

Konzeptionelle Flexibilität digitaler Audioproduktion

Methodische, pragmatische und organisatorische
Prozesse digitaler Audioproduktionsparadigmen und
deren konzeptionelle Flexibilität im Rahmen einer
rein rechnergestützten Aufnahmeumgebung

© 2008 Marco Nassenstein
<http://nassenstein.com>

*Dieses Werk ist für den privaten Gebrauch freigegeben.
Alle genannten Markennamen sind Eigentum ihrer jeweiligen Inhaber.*

Inhaltsverzeichnis

Abstract.....	5
1. Einleitung.....	5
2. Vorgehensmodelle der Audioproduktion.....	7
2.1 Sequentielle Audioproduktion.....	8
2.2 Parallele Audioproduktion.....	13
3. Qualitätsanforderungen an die digitale Technik.....	17
3.1 Zeitdiskretisierung.....	18
3.2 Amplitudenquantisierung.....	24
3.3 Resultierende Speichieranforderungen.....	27
3.4 Sinnvolle Distributionswege und -arten für digitales Audiomaterial.....	31
4. Flexibilität in der Praxis anhand eines Beispiels.....	34
4.1 Zugrundeliegendes Werk.....	34
4.2 Aufnahme.....	40
4.2.1 E-Schlagzeug.....	40
4.2.2 MIDI Bass.....	49
4.2.3 Rhythmus E-Gitarren.....	58
4.2.4 Solo E-Gitarren.....	67
4.2.5 Ambient.....	68
4.2.6 Chor.....	70
4.2.7 Gesang.....	74
4.3 Effekte der Einzelspuren.....	79
4.3.1 Dynamiken.....	80
4.3.2 Frequenzen.....	86
4.3.3 Pitch Shifter / Time Stretcher.....	92
4.3.4 Hall / Roomsimulator / Pre-Delay.....	95
4.3.5 Stereoeffekte.....	97
4.3.6 Weitere Effekte.....	99
4.4 Mix.....	100
4.5 Mastering.....	110
5. Bestandsaufnahme und Qualitätsbewertung.....	115
6. Musiker- und Zeitmanagement.....	118
7. Persönliches Fazit und Schlußwort.....	121
A. Glossar.....	124
B. Literaturverzeichnis.....	125

Abbildungsverzeichnis

Abbildung 1: Ablaufdiagramm der Audioproduktion.....	8
Abbildung 2: Aufnahmereihenfolge: Rhythmus, Harmonie und Gesang.....	10
Abbildung 3: Sequentieller Ablauf der Aufnahme.....	11
Abbildung 4: Vorgehen mit zwischengelegten Mischphasen.....	12
Abbildung 5: Parallele Aufzeichnungen an einer Grundversion.....	14
Abbildung 6: Ablaufdiagramm der Audioproduktion mit Demoversion.....	16
Abbildung 7: Parallele Aufnahmeprozesse mit Demoversion.....	17
Abbildung 8: Analoges Signal $f(t)$	18
Abbildung 9: Dirac Impulsfolge.....	19
Abbildung 10: Ideale Abtastung durch eine Dirac Impulsfolge.....	20
Abbildung 11: Alias Effekt.....	21
Abbildung 12: Zoom auf ein Audiosample.....	23
Abbildung 13: Quantisierungsrauschen.....	25
Abbildung 14: Effizienter Datenaustausch.....	31
Abbildung 15: MIDI Schlagzeugsur in Cubase SX3.....	43
Abbildung 16: Aufnahme einer Schlagzeug MIDI-Datei.....	44
Abbildung 17: Aufnahme einer Schlagzeug Audiodatei aus MIDI Daten.....	45
Abbildung 18: Audioschlagzeugsur erzeugt aus MIDI Daten.....	46
Abbildung 19: Quantisierung einer MIDI-Datei.....	48
Abbildung 20: Zusammengesetzte Schlagzeugsur.....	49
Abbildung 21: MIDI Datei mit Softwareklangerzeuger.....	50
Abbildung 22: Basslauf im MIDI Editor.....	55
Abbildung 23: Benutzerschnittstelle Trilogy VSTi.....	57
Abbildung 24: Benutzerschnittstelle Virtual Guitarist VSTi.....	59
Abbildung 25: E-Gitarre im MIDI Editor.....	61
Abbildung 26: Audioquellen Aufnahme.....	63
Abbildung 27: Positionierung des Mikrofons vor dem Verstärker.....	65
Abbildung 28: Spiccato Streicher im Refrain.....	69
Abbildung 29: Gesamtprojekt im Cubase Projektfenster.....	78
Abbildung 30: Gate.....	81
Abbildung 31: Kompressor.....	82
Abbildung 32: Limiter.....	84
Abbildung 33: Neutraler Equalizer.....	86
Abbildung 34: Frequenzbereiche klassischer Instrumente (aus [8]).....	87
Abbildung 35: Tiefpassfilter.....	89
Abbildung 36: Hochpassfilter.....	90
Abbildung 37: Notchfilter.....	91

Abbildung 38: Pitch Shifter.....	93
Abbildung 39: Time Stretcher.....	94
Abbildung 40: Pan und Pre-Delay.....	98
Abbildung 41: Mixereinstellungen des Beispielsongs.....	109
Abbildung 42: Dynamikschwund bei Überkompression.....	113
Abbildung 43: Zeitlicher Ablauf des Projektes.....	119

Abstract

The improvements in digital data processing had an enormous impact on most human endeavours. Ideally effort decreases and flexibility increases in relation to task completion. That is also the case within the framework of the extensive domain of audio production, where a paradigm shift occurred, which leaves an ideational value to the analog recording machines of the past at best. How (and to what extent) this new digital flexibility affects the final audio product will be shown in this thesis with the help of an example song.

But new innovations in technology raise new problems. Thanks to new technologies, musicians do not necessarily need to work at the same time or place to record their music. Instead, people around the globe collaborate, only connected by their love of music - and the Internet. Which problems occur through time zone differences, written communication, and inferior quality recordings and how these problems can be solved – sometimes only with the help of modern audio production software – will be discussed.

1. Einleitung

Die digitale Datenverarbeitung hat auf fast alle Bereiche des modernen Lebens einen entscheidenden Einfluß ausgeübt. Im Idealfalle verringern sich durch sie die Mühen des Anwenders und seine Flexibilität in Bezug auf die Erledigung seiner Arbeitsaufgabe steigt. Auch im Rahmen des umfangreichen Gebietes der Audioproduktion fand in den letzten Jahren ein Paradigmenwechsel statt, der den

klassischen, mechanischen und elektronischen Maschinen der frühen Pop-/Rockmusikaufzeichnung heutzutage im besten Falle noch ideale oder klangliche Werte zuspriht. In welchem genauen (und wie sich zeigen wird immensen) Ausmaß die Flexibilität des Musikproduzenten bei der Anwendung digitaler Audioproduktionsprozesse steigt und welchen positiven Einfluß sie auf das Endprodukt haben, wird diese Arbeit unter anderem an einem Beispiel zeigen.

Allerdings eröffnen sich stets durch neue Möglichkeiten auch neue Probleme, deren Lösung mehr von einem Musikproduzenten abverlangt, als lediglich das Mischen des musikalischen Basismaterials zu einem wohlklingenden Endprodukt. Musiker müssen beispielsweise dank der digitalen Technik und des Internets nicht mehr am selben physikalischen Ort und zur selben Zeit zusammenkommen, um die Aufnahme ihres Instrumentes zu absolvieren. Stattdessen ist es möglich, dass Musiker zusammenarbeiten, die auf verschiedenen Seiten des Erdballs lediglich durch ihre Liebe zur Musik verbunden werden – und durch das Internet. Welche Probleme sich durch Zeitverschiebung, schriftliche Kommunikation und suboptimale Aufnahmen ergeben und wie diese zufriedenstellend - auch und manchmal auch ausschließlich durch die Hilfe der Möglichkeiten in moderner Audioproduktionssoftware - lösen lassen, wird ebenso am Beispiel in Kapitel 4 erläutert.

Zur Anwendung kommen heterogene Systeme, die es letztlich der Präferenz des Musikers überlassen, für welche Software er sich entscheidet, wenn er seine Aufnahme vornehmen möchte. Das in Kapitel 4 beschriebene Beispiel beginnt und endet mit der Software Cubase SX3 der Firma Steinberg unter dem Betriebssystem Microsoft Windows Vista Business.

Kapitel 2 stellt die verschiedenen grundsätzlichen Vorgehensmodelle der

Audioproduktion gegenüber, die bei einem typischen Song der Populärmusik Anwendung finden können. Kapitel 3 erörtert die praktischen Anforderungen an die digitale Aufnahmetechnik, sowie sinnvolle Distributionswege zur Übermittlung von Audiodateien, die sich in Arbeit befinden. Kapitel 4 beschreibt die Methoden zur Instrumentenaufzeichnung in einem rein digitalen Audioproduktionsumfeld an einem Beispielsong. Darüberhinaus werden typische Audioeffekte und deren Anwendung erläutert. Die abschließende Mixtätigkeit, sowie das Mastering werden ebenso beschrieben und kritisch betrachtet. Kapitel 5 ordnet das Endprodukt gemäß seiner Qualitätskriterien in den Kontext von kommerzieller Musik ein und beleuchtet die noch offenen Probleme, die die digitale Technik derzeit noch nicht zu lösen im Stande ist. Die Probleme die sich in der Praxis aus der Abhängigkeit von Fremdmusikern in Bezug auf den Erfolg, sowie den Zeitaufwand des Projektes ergeben, werden in Kapitel 6 diskutiert. Kapitel 7 schließt mit einem persönlichen Fazit.

2. Vorgehensmodelle der Audioproduktion

Um aus einer Idee einen fertigen Song zu machen, sind gewisse Schritte nötig, die auch in der digitalen Audioproduktion unerlässlich sind. Wo genau sich die klassische Vorgehensweise, die in aller Regel als sequentiell zu betrachten ist, und ihr digitales Pendant, das einfache parallele Arbeitsabläufe ermöglicht, unterscheiden, beleuchtet dieses Kapitel.

2.1 Sequentielle Audioproduktion

Audioproduktionen können naturgemäß nur auf Grundlage eines vorhandenen musikalischen Werkes stattfinden, das gewöhnlicherweise durch einen Künstler komponiert und anschließend arrangiert wurde. Obwohl es grundsätzlich möglich wäre, einem Song auch im Nachhinein Aspekte in seiner zeitlichen Dimension hinzuzufügen, wäre es dennoch ratsam, mit der Aufnahme erst nach dem Vorhandensein eines fertigen Arrangements zu beginnen, da ein nachträgliches Einfügen von Material in aller Regel einen zu hohen Arbeitsaufwand nach sich zieht. Abbildung 1 zeigt das zugrundeliegende Ablaufdiagramm für eine Audioproduktion.

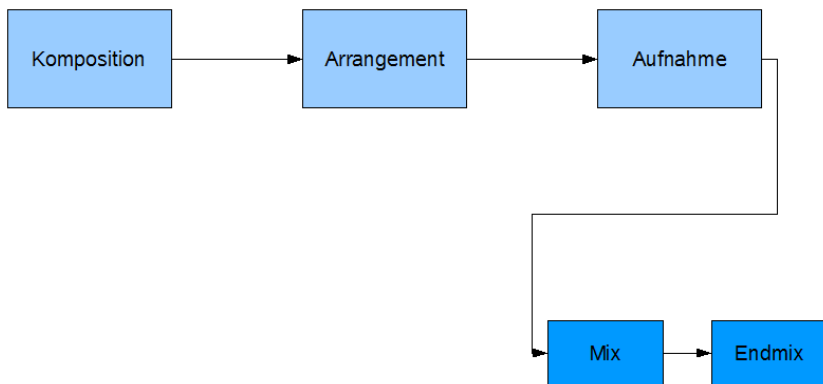


Abbildung 1: Ablaufdiagramm der Audioproduktion

Die Punkte „Mix“ und „Endmix“, auch Mastering, wurden in der Kette der Arbeitsschritte abseits dargestellt, da sie keinen produktiven, sondern in der Regel eher einen alternierenden Charakter haben, der lediglich das Klangbild der Aufnahmen im Nachhinein verändert. Im Rahmen dieser Arbeit wird auf die Punkte „Komposition“ und „Arrangement“ nicht in aller Tiefe eingegangen, da es

sich hierbei um künstlerische Aspekte der Audioproduktion handelt, die je nach Stilrichtung oder Intention mannigfaltig sein können. Stattdessen liegt das Augenmerk auf der Aufnahme und dem nachfolgenden (End)mix.

Die Aufnahme selbst unterteilt sich ihrerseits sowohl in klassischer als auch in moderner, digitaler Audioproduktion in drei kategorisierbare Unterschritte, die gewöhnlicherweise nacheinander und aufeinander aufbauend aufgezeichnet werden: Rhythmus, Harmonie und Gesang. Die hier angegebene Aufnahmereihenfolge bietet sich zwingend für jede Komposition an, da ein Sänger schlecht ohne die Harmonien der Instrumente singen kann und die Instrumente nicht ohne zugrundeliegenden Rhythmus spielen können. An dieser Stelle sei angemerkt, dass dies nicht unmöglich ist, aber aufgrund des stark ansteigenden Arbeitsaufwandes wohl kaum als empfehlenswert erachtet werden kann. Entsprechend der genannten Unterschritte ergänzt sich die oben gezeigte Abbildung 1 innerhalb des Punktes „Aufnahme“ wie in Abbildung 2 gezeigt.

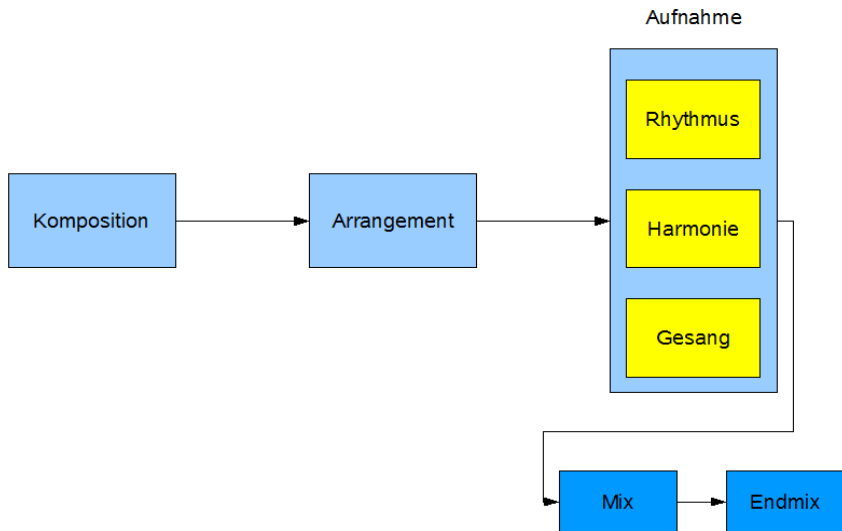


Abbildung 2: Aufnahmereihenfolge: Rhythmus, Harmonie und Gesang

Unterstellt man als Aufnahmeobjekt ein klassisches Rockstück, das sich zusammensetzt aus den Instrumenten Schlagzeug, Bass, Rhythmusgitarre, Keyboards, Sologitarren, Lead- und Backgroundgesang, dann ergibt sich als sequentielles und klassisches Audioaufnahmemodell das in nachfolgender Abbildung 3 dargestellte Vorgehen (Keyboards und ähnliche Synthesizereffekte werden im nachfolgenden Verlauf dieser Arbeit als „Ambient“, englisch für „Umgebung“, bezeichnet. Dies kann aber gegebenenfalls auch echte Instrumente wie Streicher oder Bläser beinhalten und richtet sich in erster Linie natürlich nach dem zugrundeliegenden Werk):

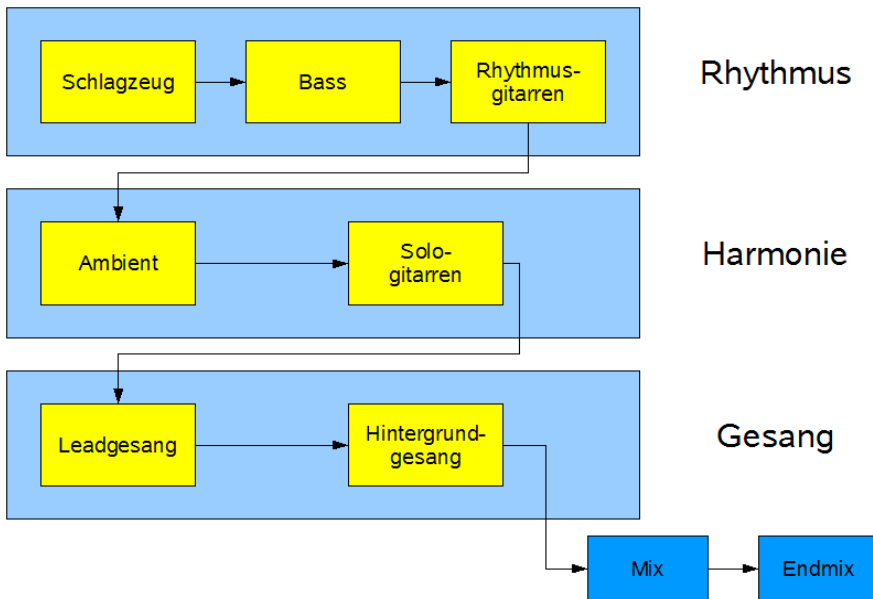


Abbildung 3: Sequentieller Ablauf der Aufnahme

Der hier gezeigte sequentielle Ablauf ist die Grundlage für ein methodisches Vorgehen, das wohl für die meisten Songs der Moderne zur Anwendung gekommen ist. Aber selbstverständlich ist es grundsätzlich möglich, gewisse Aufnahmereihenfolgen zu verändern, vor allem innerhalb der horizontal dargestellten Dimensionen Rhythmus, Harmonie und Gesang. Zusätzlich wäre es auch möglich, die bislang stets ans Ende gestellte Dimension des Mixes, also des Abmischens der zu einer bestimmten Zeit zur Verfügung stehenden Audiodaten, nach jeder erfolgten Aufzeichnung erneut vorzunehmen. Abbildung 4 verdeutlicht das:

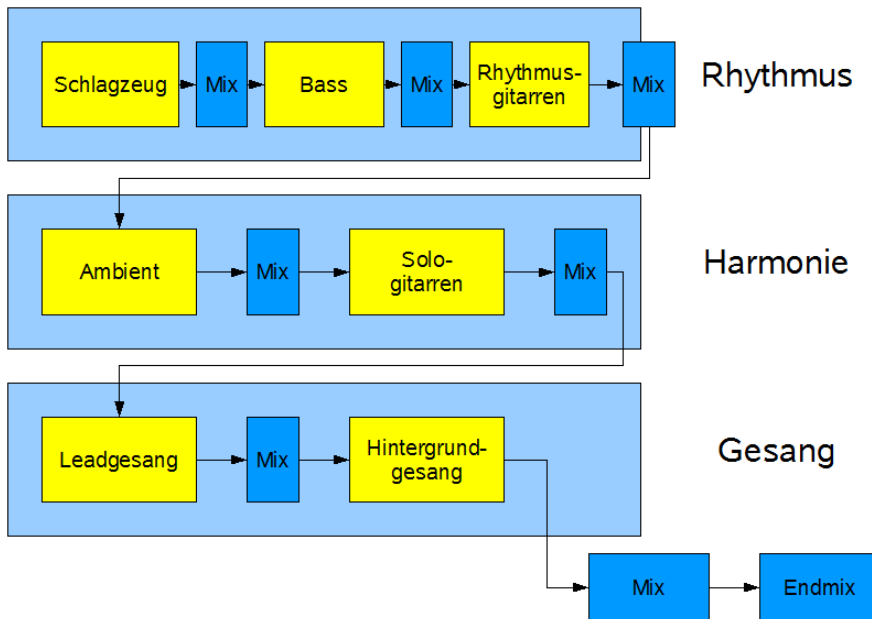


Abbildung 4: Vorgehen mit zwischengelegten Mischphasen

Ob sich dieses Vorgehen lohnt, hängt von den Zielen ab, die man damit erreichen möchte. Wenn einem Tontechniker Musiker zur Verfügung stehen, die gut mit rohen Aufnahmen umgehen können, die sich zum Teil eklatant von dem zu erwartenden Endprodukt unterscheiden, dann stellt der Zwischenmix in jedem Fall einen unnötigen Arbeitsschritt dar, da sich, wie in späteren Kapiteln noch gezeigt wird, die Hinzunahme eines einzelnen Instrumentes gewöhnlicherweise auf das gesamte (und begrenzte) Frequenzspektrum, bzw. die Verteilung der Instrumente auf demselben, auswirkt. In diesem Sinne müssen Arbeitsschritte, die in vorherigen Mischversuchen vorgenommen wurden, gegebenenfalls sogar zurückgenommen werden. Ein anschauliches Beispiel dafür ist die Rhythmusgitarre, die alleingestellt in vielen Fällen so gemischt wird, dass sie

nicht nur hohe, sondern auch tiefe Frequenzbereiche sättigt, um zu einem volleren Klang zu gelangen. Wenn nun ein Bass zusätzlich aufgezeichnet wird, der vor allem in tiefen Frequenzen agiert, dann stehen diese in Konflikt mit den tiefen Frequenzen der Rhythmusgitarre. Die Frequenzen überlagern sich und führen zu einem wenig differenzierten Klangbild, das verwaschen und unprofessionell klingt. In einem solchen Falle müßten die tiefen Frequenzen nachträglich wieder aus der Rhythmusgitarrenspur entfernt werden. Ein Zwischenmix erscheint hier wenig sinnvoll.

