate von 44,1 kHz gewandelt. Die Bittiefe bleibt dabei erhalten. Danach wird die Wave-Datei in vier Verlust-behaftete Formate codiert. Diese sind: Ogg und Vorbis mit einer Bitrate von 96, 160 und 320 kbps, AAC mit einer Bitrate von 128 und 256 kbps und HE-AACv2 mit einer Bitrate von 24 kbps. (Vgl. Spotify 2020b) Nach der Wandlung wird die Lautheit mithilfe von Replaygain analysiert. Die analysierten Daten werden dann in den Containerformaten festgehalten. (Spotify 2020a)

Der Spotifyplayer normalisiert, mit der Hilfe der von ReplayGain analysierten Lautheitswerte, alle einzeln abgespielten Lieder. Es wird daher empfohlen Lieder auf -14 dB LUFS zu Mastern, da dieser Wert näherungsweise der durchschnittlichen Lautheit der Lieder entspricht, welche durch ReplayGain adjustiert wurden. Spotify schreibt: „ReplayGain doesn't specify a measurement unit for loudness, so we're unable to give an exact measure in LUFS used by ITTU 1770. However, we adjust tracks to 3 dB higher than ReplayGain algorithm specifies, which is roughly equivalent to -14 dB LUFS, according to the ITU 1770 standard.“ (Spotify 2020a) Während der Messung durch den Autor konnte eine Durchschnittliche Lautheit von -14,1 dB LUFS gemessen werden.

Titel	Interpret	Integrated LUFS	True Peak dBTP	Integrierte LUFS - TP
Santa Monica	Theory of a Deadman	-14,1	-4,1	-10
Dark Days	Parkway Drive	-13,7	-6,4	-7,3
Einer muss den Job ja machen	Udo Lindenberg	-14,3	-5,5	-8,8
Hoch	Tim Bendzko	-14,1	-7,1	-7
Die Moldau	Uwe Mund	-16,8	-4,8	-12
Hey Laura	Gregory Porter	-16,4	-4	-12,4
Alphagenetik	Kollegah	-14,4	-6	-8,4
Animals	Martin Garrix	-13,3	-4,6	-8,7
Roller	Apache 207	-14	-5,1	-8,9
Not Kings	Candy Says	-15	-1,9	-13,1
Dich zu Lieben (1982)	Roland Kaiser	-15,8	-1,8	-14
Hulapalu	Andreas Gabalier	-14,3	-6,3	-8
Ø in LUFS		-14,7		-9,9
Ø in LUFS(korrigiert) ^a		-14,1		
Ø der lautesten Master ^b				-8,2

Tabelle 1: Messung des Streaming-Dienstes Spotify ^a Ausschluss der Lieder: „Die Moldau“, „Hey Laura“, „Dich zu Lieben“ ^b Durchschnitt der ursprünglichen Lautheit von den am lautesten ausgespielten Liedern

Musikstücke mit einer hohen integrierten Lautheit werden bei Spotify trotz Lautheitsnormalisierung lauter ausgespielt. Erkennbar ist dies an einer geringen Differenz zwischen integrierter Lautheit und Truepeak-Pegel. Berechnet man den Durchschnitt aus der Differenz von integrierter Lautheit und dem TP Level der einzelnen Lieder und wählt dann die Lieder aus, welche unter dem gesamten Durchschnitt liegen, kann man feststellen, dass diese gleichzeitig die am lautesten ausgespielten Lieder sind. Wird davon ausgegangen, dass Lieder im Masteringprozess gegen 0 dBFS ausgesteuert wurden, lässt sich durch den ermittelten Truepeak-Wert, die originale integrierte Lautheit des Masters berechnen. Der Durchschnitt der ursprünglichen Lautheiten der am lautesten ausgespielten Titel beträgt –8,2 LUFS.

Als maximalen Aussteuerungspegel empfiehlt Spotify -1dBTP. Weisen Lieder während der Analyse durch Replaygain eine zu hohe Lautheit auf, wird der Pegel bei der Wiedergabe abgesenkt. Ist das analysierte Musikstück zu leise, wird der Pegel während der Wiedergabe angehoben bis er den Lautheitsvorgaben entspricht. Dabei kommt ein Limiter zum Einsatz, welcher verhindert, dass durch die Anhebung besonders dynamischen Audiomaterials eine Übersteuerung stattfindet. Der Limiter besitzt eine Obergrenze von -1 dBTP. Die Einregelzeit des Limiters beträgt 5 ms und die Ausregelzeit 100 ms. (Vgl. Spotify 2020a)

Spotify arbeitet zusätzlich mit einer speziellen Lautheitsnormalisierung für Alben. Wird ein Album wiedergegeben, so wird die Lautstärke des Albums auf näherungsweise -14

LUFS normalisiert. Tatsächlich wird für die Album-Normalisierung durch Replaygain ein separater Wert in den Dateien gespeichert. Dieser Wert bestimmt die Adjustierung der Verstärkung auf Grund des Durchschnittwertes des Albums. So können bewusst eingesetzte Lautheitsunterschiede innerhalb des Albums erhalten bleiben und von dem Zuhörer wahrgenommen werden. Wird das Album im „Shuffle“ abgespielt, wird wieder jedes einzelne Lied normalisiert. Shuffle ist die zufällige Wiedergabe von Liedern innerhalb einer Ansammlung. (vgl. Spotify 2020a)

Premium Nutzer von Spotify haben dabei die Wahl, verschiedene Einstellung vorzunehmen. Sie können zum Beispiel zwischen drei Lautheitsprofilen wählen. Die Lautheitsprofile werden als Loud, Normal und Quiet bezeichnet. Dabei wird die Lautheitsnormalisierung die durch Replaygain erfolgt, mit einem Summanden beaufschlagt. Wird das Profil Normal angewählt, wird jeder Song um 3 dB gegenüber dem Replaygain-Wert erhöht. Dabei ist die ausgegebene Lautheit vergleichbar mit -14 dB LUFS. Für das Profil Loud werden weitere 3 dB addiert, weshalb die Ausgabelautheit auf -11 dB LUFS ansteigt. Das Profil Quiet addiert -5 dB, was dazu führt, dass die ausgegebene Lautheit -23 dB LUFS entspricht. (Vgl. Spotify 2020a)

Apple Music

Apple Music verwendet standardmäßig Lautheitsnormalisierung nur während der Wiedergabe von seinen Radiokanälen. Nutzer von iTunes können eine Lautheitsnormalisierung für alle wiedergegebenen Lieder in den Einstellungen unter der Funktion Sound Check aktivieren.

Bob Katz beschreibt: “However iTunes Sound Check, with a low target level of about -16,5 dB LUFS, accommodates the vast majority of recorded music without causing it to hit full scale peak”. (Katz 2015, S. 256) Im Gegensatz zu Bob Katz Aussage konnte der Autor während seiner Messung andere Werte ermitteln. Der Durchschnitt aller gemessenen Lieder liegt bei -15 LUFS. Werden die, in Spotify und Amazone Music, auffällig gewordenen Musikstücke aus der Rechnung ausgeschlossen, verschiebt sich der Wert nur marginal auf -14,8 LUFS.

Titel	Interpret	Integrated LUFS	True Peak dBTP	Integrierte LUFS - TP
Santa Monica	Theory of a Deadman	-14,8	-6,3	-8,5
Dark Days	Parkway Drive	-17	-9,4	-7,6
Einer muss den Job ja machen	Udo Lindenberg	-15,2	-6,6	-8,6
Hoch	Tim Bendzko	-14,9	-7,7	-7,2
Die Moldau	Uwe Mund	-15,5	-3,4	-12,1
Hey Laura	Gregory Porter	-16,3	-2,1	-14,2
Alphagenetik	Kollegah	-14,9	-6,1	-8,8
Animals	Martin Garrix	-13	-3,8	-9,2
Roller	Apache 207	-13,2	-3	-10,2
Not Kings	Candy Says	-15,7	-2,5	-13,2
Dich zu Lieben (1982)	Roland Kaiser	-15,5	-3,2	-12,3
Hulapalu	Andreas Gabalier	-14,4	-6,5	-7,9
Ø in LUFS		-15,0		-10,0
Ø in LUFS(korrigiert) ^a		-14,8		
Ø der lautesten Master ^b				-8,3

Tabelle 2 Messung des Streaming-Dienstes Apple Music^a Ausschluss der Lieder: „Die Moldau“, „Hey Laura“, „Dich zu Lieben“^b Durchschnitt der ursprünglichen Lautheit von den am lautesten ausgespielten Liedern

Ähnlich wie Spotify verhält sich die Lautheitsnormalisierung bei iTunes. Die Lautheit wird beim Kodieren der Formate analysiert und in das Containerformat geschrieben. Während des Abspielens wird durch die Anwendungssoftware die Lautheit entsprechend der Vorgaben adjustiert. Werden bei iTunes während der Wiedergabe von Alben einzelne Tracks wahlweise oder mit Zufallswiedergabe abgespielt, wird jeder Track für sich normalisiert. Wird ein Album linear gehört, findet eine Lautheitsnormalisierung des Albums statt. (Vgl. Fremon 2016, S. 5) Im Gegensatz zu Spotify werden laute Master nicht lauter ausgespielt. Erkennbar ist das daran, dass die Lieder mit der geringsten Differenz zwischen integrierter Lautheit und TP-Pegel nicht gleichzeitig den Liedern entsprechen, welche lauter als der Durchschnitt ausgespielt worden sind.