Wenn man aber – aus welchen Gründen auch immer – Zwischenversionen seiner Arbeit vorzeigen muß, dann sind wie in Abbildung 4 gezeigte Zwischenmixe ein notwendiges Übel. Grundsätzlich jedoch ist es aufgrund des ansonsten erhöhten Arbeitsaufwandes und auch der verringerten Dynamik ratsam, auf aufwendige Mischversuche zu verzichten, bis alle Instrumente aufgezeichnet vorliegen.

2.2 Parallele Audioproduktion

Die digitale Audioproduktion stellt im Rahmen eines bislang dargestellten sequentiellen Ablaufes keine nennenswerte Verbesserung gegenüber den klassischen Aufnahmemethoden dar. Tatsächlich schränkt auch die klassische Aufnahmetechnik ein paralleles Arbeiten nicht zwangsläufig ein. Abbildung 5 zeigt das Ablaufmodell für parallele Arbeiten ab einem geeigneten Punkt im Projekt:

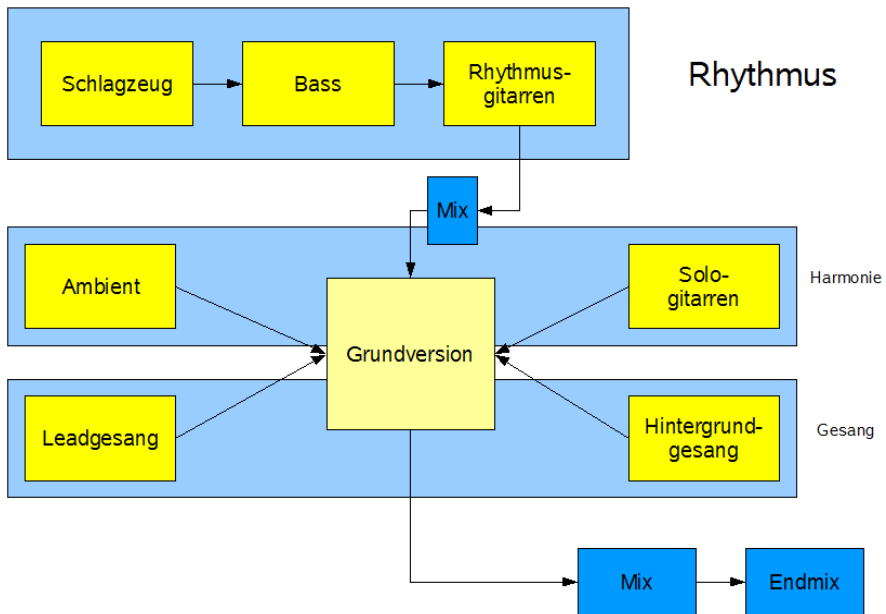


Abbildung 5: Parallele Aufzeichnungen an einer Grundversion

Nachdem mit Schlagzeug, Bass und Rhythmusgitarren eine Grundversion erstellt und gemixt wurde, können die verbliebenen Instrumente in beliebiger und voneinander unabhängiger Reihenfolge auf ihrer Basis erstellt werden. Je nach zugrundeliegendem Werk kann es auch nötig sein, dass gewisse Instrumente vor anderen aufgezeichnet werden müssen. Der Nachteil der klassischen Audioproduktion liegt hierbei im Wesentlichen auf der Distribution der Grundversion und der aufzuzeichnenden Instrumente, sowie einem präzisen Zeitgeber, der verhindern muß, dass die Grundversion oder die Aufzeichnungen in falschen Tempi abgespielt oder aufgezeichnet werden.

Besonders der parallele Ablauf in Abbildung 5 setzt voraus, dass den beteiligten Musikern bekannt ist, in welcher Art sie ihre Instrumente zur Grundversion

hinzufügen müssen, um das vom Komponisten gewünschte Endergebnis zu erreichen. Insbesondere dann, wenn sie, wie gezeigt, unabhängig voneinander arbeiten sollen bzw. müssen. Dies bedeutet, dass die Musiker entweder schon bereits beim Arrangement des Stückes involviert waren, oder eine präzise (und leider auch überaus aufwendige) schriftliche Notation bekommen, bevor sie mit ihrer Aufnahme beginnen können. Letztere Methode sichert dabei jedoch noch lange nicht das Erreichen des gewünschten Ergebnisses, da es in der Musik, vor allem beim Gesang, wesentlich mehr Nuancen gibt, als schriftlich je fixierbar wären. Vor allem im Rahmen einer verteilten und höchst heterogenen Aufnahmeumgebung, wie sie in Kapitel 4 beschrieben wird, ist es nötig, allen beteiligten Akteuren eine präzise Vorstellung vom Endprodukt zu liefern, an der sie sich stets orientieren können. Bestimmte Feinheiten können telefonisch oder gar schriftlich schlicht nicht effektiv genug kommuniziert werden. Wären der Komponist und der Musiker zur gleichen Zeit im gleichen Studio, wäre die Klarstellung bestimmter Wunschaspekte relativ einfach. Bei zeitlich und räumlich getrennten Aufnahmen muß eine Demoversion des Materials vorliegen. Entsprechend ergänzt sich Abbildung 1 um die Demoversion, wie in Abbildung 6 dargestellt:

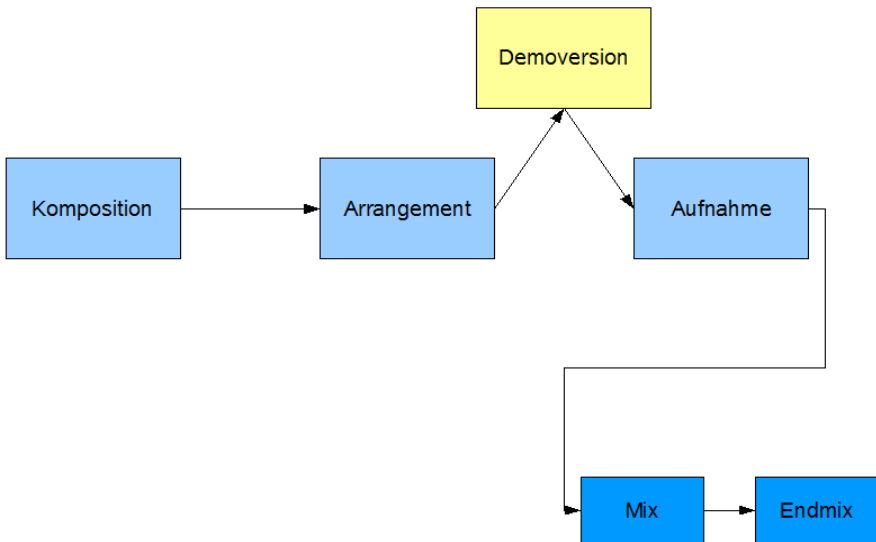


Abbildung 6: Ablaufdiagramm der Audioproduktion mit Demoversion

Die Erstellung einer Demoversion verdoppelt zwar auf den ersten Blick die Arbeit, aber beseitigt danach vor allem in räumlich und zeitlich getrennten Aufnahmeprozessen die meisten Unklarheiten, ohne eine zusätzliche Kommunikation nötig zu machen. Die Demoversion unterscheidet sich vor allem in Sachen Klangbild und Instrumentendichte von der fertigen Endversion. Sie kann als schnell angefertigter Prototyp betrachtet werden, der aber immernoch hochwertig genug ist, um bis dato unbeteiligten Dritten zu verdeutlichen, wie das Endprodukt später klingen soll. Das in Abbildung 5 gezeigte Vorgehensmodell der parallelen Audioproduktion ergänzt sich dementsprechend um die in Abbildung 7 gezeigte Version, in der die zuvor erstellte Demoversion immer als Grundlage für die Musiker hinzugezogen wurde.

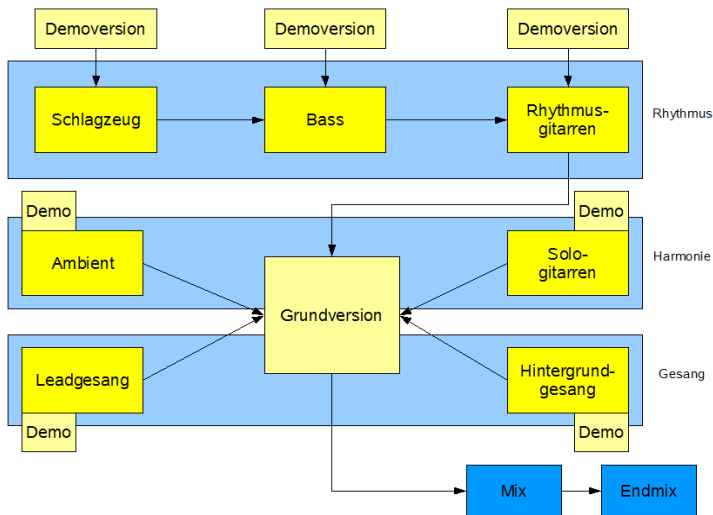


Abbildung 7: Parallele Aufnahmeprozesse mit Demoversion

Abbildung 7 verdeutlicht die Wichtigkeit und Allgegenwärtigkeit der Demoversion im eigentlichen Produktionsprozess und ist darüberhinaus die Grundlage für das später in Kapitel 4 beschriebene Beispiel.

3. Qualitätsanforderungen an die digitale Technik

Analoge Audioinformationen müssen, um im Computer verarbeitet werden zu können, digitalisiert werden. Eine Digitalisierung ist prinzipbedingt mit Informationsverlusten behaftet, da es sich um eine Diskretisierung eines sonst kontinuierlichen, analogen Signals handelt. Welche Anforderungen an das gewählte Audioformat gestellt werden und welche Einstellungen im Rahmen

einer digitalen (verteilten) Audioproduktion einen möglichst geringen Verlust an Informationen bedeutet, wird dieses Kapitel in den einzelnen Unterkapiteln darstellen. (Quelle, sofern nicht anders angegeben: [2]) Darüberhinaus wird betrachtet, welche Speicherplatzanforderungen durch den gewählten Qualitätsmaßstab nötig werden und wie diese Datenmengen vorteilhaft unter den Musikern distribuiert werden können.

3.1 Zeitdiskretisierung

Das analoge, kontinuierliche Audioquellsignal $f(t)$ (siehe Abbildung 8) wird im Analog-Digital Umsetzer der Soundkarte mit einer durch den Anwender festlegbaren Abtastrate (nachfolgend als Samplingrate bezeichnet) abgetastet.

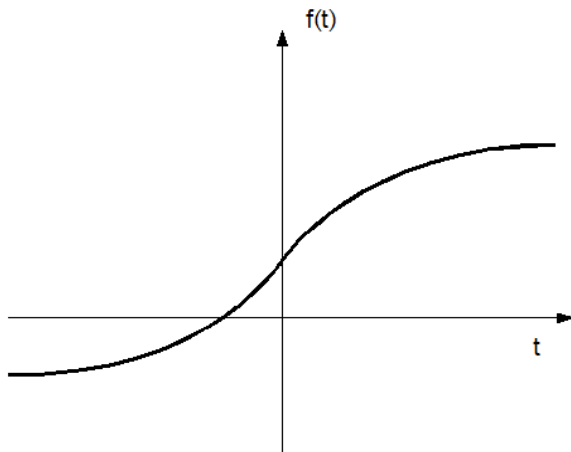


Abbildung 8: Analoges Signal $f(t)$

Mathematisch betrachtet wird das analoge Signal zu Zeitpunkten auf t mit Dirac

Impulsen im Abstand T_a multipliziert. Der Dirac Kamm ist demnach definiert als:

$$\delta_T(t) = \sum_{n=-\infty}^{+\infty} \delta(t - nT_a) \quad (1)$$

Man spricht auch von einer idealen Abtastfunktion. Abbildung 9 zeigt einen solchen Dirac Kamm (oder auch „Dirac Impulsfolge“).

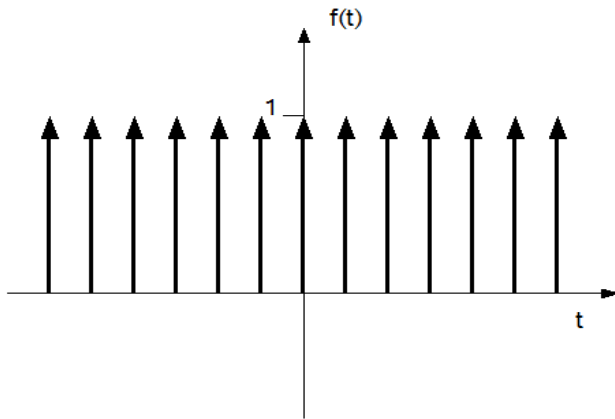


Abbildung 9: Dirac Impulsfolge

Das zeitdiskrete Signal $f_a(t)$ ergibt sich demnach als Produkt aus dem analogen Signal $f(t)$ mit der Dirac Impulsreihe $\delta_T(t)$:

$$f_a(t) = f(t) \delta_T(t) = \sum_{n=-\infty}^{+\infty} f(nT_a) \delta(t - nT_a) \quad (2)$$

Die resultierende Funktion ist, wie auch der Dirac Kamm selbst, an allen Stellen 0, die nicht diskret sind, also mit einem Dirac Impuls multipliziert wurden. Aus diesem Grunde kann man $f_a(t)$ auch als diskrete Folge beschreiben:

$$f_a[n] = \sum_{n=-\infty}^{+\infty} f(nT_a) \delta[n] \quad (3)$$

Abbildung 10 zeigt das Ergebnis der Faltung des analogen Signal mit dem Dirac Kamm.

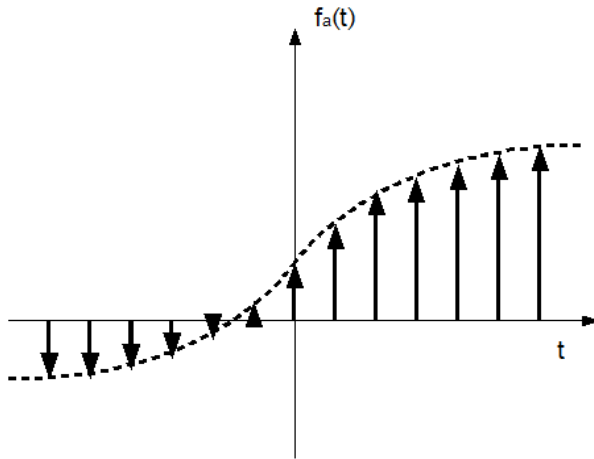


Abbildung 10: Ideale Abtastung durch eine Dirac Impulsfolge

Der Abstand der einzelnen Abtastwerte ist nun entscheidend für die Qualität des digitalisierten Audiosignales und wird in Hertz (Hz) oder Kilohertz (kHz) angegeben. Qualität kann an dieser Stelle verstanden werden als die Ähnlichkeit zwischen dem digitalen Zielsignal und dem analogen Quellsignal im Ohr eines Hörers.

Das sogenannte Nyquist-Shannon-Abtasttheorem (auch Shannonsches Abtasttheorem oder Nyquist Theorem) besagt nun, dass ein kontinuierliches, bandbegrenztes Signal mit einer Minimalfrequenz von 0 Hz und einer Maximalfrequenz von f_{max} , mit einer Frequenz größer als $2 \cdot f_{max}$ abgetastet werden

muß, damit man aus dem so erhaltenen zeitdiskreten Signal das Ursprungssignal ohne Informationsverlust exakt rekonstruieren kann (allerdings mit einem unendlich großen Aufwand), bzw. (mit endlichem Aufwand) beliebig genau approximieren kann.

Anders ausgedrückt: wenn ein analoges Signal mit einer Bandbreite von 20 kHz vorliegt (20 Hz bis 20 kHz entspricht dem vom Menschen hörbaren Bereich), dann muß dieses Signal mit mindestens 40 kHz ($2 \cdot 20 \text{ kHz}$) abgetastet werden, um sogenannte Alias Effekte zu vermeiden. Andernfalls kann es dazu kommen, dass wichtige, relativ hochfrequente Passagen im analogen Signal zwischen zwei Abtastzeitpunkten des diskreten Signals stattfinden und bei der Rekonstruktion nicht mehr auftauchen, bzw. zu einer Fehlinterpretation führen (dem besagten Alias Effekt). Abbildung 11 zeigt einen Alias Effekt (von [1]):

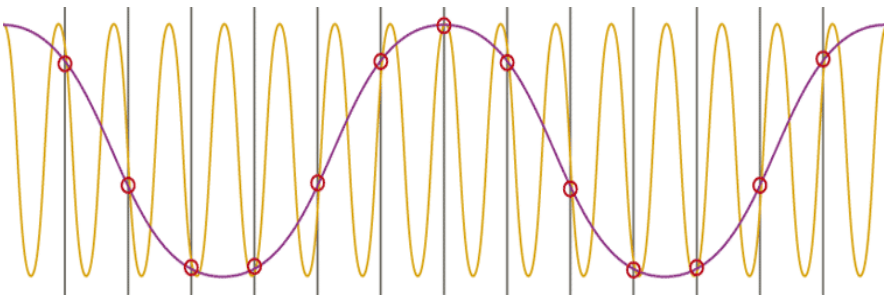


Abbildung 11: Alias Effekt

Die grauen, vertikalen Linien stellen hierbei die diskreten Abtastpunkte dar. Das eigentliche Originalsignal in gelb ist höher frequentiert als die Abtastung. Die Folge ist, dass bei der Rekonstruktion ein Alias Signal (purpur) auftritt, das mit dem Quellsignal wenig Ähnlichkeit hat, aber dieselben Abtastwerte besitzt.

Signale einer herkömmlichen Audio CD wurden mit 44,1 kHz gesampelt,

entsprechen also dem Anspruch des Nyquist Theorems zu Genüge ($44100 \text{ Hz} > 2 \cdot 20000$). Weitere übliche Samplingraten sind beispielsweise 8 kHz, die bei ISDN Telefonie benutzt werden oder 48 kHz, die bei qualitativ höherwertigen Medien (wie zum Beispiel DVD-Audio) zur Anwendung kommen.

Im Audioproduktionsbereich bietet es sich an, während des Produktionsprozesses mit einer möglichst hohen Samplingrate aufzuzeichnen, da jeder vorgenommene Schritt nach der Aufzeichnung die Qualität höchstens erhalten kann. Tatsächlich, so wird das Praxisbeispiel in Kapitel 4 zeigen, kann es darüberhinaus nötig sein, einzelne Passagen dynamisch schneller oder langsamer abspielen zu lassen, um ein musikalisches Tempo einzuhalten, wenn dies dem Quellinstrument (aus verschiedenen Gründen) nicht möglich war. Spätestens in Passagen, die nachträglich langsamer gespielt werden müssen, zahlt sich eine höher gewählte Bitrate aus. Im Rahmen des Beispiels, das in Kapitel 4 behandelt wird, wurde eine Samplingrate von 96 kHz gewählt.

Zur Veranschaulichung zeigt Abbildung 12 einen Zoom auf ein Audiosample in Cubase SX, das zu Anfang (1) in etwa 500 ms lang ist. Die invertierten Bereiche deuten jeweils die Größe des nachfolgenden Zoomfensters an.