Apple Music lässt nur Titel mit bestimmten Spezifikationen zu. Diese Spezifikationen findet man auf der Apple Website mit dem Namen Mastered for iTunes. Apple gibt an, dass für Distribution Musik in das AAC-Format codiert wird. Dabei wurden im Jahr 2003 die meisten Lieder in einer Datenrate von 128 kbps archiviert. Seit 2007 nutzt Apple Encoder die 256 kbps AAC kodieren können. (Vgl. Fremon 2016, S. 2)

Akzeptiert werden alle Sampleraten von 44,1 kHz und mehr. Fremon Christie schreibt: „An ideal master will have 24-bit 96kHz resolution.“ (Fremon 2016, S. 3) Bei Eintreffen des Masters wird dieses mit einem Konverter in ein File mit einer Samplerate von 44,1 kHz und einer Bittiefe von 32 Bit Gleitkomma umgewandelt. Das 32 Bit Gleitkomma

Format wird verwendet, um Artefakte von Intersamplepeaks zu vermeiden. Danach wird es durch die Encoder in ein 256 kbps AAC Format kodiert.

iTunes vermarktet Songs ohne Übersteuerung und bestimmten Eigenschaften mit der Auszeichnung „Mastered for iTunes“. „Although iTunes doesn't reject files for a specific number of clips, tracks which have audible clipping will not be badged or marketed as Mastered for iTunes.“ (Fremon 2016, S. 4)

Indirekt wird in dem Dokument von Apple ein Spitzenwert von -1 dB angegeben. „A growing consensus is emerging that digital masters should have a small amount of headroom (roughly 1dB) in order to avoid such clipping.“ (Fremon 2016, S. 5)

Amazon Music

Amazon bietet unterschiedliche Profile für den Konsum von Musik an. Amazon Music Free ist ein Streaming-Dienst mit Werbeeinhalten. Prime Music ist für Nutzer des Amazon Prime Dienstes. Amazon Music Unlimited ist ein kostenpflichtiger Streamingdienst der mehr Songs als prime Music anbietet. Zusätzlich kann eine Erweiterung erworben werden, welche das Streamen von Liedern in höheren Auflösungen ermöglicht.

Musik kann nur über Aggregatoren auf Amazon Music Unlimited veröffentlicht werden. Amazon stellt keine Anforderungen für die Distribution öffentlich zu Verfügung. Aggregatoren wie Tunecore fordern ein Master mit einer Samplerate von 44,1 kHz und einer Bittiefe von 16 Bit. (Vgl. Tunecore 2020)

Werden Songs mit dem HD-Erweiterung gestreamt, so können diese in zwei verschiedenen Qualitätsstufen wiedergegeben werden. „HD Songs haben eine Bittiefe von 16-Bits, eine minimale Abtastrate von 44,1 kHz [...] Ultra HD Songs haben eine Bittiefe von 24-Bits, Abtastraten von 44,1 kHz bis zu 192 kHz [...].“ (Amazon 2020). Widersprüchlich ist, dass Musik, welche durch Aggregatoren veröffentlicht wurde, durch deren Anforderungen nicht den Wiedergabeeigenschaften von Amazon entspricht. Amazon macht keine Angaben zu den verwendeten Dateiformaten.

Amazon Music Unlimited wird lauthheitsnormalisiert. Der Durchschnitt des korrigierten Wertes liegt bei -14 LUFS. Besonders auffällig ist die sehr geringe Abweichung der Musikstücke von -0,1 LU. Es ist zu vermuten, dass einer der ITU 1770 ähnlichen oder gleichen Algorithmus zum Einsatz kommt. Durch die sehr geringen Unterschiede zwischen den Werten der integrierten Lautheit ist ein Vorteil von lauten Mastern wie bei Spotify auszuschließen. Wird Klassik oder Jazz wiedergegeben, so sind LUFS größer -14 dB LUFS zu verzeichnen. Truepeak Werte waren während der Wiedergabe bei dem Lied

„Not Kings“ höher als 0 dBTP. Es ist zu vermuten, dass kein Begrenzer zum Einsatz kommt.

Titel	Interpret	Integrated LUFS	True Peak dBTP
Santa Monica	Theory of a Deadman	-14,1	-1,5
Dark Days	Parkway Drive	-14,1	-3,9
Einer muss den Job ja machen	Udo Lindenberg	-14,1	-3,2
Hoch	Tim Bendzko	-14	-3,2
Die Moldau	Uwe Mund	-17	-2
Hey Laura	Gregory Porter	-17,3	-2
Alphagenetik	Kollegah	-14	-2,6
Animals	Martin Garrix	-14,1	-1,3
Roller	Apache 207	-14	-2,2
Not Kings	Candy Says	-14	0,3
Dich zu Lieben (1982)	Roland Kaiser	-16,3	-2,5
Hulapalu	Andreas Gabalier	-14	-3,1
∅ in LUFS		-14,8	
∅ in LUFS(korrigiert) ^a		-14,0	

Tabelle 3: Messung des Streaming-Dienstes Amazon Music Unlimited^a Ausschluss der Lieder: „Die Moldau“, „Hey Laura“, „Dich zu Lieben“

Soundcloud

Soundcloud ist eine Streaming- und Download-Plattform, bei der jeder registrierte Nutzer kostenlos Musik hochladen kann. Dabei bietet sich diese Plattform besonders für Bewerbung für neue Musik und Künstler an.

Soundcloud gibt an: „Wir unterstützen AIFF, WAVE (WAV), FLAC, ALAC, OGG, MP2, MP3, AAC, AMR, und WMA Dateien.“ (Soundcloud 2020) Es werden dabei keine Angaben über Bittiefe, Samplerate oder Bitrate gemacht. Nach dem Hochladen der Dateien wird jedes Audiomaterial in 128 kbps MP3 kodiert. (Vgl. Soundcloud 2020)

Soundcloud nutzt keine Lautheitsnormalisierung. Es konnte festgestellt werden, dass die meisten Stücke eine integrierte Lautheit von -10 dB LUFS besitzen.

Titel	Interpret	Integrated LUFs	True Peak dBTP
Santa Monica	Theory of a Deadman	-8,3	1,1
Dark Days	Parkway Drive	-7,2	0,3
Einer muss den Job ja machen	Udo Lindenberg	-8,7	0,5
Hoch	Tim Bendzko	-7,3	0,5
Die Moldau	Uwe Mund	-12,4	0,1
Hey Laura	Gregory Porter	-14,2	-1,3
Alphagenetik	Kollegah	-8,4	0,8
Animals	Martin Garrix	-8,3	0,8
Roller	Apache 207	-9,2	0,7
Not Kings	Candy Says	-13,4	0,4
Dich zu Lieben (1982)	Roland Kaiser	-9,5	1,4
Hulapalu	Andreas Gabalier	-7,5	1,1
∅ in LUFs		-9,5	
∅ in LUFs(korrigiert) ^a		-8,7	

Tabelle 4: Messung des Streaming-Dienstes Soundcloud ^a Ausschluss der Lieder: „Die Moldau“, „Hey Laura“, „Dich zu Lieben“