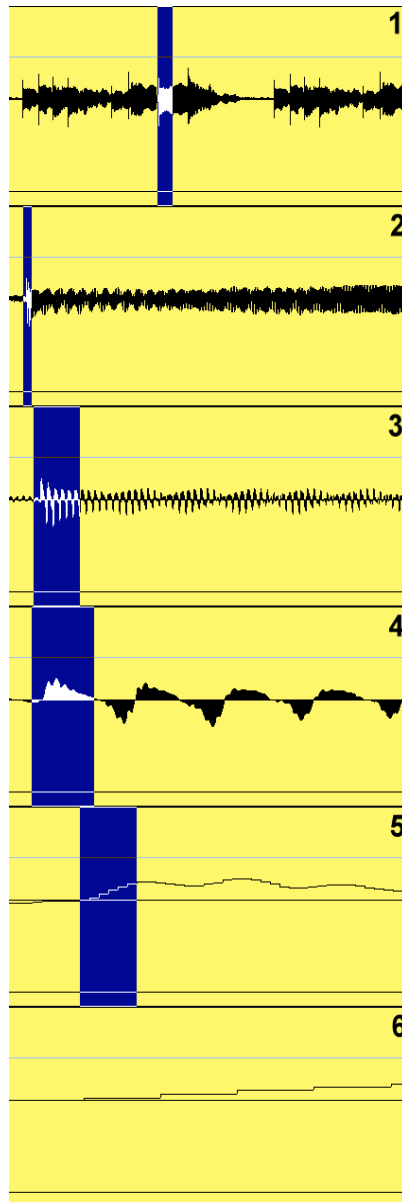


Abbildung 12: Zoom auf ein Audiosample

Zusammenfassend kann man die erforderliche Qualität bei Audioproduktionen auf mindestens 44,1 kHz festlegen, da das fertige Produkt CD Qualität besitzen soll. Deshalb sind für Zuarbeiten anderer Musiker *mindestens* 44,1 kHz nötig, um nicht unterhalb der bei Audio CDs üblichen Samplingrate zu kommen. Trotzdem gilt: je höher die Samplingrate in der Produktion ist, desto besser. Allerdings bedeutet die doppelte Samplerate auch die doppelte Datenmenge und stellt so gewisse Anforderungen an Festplatten- und Computergeschwindigkeit. Die Datenmenge errechnet sich nun aus dem Produkt aus Samplerate und Bitrate.

3.2 Amplitudenquantisierung

Zusätzlich zur Samplerate, die bislang das analoge Audiosignal lediglich in eine bestimmte Anzahl gleichgroßer Abschnitte unterteilt hat, müssen natürlich auch die entsprechenden Werte der Amplitude digital gespeichert werden. Das bedeutet, dass auch hier diskretisiert werden muß. Da die Amplitude (also die Stärke des Ausschlags der Schwingung) innerhalb des Computers ebenfalls nur durch eine endliche Anzahl von Digitalwerten ausgedrückt werden kann, entstehen Rundungsfehler, die unter Umständen bei einer Wiedergabe des Signals als sogenanntes Quantisierungsrauschen hörbar werden. Wie hoch die endliche Anzahl von Digitalwerten ist, ermittelt sich über die Bitrate, die jedem Sample eine bestimmte Anzahl Bits zugesteht.

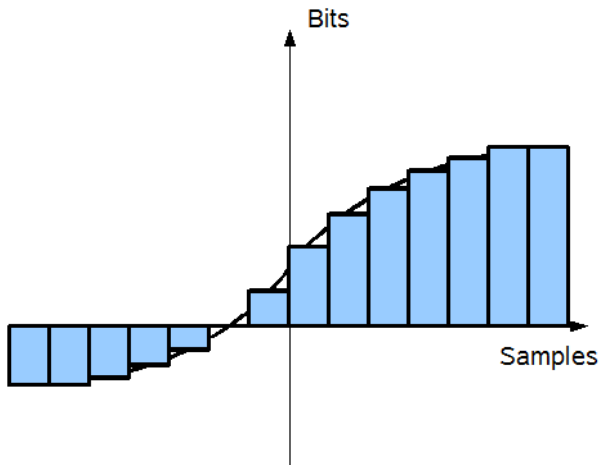


Abbildung 13: Quantisierungsrauschen

Abbildung 13 veranschaulicht die Problematik: für jedes Sample mit einer festen Breite (je breiter, desto kleiner die Samplerate und umgekehrt) muß ein digitaler Wert gefunden werden (hier in der Höhe der Samples dargestellt), der dem Originalsignal (als Linie dargestellt) am nächsten kommt. Die maximale Anzahl verschiedener Digitalwerte ermittelt sich über die Anzahl Bits, die jedem Samples zur Verfügung stehen. 8 Bit bedeutet beispielsweise, dass es pro Sample maximal $256 (= 2^8 = 8 \text{ Bit})$ verschiedene Zustände geben kann. Je nach Analogsignal kann es sein, dass diese Granularität nicht ausreicht, um das Signal auf der digitalen Basis zufriedenstellend zu reproduzieren. Entsprechend besser wären 16 Bit, denn dies würde eine doppelt so feine Auflösung bedeuten. Die qualitätsmindernden Unterschiede zwischen Analog- und Digitalsignal nennt man Quantisierungsrauschen und diese kann qualitativ durch die Signal-to-Noise Ratio angegeben werden.

Die Signal-to-Noise Ratio (SNR), zu deutsch etwa das Signal-Rausch Verhältnis,

ist definiert als das Verhältnis der vorhandenen mittleren Signalleistung σ_X zur vorhandenen mittleren Rauschleistung σ_E . Als Verhältnis von Größen gleicher Maßeinheit ist das Signal-Rausch Verhältnis dimensionslos.

$$SNR = \frac{\text{Nutzsignalleistung}}{\text{Rauschleistung}} = \frac{\sigma_X}{\sigma_E} \quad (4)$$

Da die Signalleistung aber gewöhnlicherweise weitaus größer ist, als die Rauschleistung, wird die SNR oft im logarithmischen Maßstab dargestellt, mit der Einheit Bel, bzw. Dezibel.

$$SNR_{dB} = 10 \log_{10} \left(\frac{\text{Nutzsignalleistung}}{\text{Rauschleistung}} \right) = 10 \lg \left(\frac{\sigma_X^2}{\sigma_E^2} \right) \quad (5)$$

Um nun errechnen zu können, wie groß das Rauschen (in dB) ist, das durch eine Quantisierung mit N Bit entsteht, muß man zunächst den Quantisierungsschritt q des A/D Wandlers bestimmen. Die Vollaussteuerung entspricht der doppelten Amplitude A . Ein Umsetzer mit N Bit hat somit einen Quantisierungsschritt von

$$q = \frac{2A}{2^N} \quad (6)$$

Der Quantisierungsfehler pro Schritt wird angenommen als gleichmäßig verteilt über das Intervall $-q/2$ bis $+q/2$. Damit ist die Leistungsdichte zufolge der Quantisierung gegeben als:

$$\sigma^2 = \int_{-q/2}^{q/2} e^2 P(e) de = \frac{1}{q} \int_{-q/2}^{q/2} e^{2de} = \frac{q^2}{12} \quad (7)$$

Sigma drückt das Störsignal der Quantisierung aus, welches das Rauschen verursacht. Das Eingangssignal mit Vollpegel hat eine mittlere Leistung von:

$$P_A = \frac{A^2}{2} \quad (8)$$

Diese Werte kann man nun in die SNR Gleichung einsetzen:

$$SNR = 10 \log\left(\frac{P_A}{\sigma^2}\right) = 10 \log\left(\frac{A^2/2}{q^2/12}\right) = 10 \log\left(\frac{3 \cdot 2^{2N}}{2}\right) = N \cdot 6,02 \quad (9)$$

Verkürzt ausgedrückt: um den Rauschabstand in einem Signal mit N Bit zu berechnen, benötigt man folgende Gleichung:

$$SNR = N \cdot 6,02 \text{ dB} + 1,76 \text{ dB} \quad (10)$$

Der Signal-Rausch Abstand bei 8 Bit beträgt also $8 \cdot 6,02 + 1,76 = 49,92 \text{ dB}$. Bei 16 Bit sind es entsprechend 98,08 dB und bei 24 Bit 146,24 dB. (Pro Bit also in etwa 6 dB.)

Da der Mensch bei etwa 120 bis 140 dB seine Schmerzschwelle in Bezug auf Lautstärke hat, kann man annehmen, dass bei einem mit 24 Bit gesampelten Signal in keinem Fall ein Rauschen erkennbar sein wird.

3.3 Resultierende Speicheranforderungen

Wie in den letzten beiden Kapitel gezeigt wurde, lässt sich die Qualität eines digitalisierten Audiosignales anhand der beiden Größen Sampling- und Bitrate beschreiben. Dabei gilt, dass höhere Werte grundsätzlich und prinzipbedingt eine höhere Qualität bedeuten (wobei Qualität in diesem Falle die Ähnlichkeit zwischen dem analogen Quellsignal und dem digitalen Zielsignal im Ohr des Hörers beschreibt). Wie aber ebenfalls gezeigt wurde, machen gewisse Raten ab bestimmten Punkten keinen Sinn mehr, wenn man unterstellt, dass das

Audiosignal für menschliche Ohren digitalisiert wurde, die ihrerseits gewissen Beschränkungen unterliegen. Diese sind nach [3] die maximal noch schmerzfrei zu ertragenden Lautstärke, die mit etwa 130 dB angegeben wird und die für den Menschen hörbare Bandbreite von etwa 20 Hz bis 20 kHz.

Höhere Bitraten als 24 Bit sind also unnötig, da der Abstand zwischen Nutzsignal und Quantisierungsrauschen größer als die akustische Schmerzschwelle des Menschen ist (man muß bedenken, dass 10 dB mehr aufgrund des logarithmischen Maßstabes im subjektiven menschlichen Hörempfinden einer Verdopplung der Lautstärke entspricht). In einem mit 24 Bit aufgezeichneten Audiosignal kann der Mensch demnach kein kodierungsbedingtes Rauschen mehr hören. Voraussetzung ist natürlich, dass das Audiosignal an seiner lautesten Stelle mit genau 0 dB (Vollaussteuerung) aufgezeichnet wurde. In der Praxis lässt man gewöhnlicherweise das lauteste Signal bei etwa -5 dB ausgereut (der sogenannte *Headroom* oder zu deutsch auch *Übersteuerungsreserve*). Aber auch 141 dB sind noch weit über der Schmerzschwelle anzusiedeln. Wenn nun allerdings mehrere Spuren gleichzeitig (schwach) rauschen, kann das zusammengemischte Endprodukt durchaus ein wahrnehmbares Rauschen produzieren. Entsprechend bieten sich die 24 Bit bei der Aufnahme an, um diese Rauschsumation so gering wie möglich zu halten. Wenn es nur darum geht, eine einzelne Audiospur wiederzugeben, reichen 16 Bit (98,08 dB) vollkommen aus, denn selbst die beste analoge Endstufe kann wohl kaum mehr als 100 dB Rauschabstand wiedergeben.

Ebenso ist eine Samplingrate im fertigen Endprodukt, die mehr als doppelt so groß ist, wie der vom Menschen insgesamt hörbare Bereich, aufgrund des Nyquist Theorems nicht sinnvoll. Während des Produktionsprozesses jedoch kann es sehr wohl Sinn machen, die Samplingrate höher zu wählen, um zum Beispiel unter

Umständen die Möglichkeit zu besitzen, Signale langsamer abspielen zu lassen, ohne dass darunter die Qualität deutlich hörbar leidet.

Der einzige Punkt, der gegen hohe Bit- oder Samplingraten spricht, ist der Speicherplatzbedarf und die gegebenenfalls nicht ausreichende Geschwindigkeit des Studiocomputers, der in vielen Fällen eine Vielzahl von voneinander unabhängigen Audiosignalen in Echtzeit abspielen können muß.

Eine Minute eines mit 44,1 kHz und 16 Bit kodierten Audiosignals summiert sich auf 42336000 Bit Speicherbedarf. Das sind 5292000 Byte oder 5,04 Megabyte. Eine 8 Bit Kodierung würde entsprechend genau die Hälfte, nämlich 2,52 Megabyte beanspruchen. Eine Minute 96 kHz, 24 Bit benötigen schon 16,48 Megabyte.

Ein dreiminütiger Durchschnittssong in der Produktion benötigt also bei 7 verschiedenen Instrumenten (siehe Abbildung 3 auf Seite 11) 346,07 Megabyte ($7 \cdot 3 \cdot 16,48$ MB) Speicherkapazität. Zusätzlich bedeutet dies, dass die Festplatte, auf der die Daten gespeichert sind, mindestens 1,9 Megabyte pro Sekunde (15,38 Mbit/sec) lesen können muß, um alle Spuren in Echtzeit wiederzugeben. Bei entsprechender Fragmentierung der Festplatte müssen ebenso die Zugriffszeiten entsprechend hoch sein. In der Praxis sind die Anforderungen an den Rechner meist noch höher, da es mehrere Takes zu verschiedenen Instrumenten gibt und besonders Echtzeiteffekte und VST-Instrumente (dazu später mehr) die Kapazitäten der CPU und der Festplatte beanspruchen.

Für moderne Rechner stellen diese Werte keine nennenswerte Hindernisse dar, jedoch sehr wohl für die Distribution der Zwischenprodukte zwischen beteiligten Musikern und dem Produzenten, wie in einem Szenario von Abbildung 5 auf Seite 14 angedeutet. Während es zwar sinnvoll und auch empfehlenswert ist, die

Audiodaten, die bereits im Studio hochwertig vorliegen, mit einem verlustbehafteten Format (zum Beispiel MP3 oder Ogg Vorbis) zu kodieren und zu verschicken, damit beteiligte Musiker dazu ihren Teil aufnehmen können, darf der umgekehrte Weg der Aufnahme vom Musiker zum Studio nicht verlustbehaftet sein.

Verlustbehaftete Audioformate sind qualitativ gut genug, um einem Musiker die nötige Begleitung für seine Arbeit zu liefern und ebenso das Tempo präzise vorzugeben, allerdings wäre es wohl kaum wünschenswert, eine MP3 Datei in eine Spur eines Audioprojekt zu importieren, da stets das qualitativ schlechteste Audiosample die höchstmögliche Qualität des Endproduktes determiniert. (Eine Kette ist so stark wie ihr schwächstes Glied!)

Aus diesem Grunde sollte die Zuarbeit der Musiker im WAV-, im AIFF- (beide jeweils mit unkomprimierten PCM-Rohdaten) oder im FLAC-Format (komprimiert aber nicht verlustbehaftet) ans Studio zurückgeschickt werden. Welchen verlustfreien Audiocodec man dafür verwendet, ist letztlich eine Geschmacksfrage, allerdings sollten sowohl Bitrate als auch Samplingrate dem restlichen Projekt aus oben angeführten Gründen entsprechen.

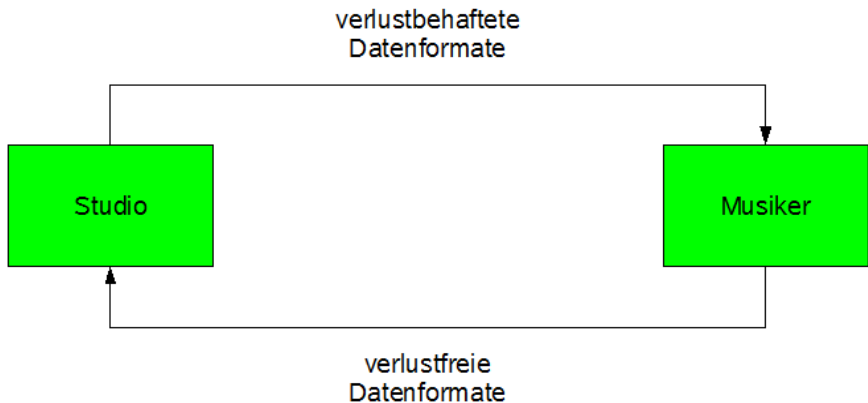


Abbildung 14: Effizienter Datenaustausch

Abbildung 14 stellt den optimalen Datenaustausch zwischen Studio und Musiker dar: während das Studio die Sounddateien, die zur Arbeit des Musikers nötig sind, im ressourcenschonenden, aber verlustbehafteten Format verschickt, versendet der Musiker seiner fertige Arbeit, die ins Endprodukt einfließen soll, in jedem Fall verlustfrei.

3.4 Sinnvolle Distributionswege und -arten für digitales Audiomaterial

Im Idealfalle benutzt ein Musiker, der seinen Teil zum Audioprojekt hinzufügen muß, dieselbe Software, dasselbe Betriebssystem und arbeitet mit demselben Audioprojekt, das dem letzten Stand aus dem bearbeitenden Studio entspricht. Dies ist in der Praxis jedoch erstens höchst selten der Fall und zweitens erscheint es auch sehr aufwendig, das gesamte Projekt zu verschicken. Jedem beteiligten

Musiker eine eigene gebrannte CD zukommen zu lassen (oder je nach Projektgröße sogar eine DVD) und diese in regelmäßigen Abständen durch neue Versionen zu aktualisieren, verschwendet nicht nur Ressourcen, sondern ist auch höchst arbeitsaufwendig. Auch wiederbeschreibbare Wechselmedien wie USB-Sticks oder Speicherkarten besitzen den Nachteil, dass sie immernoch physikalisch ausgetauscht werden müssen, d.h. entweder von Angesicht zu Angesicht überreicht oder von der gelben Post nach einigen Tagen zugestellt werden.

Von daher bietet sich das Internet als geeignet erscheinender Kanal für die Verteilung von Arbeitsversionen in besonders hohem Maße an: die Zustellung geschieht weltweit sofort und gegebenenfalls für eine kaum beschränkte Anzahl von beteiligten Musikern. Der Nachteil der Speicherplatzanforderungen, der im letzten Kapitel angedeutet wurde, besteht aber nach wie vor, wodurch Verteilungen im Anhang von E-Mails in der Regel nur bei einzelnen MP3 Dateien mit wenig Megabyte sinnvoll erscheinen. Größere Datenmengen, wie insbesondere Audiodateien, die in das Projekt einfließen sollen und von daher verlustfrei sein müssen, können nur in wenigen Fällen per E-Mail verschickt werden, da die meisten Mailprovider Beschränkungen der Anhangsgröße technisch vorgeben und darüberhinaus das SMTP (Simple Mail Transfer Protocol) aufgrund seiner Kodierung von 7-bit ASCII [4] wenig geeignet für ressourcenschonende Datendistribution ist.

Eine Alternative stellt ausreichender Webspace (sofern vorhanden) dar, der allerdings vor unbefugten Zugriffen geschützt sein sollte. Die Daten können dann per FTP auf einen Server hochgeladen und zeitversetzt vom Musiker, bzw. dem Studio heruntergeladen werden.

Wenn man bereit ist, seine Daten einem anderen Unternehmen anzuvertrauen, bieten sich die zahlreichen Sharehoster¹ (auch bekannt als *One-Click Hoster* oder *Filehoster*) an, bei denen man gewisse Datenmengen ohne vorherige Anmeldung hochladen und diese danach über einen (prinzipiell geheimen) URL wieder beziehen kann. Diese müßte dann lediglich mit den Musikern oder dem Studio kommuniziert werden.

Der im Jahre 2007 von Microsoft gestartete kostenlose Dienst *Skydrive*² bietet seinen Nutzern 5 Gigabyte Speicherplatz. Die hochgeladenen Daten können entweder für Freunde oder für alle freigegeben werden. Aufgrund der Speicherkapazität und dem Schutz der Daten auf Benutzernamenbasis erscheint dieser Dienst derzeit als besonders geeignet für asynchronen Austausch großer Datenmengen (5 GB entsprechen immerhin mehr als 5 Stunden unkomprimiertem Audiomaterial, das mit 96 kHz und 24 Bit gesampelt ist).

Als synchrone Variante wäre an dieser Stelle noch die üblicherweise in fast allen Clients enthaltene Dateitransfermöglichkeit der Instant Messenger³ zu nennen. Hier werden die Daten direkt zwischen zwei Parteien ausgetauscht, ohne sie auf einem Server zwischenzulagern. Der Nachteil besteht hauptsächlich in der Funktionalität, die oft durch Firewall- oder Routereinstellungen beeinflusst wird und in der Notwendigkeit, den Start des Transfers mit dem Empfänger abzusprechen.

Es ist zu erwarten, dass in Zukunft (und mit steigender Bandbreite für Internetanschlüsse) noch weitere komfortable Möglichkeiten zum direkten Datenaustausch geschaffen werden. Für die im Rahmen der Audioproduktion

1 RapidShare.com, Megaupload.com, etc.

2 <http://skydrive.live.com/>

3 ICQ, Google Talk, AIM, Live Messenger, etc.

nötigen Schritte sind heutige Techniken aber schon mehr als geeignet, wie auch das folgende Kapitel zeigen wird.