YouTube

Wird ein Music Video für YouTube erstellt, kann man die Anforderungen den „recommended upload encoding settings“ entnehmen. Das Audiomaterial wird beim Rendern des Videos eingeschlossen. Das Master sollte in der besten Qualität an die Firma überliefert werden, die das Video erstellt. YouTube empfiehlt Videos im Containerformat MP4 hochzuladen. Das Containerformat MP4 beinhaltet den Audio codec AAC. Das Format soll eine Samplerate von 96 oder 48 kHz haben. Die Bitrate sollte für Stereo 384 kbps entsprechen. (Vgl. Youtube 2020)

YouTube macht selber keine Angaben zur Lautheitsnormalisierung von Videos. Die Messung des Video-Streaming-Dienstes zeigt, dass Lautheitsnormalisierung verwendet wird. YouTube normalisiert dabei auf einen Wert von durchschnittlich -14,1 LUFs. Schließt man das Lied „Hey Laura“ von der Berechnung aus, ergibt sich ein Durchschnitt von -14 LUFs. Es kann davon ausgegangen werden, dass „Hey Laura“, ähnlich wie bei anderen Streaming-Diensten, anders normalisiert wird. Auf Grund der geringen Abweichung der Lautheit ist zu vermuten, dass YouTube eine ähnliche Methode oder dieselbe verwendet, wie in der ITU BS 1770 Teil 4 beschrieben wird. Die Adjustierung der Lautheit kann dabei in den Videos abgelesen werden. Im Kontext Menü lassen sich „Statistik für Nerds“ aufrufen. In den Infos kann man unter content Loudness den Differenzwert ablesen. Wird zum Beispiel ein Wert von -6 dB angezeigt, wird das Video um 6 dB verstärkt, um die Lautheit dem Referenzwert anzugleichen.

Titel	Interpretet	Integrated LUFs	True Peak dBTP
Santa Monica	Theory of a Deadman	-14	-3,9
Dark Days	Parkway Drive	-14	-5,5
Hoch	Tim Bendzko	-14,1	-5,3
Hey Laura	Gregory Porter	-15	-1,7
Alphagenetik	Kollegah	-13,8	-3,3
Animals	Martin Garrix	-14	-3,9
Roller	Apache 207	-14	-3,3
Hulapalu	Andreas Gabalier	-14	-5
∅ LUFs		-14,1	-4,0
∅ in LUFs(korrigiert) ^a		-14,0	-4,3

Tabelle 5: Messung von Music-Videos des Video-Streaming-Dienstes YouTube ^a Ausschluss des Liedes: „Hey Laura“

YouTube bietet zusätzlich einen separaten Dienst zum Streamen von Musik an. Die Distribution von Musik erfolgt wie bei Amazone Music über Aggregatoren. Die Anforderung werden deshalb durch den Aggregator bestimmt. YouTube Music ist zudem nicht lautheitsnormalisiert, was man an den Messwerten verschiedener YouTube Videos erkennen kann. Wird der korrigierte Wert betrachtet, kann eine durchschnittliche Lautheit von -8,7 LUFs festgestellt werden.

Titel	Interpret	Integrated LUFs	True Peak dBTP
Santa Monica	Theory of a Deadman	-8	0,4
Dark Days	Parkway Drive	-7,4	0,4
Einer muss den Job ja machen	Udo Lindenberg	-8,4	0,5
Hoch	Tim Bendzko	-7,9	0,5
Die Moldau	Uwe Mund	-12,1	0,1
Hey Laura	Gregory Porter	-13,3	-0,1
Alphagenetik	Kollegah	-8,4	0,7
Animals	Martin Garrix	-8,7	1,3
Roller	Apache 207	-9,2	0,7
Not Kings	Candy Says	-13	0,3
Dich zu Lieben (1982)	Roland Kaiser	-16,3	-2,6
Hulapalu	Andreas Gabalier	-7,6	1
∅ in LUFs		-10,0	
∅ in LUFs(korrigiert) ^a		-8,7	

Tabelle 6: Messung des Streaming-Dienstes YouTube Music ^a Ausschluss der Lieder: „Die Moldau“, „Hey Laura“, „Dich zu Lieben“

4.3.3 CD

Die CD gehört zu den Massenmedien, welche keine Lautheitsnormalisierung besitzen. Die CDDA besitzt eine Samplerate von 44,1 kHz und eine Bittiefe von 16 Bit. Die CDDA besitzt dabei einen PCM kodierten Stereo Kanal. Die maximale Spieldauer beträgt 74 min. (Vgl.Görne 2015, S. 335)

4.4 Eigenschaften von Wiedergabesystemen

Die Eigenschaften der Wiedergabesysteme bestimmen die Übertragung der Musik vom Massenmedium zum Konsumenten. Dabei müssen vor allem die Eigenschaften Beachtung im Masteringprozess finden, welche Veränderungen an Frequenzgang und Dynamik vornehmen. Kleine Lautsprecher verursachen dabei häufig diese Probleme.

4.4.1 Kleine Lautsprecher

Kleine Lautsprecher sind in Smartphones und tragbaren Lautsprechern zu finden. Tragbare Lautsprecher sind meist batteriebetriebene Lautsprecherboxen bei denen das Signal durch Bluetooth oder ein 3,5 mm TRS Kabel eingespeist werden kann.

Um tiefe Frequenzen wiederzugeben muss ein größeres Luftvolumen bewegt werden um ein Hörereignis auszulösen. Das kann durch eine größere Oberfläche oder durch eine größere Elongation erreicht werden. Die Elongation des Lautsprechers ist durch seine Baueigenschaften begrenzt, weshalb nur durch eine größere Membran, tiefe Frequenzen wiedergegeben werden können. Auf Grund der geringen Einbaugröße der Treiber, ist der Frequenzgang dieser Boxen meist zu tiefen Frequenzen begrenzt. Die Wiedergabe von Frequenzen unter 60 Hz ist nicht immer möglich. Der meist verkaufte tragbare Lautsprecher auf Amazon ist der Soundcore 2 von Anker. Aus dessen Spezifikationen kann man entnehmen, dass er einen Frequenzgang von 70 Hz bis 20 kHz abdeckt. (Vgl. Anker 2017, S. 7)

Um das kleine Lautsprechergehäuse, die kleinen Lautsprecher und die geringe Leistung zu kompensieren, greifen viele Hersteller häufig auf AGC Tools oder fest verbaute Begrenzer zurück, welche die Lautsprecher schützen und die Lautheit steigern sollen. Bei Musik mit sehr hoher Dynamik kann die Dynamikbearbeitung zu hörbaren Artefakten führen. Diese Artefakte werden erfahrungsgemäß als Pumpen oder abrupte Pegelsprünge wahrgenommen.

Auf Grund der Größe von Smartphones müssen die Lautsprecher klein gehalten werden um noch in ein kompaktes Gehäuse zu passen. Für die Lautsprecher gelten in Bezug

auf ihre Größe die gleichen Aspekte wie bei tragbaren Lautsprechern. Der Frequenzgang der internen Lautsprecher ist deshalb durch ihre Größe extrem limitiert.

Die internen Lautsprecher eines Smartphones werden ähnlich wie tragbare Lautsprecher häufig durch DSP davor geschützt zu viel Leistung aufzunehmen. Um eine möglichst hohe Lautheit zu erzielen, wird dabei ein Kompromiss aus gesteigerter Lautheit und erhöhtem Klirrfaktor eingegangen. Erfahrungsgemäß kommt es bei Vollaussteuerung der Lautsprecher zu einer Verzerrung des Klangbildes.

4.4.2 Kopfhörer

Kopfhörer verfügen über den größten Frequenzbereich von allen aufgeführten Wiedergabesystemen. Sie haben den Vorteil, dass durch die geringe Distanz zum Ohr der Schalldruck nur geringfügig abfällt. Daraus resultiert, dass weniger Schalldruck erzeugt werden muss, um auch bei tiefen Frequenzen ein Hörereignis auszulösen. In der Betriebsanleitung des Kopfhörer MDR-ZX310 von Sony findet man einen angegebenen Frequenzbereich von 10 Hz bis 24 kHz. (Vgl. Sony 2018) Allgemein kann man sagen, dass alle Kopfhörer zumindest einen Frequenzgang von 20 Hz bis 20 kHz abbilden können. Kopfhörer, die für den Konsum von Musik bestimmt sind, wie der MDR-ZX310 von Sony, haben häufig einen verstärkten Bass- und Hochtonbereich. Das entspricht in etwa der negativen Abbildung der Kurven gleicher Lautstärke. Der Konsument nimmt die Musik als voller und scheinbar detaillierter wahr. Einzelne Klangobjekte können mit Kopfhörern besser gehört werden, da sie nicht durch Diffussschall maskiert werden.