4. Flexibilität in der Praxis anhand eines Beispiels

In welchem Umfang die digitale Technik (nachträglichen) Einfluß auf den Klang eines Musikstückes hat und wie immens die Flexibilität bei dessen Produktion ist, wird das folgende Kapitel darlegen. Als roter Faden dient die Erstellung und Aufnahme eines Beispielsongs namens *Chapter 2 – The Quest for Gold* aus der Metal Oper *Enuma Elish*⁴.

4.1 Zugrundeliegendes Werk

Das zugrundeliegende Werk, anhand dessen in den nachfolgenden Kapiteln eine rein digitale Audioproduktion dargestellt wird, ist ein Song eines Konzeptalbums, das der Autor dieser Arbeit seit dem Sommer 2006 komponiert. Dieses Konzeptalbum behandelt in zwölf Songs (referenziert als Kapitel) den babylonischen Schöpfungsepos, der ca. im 8. Jahrhundert vor Christi Geburt auf 7 Tontafeln in Keilschrift niedergeschrieben wurde. Der Name dieses Epos ist „Enûma elîsch“ und diente – festzumachen anhand seiner deutlichen Parallelen – einst den Autoren der hebräischen Schöpfungsgeschichte Genesis als Vorlage. Anders als die Bibel unterscheidet aber die Enûma elîsch verschiedene agierende

4 Enuma Elish – The Metal Opera – <http://www.enuma-elish.de/>

Gottheiten und nennt darüberhinaus die Beweggründe für ihre Entscheidung, den Menschen zu schaffen. Diese Geschichte bildet den Rahmen für die 12 Kapitel der Metal Oper „Enuma Elish“. Die 12 Kapitel sind im Einzelnen:

1. A Blue Planet
2. The Quest for Gold
3. A New Order
4. Rebellion
5. Let Us Make Man
6. Abduction to Paradise
7. The Gift of Sapience
8. Independence
9. Scavenging
10. Tears of a God
11. Sodom and Gomorrah
12. Creation Forsaken

Die Handlung wird in der Oper ausschließlich durch gesungene Dialoge oder gegebenenfalls Monologe fortgetragen, was eine Notwendigkeit für verschiedene Sänger nach sich zieht. Kapitel 1 beschreibt das Anfangsstadium des Planeten Erde, als er noch ausschließlich mit Wasser bedeckt ist. Kapitel 2 behandelt die Notwendigkeit für ein Volk vom Planeten Nibiru, Gold zu fördern, um tödliche Strahlung von ihrem Planeten fernzuhalten. Darum schickt König Anu seinen Sohn Enki und 60 Männer zur Erde. Diese arbeiten nicht schnell genug,

weswegen Enkis sehr viel strengerer Halbbruder Enlil zur Erde gesandt wird. Dieses Geschehnis wird in Kapitel 3 beschrieben. Kapitel 4 beschreibt die Rebellion der ausgebeuteten Arbeiter, die Enki schließlich als Geisel nehmen, um ihre Forderung nach gerechteren Arbeitsbedingungen durchzusetzen. In Kapitel 5 spricht Enki dann auf einer Versammlung die berühmten biblischen Worte „Lasst uns den Menschen machen“ (einen primitiven Arbeiter). Zusammen mit seiner Halbschwester Nindmah erschafft Enki einen Hybriden aus der DNA seines Volkes und einem auf der Erde lebenden Menschenaffen. Fortan unterstützt der neu geschaffene Mensch die Goldschürfarbeiten. Kapitel 6 beschreibt die mißbräuchliche Behandlung des primitiven Arbeiters als Sklave durch Enlil in seinem Garten (Eden). Da Enki dieses Verhalten nicht gutheissen kann, sorgt er dafür, dass der bis dato unfruchtbare Hybrid sich fortpflanzen kann und „Erkenntnis“ erlangt (in der Bibel bedeutet die Vokabel „Erkenntnis“ Geschlechtverkehr). Dieses Ereignis liegt dem Kapitel 7 zugrunde. Kapitel 8 beschreibt, wie Enlil voller Zorn über Enkis Tat seine Sklaven in Afrika aussetzt und sich selbst überlässt. Kapitel 9 beschreibt die biblische Sintflut, die über die Erde kommt und das Versprechen der Außerirdischen, den Menschen, der verkommen und böse ist und sich viel zu stark vermehrt, nicht vor der Flut zu warnen. Enki jedoch bricht sein Versprechen und warnt den menschlichen Noah, der daraufhin die Flut überlebt. Kapitel 10 ist ein Monolog Enkis, in dem ihn sein Gewissen über die Behandlung des Menschen als Instrumentarium plagt. Kapitel 11 beschreibt schließlich nochmal die ganze Verkommenheit und Boshaftigkeit des Menschen und seine Rolle als Lebewesen ohne ökologische Nische auf diesem Planeten. In Kapitel 12 verlassen die Außerirdischen die Erde endgültig und lassen den Menschen zurück. Nur Religionen und Priester erzählen noch vage Geschichten von den Ereignissen der fernen menschlichen Vergangenheit und

von den Göttern, die einst auf der Erde wandelten und den Menschen schufen.

Entsprechend dieser Rahmenhandlung sind folgende Gesangsrollen geplant::

- Enki (*Sohn von König Anu, Fürsprecher der Menschen*)
- Enlil (*pflichtbewußter Sohn von König Anu, der den Menschen hasst*)
- Nindmah (*Tochter von König Anu, medizinisch bewandert, für das Klonen des Menschen verantwortlich*)
- Adamus (*der Hybrid, der im Garten Eden arbeiten muß*)
- Noah (*der Hybrid, der vor der Flut gewarnt wird*)
- Cherub 1 (*eine weitere Person aus Enkis Volk*)
- Cherub 2 (*noch eine Person*)
- Priester (*eine Rolle im letzten Kapitel 12*)

Um die digitalen Audioproduktionsprozesse darzustellen, wird in den folgenden Kapiteln die Entstehung des zweiten Kapitels „*The Quest for Gold*“ beschrieben.

Die Dialoge für diesen Song lesen sich wie folgt:

Cherub 1: A cold wind blows; Rain keeps falling down and down and down again

Cherub 2: We can't deny; That all we know and all we loved will die

Cherub 1: I'm afraid; Our king has lost his last chance as he failed

Cherub 2: Hail the king! The old king bears a new hope as he

Both: Found what shines so bright; Will save our lives; Hail king Anu! Hail king Anu! Hail King Anu!

Chorus, alle: The quest for gold; To planet earth; What are its means? What is it worth?

But for my home; I'll be bold; I'll give my life; - I've been told

Enki: I need men to join my ride; I need men to stand beside; Join me on my way to earth; Join to save our home; The journey's long; The voyage hard; I'll miss my home; I am to guard; It's sparkle seems; So far away; I'll be back; One precious day

Enki: I see this planet dying, it's end has come to sight; We wonder why this happened, why it rains all the night; Yes, father I will go and try to save all of our kind; I'll lead these sixty men and I'll leave my home behind; If gold is the answer, I'll bring it here; To stop the radiation, while death is near; Our next stop is named Eridu

Cherubs: Hail king Anu! Hail king Anu! Hail king Anu!

Chorus

Solo

Chorus

Ki you ancient Tiamat; Mutilated by impact; Carry in your oceans sea; The remedy to our disease

Auf dieser Grundlage lassen sich 3 verschiedene Gesangsrollen identifizieren, nämlich Cherub 1, Cherub 2 und Enki. Der Song hat eine Geschwindigkeit von 145 BPM (Beats pro Minute) und wird stilistisch als Power Metal Song kategorisiert. Die Musikrichtung des Power Metal zeichnet für gewöhnlich aus

durch relativ hohes Tempo, eingängige Melodien und verhältnismäßig hohen Gesang aus (Tenor bis hin zu Sopran im Falsett). Der Klang ist allgemein recht voluminös und wird oftmals mit zusätzlichen Keyboards aufgefüllt. Hinzu kommen eigentlich immer stark verzerrte Gitarren und anspruchsvolle Soli.

Das Arrangement des Songs „*The Quest for Gold*“ zielt auf die komplette in Kapitel 2 dieser Arbeit aufgeführte Palette von Instrumenten ab, als da wären: Schlagzeug, Bass, Rhythmusgitarre, Streicher, Sologitarren, 3 Sänger, sowie ein Chor. Die Tonlage des Songs ist A#-Moll, was dadurch zustande kommt, dass die Gitarren um einen Ganztonschritt heruntergestimmt wurden, um das tiefe D spielen zu können. Ursprünglich komponiert wurde der Song mit normal gestimmten Gitarren demnach 2 Halbtonschritt höher, in C-Moll.

Der Ablauf des Songs lässt sich textuell folgendermaßen zusammenfassen: der Song beginnt mit einem Schlagzeugsolo. Danach steigen Bass, Rhythmus- und Sologitarre zusammen mit dem Schlagzeugtakt ein. Es folgt der Gesangspart von Cherub 1 und Cherub 2 zusammen mit dem Bass, der Rhythmusgitarre und dem Schlagzeug. Nach einem kurzen Chorteil bei gleichbleibender Instrumentierung startet der Refrain, der von Enki gesungen wird und von den Cheruben begleitet. Im Hintergrund spielen dazu Streicher Triolen, um dem Song Dramatik zu verleihen. Zusätzlich spielt kaum hörbar eine Sologitarre eine Melodie. Danach folgen Breaks, die von Bass und Gitarre begleitet werden, dazu singt Enki alleine, bis zu einem refrainartigen Teil, der von einer Sologitarre begleitet wird. Dann folgt eine Bridge mit Triolen, die ebenfalls von Enki alleine gesungen wird und nach einem weiteren kurzen Chorteil wird erneut in den Refrain übergegangen. Gefolgt wird dieser von dem Solo, das mit dem Ende der schon bekannten Intro endet. Schlußendlich wird der Refrain sechsmal wiederholt, wobei ab Refrain 2 eine weitere Stimme hinzukommt. Ab Refrain 4 werden die Instrumente

ausgeblendet, so dass nur der Gesang und ein Chor stehenbleiben. Refrain 6 enthält nicht mehr die zusätzliche Stimme. Damit endet das Stück.

Mit der Band *Powerslave* wurde auf Grundlage dieses Arrangements eine Demoversion erstellt, die später allen Musikern als Anhaltspunkt diene (vgl. Abbildung 7). Diese Demoversion wurde mit ausschließlich natürlichen Instrumenten eingespielt und später vom Autor dieser Arbeit und dem Bassisten eingesungen.

4.2 Aufnahme

In diesem Kapitel wird die Aufnahme der verschiedenen Instrumente für das zugrundeliegende Werk beschrieben. Unter Aufnahme versteht sich die Übergabe der einzelnen Instrumentenspielinformationen an das Programm Cubase SX3, das auf dem Studiocomputer läuft. Grundsätzlich wird zwischen zwei verschiedenen Arten von Spielinformationen unterschieden: Audiodaten in Form von unkomprimierten PCM-Rohdaten, die wie in Kapitel 3.1 und 3.2 beschrieben gewonnen werden, sowie MIDI Daten, die Musik lediglich deskriptiv speichern und die Klangerzeugung jeweiliger dafür geeigneter Software oder Hardware überlässt. Da letztere Option eine immense Flexibilität mit sich bringt, wurde versucht, so oft wie möglich mit MIDI-Daten zu arbeiten, statt mit Audiodaten.

4.2.1 E-Schlagzeug

Die Aufnahme eines natürlichen Schlagzeuges beansprucht hochgesteckte

Rahmenbedingungen, um einen akzeptablen Klang zu erzeugen, der sich mit anderen professionellen Musikproduktionen messen kann. Zuerst muß das Schlagzeug selbst entsprechend hochwertig sein, um überhaupt einen guten und satten Sound erzeugen zu können. Dann muß das Schlagzeug in einem angemessenen Raum aufgezeichnet werden, der möglichst reflektionsarm sein sollte. Dies hängt mit der Tatsache zusammen, dass Hall ohne Probleme der Aufnahme nachträglich hinzugefügt werden kann, aber bereits vorhandener Hall, der durch die Aufnahmeumgebung bedingt ist, kaum zu entfernen ist. Außerdem muß jede einzelne Trommel und jedes einzelne Becken mit hochwertigen Mikrofonen aufgezeichnet werden, die in einem entsprechend großen Mischpult auflaufen, um dort entweder direkt gemischt zu werden oder als Einzelspuren an ein entsprechend fähiges Aufzeichnungsgerät weitergegeben werden. Selbst wenn diese sehr aufwendigen Rahmenbedingungen gegeben sind, bedeutet dies noch nicht, dass die Aufnahme eines natürlichen Schlagzeuges befriedigend vollzogen wurde. Vielmehr sind die exakten Mikrofonpositionierungen in Relation zu ihren Schallquellen, sowie deren Ausrichtung in großem Maße entscheidend für den Klang. Eine wirklich gute Aufnahme eines natürlichen Schlagzeuges ist in der Regel erst nach vielen Jahren Erfahrung und mit entsprechend geschulten Ohren machbar.

Aus diesem Grunde wurde bei der Produktion des Songs „*The Quest for Gold*“ auf ein elektronisches Schlagzeug zurückgegriffen, das neben dem heutzutage sehr guten Klang auch noch andere Vorteile mit sich bringt, die im Rahmen dieses Kapitels noch angesprochen werden und ein hohes Maß an Flexibilität ermöglichen. Häger[5] drückt den Unterschied zwischen natürlichem Schlagzeug und E-Drums folgendermaßen aus:

„Im Gegensatz zum natürlichen Schlagzeug entstehen bei einem Edrum die

Klänge nicht unmittelbar dort wo der Stock die jeweilige Spielfläche berührt. Vielmehr wird bei dem elektrischen Schlagzeug über spezielle Sensoren (so genannte Trigger) aufgezeichnet, wann, wie stark und wo welche Trommel (Edrumpads) angeschlagen wurde. Diese Information wird an ein so genanntes Steuerteil übermittelt, in dem aus einer Klangbibliothek ein abgespeicherter Klang geladen und abgespielt wird. Edrumpads und Edrumsteyerteile sind in sehr unterschiedlicher Qualität erhältlich und bieten deswegen sehr unterschiedliche Möglichkeiten.“

Diese Beschreibung erinnert stark an das MIDI-Format, das ebenfalls lediglich Informationen über Tonhöhe, Anschlagsstärke und Anschlagsdauer (sowie viele, viele weitere Instrumentenspieleigenarten) speichert und die eigentliche Klangerzeugung der jeweiligen Software auf dem entsprechenden Rechner überlässt. (Dies ist auch der Grund, wieso viele Menschen mit dem MIDI Format kaum gut klingende Pieptöne assoziieren: der mit den meisten Betriebssystemen mitgelieferte MIDI Synthesizer ist von einer sehr schlechten Qualität.)

Tatsächlich ist es mit modernen E-Drum Steuerteilen möglich, diese Daten direkt an ein Audioaufzeichnungsprogramm wie Cubase zu übergeben. Abbildung 15 zeigt einen Ausschnitt der MIDI Schlagzeugspur in Cubase SX 3.

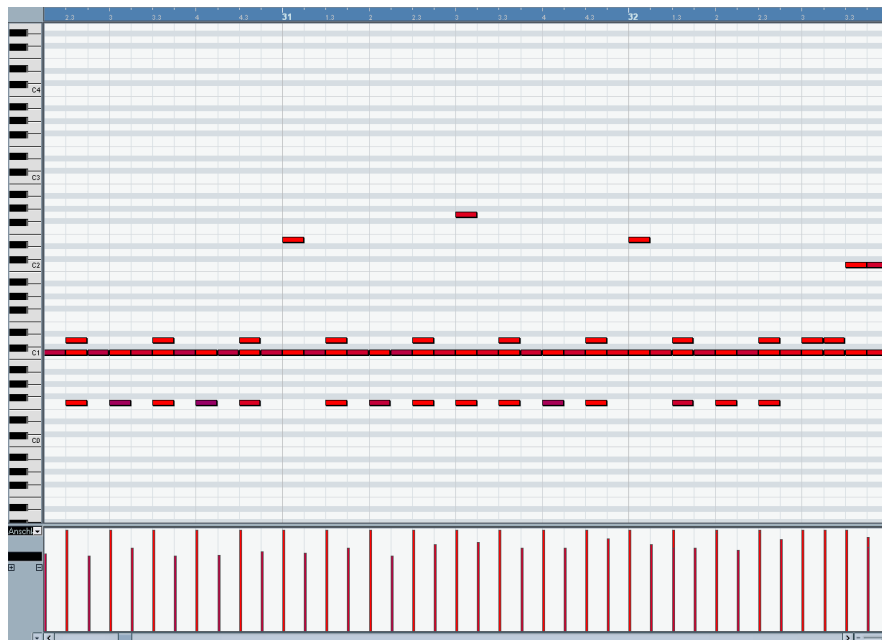


Abbildung 15: MIDI Schlagzeugspur in Cubase SX3

Im Hauptfenster des Cubase MIDI Editors sind die Aktivierungen (und aufgrund ihrer jeweiligen Länge ihre Deaktivierungen) der verschiedenen Trommeln in ihrem zeitlichen Verlauf zu erkennen. Verschiedene Trommeln korrespondieren mit verschiedenen Tonhöhen, wie die links angezeigte Klaviatur errahnen lässt. Beispielsweise liegt die Basedrum, die in diesem Beispiel (in der Strophe von „*The Quest for Gold*“) in durchgängigen Sechzehnteln gespielt wird, auf dem Ton C1 der Klaviatur und die Snare auf D1. Die jeweilige Stärke des Anschlages wird über die Farbe dargestellt: rot ist ein harter Anschlag. Je blauer, desto schwächer. Die Anschlagstärke, auch als Velocity bezeichnet, ist ebenso im unteren Fenster mit Balken dargestellt, die neben der Farbinformation auch durch

ihre Höhe Aufschluß über die Anschlagstärke geben.

In einer MIDI Datei wären nun beispielsweise für Takt 31, Beat 1.1 lediglich eine Aktivierung der Basedrum (C1) und des Beckens (E2) gespeichert, sowie deren Deaktivierung bei Beginn von Beat 1.2. Zusätzlich würde die MIDI Datei noch Informationen über die Anschlagstärke besitzen. Diese Informationen werden später an einen Klangerzeuger geschickt, der das Schlagzeug dann hörbar macht.

In vielerlei Hinsicht ähnelt das MIDI Format also den Lochstreifen aus Papier, die bereits im 19. Jahrhundert Kunstspielklaviere (zum Beispiel Pianolas) zum Erklängen brachten.

Abbildung 16 zeigt das Prinzip der Aufzeichnung einer MIDI-Datei in einem Hostprogramm (in unserem Falle Cubase SX3) mit Hilfe eines E-Drumsteuerteils:

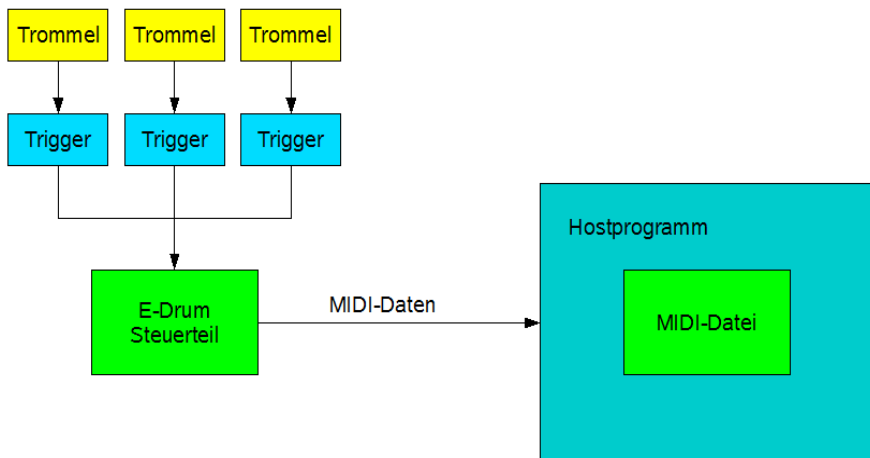


Abbildung 16: Aufnahme einer Schlagzeug MIDI-Datei

Das Schlagzeug, das aus verschiedenen Trommeln besteht, wird von einem Schlagzeuger gespielt. Die Anschlaginformationen werden von Triggern

aufgezeichnet, die an ein E-Drum Steuerteil angeschlossen sind und dort ausgewertet werden. Das Steuerteil erzeugt MIDI-Daten, die das Hostprogramm entgegennimmt und als MIDI-Datei auf die Festplatte schreibt.