4.4.3 Lautsprecher im PKW

Die Akustik im Innenraum eines PKWs hat sehr spezielle Eigenschaften. Durch das geringe Raumvolumen im Innenraum und die Dämpfung der gepolsterten Sitze, sind die Nachhallzeiten sehr kurz. Das sorgt für die subjektive Empfindung von Direktheit, da sich Quellsignal und Hallfahne nicht stark überlagern. Es treten ähnliche Effekte wie bei Kopfhörern auf.

4.4.4 Lautsprecherboxen

Die meisten Lautsprecherboxen können häufig einen breiten Frequenzbereich wiedergeben. Dieser liegt in der Regel bei 20 Hz bis 20 kHz. Größere nicht transportable Lautsprecherboxen kommen meistens in Abhörumgebungen zum Einsatz, wo bewusst Musik konsumiert wird.

5 Anforderung an den Masterprozess

5.1 Ausrichtung des Masterprozesses

Die Ausrichtung des Mastering Prozesses muss so gewählt werden, dass die Intention des Künstlers bestmöglich umgesetzt werden kann. Das bedeutet, dass durch den Masteringprozess die Qualität unabhängig von Massenmedium und Wiedergabesystem erhalten werden muss. Durch die gezielte Ausrichtung des Masterprozesses, abhängig von den Anforderungen und Eigenschaften der marktrelevanten Massenmedien und Wiedergabesysteme, lässt sich im Mittel eine höhere Qualität erzielen.

Eine wichtige Rolle für die Ausrichtung des Masteringprozesses spielen vor allem Streaming-Dienste. Wie in Kapitel 4 beschrieben, werden Streaming-Dienste weiter an Bedeutung gewinnen, während das Radio an Bedeutung verlieren wird, obwohl es aktuell noch eines der am meisten genutzten Massenmedien ist. Sowohl Radio als auch die Streamingdienste sind lautheitsnormalisiert, weshalb der Übergang von Radio zu Streamingdiensten keine Auswirkung auf den Aspekt der Lautheitsnormalisierung als Teil des Masteringprozesses hat.

Die Ausrichtung des Masteringprozesses anhand der Wiedergabesysteme gestaltet sich ähnlich wie die Ausrichtung anhand der Massenmedien. Um die bestmögliche Klangqualität beim Kunden zu erreichen, ist es notwendig die Qualität der Wiedergabe für die meistgenutzten Wiedergabesysteme zu optimieren. Zum Beispiel ist eines der am meisten genutzten Wiedergabesysteme das Smartphone. Optimiert man den Klang des Masters so, dass es ohne Probleme auf einem Smartphon wiedergegeben werden kann, ist es möglich, einem Großteil der Hörerschaft einen, den Intentionen des Künstlers entsprechenden, Klang zu vermitteln. Kombiniert man die Eigenschaften der Endgeräte mit dem Nutzungsverhalten, so lässt sich daraus ableiten, dass die Wiedergabe in der Regel über kleine Lautsprecher, Kopfhörer und Lautsprecher der PKW erfolgt. Während Kopfhörer und die Lautsprecher eines PKWs die Wiedergabe von Musik wenig einschränken, kann die Wiedergabe über kleine Lautsprecher zu den in Kapitel 4.4 beschriebenen Problemen führen. Da kleine Lautsprecher sehr marktrelevant sind, muss dieser Aspekt in die Anforderung an den Masterprozess mit einfließen.

Eine Ausrichtung des Masteringprozesses für Massenmedien und Wiedergabesysteme ist immer ein Kompromiss. Werden die Anforderungen an den Masteringprozess durch die marktrelevantesten Massenmedien und Wiedergabesysteme bestimmt, so ergeben sich daraus Nachteile für andere Systeme. Durch das Einbeziehen der Marktverteilung lässt sich jedoch eine Gewichtung erzielen, welche eine höhere Qualität für alle Konsumenten erzielt. Wird im Gegensatz dazu eine Optimierung für alle Geräte durchgeführt,

genießt der Großteil der Konsumenten nur die Hälfte des möglichen Qualität-Zuwachses.

5.2 Anforderung an den Masterprozess

5.2.1 Eigenschaften der Dynamik und Lautheit

Die Anforderungen an die Dynamik und Lautheit haben sich durch die Einführung der lautheitsnormalisierten Massenmedien geändert. Bis auf wenige Ausnahmen wie YouTube Music und Soundcloud hat eine unverhältnismäßige hohe Steigerung der Lautheit keine Marktrelevanz mehr. Im Gegenteil sollte eine allgemeine, den meistverwendetsten Massenmedien entsprechende Lautheit eingestellt werden, um Lautheitssprünge zwischen den Medien zu verringern. Durch moderate Lautheitspegel ist es möglich, mit größeren Dynamiken zu arbeiten, da die Spitzenpegel im Vergleich zu der durchschnittlichen Lautheit größer sein können.

Während der Vermessung des Streaming-Dienstes Spotify wurde festgestellt, dass laute Master lauter übertragen werden. Es ist jedoch ratsam, diesen Faktor zu vernachlässigen. Spotify möchte in Zukunft eine Lautheitsmessung nach der ITU BS 1770 Vorgaben einzuführen. (Vgl. Spotify 2020a) Die aktuellen Vorteile eines lauten Masterings könnten sich, wie in Kapitel 3.1.5 beschrieben, in Nachteile wandeln. Um zu vermeiden, dass Musikstücke mit geringer Lautheit durch den Begrenzer verändert werden, sollte während des Masterprozesses zumindest näherungsweise eine Lautheit von -14 LUFS und einem TP-Wert von -1 dBTP eingehalten werden.

Bei Apple Music konnte kein Zusammenhang zwischen einem lauterem Master und einer höheren ausgespielten Lautheit festgestellt werden. Die Werte der Lautheit von den Masterprozessen entsprechen mit geringen Abweichungen denen von Spotify. Größere Abweichungen können durch das Veröffentlichen von zwei unterschiedlich gemasterten Titeln entstehen. Allgemein sollten Lieder für Apple Music während des Masterprozesses nicht leiser als -16 LUFS sein und keine TP-Werte oberhalb von -1 dBFS aufweisen.

In dem korrigierten Mittelwert, bei dem drei Lieder auf Grund ihrer großen Differenz ausgeschlossen wurden, wurde erkannt, dass Amazon Music Unlimited nahezu alle Titel auf -14 LUFS normalisiert. Daraus kann geschlussfolgert werden, dass Lieder, welche auf -14 LUFS angepasst werden, bei Amazon Music Unlimited ohne Adjustierung ausgespielt werden. Da anzunehmen ist, dass Amazon Music kein Begrenzer verwendet, sollten Lieder, welche für Amazon Music Unlimited gemastert werden einen Truepeak-Wert von -1 dBTP bei einer Lautheit von -14 LUFS oder größer aufweisen. Ist die integrierte Lautheit größer als -14 LUFS, so ist davon auszugehen, dass nach unten adjustiert, und

somit ein größerer Aussteuerungsbereich geschaffen wird. Wird zum Beispiel ein Lied aus dem Bereich Pop mit dem Werte kleiner -14 LUFS gemastert, ist anzunehmen, dass es zu einer Übersteuerung nach der Lautheitsnormalisierung kommt.

Während des Masterprozesses sollte eine ähnliche Lautheit für die CD erreicht werden, wie für den Durchschnitt der lautheitsnormalisierten Streamingdienste. Der Grund dafür ist, dass immer mehr Konsumenten Musik über Streamingdienste konsumieren und Streaming-Dienste auf immer mehr Endgeräten verfügbar werden. Es wird deshalb immer wahrscheinlicher, dass vor oder nach dem Konsum über das Massenmedium CD die Ausspielung von Musik über einen lautheitsnormalisierten Streamingdienst stattgefunden hat. Es sollte beim Umschalten von einem lautheitsnormalisierten Streamingdienst auf eine CD keine Unterschiede in der Lautheit geben, um den Konsumenten den größtmöglichen Komfort zu bieten. Die durchschnittliche Lautheit der lautheitsnormalisierten Streamingdienste beträgt -14,32 LUFS. Der Wert wurde aus dem Durchschnitt des korrigierten Mittelwertes berechnet.