Diese MIDI-Datei muß später nun nicht zwangsläufig von einem Softwareklangerzeuger auf dem Computer in hörbare Toninformationen umgewandelt werden. In der Regel ist der Tonerzeuger in einem guten E-Drumsteuerteil sehr hochwertig, da er auch bei Liveauftritten zur Anwendung kommen muß. Entsprechend kann man nun die zufriedenstellend bearbeitete MIDI-Datei später wieder an das E-Drumsteuerteil zurückschicken und so den eingebauten Klangerzeuger benutzen, um das Schlagzeugspiel hörbar zu machen. Abbildung 17 zeigt das Prinzip:

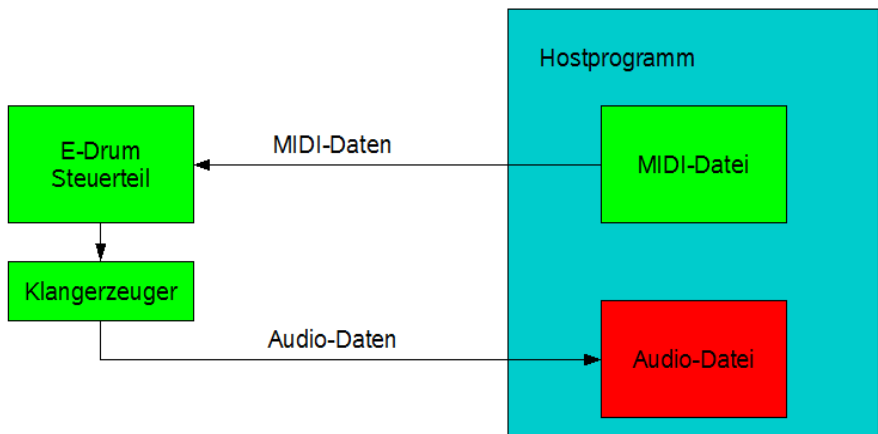


Abbildung 17: Aufnahme einer Schlagzeug Audiodatei aus MIDI Daten

Das Hostprogramm sendet in diesem Fall nicht nur die MIDI Daten, sondern zeichnet auch die durch das E-Drumsteuerteil erzeugten Klänge auf eine eigene Audiospur auf. Dies muß deswegen gemacht werden, weil nicht jeder beteiligte

Musiker ein E-Drumsteuerteil zur Verfügung stehen hat und Softwareklangerzeuger sich zu sehr in ihrem Klang unterscheiden können. Abbildung 18 zeigt die Audiodaten im Hostprogramm, als Ergebnis aus MIDI-Datei (Abbildung 15) und Klangerzeugung durch das E-Drumsteuerteil:



Abbildung 18: Audioschlagzeugspur erzeugt aus MIDI Daten

Die MIDI-Daten, die bereits im Studiorechner gespeichert waren, wurden somit in Audiodaten überführt, die auf jedem Musikcomputer gleich klingen. Dies wäre bei der MIDI-Datei aufgrund der unterschiedlichen Schlagzeugklangerzeuger nicht der Fall gewesen.

Die große Flexibilität der digitalen Audioproduktion liegt nun insbesondere in den dynamischen Eigenschaften der MIDI-Datei, die niemals einen fixen Status repräsentiert, sondern jederzeit mannigfaltig manipuliert werden kann. Insbesondere können in MIDI-Dateien explizit falsche Töne nachträglich (das heißt nach der Aufnahme) mit der Maus geradegerückt oder Teile, die nicht genau

im Tempo liegen, angepasst werden. Dies bedeutet nicht nur weniger Fehlversuche bei den Aufnahmen (Audioaufnahmen müssen bei geringen Fehlern abgebrochen werden, da es kaum Möglichkeiten zur nachträglichen Korrektur gibt), sondern stellt auch weniger Ansprüche an die beteiligten Musiker.

Die Schlagzeugspur im Song „*The Quest For Gold*“ wurde so zum Beispiel gar nicht im Tempo 145 BPM eingespielt, sondern bei Tempo 110, was eine immense Erleichterung für den Schlagzeuger bedeutete. Die Geschwindigkeit wurde dann erst nachträglich erhöht, was auf den Klang der MIDI-Datei keinerlei Einfluß hat.

Trotz langsamerer Einspielgeschwindigkeit sind Menschen nicht präzise genug, um mehrere Minuten mit einem fixen BPM-Wert einzuspielen. Das ist auch der Fall, wenn man seine Aufnahme durch einen Klick (ein Metronom) begleiten lässt. Tatsächlich gibt es immer kleine Variationen im Tempo, die für den Menschen kaum hörbar sind, aber in einer MIDI Notation (zumindest optisch) auffallen. Diese Ausreisser kann man entweder von Hand auf die richtigen Taktstriche ziehen oder man quantisiert die Aufnahme, d.h. der Computer rückt die Noten automatisch auf die nächstliegenden Taktstriche. In Cubase SX3 startet man diese Funktion mit der Taste „Q“ im MIDI-Editor. Auf Grundlage der beiden Werte „Quantisierung“ und „Längenquantisierung“, die der Tontechniker vorgeben muß, kann der Computer so eine Abschätzung über die eigentlichen Positionen der Ereignisse innerhalb des Taktes vornehmen. Abbildung 19 zeigt eine MIDI-Datei vor und nach der Quantisierung (Quantisierung und Längenquantisierung stehen in diesem besonderen Fall beide auf 1/16):

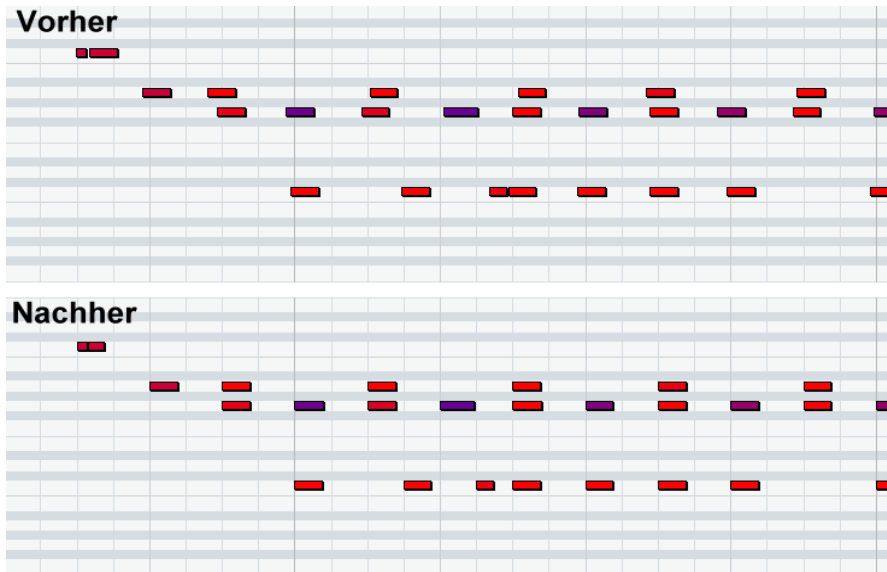


Abbildung 19: Quantisierung einer MIDI-Datei

Erkennbar sind in der MIDI-Datei leichte zeitliche Variationen, die in der quantisierten Version nicht mehr auftreten. Es handelt sich also nach Anwendung einer Quantisierung um eine vom Timing her perfekte Aufnahme, die eine ausreichende taktgebende Grundlage für alle anderen Instrumente darstellt.

In Kapitel 4.1 wurde der Ablauf des Songs umrissen. Es bietet sich an, nicht den kompletten Song in einem Stück aufzunehmen, sondern immer nur Teile, die danach entsprechend dem Ablauf zusammengesetzt werden. Der Vorteil liegt in erster Linie in der Ersparnis des Arbeitsaufwandes. Außerdem müssen wiederkehrende Songpassagen, wie zum Beispiel der Chorus, nicht 9 mal gespielt werden, sondern nur einmal. Danach werden sie schlicht an die Stellen kopiert, an denen sie noch gebraucht werden. Abbildung 20 zeigt die zusammengesetzte Schlagzeug Midispur.

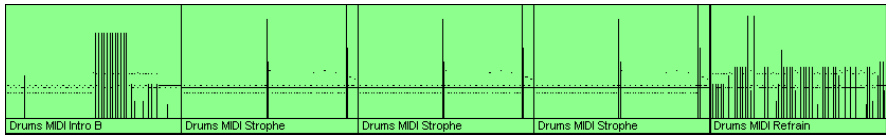


Abbildung 20: Zusammengesetzte Schlagzeugspur

Diese Sicht auf die verschiedenen MIDI-Dateien bietet das Cubase Projektfenster, in dem die verschiedenen Spuren untereinander angezeigt werden. Man erkennt das Ende der Intro, die Strophe, die 3 mal kopiert wurde und den Chorus/Refrain. Die Kopie von Songpassagen an andere Stellen ist nur dann möglich, wenn diese Passagen exakt im vorher festgelegten Tempo eingespielt wurden, so dass Übergänge flüssig genug sind, um nicht aufzufallen.

In diesem Kapitel wurde das Prinzip gezeigt, wie mit einem elektrischen Schlagzeug eine MIDI-Datei im Computer erzeugt wird und wie diese nachbearbeitet und dupliziert wird, um die komplette Trommelspur zu erhalten. Ebenfalls wurde gezeigt, nach welchem Prinzip aus dieser MIDI-Datei eine Audiodatei erzeugt wird, mit Hilfe des Klangerzeugers in einem E-Drum Steuerteil. Im nächsten Kapitel wird der Bass behandelt, der ohne externe Klangerzeuger auskommt - und ohne Bass.

4.2.2 MIDI Bass

Steht kein externer Klangerzeuger zur Verfügung, müssen die MIDI Daten auf andere Weise hörbar gemacht werden. Wie bereits erwähnt, sind die MIDI Synthesizer, die bei Betriebssystemen mitgeliefert werden, für gewöhnlich von sehr niedriger Qualität. Qualität versteht sich in diesem Fall als akustische

Ähnlichkeit zwischen Klanggeber und dem natürlichen Instrument in einer optimalen Aufnahmeumgebung.

Für quasi alle Instrumente bietet der Markt Klangerzeuger auf Softwarebasis, die über das VST-Protokoll mit Hostprogrammen wie Cubase kommunizieren können. VST steht für *Virtual Studio Technology* und ist eine von der Firma Steinberg im Jahre 1996 veröffentlichte[6] Schnittstelle zur Kommunikation mit den Klangerzeugern, den sogenannten VST-Instrumenten (VSTi). Neben Klangerzeugung auf Grundlage einer MIDI-Datei können VSTi auch Audioeffekte aller Art sein, die in Software nachgebildet wurden. Dazu in späteren Kapiteln mehr.

Abbildung 21 zeigt den Signalweg innerhalb eines Hostprogrammes wie Cubase.

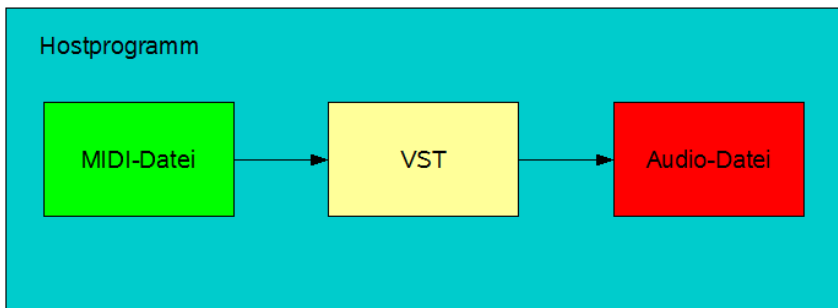


Abbildung 21: MIDI Datei mit Softwareklangerzeuger

Die MIDI Daten werden in Echtzeit per VST Schnittstelle vom VST Instrument entgegengenommen (VST Schnittstelle und VST Instrument sind in der Abbildung als „VST“ zusammengefasst) und dort in hörbare Audioinformationen umgesetzt (hier als „Audio-Datei“ beschriftet, was nicht ganz der Realität entspricht, denn es besteht keine Notwendigkeit, die Ausgangsinformationen des

VST direkt in eine Datei zu schreiben).

Im Prinzip funktionieren alle Klangerzeuger gleich: für einen bestimmten MIDI Ton rufen sie in ihrer Datenbank einen entsprechenden Klang ab. Es gibt im Rahmen von VST Instrumenten, die echte Instrumente simulieren sollen, grundsätzlich zwei verschiedene Ansätze, die nicht unbedingt getrennt voneinander existieren müssen: zum einen basieren die Klänge auf tatsächlichen Aufnahmen von echten Instrumenten (sogenannten Samples), die andere Möglichkeit ist, die Klänge der Instrumente auf synthetische Weise nachzubilden. In der Regel gilt dabei, dass echte Sounddatenbanken im Gegensatz zu synthetisch erzeugten Klängen sehr viel Speicherplatz benötigen, aber auch sehr viel authentischer klingen.

Ein einfaches Rechenbeispiel veranschaulicht den benötigten Speicherplatz: würde man ein Konzertflügel mit 88 verschiedenen Tönen auf 7,25 Oktaven aufzeichnen wollen, so müßte man jeden einzelnen Anschlag einer Taste bis zum kompletten Ausklingen aufzeichnen. Nimmt man an, dass die Gesamtlänge der Töne von Anschlag bis Ausklingen ca. 5 Sekunden beträgt (außer acht gelassen, dass tiefe Töne naturgemäß länger ausklingen, als hohe), dann würde dies für eine Taste bei 96 kHz und 24 Bit ca. 1,4 MB Speicher in Anspruch nehmen. 88 Tasten entsprechen demnach 120 MB. In diesem Falle hat man aber noch kein brauchbares Klavier, da es in einem solchen Falle keine Anschlagdynamik gibt und jeder Ton gleich stark angeschlagen wird. Eine einfache Reduktion der Lautstärke für leichtere Anschläge mag zwar praktikabel sein, entspricht aber nicht dem Klang eines echten Flügels. Also muß jede der 88 Tasten noch mit unterschiedlichen Stärken angeschlagen werden. Nimmt man dafür nur drei verschiedene an (was sicherlich zu wenig wäre), errechnet sich schon eine Größe von 362 MB für die Sounddatenbank des VST Instruments. In dieser Rechnung

sind die verschiedenen Klänge bei der Betätigung des Una Corda Pedals noch gar nicht eingerechnet. Ebenso müßte man das Sostenuato Pedal bei der Aufnahme beachten und gegebenenfalls weit länger als 5 Sekunden aufzeichnen. Viele VST Instrumente bieten darüberhinaus zusätzlich noch verschiedene Mikrofonpositionen und Klangfarben und erlauben den Wechsel zwischen wohltemperiertem und gleichstufigem Klavier. Ein Konzertpiano VSTi der Firma Steinberg namens „The Grand“ wurde bereits in seiner ersten Version auf 3 CDs ausgeliefert, was in etwa 2 GB Speicher entspricht. Das Chor VSTi „Symphonic Choirs“ der Firma East West, mit dem nicht nur Gesangsstimmen wie Bässe, Tenöre, Altos und Sopranos dargestellt, sondern sogar Wörter und Sätze geformt werden können, bringt es sogar auf 9 DVDs und damit 38 GB Speicherplatzanforderung.

Demgegenüber stehen VST Instrumente wie Edirols „Orchestral“, das mit 120 MB auskommt, aber ein komplettes Symphonie Orchester enthält (einschließlich Konzertpiano) oder „Slayer2“, das elektrische Gitarren mit nur 3 MB Speicherplatz simuliert. Hierbei handelt es sich naturgemäß um Synthesizer, die lediglich Klangcharakteristiken der entsprechenden Instrumente nachahmen.

Zusätzlich gibt es eine in der Praxis übliche hybride Form, in der besonders charakteristische Klänge (bei Instrumenten in der Regel der Anschlag) als Sample vorliegen, andauernde, gleichförmige Klanginformationen aber dann entweder synthetisch angefügt oder in einer Schleife abgespielt werden.

Auch für den Bass gibt es eine Vielzahl von möglichen VST Instrumenten, so zum Beispiel Steinbergs „Virtual Bassist“ oder Spectrasonics „Trilogy“. Letzterer, der auch bei dem Beispiel zur Anwendung kam, benötigt 3 GB Speicher und bietet dafür neben einem akustischen und einem elektrischen Bass

auch einen Synthesizerbass an. Insgesamt kann man in über 900 verschiedenen klingenden Bässen aller Arten wählen.

Im Beispielsong wurde der Bass nicht von Hand gespielt, da es schwierig ist, von Saiteninstrumenten MIDI Informationen zu erhalten. Zwar gibt es tatsächlich inzwischen entsprechende Effektgeräte für den E-Bass oder die E-Gitarre, die zusätzlich auch einen MIDI Ausgang besitzen, allerdings ist fraglich, ob sich dieser Aufwand im Rahmen der normalerweise überschaubaren Komplexität eines Basslaufes lohnt und inwiefern das MIDI Pendant nach seiner Digitalisierung noch Ähnlichkeit mit dem tatsächlichen Spiel des Bassisten hat. Sowohl beim Bass als auch bei der E-Gitarre gibt es über die reinen Toninformationen hinaus Geräusche, die von einem MIDI Umwandler wohl kaum umgesetzt werden dürften. Dazu gehören alle Arten von kurzen Quietschlauten beim Verschieben der Finger über die Saiten, sowie unterschiedliche Anschlagtechniken, wie zum Beispiel das Slapping, bei dem der Bassist die Saiten zum Schwingen bringt, indem er mit der Daumenwurzel daraufschlägt. Gute Bass VSTi stellen diese Geräusche selbstverständlich zur Verfügung, jedoch ist es unsicher, ob sie auch aufgerufen werden, wenn sie durch die MIDI Datei repräsentiert werden. Und wenn dies der Fall ist, so ist trotzdem nicht gewährleistet, dass das ursprüngliche Spiel des Bassisten Ähnlichkeit mit dem MIDI Endergebnis hat, das durch das VST Instrument geroutet wird. Alternativ könnte man wohl das entsprechende Effektgerät, ähnlich wie zuvor das E-Drumsteuerteil, als Klangerzeuger benutzen, jedoch wurde im Beispielsong ein VST Instrument benutzt.

Aufgrund dessen und weil der Aufwand in keinem Verhältnis zum Nutzen steht, wurde der Bass für den Beispielsong „*The Quest For Gold*“ mit der Maus im MIDI Editor des Hostprogrammes Cubase SX3 erstellt. Neben der Manipulation

von vorhandenen MIDI Informationen bietet das Programm nämlich über eine entsprechende Schaltfläche (ein sogenannter „Stift“) die Möglichkeit, neue Töne zu erstellen. Diese werden an einer beliebigen Stelle mit einem Klick begonnen und so lange in der Partitur nach rechts hin vergrößert, bis sie die Wunschkdauer haben. Wenn man dann die Maustaste loslässt, wird die Note mit einer festen Anschlagsstärke eingezeichnet. Da ein ständig gleichstarker Anschlag wenig natürlich erscheint, ist es empfehlenswert, einige bestimmte Töne mehr oder weniger zu akzentuieren, indem man im entsprechenden Manipulationsfenster am unteren Rand des MIDI-Editors die Höhen der Balken durch Drag and Drop einstellt. Abbildung 22 zeigt einen Ausschnitt des Basslaufes im Solo des Beispielsongs, der mit der Maus am Computer und ohne Instrument konstruiert wurde.

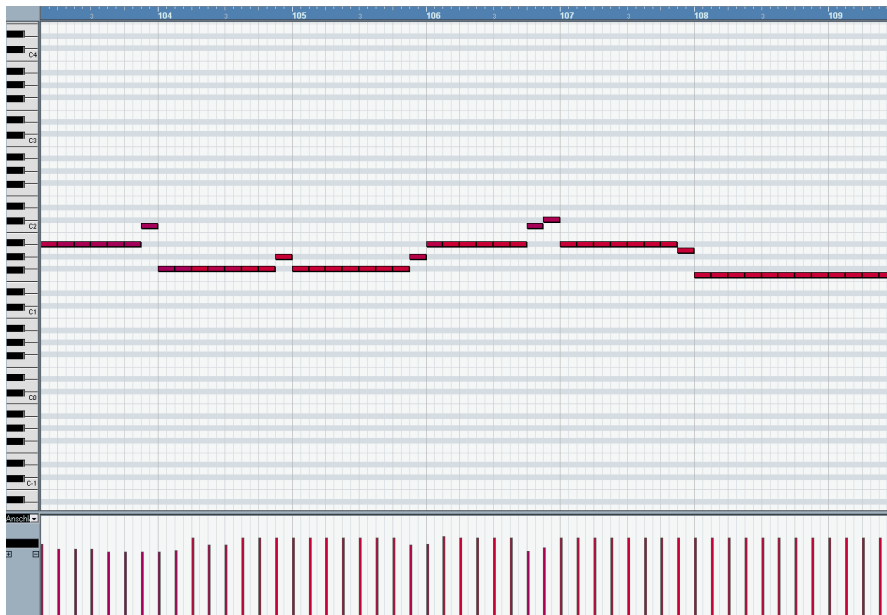


Abbildung 22: Basslauf im MIDI Editor

Neben dem Erstellen von Einzelnoten bietet der MIDI Editor von Cubase auch Arbeitserleichterungen, wie das Einzeichnen von Linien, die dann aus Einzelnoten der ausgewählten Quantisierungslänge bestehen. Aber auch Parabeln, Sinuskurven, Dreiecke und Rechtecke können in die Partitur eingezeichnet werden, auch wenn diese aber in der Praxis eher von wenig Nutzen sein dürften.