Weil Radioprogramme nach der EBU R 128 Empfehlungen normalisieren werden, kann man aus einer Steigerung der Lautheit während des Masteringprozesses keinen ökonomischen Mehrwert erzielen. Im Gegenteil kann durch die Bearbeitung durch die im Radio verwendeten Filter und Dynamikprozessoren, eine weitere Degradierung des zu überkomprimierten und übersteuerten Audiomaterials stattfinden. Wie in Kapitle 4.3.1 Beschrieben sind viele der Rundfunkbetreiber mit Dynamikprozessoren ausgestattet, welche Gehörermüdung bei zu stark komprimierten Audiomaterial vermeiden sollen. Ein weiteres nichtsteuerbares Eingreifen durch Dynamikprozessoren kann durch eine Transparente und dem Lied entsprechend angepasste Dynamik verhindert werden.

Während des Masterprozesses ist es sinnvoll ein Master anzulegen, welches in allen nicht normalisierten Streaming-Diensten und für Werbezwecke auf Instagram verwendet werden kann. Wichtig ist es, einen Truepeak von gleich oder kleiner -1 dBTP einzuhalten, um Übersteuerungen zu vermeiden. Die Lautheit der nicht lautheitsnormalisierten Streaming-Dienste beträgt -8,72 LUFS. Wird gleichzeitig eine Veröffentlichung auf Spotify in Betracht gezogen, mit der Absicht lauter ausgespielt zu werden, ist es sinnvoll eine integrierte Lautheit von -9 LUFS zu erreichen.

Auf Grundlage der Ausrichtung des Masteringprozesses ist es wichtig, die Dynamik hinsichtlich der Abspielbarkeit auf Wiedergabesystemen mit kleinen Lautsprechern zu optimieren. Der pumpende Effekt von tragbaren Lautsprechern bei zu großen Lautheitssprüngen kann umgangen werden, indem laute und leise Passagen im Lied zum Beispiel mit manueller Automation der Verstärkung angeglichen werden. Dabei sollte auch beachtet werden, dass tragbare Lautsprecher hauptsächlich dann eingesetzt werden, wenn der Konsum der Musik nicht im Fokus steht, sondern nur als Unterhaltung

dient. Es ist daraus zu schlussfolgern, dass Lautheitsunterschiede geringer ausfallen sollten, um zu vermeiden, dass bei schwankenden Umgebungsgeräuschen die Musik trotzdem noch hörbar ist. Angesichts der geringen Nutzung von tragbaren Lautsprechern ist diesem Aspekt wenig Gewichtung zuzuweisen.

Im Gegensatz zur Wiedergabe über kleine Lautsprechersysteme beeinflusst die Wiedergabe eines großen Dynamikbereiches über Kopfhörer oder Nahfeldlautsprecher das Hörerlebnis nicht negativ, sondern positiv. Wird Musik auf diesen Wiedergabesystemen abgespielt, kann davon ausgegangen werden, dass Störgeräusche nur wenig Einfluss nehmen und eine größere Dynamik dadurch ein besseres Hörerlebnis hervorruft. Im Gegensatz dazu können Lieder mit wenig Dynamik negativ auffallen. Sie fallen dem Konsumenten meist als zu intensiv und flach auf. Durch das intensive Bearbeiten mit Dynamikprozessoren werden Frequenzen hörbar, welche zuvor nur leise wahrnehmbar waren. Werden immer mehr Hörereignisse ausgelöst, führt das zu einem Hörerlebnis, welches man als anstrengend beschreiben kann, weil mehr Informationen im Gehirn verarbeitet werden müssen.

In Bezug auf die Lautheit und Austeuerung, lässt sich ableiten, dass es sich lohnt zwei Versionen mit unterschiedlichen Anforderungen anzulegen. Beide Versionen sollten bei -1 dBTP ausgesteuert werden um Übersteuerungen zu vermeiden und Anforderungen von den speziellen Massenmedien zu erfüllen. Ein Master sollte mit bestmöglicher Dynamik erstellt werden, wobei die Lautheit einen Wert von -15 LUFS nicht überschreiten sollte um eine Übersteuerung bei der Adjustierung zu vermeiden. Ein weiteres Master sollte mit -9 LUFS für die nicht lautheitsnormalisierten angelegt werden.

5.2.2 Eigenschaften des Frequenzgangs

Vor allem die Übertragung über das Massenmedium UKW-Radio beschneidet den Frequenzbereich stark. Es muss deshalb beim Masteringprozess darauf geachtet werden, dass keine wichtigen Informationen ober- und unterhalb dieser Frequenzbereiche enthalten sind. Sind wichtige Informationen im Bassbereich unter 30 Hz enthalten, können diese zum Beispiel saturiert werden. Durch die Saturation entstehen Harmonische in dem Frequenz-Abschnitt über dem kritischen Frequenzbereich. Die musikalischen Informationen bleiben in den Harmonischen enthalten. Der Verlust der hohen Frequenzen kann zum Beispiel durch Einsatz eines Entzerrers kompensiert werden. Durch einen Entzerrer können die Frequenzen unterhalb von 15 kHz mit einem Kuhschwanzfilter angehoben werden. Dadurch kann der Klang als detaillierter wahrgenommen werden, ohne dass er über sehr hohe Frequenzanteile verfügt.

Streaming-Dienste bieten im Gegenteil zum UKW-Radio eine höhere Frequenzbandbreite. Auf der Einstellung hohe Audioqualität gibt Spotify einen Frequenzbereich von 20

Hz bis 20 kHz wieder. Wird für lauthheitsnormalisierte Streaming-Dienste ein Master angefertigt, so besteht die Möglichkeit, den Bassbereich stärker zu betonen, ohne dadurch in der Lautheit benachteiligt zu werden. Zwei identisch laute Songs können unterschiedlich starke Bassbereiche aufweisen. Es muss beachtet werden, dass viele Konsumenten Musik über kleine Lautsprecher, wie zum Beispiel dem Smartphone hören. Ein Master, welches überdurchschnittlich im Bassbereich verstärkt wurde, könnte zur nicht optimalen Wiedergabe auf den meisten Endgeräten führen.

Ein Tiefpass sollte während des Masteringprozesses zum Einsatz kommen. Um gerade bei kleinen Lautsprechern starke Verzerrungen zu vermeiden und bei nicht normalisierten Massenmedien eine höhere Lautheit zu erreichen. Gleichzeitig sollten Nutzer von Kopfhörern die Möglichkeit besitzen, die wiedergebbaren Frequenzen bestmöglich zu nutzen. Als Grenzfrequenz bietet sich ein Bereich von 20-30 Hz an. Die meisten Lautsprecher können tiefere Frequenzen nicht wiedergeben. Bei Kopfhörern sind zwar tiefere Frequenzen als 20 Hz möglich, jedoch ist hier das Ohr mit 16 Hz der begrenzende Faktor. Die Flankensteilheit des Filters kann nach Qualität des Bassbereiches gewählt werden. Es sollte bei der Filterung ein digitaler Filter verwendet werden, da vor allem bei hoher Flankensteilheit analoge Filter wesentlich mehr Verzerrungen verursachen.

Die Balance des Frequenzgangs von Musik ist dabei immer einer subjektiven Einschätzung unterworfen. Die klangästhetische Anforderung an einen Master können sich wie auch die Hörgewohnheiten der Konsumenten, ständig ändern. Allgemein ist für viele Masterprozesse, ein zu hohen Frequenzen gleichmäßig abfallender Frequenzgang dem allgemeinen Konsumenten zuträglich. Diese Behauptung kann durch eine Messung des durchschnittlichen Frequenzganges der zuvor verwendeten Lieder gestützt werden. Für die Messung wurde der die Mittelungszeit auf unendlich gestellt. Danach wurden die Lieder nach einander über Amazone Music Unlimited abgespielt. Am Ende der Messung wurde der Frequenzgang durch einen Screenshot übernommen. Ähnlich wie die Dynamik ist der Frequenzgang auch sehr vom Musikstil abhängig. Modifikationen des Frequenzganges können mit statischen Entzerrern erreicht werden. Treten Probleme, wie störende Resonanzen oder überrepräsentierte Frequenzen auf, können diese mit dynamischen Entzerrern bedämpft oder entfernt werden.

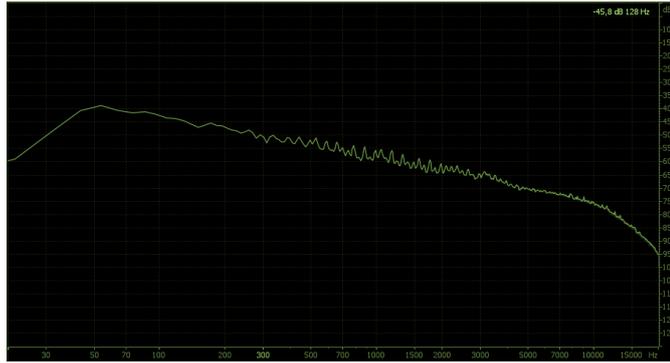


Abbildung 1: Durchschnittlicher Frequenzgang ausgewählter Lieder des Streaming-Dienstes Amazon Music Unlimited

5.2.3 Samplerate und Bittiefe

Möglich ist es, eine Einheitliche Bittiefe für den Export zu nutzen. 16 Bit ist die kleinste gemeinsame Bittiefe. Soll als Massenmedium nur Spotify und Apple Music zum Einsatz kommen, kann eine Bittiefe von 24 Bit gewählt werden. Die Erzeugung unterschiedlicher Master für unterschiedliche Medien in Bezug auf die Bittiefe ist sehr sinnvoll. Die Encoder von Apple und Spotify profitieren von der höheren Bittiefe. Sie können dadurch ein genaueres Abbild, des durch den Kunden bewilligten Klanges schaffen.

Soll die maximale Qualität bei der Wiedergabe erreicht werden, müssen bezüglich der Samplerate unterschiedliche Master angefertigt werden. Eine höhere Quell-Samplerate sorgt bei den Streaming-Diensteanbietern Spotify und Apple Music für eine bessere Qualität. Mögliches Aufaddieren von Fehlern, welche durch die Konzipierung der antialiasing-Filter entstehen, werden reduziert. Soll ein Master für alle Massenmedien verwendet werden, muss die kleinste gemeinsame Samplerate gewählt werden. Diese entspricht bei allen Medien 44,1 kHz.

Wird nur eine Version des Masters angefertigt, so setzen sich die Anforderungen aus den übereinstimmenden Eigenschaften der unterschiedlichen Medien zusammen. Das Master sollte dann mit einer Samplerate von 44,1 kHz und einer Bittiefe von 16 Bit angefertigt werden.

5.2.4 Dithering

Für Spotify und Apple Music muss das Master auf 24 Bit gedithert werden. Wird für YouTube und Amazon Music ein Master angefertigt, sollte dies entsprechend den Vorgaben des Aggregators auf 16 Bit gedithert werden. Quantisierungsfehler können dadurch verringert und die möglichst genaue Wiedergabe des gemasterten Audiomaterials ermöglichen werden. Um eine möglichst rauschfreie Wiedergabe von besonders

dynamischen Musikstücken zu ermöglichen, sollte Noiseshaping beim Dithering auf 16 Bit verwendet werden. Bei Musikstücken, die eine sehr hohe Lautheit im Masterprozess erreicht, kann auf ein Noiseshaping während des Dithering verzichtet werden. Durch die hohe Lautheit maskiert das Nutzsignal das Rauschen völlig. Allerdings sind keine Nachteile bekannt, welche durch die Verwendung von Noiseshaping entstehen könnten.

5.2.5 Daten Distribution

Die Überprüfung der Datendistribution kann fast immer gewährleistet werden. Spotify und Apple Music überprüfen die Daten nach der Distribution auf Korrumpierung. Bei der Distribution zu Soundcloud, kann das Format FLAC genutzt werden, da es neben hoher Qualität auch eine Korruptionsschutz bietet. Allgemein besitzen viele Formate Schutzmechanismen um Korruption zu vermeiden.

6 Schlussbetrachtung

Schlussendlich lässt sich der Masterprozess sehr gut auf die Eigenschaften der verschiedenen Medien ausrichten. Um die bestmögliche Qualität in allen Medien zu erreichen, müssen mehrere Master angefertigt werden. Ist dies aus ökonomischen Gründen nicht möglich, muss sich an den Anforderungen der Medien orientiert werden, welche für die Veröffentlichung genutzt werden. Die Arbeit hat gezeigt, dass es trotz der Lautheitsnormalisierung vieler Massenmedien immer noch notwendig sein kann ein Master sehr laut anzufertigen. Trotz der entstandenen Referenzwerte für die Lautheit sollte beachtet werden, dass verschiedene Musikrichtungen auch die Manipulation von Makro- und Mikrodynamiken als Stilmittel verwenden. Gerade bei Musikrichtungen wie zum Beispiel Metal ergibt es Sinn, die Verdichtung des Klanges und damit die Einstellung der Lautheit nur nach ästhetischen Aspekten zu regulieren.

Während der Erstellung der Arbeit ist dem Autoren aufgefallen, dass vor allem die detaillierte Ausrichtung des Masterprozesses in Hinsicht auf die Wiedergabesysteme sehr aufwändig ist. Das Problem ergibt sich aus dem Mangel an Daten. Um eine möglichst genaue Abstimmung zu ermöglichen, müssten alle Eigenschaften der am häufigsten verwendeten Wiedergabesysteme bekannt sein. Zum Beispiel müssten die Frequenzgänge der verschiedenen Wiedergabesysteme gemittelt werden, um Trends ersichtlich zu machen und so in die Ausrichtung des Masteringprozesses einfließen zu lassen. Eine Forschung in diesem Bereich würde sich vor allem dann anbieten, wenn sich Monopollösungen von einzelnen Wiedergabesystemen weiter ausbauen. Dadurch würden sich noch spezifischere Anpassungen des Masterprozesses realisieren.

Letztlich ist noch aufzuführen, dass diese Arbeit eine Ermittlung der aktuellen Anforderungen an den Masterprozess ist. Mit der Weiterentwicklung der Medien werden sich auch die Anforderungen an den Masterprozess ändern.

Literaturverzeichnis

Altenmüller, Eckart; Kopiez, Reinhard (2012): Starke Emotionen und Gänsehaut beim Musikhören: Evolutionäre und musikpsychologische Aspekte. In: *Zeitschrift für Audiologie* 2013, 14.02.2012 (19). Online verfügbar unter <https://www.kind-hoerstiftung.de/wp-content/uploads/2014/01/Band-19.pdf>.

Amazon (2020): Amazon.de: Amazon Music HD FAQs: Musik-Downloads. Welche Klangqualität unterstützt Amazon Music HD? Hg. v. Amazon. Online verfügbar unter <https://www.amazon.de/b?ie=UTF8&node=8322164031>, zuletzt aktualisiert am 01.06.2020, zuletzt geprüft am 01.06.2020.

Anker (2017): Owners Manual. Soundcore 2. Hg. v. Anker. Online verfügbar unter https://d2211byn0pk9fi.cloudfront.net/spree/accessories/attachments/70877/A3105_Manual_artwork__SM-A508-V01_20170314_65X105mm_for_reviewing.pdf?1533028050, zuletzt geprüft am 02.06.2020.

Bock, Stefan (2014): Mastering. Definition des Mastering. In: Michael Dickreiter, Volker Dittel, Wolfgang Hoeg und Martin Wöhr (Hg.): *Handbuch der Tonstudioteknik*. 8., 8., völlig überarbeitete und erweiterte Aufl. Berlin: De Gruyter.

Bundesverband Musikindustrie e.V. (2018): Musikindustrie in Zahlen 2017 2018. Online verfügbar unter http://www.miz.org/downloads/dokumente/932/2018_Musikindustrie-in-Zahlen_BVMI.pdf, zuletzt geprüft am 20.05.2020.

Bundesverband Musikindustrie e.V. (2020): Musikindustrie in Zahlen 2019 2020. Online verfügbar unter https://www.musikindustrie.de/fileadmin/bvmi/upload/06_Publikationen/MiZ_Jahrbuch/2019/Musikindustrie_in_Zahlen_2019.pdf, zuletzt geprüft am 20.05.2020.

Bundeszentrale für politische Bildung (2011): Massenmedien | bpb. Bundeszentrale für politische Bildung. Internet. Online verfügbar unter <https://www.bpb.de/nachschlagen/lexika/pocket-politik/16493/massenmedien>, zuletzt aktualisiert am 17.07.2005, zuletzt geprüft am 24.06.2020.

Büsching, Thilo; Goderbauer-Marchner, Gabriele (2014): *E-Publishing-Management*. Wiesbaden: Springer Gabler.