Es bietet sich an, den Bass in ebenso kleinen Abschnitten zu erstellen, wie zuvor das Schlagzeug (siehe Abbildung 20 auf Seite 49), um auch hier einzelne Passagen nachträglich kopieren zu können.

Neben der beschriebenen manuellen Eingabe der Bassnoten, die dann über das VSTi Trilogy einem echten Bass zum Verwechseln ähnlich klingen, ist es

natürlich auch möglich, einen MIDI Controller zur Eingabe zu benutzen. Ein MIDI Controller ist eine Hardwareklaviatur, die einen Ausgang für MIDI Signale besitzt und mit der Soundkarte des Studiocomputers auf dieselbe Weise verbunden werden kann, wie das E-Drum Steuerteil aus dem letzten Kapitel. Als Verbindungen haben sich für MIDI Datenkabel mit DIN Steckern der Norm 41524[7] durchgesetzt, in moderneren Geräten kommen aber auch USB Anschlüsse zum Einsatz, die einen entsprechenden MIDI Eingang am Computer überflüssig machen. Der Ablauf einer solchen MIDI Aufnahme wird in den nachfolgenden Kapiteln noch genauer beschrieben.

Der große Vorteil des MIDI Formats und der Klangerzeugung in Echtzeit ist nicht nur die nachträgliche Korrektur von Einzeltönen, wie bereits zuvor beschrieben, sondern auch die Manipulation des Klanges in vielen unterschiedlichen Arten. Abbildung 23 zeigt⁵ exemplarisch die mannigfaltigen Einstellungen, die im Bass VSTi Trilogy möglich sind und die Charakteristik des hörbaren Audiosignals signifikant ändern können.

5 von <http://www.spectrasonics.net/instruments/trilogy.html>



Abbildung 23: Benutzerschnittstelle Trilogy VSTi

Die Beschreibung der einzelnen Einstellmöglichkeiten würde den Rahmen dieser Arbeit sprengen und kann der Anleitung des Herstellers entnommen werden. Die meisten VSTi bringen eine derartige oder sogar höhere Flexibilität mit sich. Der Benutzer kann während dem Abspielen der MIDI Daten Einstellungen vornehmen und die Ergebnisse sofort hören. Wie sich in späteren Kapiteln noch zeigen wird, kann das Hinzufügen oder die Änderung der Klangcharakteristik eines bestimmten Instrumentes Auswirkungen auf den Klang aller anderen Instrumente haben. So kann es möglich sein, bestimmte Frequenzen zu dämpfen oder zu verstärken. Dies ist auch mit entsprechenden Equalizereinstellungen in fest aufgenommenen Audiospuren machbar (siehe spätere Kapitel), jedoch nicht die nachträgliche Änderung des Instrumentes, des Verstärkers oder der Anschlagtechnik. Dies bleibt den VSTi vorbehalten.

Im Gegensatz zum Schlagzeug ist es bei der Bass MIDI Spur nicht nötig einen sogenannten Downmix (also eine Umwandlung von MIDI & VSTi zu Audio)

vorzunehmen, da der Klangerzeuger nicht als Hardware, sondern als Software vorliegt und so stets bei Bedarf vom Computer angesprochen werden kann. Das E-Drumsteuerteil müßte zum Abspielen zunächst immer erst mit dem Computer verbunden werden, was spätestens dann Probleme verursacht, wenn mehrere externe Klangerzeuger zum Einsatz kommen.

4.2.3 Rhythmus E-Gitarren

Die Rhythmus E-Gitarren sind im Beispielsong sowohl VSTi, als auch echte Audioaufnahmen von E-Gitarren aus dem Studio. Die MIDI Daten für die VSTi wurden in ähnlicher Weise wie die Bass MIDI Daten erzeugt (also manuell per Maus). Das hier zur Anwendung gekommene VSTi nennt sich Virtual Guitarist und wurde von der Firma Steinberg entwickelt, ist aber inzwischen nicht mehr käuflich auf der Homepage des Herstellers beziehbar. Abbildung 24 zeigt⁶ das User Interface von Virtual Guitarist:

6 von <http://messe.harmony-central.com/Musikmesse02/Content/Steinberg/PR/Virtual-Guitarist.html>



Abbildung 24: Benutzerschnittstelle Virtual Guitarist VSTi

Gitarren sind, im Gegensatz zu monophonen Instrumenten wie einem Bass, kaum bis gar nicht synthetisch nachzuahmen. Während außerdem die kleinsten Klangeinheiten von Instrumenten wie Klavier oder Schlagzeug grundsätzlich auch in Sampleform kombiniert werden können und daraufhin einen dem Originalinstrument sehr ähnlichen Ton erzeugen, ergeben zwei kombinierte E-Gitarrentöne bei einer echten E-Gitarre einen ganz eigenen Sound, der derzeit nicht an einem Computer erzeugt werden kann. Darüberhinaus gibt es, wie weiter oben bereits erwähnt wurde, sehr viele Zusatzgeräusche beim Spiel einer E-Gitarre, die sich ebenfalls künstlich nur sehr unzufriedenstellend reproduzieren lassen. Schlußfolgernd bedeutet dies, dass (E-)Gitarren VSTi wie „Virtual Guitarist“, sein rein elektrisches Pendant „Virtual Guitarist Electric Edition“ oder sein Nachfolger „Virtual Guitarst 2“ (alle von Steinberg), lediglich auf Samples

basieren können. Dies bedeutet ebenfalls, dass Akustik Pickings oder E-Gitarrenriffs, die bei tatsächlichen Instrumenten aus mehreren Einzeltönen oder sogar schnell aneinandergereihten Akkorden bestehen, bei einem Gitarren VSTi als fertige Samples vorliegen, die zuvor vom Tontechniker ausgewählt werden und dann über einzelne MIDI Noten aufgerufen werden. Solche VSTi sind also nicht dazu geeignet, um beispielsweise in MIDI Noten einzelne Akkorde zu konstruieren und diese wiederzugeben, sondern enthalten stattdessen diese Akkorde schon in Sampleform. Anders ausgedrückt: MIDI Daten zur Steuerung eines VSTi wie Virtual Guitarist sind ausschließlich monophon. Die einzige Ausnahme bilden spezielle Kombinationen von parallelen Tönen, die lediglich signalisieren sollen, dass der entsprechende Akkord moll und nicht dur ist.

Zur Verdeutlichung zeigt Abbildung 25 eine solche simple MIDI Information, die aus einer Spur des Beispielsongs stammt:

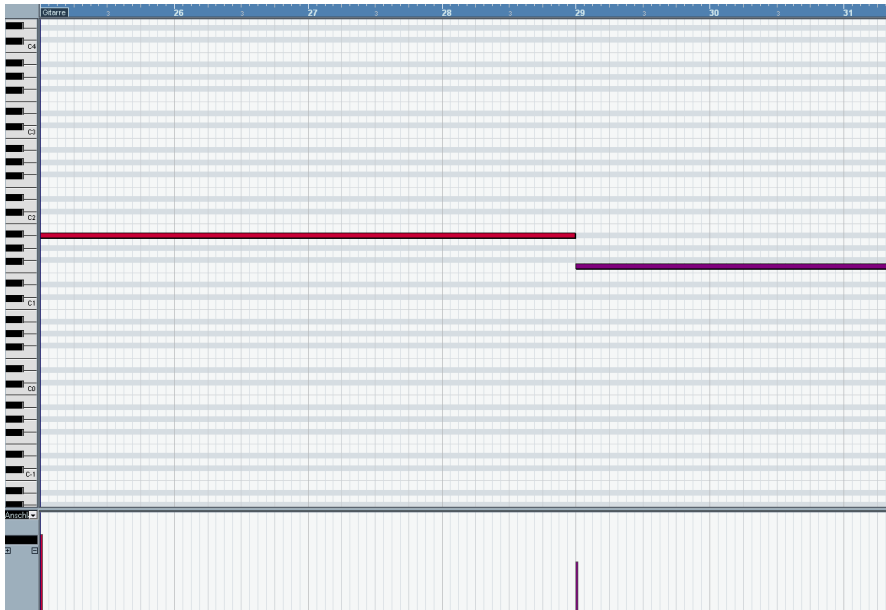


Abbildung 25: E-Gitarre im MIDI Editor

Hierbei handelt es sich um die ersten offenen (d.h. nicht abgedämpften) Akkorde des Songs. Man erkennt nur zwei lange Töne, nämlich A# und F. A# ist dabei stärker angeschlagen (rot), als das F (blau). Hörbar ist aber bei entsprechender Einstellung des VSTi ein A# Power Chord, der normalerweise aus den verzerrten E-Gitarrentönen A#, F und A#' besteht (respektive bei dem F die Töne F, C und F'). Interessant ist, dass präzise dieselben MIDI Informationen, die in Abbildung 25 zu sehen sind, später im Song dazu dienen, um gedämpfte 16tel Anschläge zu erzeugen. Dies ist deshalb möglich, weil eine andere Einstellung an dem VST Instrument vorgenommen wurde. Bei dieser Einstellung wird die MIDI Information lediglich anders interpretiert, nämlich als gedämpfte 16tel, die einmal mit einem akzentuierten Ton beginnen (wegen dem roten, harten Anschlag) und

einmal nicht (blau). MIDI Daten gelten hier also nicht mehr als Noten für Klangerzeuger, sondern eher als einfachere Steuerdaten für ein vorgefertigtes Arsenal an verschiedenen Riffs. Dieselben MIDI Daten würden bei einem Bass oder Klavier VSTi nur simple Einzeltöne erzeugen.

Entsprechend dieser Tatsachen sind die Samples eines Gitarren VSTis grundsätzlich sehr viel länger, denn sie enthalten nicht Einzeltöne, sondern ganze Spielarten. Andersherum sind die Daten, die sie erklingen lassen, viel einfacher, aber repräsentieren nicht die tatsächliche dahintersteckende Musik. Eine weitere Folgerung ist, dass die Sampledatenbank sehr viel größer ist, weswegen in Virtual Guitarist nur eine Oktave erzeugt werden kann. Dies ist allerdings wenig tragisch, denn auch Akkorde auf der E-Gitarre werden in aller Regel innerhalb der tiefsten Oktave gespielt.

Da ein solches VSTi nur eine begrenzte Datenbank aus einer überaus endlichen Anzahl von herkömmlichen Riffs und Pickings enthalten kann, ist ebenfalls offensichtlich, dass man bei ausschließlicher Benutzung von virtuellen Gitarren wohl kaum originelle Musik produzieren kann. Die Gitarren MIDI Spuren erfüllen also in erster Linie den Zweck, den Gesamtsound breiter zu machen und das Gitarrenspiel präziser wirken zu lassen. Tatsächlich ist es so, dass beispielsweise die oben erwähnten 16tel Anschläge von der Lautstärke her so weit runtergeregelt wurden, dass lediglich die zeitliche Präzision noch zu hören ist. Der Rest des Gitarrensounds für den Beispielsong „*The Quest for Gold*“ wurde nach herkömmlicher Weise aufgenommen und bildet die eigentliche Charakteristik des Gitarrenspiels.

Im Beispielsong wurden insgesamt zwei Audiospuren für die Rhythmusgitarre

aufgezeichnet. Abbildung 26 zeigt den grundsätzlichen Ablauf der Aufnahme einer Audioquelle im Studio:

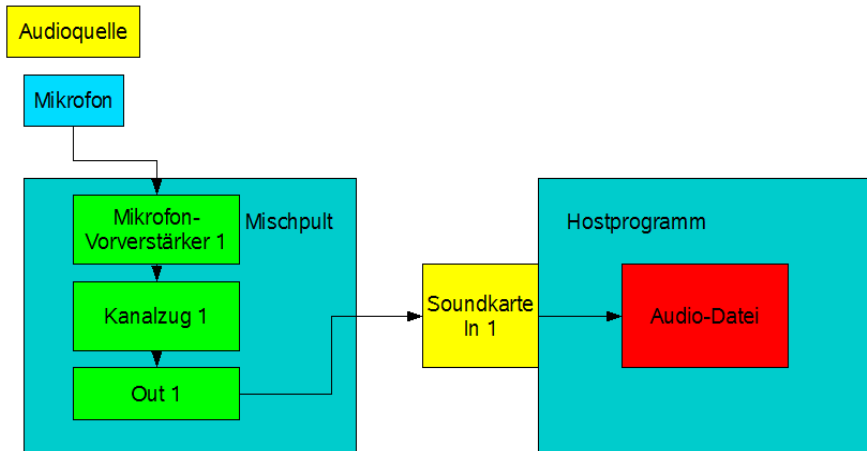


Abbildung 26: Audioquellen Aufnahme

Eine Audioquelle wird über ein Mikrofon aufgezeichnet, das an einem Mischpult betrieben wird. Ein entsprechender, in jedem Kanalzug eines Mischpultes eingebaute Vorverstärker, versorgt das Mikrofon mit der dazu nötigen Energie. Das Signal liegt im entsprechenden Kanalzug an und muß nun gemäß den in Kapitel 3.2 gemachten Vorgaben gepegelt werden, d.h. es muß so ausgepegelt sein, dass es eine Übersteuerungsreserve von etwa -5 dB gibt, aber ansonsten die von der Soundkarte zur Verfügung gestellte Bittiefe optimal in Bezug auf die aufzuzeichnende Maximalamplitude ausgenutzt wird. Weiterhin verfügt das Mischpult über einen Signalausgang, der entweder pro Kanalzug vorhanden ist oder lediglich als Gesamtsumme des Pultes. Dieses Signal wird an den Eingangskanal der im Studiorechner eingebauten Soundkarte weitergegeben, die ihrerseits durch das Hostprogramm gesteuert wird. Dieses nimmt die Signale der

Soundkarte am entsprechenden Eingangskanal entgegen und schreibt sie in eine Audiodatei. Je nach Audioquelle kann man auch mehrere Mikrofone parallel betreiben (wie zum Beispiel bei einer Schlagzeugaufzeichnung). In dem Falle würden die anderen Kanäle des Mischpultes zur Anwendung kommen und die einzelnen Audiosignale würden optimalerweise in separaten Soundkarteneingängen an das Hostprogramm übergeben werden. Anders ausgedrückt: würden 3 Mikrofone betrieben werden, so wäre es für spätere Mischarbeiten von Vorteil, wenn auch 3 separate Audiodateien vom Hostprogramm erzeugt werden würden. Dies setzt allerdings eine entsprechende Hardware voraus. Insbesondere die Soundkarte benötigt mehrere Eingänge und das Mischpult separate Ausgänge für jeden einzelnen Kanalzug.

Abbildung 27 zeigt die Positionierung des Mikrofons vor dem Gitarrenverstärker.



Abbildung 27: Positionierung des Mikrofons vor dem Verstärker

Neben der Aufzeichnung mittels Mikrofon und Gitarrenverstärker gibt es natürlich auch die Möglichkeit, die E-Gitarre (oder besser: den Ausgang ihres Effektgerätes) direkt über ein entsprechendes Kabel entweder mit dem Mischpult oder der Soundkarte zu verbinden. Allerdings geht dabei die spezielle Klangcharakteristik des Verstärkers (und dessen Lautsprechers) verloren, die in aller Regel erhalten bleiben soll.

Die Gitarrenspur wurde nun begleitend zu den bislang vorliegenden Instrumenten Schlagzeug, Bass und Gitarre (VSTi) insgesamt zweimal komplett aufgezeichnet. Zum einen, um einen volleren Sound zu erhalten und zum anderen, um die beiden Spuren etwas im Stereopanorama zu verteilen (dazu später mehr). Darüberhinaus

können eventuelle Fehler, die bei der Aufnahme nicht aufgefallen sind, mit Hilfe einer zweiten Gitarrenspur gegebenenfalls retouchiert werden, ohne dass eine weitere Aufnahme im Studio nötig wird.

Das wichtigste Argument aber für eine zusätzliche E-Gitarrenspur, die mit einer echten E-Gitarre von einem richtigen Gitarristen eingespielt wurde, ist die oben genannte Tatsache, dass Virtual Guitarist und ähnliche Programme keinen Spielraum für Nuancen bieten. So fehlen spezielle Akzentuierungen des Anschlages, die auf das Schlagzeug an einer bestimmten Stelle passen ebenso, wie nicht herkömmliche Akkorde, außergewöhnliche Anschlagetechniken und die vormals schon genannten Zusatzgeräusche. Außerdem ist ein bestimmter Gitarrenklang oftmals eine Individualität, die natürlich mit einem VSTi ebenfalls nicht zu erreichen ist, da es sich um bereits vorher aufgezeichnete Samples handelt. Eine echte Gitarre unterscheidet sich aber im Klang anhand vieler Aspekte von einer anderen: die benutzten Saiten, das Alter dieser Saiten, die Wahl des Tonabnehmers, die Art des Tonabnehmers (Humbucker oder Single Coil), die Charakteristik des Halses und der Bünde, die Stimmung (die E-Gitarre im Beispielsong wurde auf D runtergestimmt), das verwendete Effektgerät mit unzähligen Einstellungen, sowie der benutzte Verstärker und dessen Lautsprecher. Diese speziellen Charakteristiken können naturgemäß von einem VST Instrument nicht erreicht werden. (Dies gilt selbstverständlich für alle Instrumente, jedoch wohl für kaum eines in einem derartigen Umfang, wie bei der E-Gitarre.)

Durch das Zusammenmischen sowohl der VSTi Gitarre, als auch der echten, erreicht man nun aber die Vorteile von beiden: die Präzision der MIDI-Gitarre und die Charakteristik der Audiogitarre. Dieser spezielle Sound muß nur noch in seinen Lautstärkeverhältnissen angepasst werden, was im Mix geschieht und in

späteren Kapiteln beschrieben wird.

4.2.4 Solo E-Gitarren

E-Gitarren Soli lassen sich ebenfalls nicht authentisch per VST Instrument darstellen. Dem Autor sind keine Sampler oder Synthesizer bekannt, die einer Solo E-Gitarre in einem Maße ähneln, das dem Hörer keinen deutlichen Unterschied auffallen läßt. Entsprechend wurden die Solo E-Gitarren auch von Hand mit einer echten E-Gitarre eingespielt. Im Unterschied zu den Rhythmusgitarren wurde auf die Aufnahme via Mikrofon und Verstärker verzichtet. Stattdessen wurde direkt der Gitarreneffektgerätausgang mit einem Eingang der Soundkarte verbunden. Hochwertige Effekgeräte versprechen, verschiedenartige Verstärkerklänge simulieren zu können. Im tieffrequenten Bereich ist das Ergebnis zumeist unzufriedenstellend, aber im eher hochfrequenten Bereich eines E-Gitarrensolos fallen Verstärkercharakteristiken ohnehin weniger auf. Zusätzlich sei angemerkt, dass Verstärker gewöhnlicherweise Klänge verschlechtern, indem sie sie zum Beispiel dumpfer klingen lassen. Dies hängt damit zusammen, dass für E-Gitarrensound die Ära der Musikgeschichte charakteristisch ist, in der es keine hochwertigen Verstärker und Boxen gab. So schwören heute viele Musikliebhaber auf den dumpfen Klang eines Röhrenverstärkers. Ein Solo sollte aber einen möglichst reinen Sound besitzen, um sich vom Rest der Instrumentierung absetzen zu können, vor allem, wenn viele Instrumente gleichzeitig ein großes Maß an Frequenzbereichen belegt. Letztendlich ist dies aber eine Geschmacksfrage. Es wäre ebenso gut möglich gewesen, die Sologitarren wie die Rhythmusgitarren aufzuzeichnen.

Es wurden wieder zwei Spuren aufgenommen, die stets gleichzeitig zu hören sind und zu 90 Grad links und rechts im Stereopanorama verteilt sind. Zum einen erhält man als Ergebnis einen breiten Klang, zum anderen ist dies aber vor allem ein charakteristisches Merkmal: insbesondere in Passagen, in denen die beiden Gitarrenspuren sich leicht im Tempo oder in der Stimmung unterscheiden, entsteht der Eindruck von tatsächlich handgemachter Musik, der sicherlich im Rahmen der Präzision der oben genannten VST Instrumente als sehr angebracht erscheint.