Dickreiter, Michael (2014a): Mikrofone und Lautsprecher. Bauformen von Lautsprechern. In: Michael Dickreiter, Volker Dittel, Wolfgang Hoeg und Martin Wöhr (Hg.): *Handbuch der Tonstudioteknik*. 8., 8., völlig überarbeitete und erweiterte Aufl. Berlin: De Gruyter.

Dickreiter, Michael (2014b): Schallschwingung. Schallformen und Begriffe. In: Michael Dickreiter, Volker Dittel, Wolfgang Hoeg und Martin Wöhr (Hg.): *Handbuch der Tonstudioteknik*. 8., 8., völlig überarbeitete und erweiterte Aufl. Berlin: De Gruyter.

Dickreiter, Michael; Goeres-Petri, Jürgen (2014): Schallwahrnehmung. Eigenschaften der Schallwahrnehmung. In: Michael Dickreiter, Volker Dittel, Wolfgang Hoeg und Martin Wöhr (Hg.): *Handbuch der Tonstudioteknik*. 8., 8., völlig überarbeitete und erweiterte Aufl. Berlin: De Gruyter.

EBU (2014): R 128. Loudness normalisation and permitted maximum level of audio signals. Online verfügbar unter <https://tech.ebu.ch/docs/r/r128.pdf>, zuletzt geprüft am 07.05.2020.

EBU (2016a): TECH 3341. Loudness Metering: 'EBU Mode' metering to supplement EBU R 128 Loudness normalization. Online verfügbar unter <https://tech.ebu.ch/docs/tech/tech3341.pdf>, zuletzt geprüft am 07.05.2020.

EBU (2016b): TECH 3342. Loudness range: A measure to supplement EBU R 128 Loudness normalization. Online verfügbar unter <https://tech.ebu.ch/docs/tech/tech3343.pdf>, zuletzt geprüft am 07.05.2020.

EBU (2016c): TECH 3343. Guidelines for production of programmes accordance with EBU R 128, zuletzt geprüft am 08.05.2020.

Ellermeier W., Hellbrück J. (2008a): Hören - Psychoakustik - Audiologie. Anatomie und Physiologie des Gehörs. In: Stefan Weinzierl (Hg.): Handbuch der Audiotechnik. [Ausgabe in 2 Bänden]. Berlin, Heidelberg: Springer.

Ellermeier W., Hellbrück J. (2008b): Hören - Psychoakustik - Audiologie. Psychoakustik. In: Stefan Weinzierl (Hg.): Handbuch der Audiotechnik. [Ausgabe in 2 Bänden]. Berlin, Heidelberg: Springer.

Feierabend, Sabine; Rathgeb, Thomas; Reutter, Theresa (2019): JIM-Studie 2019. Basisuntersuchung zum Medienumgang 12- bis 19- Jähriger. Hg. v. Medienpädagogischer Forschungsverbund Südwest (mpfs). Landesanstalt für Kommunikation Baden-Württemberg, Landeszentrale für Medien und Kommunikation Rheinland-Pfalz. Online verfügbar unter https://www.mpfs.de/fileadmin/files/Studien/JIM/2019/JIM_2019.pdf, zuletzt geprüft am 26.05.2020.

Fletcher, Harvey; Munson, W. A. (1933): Loudness, Its Definition, Measurement and Calculation. In: *The Journal of the Acoustical Society of America* 5 (2), S. 82–108. DOI: 10.1121/1.1915637.

Fremon, Christie (2016): Mastered for iTunes 5.0. Music as the Artist and Sound Engineer Intended. Online verfügbar unter https://images.apple.com/euro/itunes/mastered-for-itunes/docs/mastered_for_itunes.pdf, zuletzt geprüft am 30.05.2020.

Friesecke, Andreas (2007): Die Audio-Enzyklopädie. Ein Nachschlagewerk für Tontechniker. Berlin: De Gruyter. Online verfügbar unter <http://www.degruyter.com/doi/book/10.1515/9783110973310>.

Goeres-Petri, Jürgen; Dickreiter, Michael (2014): Schallwahrnehmung. Das Ohr. In: Michael Dickreiter, Volker Dittel, Wolfgang Hoeg und Martin Wöhr (Hg.): Handbuch der Tonstudientechnik. 8., 8., völlig überarbeitete und erweiterte Aufl. Berlin: De Gruyter.

Görne, Thomas (2015): Tontechnik. Hören, Schallwandler, Impulsantwort und Faltung, digitale Signale, Mehrkanaltechnik, tontechnische Praxis. 4., aktualisierte Auflage. München: Hanser (Medien).

Götz, Romahn (2014): Grundlagen der digitalen Tontechnik. Quantisierung. In: Michael Dickreiter, Volker Dittel, Wolfgang Hoeg und Martin Wöhr (Hg.): Handbuch der Tonstudientechnik. 8., 8., völlig überarbeitete und erweiterte Aufl. Berlin: De Gruyter.

- H.-J.Maempel, S.Weinzierl, P.Kaminski (2008a): Filter. Equalizer. In: Stefan Weinzierl (Hg.): Handbuch der Audiotechnik. [Ausgabe in 2 Bänden]. Berlin, Heidelberg: Springer.
- H.-J.Maempel, S.Weinzierl, P.Kaminski (2008b): Kompressor. Statisches Verhalten. In: Stefan Weinzierl (Hg.): Handbuch der Audiotechnik. [Ausgabe in 2 Bänden]. Berlin, Heidelberg: Springer.
- Heidrich, klaus; Maniak, Stephan (2014): Digitale Tonspeicherung. Speichersysteme. In: Michael Dickreiter, Volker Dittel, Wolfgang Hoeg und Martin Wöhr (Hg.): Handbuch der Tonstudioteknik. 8., 8., völlig überarbeitete und erweiterte Aufl. Berlin: De Gruyter.
- Heyna, Arne; Schmidt, Ulrich; Briede, Marc (2004): Datenformate im Medienbereich. Mit ... 26 Tabellen ; [digitale Signalformen, Datenreduktion, MPEG, Metadaten, Fileformate, AVI, Quicktime, MXF. 1. Aufl. s.l.: Carl Hanser Fachbuchverlag. Online verfügbar unter <http://www.hanser-elibrary.com/isbn/9783446225428>.
- Hoeg, Wolfgang (2014a): Terrestrische Rundfunksysteme. FM-Rundfunk im UKW-Band II. In: Michael Dickreiter, Volker Dittel, Wolfgang Hoeg und Martin Wöhr (Hg.): Handbuch der Tonstudioteknik. 8., 8., völlig überarbeitete und erweiterte Aufl. Berlin: De Gruyter.
- Hoeg, Wolfgang (2014b): Verlustfreie Audiocodierverfahren. Free Lossless Audio Codec (FLAC). In: Michael Dickreiter, Volker Dittel, Wolfgang Hoeg und Martin Wöhr (Hg.): Handbuch der Tonstudioteknik. 8., 8., völlig überarbeitete und erweiterte Aufl. Berlin: De Gruyter.
- ITU (2007): RECOMMENDATION ITU-R BS.1770-1. Algorithms to measure audio programme loudness and true-peak audio level. Online verfügbar unter <https://tech.ebu.ch/docs/tech/tech3341.pdf>, zuletzt geprüft am 07.05.2020.
- ITU (2017): RECOMMENDATION ITU-R BS.1770-4. Algorithms to measure audio programme loudness and true-peak audio level. Online verfügbar unter https://www.itu.int/dms_pubrec/itu-r/rec/bs/R-REC-BS.1770-4-201510-!!!PDF-E.pdf, zuletzt geprüft am 07.05.2020.
- Katz, Robert A. (2015): Mastering audio. The art and the science. Third edition. Burlington, MA, Abingdon, Oxon: Focal Press Taylor & Francis Group.
- Kunst, Alexander (2019): Kopfhörermarken in Deutschland 2019 Statista. von welcher Marke sind ihre persönlichen Kopfhörer. Hg. v. Statista. Statista. Internet. Online verfügbar unter <https://de.statista.com/prognosen/999846/umfrage-in-deutschland-zu-beliebten-kopfoerermarken>, zuletzt aktualisiert am 27.05.2020, zuletzt geprüft am 27.05.2020.
- Lauterbach, Ulrike (2019): Optimod 8700i A3. FM & DAB+/HD Audio Processor. Hg. v. Orban. Internet. Online verfügbar unter <https://static1.squarespace.com/static/58f8d954b8a79b4ccf726c3b/t/5d77ae5bcce11f578860e233/1568124535221/8700i+A3+brochure.pdf>, zuletzt geprüft am 12.06.2020.
- Maempel, Hans-joachim (2014a): Dynamik und Lautheit. Expander und Gate. In: Michael Dickreiter, Volker Dittel, Wolfgang Hoeg und Martin Wöhr (Hg.): Handbuch der Tonstudioteknik. 8., 8., völlig überarbeitete und erweiterte Aufl. Berlin: De Gruyter.

Maempel, Hans-joachim (2014b): Dynamik und Lautheit. Kompressor und Limiter. In: Michael Dickreiter, Volker Dittel, Wolfgang Hoeg und Martin Wöhr (Hg.): Handbuch der Tonstudioteknik. 8., 8., völlig überarbeitete und erweiterte Aufl. Berlin: De Gruyter.

Maempel, Hans-joachim (2014c): Klanggestaltung. Klangfarbe. In: Michael Dickreiter, Volker Dittel, Wolfgang Hoeg und Martin Wöhr (Hg.): Handbuch der Tonstudioteknik. 8., 8., völlig überarbeitete und erweiterte Aufl. Berlin: De Gruyter.

Mai, Lothar; Meinzer, Nils; Schröter Christian (2019): ARD/ZDF-Onlinestudie 2019: Radio- und Audionutzung 2019. Hg. v. ARD ZDF. Kantar. Online verfügbar unter http://www.ard-zdf-onlinestudie.de/files/2019/0919_Mai_Meinzer_Schroeter.pdf, zuletzt geprüft am 24.05.2020.

Renner, Tim (2004): Kinder, der Tod ist gar nicht so schlimm! Über die Zukunft der Musik- und Medienindustrie. Frankfurt/Main: Campus-Verl.

Robinson, David J. M. (2002): Perceptual model for assessment of coded audio. Essex, Colchester. Department of Electronic Systems Engineering. Online verfügbar unter

<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.692.6717&rep=rep1&type=pdf>, zuletzt geprüft am 29.05.2020.

Schmidt, Sebastian; Zaborowski, Katrin (2017): Video- und Musik-Streaming-Dienste aus Verbrauchersicht. Eine Untersuchung der Verbraucherzentrale - Oktober 2017 2017. Online verfügbar unter https://www.verbraucherzentrale-rlp.de/sites/default/files/2017-10/Marktw%C3%A4chter_Untersuchung%20zu%20Streaming-Diensten%20und%20AGB.pdf, zuletzt geprüft am 23.05.2020.

Schwarzenberger, Gernot Meyer (2014): Grundlagen der digitalen Tontechnik. Audio-dateiformate. In: Michael Dickreiter, Volker Dittel, Wolfgang Hoeg und Martin Wöhr (Hg.): Handbuch der Tonstudioteknik. 8., 8., völlig überarbeitete und erweiterte Aufl. Berlin: De Gruyter.

Sony (2018): Stereo Headphones. MDR-ZX310. Hg. v. Sony. Online verfügbar unter <https://www.sony.com/electronics/support/res/manuals/4463/44637764M.pdf>, zuletzt geprüft am 02.06.2020.

Soundcloud (2020): Anforderungen für das Hochladen. Hg. v. Soundcloud. Soundcloud. Internet. Online verfügbar unter <https://help.soundcloud.com/hc/de/articles/115003452847-Anforderungen-f%C3%BCr-das-Hochladen>, zuletzt aktualisiert am 30.05.2020, zuletzt geprüft am 30.05.2020.

Spikofski, Gerhard; Camerer, Florian (2014): Spitzenpegelbezogene Aussteuerung. Quasi-Spitzenwert und True Peak-Aussteuerungsmesser. In: Michael Dickreiter, Volker Dittel, Wolfgang Hoeg und Martin Wöhr (Hg.): Handbuch der Tonstudioteknik. 8., 8., völlig überarbeitete und erweiterte Aufl. Berlin: De Gruyter.

Spotify (2020a): Mastering & loudness – FAQ – Spotify for Artists. Hg. v. Spotify for Artist. Internet. Online verfügbar unter <https://artists.spotify.com/faq/mastering-and-loudness#how-does-spotify-process-my-audio-files>, zuletzt aktualisiert am 28.05.2020, zuletzt geprüft am 28.05.2020.

Spotify (Hg.) (2020b): Soundeinstellungen - Spotify. Details zur Soundqualität. Spotify. Online verfügbar unter <https://support.spotify.com/de/article/high-quality-streaming/>, zuletzt aktualisiert am 29.05.2020, zuletzt geprüft am 29.05.2020.

Steinberg Media Technologies GmbH (2018): WaveLab Pro 9.5.20 - Benutzerhandbuch. Online verfügbar unter https://steinberg.help/wavelab_pro/v9.5/de/WaveLab_Pro_9_5_Benutzerhandbuch_de.pdf, zuletzt geprüft am 18.04.2020.

Tunecore (2020): Welches Format muss meine Musik haben, damit sie auf TuneCore hochgeladen werden kann? Hg. v. Tunecore. Online verfügbar unter <https://support.tunecore.com/hc/de/articles/115006499327-Welches-Format-muss-meine-Musik-haben-damit-sie-auf-TuneCore-hochgeladen-werden-kann->, zuletzt aktualisiert am 01.06.2020, zuletzt geprüft am 01.06.2020.

VuMA Touchpoints (2019): Konsumenten punktgenau erreichen. Basisinformationen für fundierte Mediaentscheidungen. VuMA Touchpoints 2019. Hg. v. Arbeitsgemeinschaft Verbrauchs- und Medienanalyse (Berichtsband VuMA). Online verfügbar unter <https://www.vuma.de/vuma-praxis/vuma-berichtsband/>, zuletzt geprüft am 27.05.2020.

Weidenbach, Bernhard (2017): Mobiler Musikkonsum - Nutzungshäufigkeit ausgewählter Geräte 2016. Anteil der befragten musikkaffinen Internetnutzer, die am liebsten über die folgenden Geräte unterwegs Musik hören. Hg. v. Statista. Statista. [de.the-sound-board.com](https://de.statista.com/statistik/daten/studie/723316/umfrage/zur-nutzungshaeufigkeit-ausgewaehlter-geraete-beim-mobilenumusikkonsum-in-deutschland/). Online verfügbar unter <https://de.statista.com/statistik/daten/studie/723316/umfrage/zur-nutzungshaeufigkeit-ausgewaehlter-geraete-beim-mobilenumusikkonsum-in-deutschland/>, zuletzt aktualisiert am 27.05.2020, zuletzt geprüft am 27.05.2020.

Weinzierl, Stefan (2008a): Aussteuerung und Lautheit. Lautheitsmessung bei mehrkanaligem Audiomaterial nach ITU-R Bs.1770. In: Stefan Weinzierl (Hg.): Handbuch der Audiotechnik. [Ausgabe in 2 Bänden]. Berlin, Heidelberg: Springer.

Weinzierl, Stefan (2008b): Aussteuerung und Lautheit. Leq Messung. In: Stefan Weinzierl (Hg.): Handbuch der Audiotechnik. [Ausgabe in 2 Bänden]. Berlin, Heidelberg: Springer.

Weinzierl, Stefan; Lerch, Alexander (2008): Digitale Audiotechnik: Grundlagen. Dither. In: Stefan Weinzierl (Hg.): Handbuch der Audiotechnik. [Ausgabe in 2 Bänden]. Berlin, Heidelberg: Springer.

Youtube (2020): Recommended upload encoding settings - YouTube Help. Hg. v. Youtube. Internet. Online verfügbar unter <https://support.google.com/youtube/answer/1722171?hl=en>, zuletzt aktualisiert am 02.06.2020, zuletzt geprüft am 02.06.2020.

Zehentmeier, wolfgang (2015): BR optimiert Klang durch Aussteuerung nach Lautheit. Hg. v. BR. Pressemitteilung. Internet. Online verfügbar unter <https://www.br.de/presse/inhalt/pressemitteilungen/radio-r128-lautheit-klang-100.html>, zuletzt aktualisiert am 10.07.2015, zuletzt geprüft am 04.06.2020.

Eigenständigkeitserklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und nur unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt habe. Stellen, die wörtlich oder sinngemäß aus Quellen entnommen wurden, sind als solche kenntlich gemacht. Diese Arbeit wurde in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegt.

Ort, Datum

Vorname Nachname