4.2.5 Ambient

Wie in Kapitel 2.1 bereits erklärt, wird im Rahmen dieser Arbeit unter „Ambient“ jede Art von stimmungs- und klangfördernder Instrumentierung verstanden, die das Lied im Rahmen seines Arrangements nicht unbedingt nötig hat, aber dennoch entscheidend zum Klangbild beitragen kann. In der Stilrichtung des Power Metals sind dies oftmals Keyboards mit Synthesizer-Streichern oder sogar richtige Streicher, die schlicht die Grundakkorde des Songs begleiten. Eine solche Untermalung war für den Beispielsong ebenfalls geplant und es wurde sogar ein entsprechender Keyboarder beauftragt, allerdings wurden leider keine Ergebnisse abgeliefert (Musikermanagement wird in einem späteren Kapitel noch behandelt). Da der Song aber auch so bereits einen sehr breiten Klang aufweisen konnte, wurde dieser Plan später nicht weiterverfolgt. Die einzigen Streicher kommen nun im Refrain zum Einsatz.

Diese Streicher spielen keine vollen Akkorde über ganze Takte, sondern sehr

schnell gespielte Noten in der Strichart Spiccato, bzw. Sautillé. Die Noten werden in Abbildung 28 gezeigt und werden ausnahmsweise nicht in ihrer Darstellung im MIDI Editor von Cubase gezeigt, sondern in der Partituranzeige. Dies hängt damit zusammen, dass es zu viele Noten sind und in der MIDI Ansicht kaum mehr als viele Punkte erkennbar wären.

Refrainstreicher

49

Refrainstreicher

Refrainstreicher

51

Refrainstreicher

Abbildung 28: Spiccato Streicher im Refrain

Diese Noten begleiten die Akkorde des Refrains, die die anderen Instrumente spielen und wurden, wie zuvor der Bass, per Maus in den MIDI Editor von Cubase eingegeben. Wiedergegeben werden sie vom VSTi Orchestral von Edirol in seiner Einstellung für mehrere Steicher, Strichart Spiccato. Diese Ambientpassage spielt in jedem der insgesamt 9 Refrains, wobei sie, wie alle anderen Instrumente, während der letzten 3 Refrains ausgeblendet wird.

Durch die Strichart und die Geschwindigkeit (bei 145 BPM) entsteht durch diese Passage eine beeindruckende Dramatik, die entscheidend zur Qualität und zum Charakter des Beispielsongs beiträgt.

4.2.6 Chor

In Kapitel 4.1 wurde der Einsatz eines Chores angesprochen. Es wurden mehrere verschiedene Ansätze hierzu verfolgt, die letztlich alle zu keinem Erfolg führten. Deswegen sind im fertigen Beispielsong kaum Chöre zu hören. Trotzdem sollen in diesem Kapitel die Vorgehensweisen beschrieben werden und der Grund für ihr Scheitern.

Ein synthetischer Chor begleitet auch in der finalen Version die letzten 6 Refrains. Es handelt sich hierbei um das VSTi „Magnus Choir“, das aber lediglich A-Laute wiedergeben kann und dies mit den entsprechenden MIDI-Daten auch eindrucksvoll tut – vor allem in Hinblick auf seine geringe Speicherplatzanforderung von in etwa 7 MB. Für die Stimmung an dieser Stelle ist dies sicherlich vorteilhaft. Auch ist dieser Chor das einzige Instrument, das bei der Instrumentalblende am Ende des Songs zusammen mit dem Gesang stehenbleibt. Für Profis und Musikkennner wird dieser Chor nicht sonderlich authentisch klingen, allerdings fügt er sich derart in den Gesamtkontext der anderen Instrumente ein, dass dies kaum auffallen sollte.

Darüberhinaus wurde probiert, einen artikulierten Chor, der die Worte „Hail King Anu“ an insgesamt 2 Stellen im Song dreimal wiederholt, hinzuzufügen. Dazu wurden drei Ansätze verfolgt. Einer mit tatsächlichen, menschlichen Sängern und zwei mit VSTi, die über eine immense Anzahl von Samples verfügen, darunter neben A-, E-, I-, O- und U- Lauten und anderen Vokalen auch die meisten geläufigen Konsonanten.

Die Aufnahme der menschlichen Sänger wurde immer und immerwieder in verschiedenen Tonlagen vorgenommen, um möglichst viele Gesangsstimmen zu

erhalten. Diese wurden dann zusammengemischt und mit viel Hall versehen. Zwar klingt das Ergebnis durchaus zufriedenstellend, allerdings wirkt es im Gesamtkontext des Songs derart deplaziert, dass man es in der finalen Version des Beispielsongs nicht verwenden konnte.

Die beiden anderen Ansätze wurden mit den VSTi „Voices of the Apocalypse“ und „Symphonic Choirs“, beide von East West / Quantum Leap, verfolgt. Voices of the Apocalypse ist der Vorgänger der Symphonic Choirs und konnte nur kurz getestet werden. Die Bedienung, bzw. die MIDI Programmierung der einzelnen Worte ist derart kompliziert, dass in diesem Zeitraum kein brauchbares Ergebnis zustande kam.

Symphonic Choirs, das ausgiebiger getestet werden konnte, bietet ein vielversprechendes Tool, den sogenannten Word Builder, das die Formung verschiedener Worte sehr viel einfacher werden lässt. Technisch funktioniert dies folgendermaßen:

Ein (Hardware) MIDI Keyboard sendet seine MIDI Daten an den Word Builder. Dieser ist im Wesentlichen ein Eingabefeld für (englische) Wörter. Eingegebene Silben werden in die korrespondierenden Längen der Einzellaute umgewandelt und gemäß den auf dem MIDI Keyboard gedrückten Tasten in bestimmte MIDI-Signale pro Silbe umgewandelt. Die verschiedenen Oktaven des MIDI-Ausgangssignals des Word Builder stehen nicht nur für verschiedene Tonhöhen, sondern auch für die unterschiedlichen Konsonanten. Dieses MIDI Signal kann nun von einem Hostprogramm wie Cubase normal entgegengenommen werden. Die MIDI-Daten sind entsprechend der darin enthaltenen Konsonanten und unterschiedlichen Vokale entsprechend komplex und in Ermangelung eines Word Builders für Voices of the Apocalypse ist es verständlich, dass keine brauchbaren

Ergebnisse in kurzer Zeit erzielt werden konnten.

Zur Konstruktion der Wörter im Word Builder bieten die Hersteller nicht nur Englisch, sondern auch Lautschrift und Votox, eine extra für Word Builder geschaffene Sprache. Die Zeile „Hail King Anu“ würde sich in Votox folgendermaßen lesen:

HiEl KEng u nO

Worte werden also in ihre Silben heruntergebrochen und mit bestimmten Groß- oder Kleinbuchstaben umschrieben. Großgeschriebene Konsonanten enthalten beispielsweise keine Tonhöheninformation (non-pitched), kleine (wie im Beispiel das g), werden in verschiedenen Höhen artikuliert (pitched). Vokale dagegen sind natürlich alle mit Tonhöheninformationen gespeichert. Das große E im Beispiel wird wie das deutsche I ausgesprochen, das O wie ein U. Es bedarf etwas Übung, um diese Sprache eingeben zu können, die merklich von englischsprachigen Menschen geschaffen wurde. Ein späterer Versuch, die Textzeile „*Ki you ancient Tiamat; Mutilated by impact; Carry in your oceans sea; The remedy to our disease*“ umzusetzen, erforderte folgenden Votox Text:

KE yO en S!end TE a maT

mO TE lA TeT baE Em PAKT

KA rE En yor o S!An SE

t!e r!Am A dE To au dE SES

Anhand dieses Beispiels sollte die Charakteristik der Votox Sprache deutlich werden. Pro Anschlag auf dem MIDI-Keyboard (dies können sowohl Einzeltöne als auch ganze Akkorde sein), wird nun jeweils eine Silbe wiedergegeben. So

kann der Keyboarder selbst Tempo und Gesangsmelodie bestimmen. Im Word Builder kann auch entschieden werden, ob die Silben hart oder weich gesungen werden und ob die Konsonanten laut oder leise sein sollen.

Bei Symphonic Choirs hat man nicht nur die Wahl zwischen Bass, Bariton, Tenor, Alt, Mezzosopran, Sopran und Knaben, sondern kann auch Mikrofonpositionen, Gesangsvibrato und schier unzählige andere Einstellungen vornehmen. Dies alles resultiert in sehr hohen Anforderungen an den Computer, die im Rahmen der Arbeit am Beispielsong nur dann erfüllt werden konnten, wenn sonst keine anderen Programme oder VSTi geladen wurden. Symphonic Choirs wird auf 9 DVDs mit 38 GB Speicherplatzanforderung ausgeliefert und ein System mit 1,5 GB RAM war nur nach sehr langer Wartezeit fähig, die Wiedergabe in Echtzeit einigermaßen flüssig laufen zu lassen – allerdings nicht ohne den zweifachen Hinweis darauf, dass das Laden weiterer Samples auf eigenes Risiko geschieht.

Diese Tatsache macht die unbeschwerte Arbeit mit Symphonic Choirs mit der Computerhardware, die für den Test zur Verfügung stand, sehr träge. Ein Testergebnis klingt zwar alleinstehend beeindruckend und könnte durchaus als echter symphonischer Chor durchgehen, aber im Rahmen des Beispielsongs wirkte der Chor zu aufgesetzt und auch zu undeutlich in der Aussprache. Wörter können errahnt, aber nicht verstanden werden. Dies ist zwar bei Chören eher die Regel, aber nicht im Rahmen eines Musikstückes dieser Stilrichtung. Daher wurde auf den Gebrauch dieses Chores und die Anschaffung des VSTi (derzeit 500.- EUR) verzichtet.

Diese Fehlschläge haben zu dem Entschluß geführt, auf einen Chor im Beispielsong zu verzichten. Das erste Vorkommen von „Hail King Anu“ wird von

einem normalen Sänger gesungen und das zweite Vorkommen wurde komplett weggelassen, zumal es sich um eine Stelle handelt, die sich bei genauerer Betrachtung im Nachhinein als zu hektisch für den Chor erwiesen hat.

4.2.7 Gesang

Die Qualität eines Songs steht und fällt mit der Stimme und dem Talent des Sängers. Bei keinem Instrument sind so viele Nuancen möglich, wie bei der menschlichen Gesangsstimme und noch weit mehr negative Teilaspekte einer Stimme können dafür sorgen, dass dem Hörer nicht gefällt, was er hört, womit wird der ganze Song als schlecht empfunden wird.

Zunächst ist zweifelsohne wichtig, dass der Sänger die Töne trifft, die er versucht zu singen. Aber ebenso muß er die Silben genau im richtigen Takt singen und die Töne halten können. Er muß Gesangsvibrato ebenso beherrschen, wie Falsett und Bass. Er muß Emotionen transportieren können, die in der jeweiligen Passage zum Text oder zur Melodie passen, d.h. seine Stimme muß „Gefühl“ besitzen. All diese Eigenschaften sind erlernbar, aber machen noch keinen guten Sänger aus.

Der Sänger muß eine angenehme oder zumindest passende Stimme besitzen, muß deutlich sprechen bzw. singen können, und darf nicht lispeln, krächzen oder stottern. Diese Eigenschaften sind nur bedingt erlernbar und deshalb kann man fast sagen, dass ein guter Sänger zum Sänger geboren sein muß. Wie sich in einem späteren Kapitel zeigen wird, kann die Tontechnik und im Speziellen die digitale Tontechnik einige der oben genannten Aspekte korrigieren. Aber eben leider nicht alle.

Aus diesem Grunde wurde für den Beispielsong nach 3 talentierten Sängern gesucht. Eine Suche nach einem guten Vokalist ist nie einfach, vor allem im Hinblick darauf, dass für den Beispielsong kaum eine nennenswerte Bezahlung der Musiker stattgefunden hat. Insgesamt wurden 4 Sänger in Betracht gezogen, von denen lediglich 2 gut genug für die finale Version des Beispielsongs waren. Daher ist es leider unvermeidlich gewesen, dass ein Sänger zwei verschiedene Rollen belegt.

Zwei der Sänger, die später nicht in der finalen Version erschienen sind, wurden über ein soziales Netzwerk im Internet angesprochen, in dem vorher anhand ihrer Interessen der Rückschluß gezogen werden konnte, dass sie sich selbst als gute Metalsänger verstehen. Einer hatte seine Mitarbeit zwar immer fest zugesagt, aber niemals Ergebnisse geliefert, trotz einer deutlichen, anfänglichen Motivation. Nachfragen blieben unbeantwortet, wodurch sich der Sänger als unbrauchbar erwiesen hat, vor allem dann, wenn man bis zu einem gewissen Zeitpunkt die Zuarbeit der Musiker braucht, um den Beispielsong zu einem Stichtag fertigzustellen.

Der andere Sänger reiste 50 Kilometer mit dem Auto an und besang den kompletten Song. Diese Gesangsspuren, die es ebenfalls nicht in die finale Version geschafft haben, waren im Bassbereich anzusiedeln, was ganz explizit gegen die ursprüngliche Vorstellung des eher hohen Power Metal Gesangs verstoßen hat. Da zu diesem Zeitpunkt aber nicht absehbar war, ob noch bessere Sänger zur Verfügung stehen würden, wurde die Aufnahme komplett vollzogen.

Der dritte Sänger, der auf der Aufnahme in der Rolle des Cheruben 1 und 2 zu hören ist, ist der Sänger einer örtlichen Heavy Metal Band, die gewöhnlicherweise in tiefen und eher rauen Tönen spielt (und singt). Die

Aufnahme wurde vom Sänger selbst in seinem eigenen Tonstudio ohne Anwesenheit des Autors dieser Arbeit aufgezeichnet und per E-Mail übersandt. Grundlage bildete die Demoverision, wie im Ablauf in Abbildung 7 auf Seite 17 zu sehen ist. Zusätzlich wurden Aufnahmen des Refrains in verschiedenen Tonhöhen mitgeliefert, die in der Endversion dazu beitragen, das Klangbild des Gesangs an jener Stelle zu verbreitern. Die Aufnahme ist, da es sich merklich um einen Profi handelt, überaus zufriedenstellend und mußte nur noch mit entsprechenden Effekten versehen werden. Dies wird in einem späteren Kapitel beschrieben.

Der vierte Sänger, der den größten Textanteil des „Enki“ hat und somit die Hauptrolle singt, ist ein kanadischer Freizeitsänger, der nicht nur über einen beeindruckenden Tonumfang und eine perfekt passende Stimme verfügt, sondern darüberhinaus Emotionen sehr gut transportieren kann. Ein zusätzlicher positiver Aspekt ist die Tatsache, dass seine Stimme der des bekannten Heavy Metal Sängers Bruce Dickinson (Sänger der Band Iron Maiden) gleicht. Der Kontakt wurde hergestellt, als der Autor dieser Arbeit die Idee hatte, die besten Sänger auf Online Karaoke Plattformen um ihre Mitarbeit zu bitten. Glücklicherweise mußte nur dieser eine Sänger befragt werden, der sogleich enthusiastisch zusagte. Online Karaoke Plattformen bieten ihren Mitgliedern kostenlos die Möglichkeit, instrumentale Lieder der Popmusik zu besingen und dies von den anderen Mitgliedern bewerten zu lassen. Durch die sprachliche Barriere und die große Entfernung erwies sich die Formulierung der Wünsche zwischen Kanada und Deutschland als schwieriger. Die gesamte Kommunikation fand über den Instant Messenger MSN statt, der in Amerika weiter verbreitet ist, als in Europa.

Zur Verdeutlichung der Anforderungen an den Gesang wurden deshalb vom Autor einzelne Gesangspassagen provisorisch besungen und per Skydrive (siehe

Kapitel 3.4) übermittelt. Dies fand für gewöhnlich gegen Abend deutscher Zeit statt, wodurch die Sounddateien in Kanada bereits morgens zur Verfügung standen (Deutschland ist diesem Teil Kanadas 7 Stunden voraus). Auf Grundlage dieser provisorischen Gesangsfragmente konnte der Enki-Sänger seinen Teil im eigenen Tonstudio aufzeichnen und (abends kanadischer Zeit) zurücksenden, wodurch das Material früh morgens in Deutschland bereits zur Verfügung stand. In diesem Hinblick erweist sich eine Zeitverschiebung sogar als sehr effektive Arbeitsbeschleunigung.

Die verlustfreien WAV Dateien wurden dann einfach in das Cubase Projekt des Beispielsongs importiert.

Mit dem Gesang wurde die letzte Spur dem Gesamtprojekt hinzugefügt. Abbildung 29 zeigt eine Übersicht im Programm Cubase SX3.

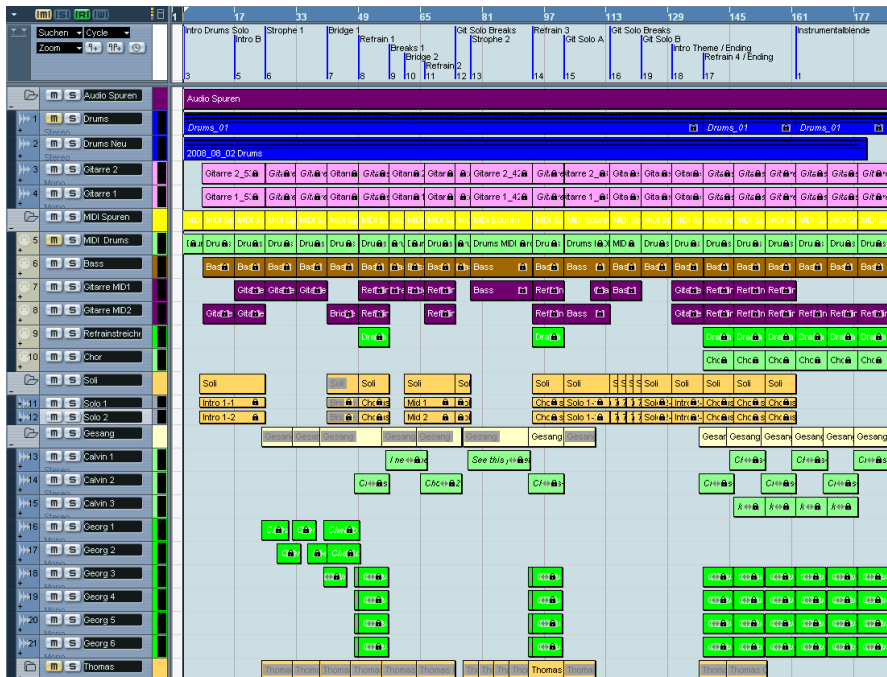


Abbildung 29: Gesamtprojekt im Cubase Projektfenster

Ganz oben befindet sich die Benennung einzelner Passagen, damit sich Musiker autark im Projekt hätten zurechtfinden können. Darunter folgt ein Überordner für Audiospuren, in dem eine erste Schlagzeugspur (im Beispiel stummgeschaltet) und die finale Schlagzeugspur zu finden sind, sowie die beiden mit Mikrofon aufgezeichneten Rhythmusgitarrenspuren.

Darunter folgt ein Unterordner mit MIDI Spuren, worin von oben nach unten enthalten sind: Trommelspur (stumm, da Audiospur vorhanden), Bass, Gitarre 1 (16tel), Gitarre 2 (volle Akkorde, Streicher (die nur im Refrain zum Einsatz kommen) und der Chor (nur in den letzten 6 Refrains).

Danach folgt ein Solo-Ordner, der, wie weiter oben beschrieben, zwei Gitarrensolospuren enthält.

Zuletzt folgt der Gesangsordner, der zunächst den kanadischen Sänger auf 3 Spuren enthält, danach den deutschen Sänger der lokalen Heavy Metal Band in 6 Spuren und schließlich der Sänger, der im sozialen Netzwerk angesprochen wurde, aber leider im Bassbereich gesungen hat (ebenfalls stummgeschaltet).

Erkennbar an dieser Ansicht ist vor allem der Aufbau einzelner Spuren, die in erster Linie aus kleinen Abschnitten zusammengesetzt wurden. Dieser konstruktivistische Ansatz unterscheidet sich deutlich von klassischen Audioproduktionsparadigmen, in denen Musiker dazu gezwungen waren, größere Abschnitte am Stück spielen zu können. Dieser Anspruch entfällt weitgehend im Rahmen des bislang gezeigten Vorgehens.

Zwar ist zu diesem Zeitpunkt das komplette Beispielsong aufgezeichnet, aber er klingt noch nicht gut genug. Deshalb wird im nächsten Kapitel beschrieben, wie und welche Postmanipulationen vorgenommen wurden.

4.3 Effekte der Einzelspuren

Die in den vorigen Kapiteln aufgezeichneten Audiosignale mußten den heutzutage üblichen Gewohnheiten von Musikhörern angepasst werden. In ihrer rohen, d.h. effektfreien Form erscheinen Tonaufnahmen den meisten Menschen als „nicht hochwertig“, „schlecht klingend“ oder „zu dünn“. Die Palette der Kritikpunkte an rohen Audioaufnahmen ist dabei unübersichtlich groß. Hinzu kommt, dass verschiedene Audiospuren in ihrer Ursprungsform in denselben

Frequenzbereichen hohe Lautstärkeanteile haben, was zu einem verwaschenen und auch in der Regel schlecht bewerteten Sound führen kann, wenn diese zusammen abgespielt werden. Wie Abbildung 29 auf Seite 78 zeigt, besteht der Beispielsong sogar aus 21 verschiedenen Spuren, die aber *gemeinsam* gut klingen sollen. In diesem Kapitel sollen verschiedene Audioeffekte beschrieben werden, die dazu dienen, Mängel der beschriebenen Art zu beseitigen. Diese Effekte manipulieren zum einen die Gesamtdynamik, sowie die Dynamik einzelner Frequenzen, oder fügen Klangcharakteristika auf Grundlage des Basismaterials hinzu. Zusätzlich wird beschrieben in welchem Umfang diese Effekte an welchen Spuren angewandt wurden.

4.3.1 Dynamiken

Ein wichtiger Bereich der postmanipulativen Effekte sind jene, die auf Grundlage der Lautstärke des Eingangssignal (in der Regel die rohe Audiospur) Einfluß auf die Lautstärke des Ausgangssignals nehmen. In Hostprogrammen wie Cubase sind diese Effekte, die ebenfalls über die VST Schnittstelle angesprochen werden, als „Dynamics“ bezeichnet und enthalten für gewöhnlich Gate, Kompressor und Limiter.

Ein Gate betrachtet zur Echtzeit die jeweils aktuelle Lautstärke des Eingangssignals. Fällt diese Lautstärke unter eine gewisse, durch den Tontechniker vorgegebene Schwelle (englisch: Threshold) angegeben in Dezibel, dann wird das Ausgangssignal auf stumm geschaltet. Abbildung 30 zeigt die Funktionsweise eines Gates:

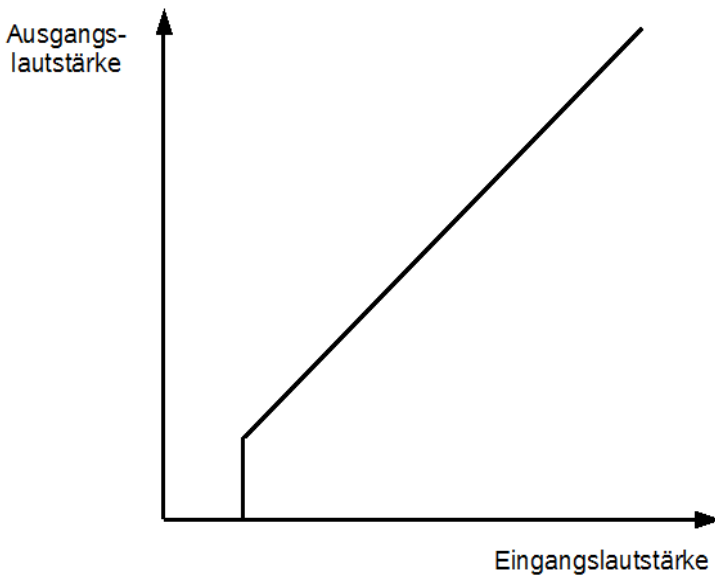


Abbildung 30: Gate

Zusätzlich zum Threshold, der in der Abbildung die Position des vertikalen Verlaufes bezüglich der X-Achse bestimmt, enthält ein Gate in aller Regel noch Attack- und Release-Regler. Diese bestimmen die Zeit in Millisekunden, nach denen das Gate schließt (Attack), bzw. wieder öffnet (Release). Ein Gate kann dazu gebraucht werden, um aufnahmebedingtes Rauschen zu unterdrücken. Das eigentliche Nutzsignal, das weitaus lauter als das Rauschen ist, kann das Gate passieren, das Rauschen selbst wird jedoch herausgefiltert, sofern es unterhalb der Schwelle liegt. Um unangenehme und abgehackt klingende Effekte bei relativ betrachtet leisen und kurzen Passagen des Nutzsignals zu vermeiden, kann die Attack-Zeit etwas größer gewählt werden. dasselbe gilt umgekehrt für die Release-Zeit: soll zum Beispiel ein Knacksen, wie auf alten LP Aufnahmen üblich, gefiltert werden, dann könnte man die Release Zeit geringfügig höher

einstellen, als die Zeit, die ein Knacksen in Anspruch nimmt, damit es das Gate nicht öffnet.

Im Beispielsong wurde vor allem die Sologitarre mit einem Gate gefiltert, da die Tonabnehmer einer E-Gitarre prinzipbedingt ein Rauschen erzeugen. Durch das Gate wurde dies nachträglich wieder entfernt.

Ein Kompressor ist, bezogen auf die Hörgewohnheiten im 21. Jahrhundert, der wohl wichtigste Effekt, um im Ohr des naiven Zuhörers einen professionellen Klang zu erzeugen. Tatsächlich wird ein Kompressor höchst ausgiebig in moderner Popmusik angewandt. Abbildung 31 zeigt die Arbeitsweise eines Kompressors:

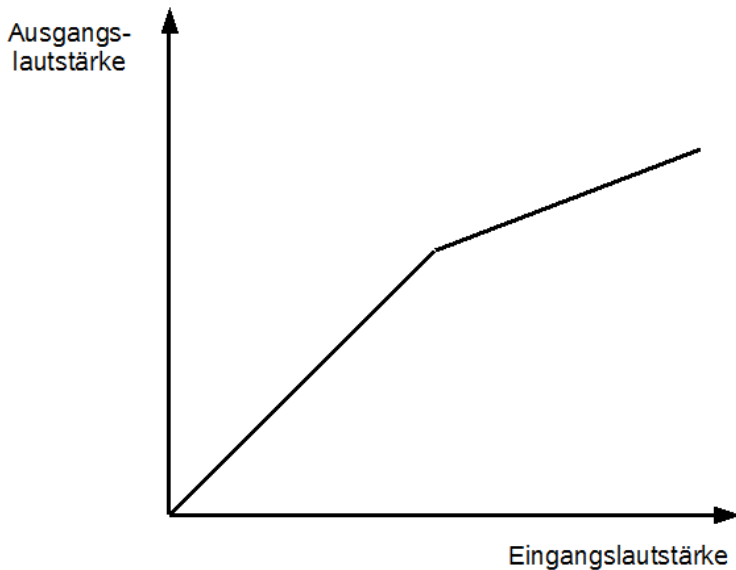


Abbildung 31: Kompressor

Wie bereits beim Gate entscheidet auch beim Kompressor ein vom Tontechniker

eingestellter Threshold über Aktivierung und Deaktivierung des Effekts. Ebenso entscheiden Attack- und Releasezeiten über die Zeiträume, nach denen bei Eintritt der Bedingung (= Überschreitung oder Unterschreitung des Thresholds) der Effekt angewandt (bzw. deaktiviert) wird. Wie in der Abbildung zu sehen ist, „knickt“ die Gleichheit von Eingangslautstärke und Ausgangslautstärke an einem gewissen Punkt (dem Threshold) ab. Anders ausgedrückt: das Ausgangssignal wird ab einer bestimmten Lautstärke weniger schnell laut, als das Eingangssignal. Das Verhältnis, in dem dies geschieht, wird über einen zusätzlichen Regler an einem Kompressor eingestellt, dem sogenannten „Ratio“. Übliche Werte sind hier 1:1, 1:2, 1:4, 1:8 und 1:∞, wobei dies je nach Kompressor auch stufenlos regelbar sein kann. 1:1 entspricht dabei keinem Kompressor (in der Abbildung also einer durchgezogenen Winkelhalbierenden) und 1:∞ verhindert konsequent, dass das Ausgangssignal über einen bestimmten Pegel gehen kann. In diesem Falle spricht man auch von einem Limiter. Siehe Abbildung 32.

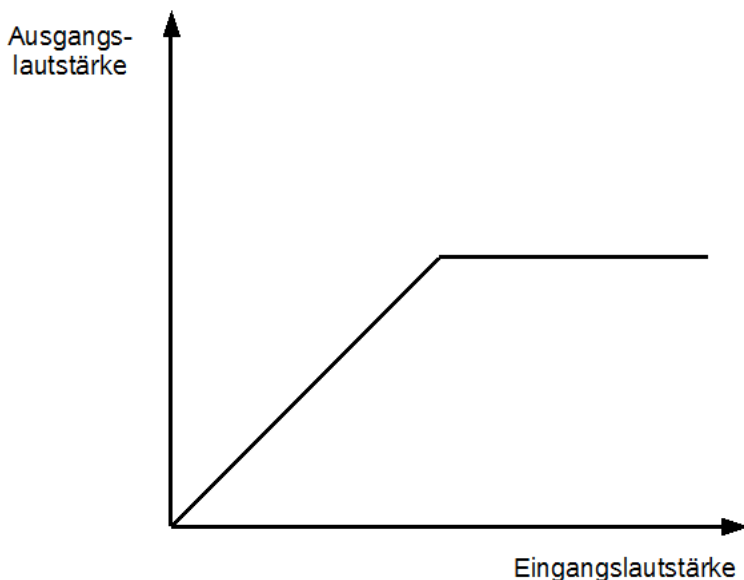


Abbildung 32: Limiter

Der Limiter ist also ein Kompressor mit einer Ratio von 1 zu Unendlich.

Praxisnah ausgedrückt regelt ein Kompressor laute Passagen herunter, ohne dabei bereits leise Passagen in ihrer Lautstärke zu reduzieren. Ein letzter, zusätzlicher Regler, oftmals beschriftet mit „Gain“ oder „Output Gain“, hebt die Gesamtlautstärke des Ausgangssignal wieder an, um die Lautheitsverluste bei der Kompression zu kompensieren. Eine Art Faustregel ist also, dass ein Kompressor laute Stellen leise und leise Stellen laut macht.

Besonders der Gesang profitiert von der Anwendung eines Kompressors, da die menschliche Stimme naturgemäß stark in ihrer Lautstärke variiert. Dies ist nicht nur der Fall, weil es in einem Lied leise und laut gesungene Passagen geben kann, sondern auch weil der Abstand des Sängers zu seinem Mikrofon selbst im Studio

keine Konstante sein kann. Es gibt in moderner Musik kaum Tonquellen, die nicht komprimiert werden. Höchstens experimentelle Jazzaufnahmen verzichten bewußt auf die Anwendung eines Kompressors.

Alte Aufnahmen von Bands wie den Beatles, Rolling Stones oder Deep Purple verwenden noch keinen oder kaum Kompressor, wodurch der Gesang für heutige Hörgewohnheiten ungewohnt bis penetrant klingen kann. Ein Kompressor gleicht die Lautstärkeverhältnisse innerhalb eines Signals an und macht vor allem hochdynamische Instrumente, zu denen man vor allem die Stimme und Blasinstrumente zählen kann, angenehmer hörbar.

Im Beispielsong kam ein Kompressor vor allem bei den Gesangsspuren zur Anwendung, um unterschiedlich laute Passagen, auch bedingt durch unterschiedliche Tonhöhen (hohe Töne müssen gewöhnlich lauter gesungen werden), einander anzugleichen. Verzerrte Gitarren, wie Sologitarre oder Rhythmusgitarre, überschreiten ein gewisses Lautstärkemaß nicht und erscheinen in Bezug auf ihrer Dynamik sehr homogen. Daher ist die Anwendung eines Kompressors hier in der Regel nicht nötig. Das Schlagzeug benutzt ebenfalls einen Kompressor, der in dem oben beschriebenen externen Klangerzeuger zu finden ist. Wie bei perkussiven Instrumenten nachvollziehbarerweise üblich, sind hier die Attack- und Releasezeiten sehr gering zu wählen.

VSTi bringen möglicherweise in ihrem Benutzerinterface eigene Kompressoren mit oder die benutzten Samples selbst sind bereits ausreichend komprimiert. Jedenfalls ergab sich im Rahmen des Beispielsongs keine Notwendigkeit für die Anwendung eines Kompressors bei MIDI Spuren.

4.3.2 Frequenzen

Eine weitere Möglichkeit, das Klangbild der Einzelspuren deutlich zu verändern, ist der Equalizer. Viele Audiogeräte für Endanwender besitzen jeweils Einstellmöglichkeiten für hohe, mittlere und tiefe Frequenzen. Im Bereich der professionellen Audioproduktion stehen dem Tontechniker selbstverständlich viel präzisere Manipulationsfenster zur Verfügung, um gezielt kleinere Frequenzbereiche innerhalb der vom Menschen hörbaren 20 kHz abzdämpfen oder hervorzuheben. Abbildung 33 zeigt die Equalizereinstellung einer Spur, die nicht verändert wurde und allen Frequenzen des Zielsignals die gleichen Lautstärken des Quellsignals zugesteht:

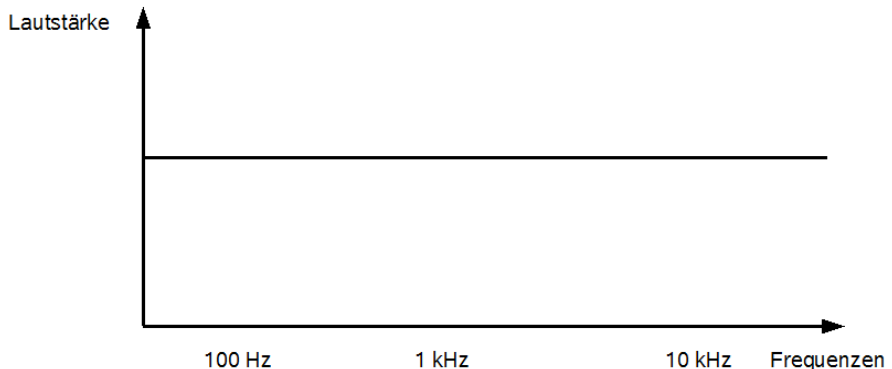
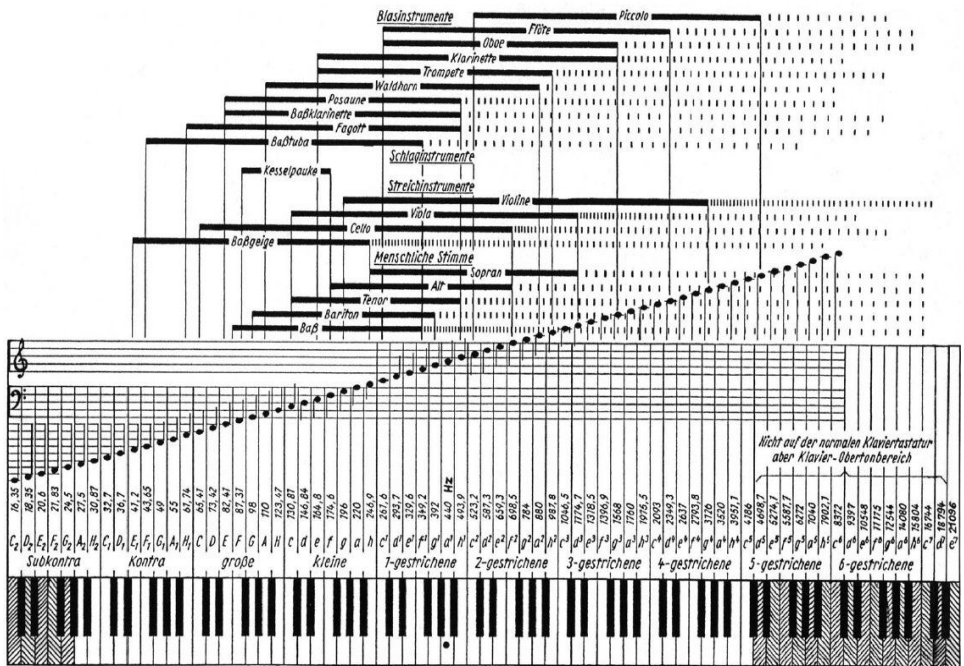


Abbildung 33: Neutraler Equalizer

Gemäß der durch den Menschen wahrnehmbaren Wichtigkeit ist die Frequenz logarithmisch dargestellt. Jedes Musikinstrument besitzt einen ungefähr festlegbaren Bereich, in dem es hauptsächlich Töne spielen kann. Laut Quelle [8] beginnt dieser Bereich bei etwa 47,2 Hertz (Ton E₁) der Bassgeige, die ein Frequenzband bis 264 Hertz (Ton h) einnimmt. Die Violine beginnt bei etwa 196

Hertz und endet bei ca. 3700 Hertz, was in etwa einem f^4 entspricht. Die Piccolo Flöte ist das klassische Instrument mit der höchsten Tonlage, welche bei ca. 4200 Hertz (c^5) endet. Die menschliche Stimme umfasst in etwa den Bereich von 87 Hertz (F) bis 1100 Hetz (c^3 ; Bass: 87 bis 350, Bariton: 98 bis 400, Tenor: 130 bis 500, Alt: 174 bis 700, Sopran: 250 bis 1100 Hertz). Siehe Abbildung 34.



Um herauszufinden, welche Frequenzen am Klang einer Audioquelle beteiligt sind, ist es nötig, das Quellsignal einer Spektralanalyse zu unterziehen. Hierbei werden auf Grundlage der Eigenarten der verschiedenen Amplituden die beteiligten Frequenzen ermittelt. Die mathematische Grundlage hierfür bildet die Fourier Analyse, die allerdings nur bei kontinuierlichen Signalen funktioniert. Die Signale, die gemäß den Kapiteln 3.1 und 3.2 dieser Arbeit digitalisiert wurden, liegen in diskreten Werten vor, weswegen eine diskrete Fouriertransformation (DFT) vorgenommen werden muß. Diese ist wie folgt definiert:

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

N steht hierbei für die Gesamtanzahl der Samples des digitalen Audiosignals und $x[n]$ ist demnach ein spezielles Sample. Alle Samples werden mit einem Oszillator mit der Frequenz k gewichtet. $X[k]$ stellt also den Anteil der Amplituden dar, die einem bestimmten Frequenzkanale k entsprechen.

In dieser Form ermittelt die DFT die Gewichtung des Frequenzkanales k über das gesamte Signal, was zur Ermittlung aller beteiligten Frequenzen auch sinnvoll ist. Ist man interessiert am Echtzeitverlauf der verschiedenen Frequenzen, muß man das Signal in kleine Fenster unterteilen und die DFT jeweils erneut anwenden. Dies kann allerdings nur mit einer gewissen Genauigkeit passieren, weil die Fensterlänge nicht die Länge einer tieffrequenten Amplitude unterschreiten darf (Unschärferelation).

In Audioproduktionsprogrammen wie Cubase oder Samplitude steht die DFT gewöhnlicherweise als Fast Fourier Transformation (FFT) zur Verfügung, die einen weitaus geringeren Rechenaufwand erfordert.

Durch eine solche Analyse und durch das Wissen, welche Frequenzen ein

bestimmtes Instrument für gewöhnlich einnimmt, kann man die Frequenzen einer Audiospur gezielt manipulieren.

Wie genau die Equalizereinstellung für eine bestimmte Audiospur vorgenommen wird, unterliegt letztlich dem Geschmack des Tontechnikers, bzw. des Hörers. Wichtig ist bei dieser Art der Manipulation vor allem, Frequenzen nicht derart anzuheben, dass sie übersteuern oder unangenehm klingen, was vor allem in hohen Frequenzen schnell der Fall sein kann.

Darüberhinaus kann es sinnvoll sein, Instrumente, die nur in einem bestimmten Frequenzbereich Töne produzieren, in den an anderen Bändern abzdämpfen. Ein Beispiel hierfür können die Base Drum oder der Bass sein, die nur relativ tiefe Frequenzen erzeugen. Die Anwendung eines Tiefpassfilter könnte in diesem Falle störende Nebengeräusche, die mit aufgezeichnet wurden und möglicherweise laut genug sind, um ein eventuelles Gate zu triggern, beseitigen. Abbildung 35 zeigt einen solchen Tiefpassfilter, der, wie der Name vermuten lässt, nur tiefe Frequenzen passieren lässt.

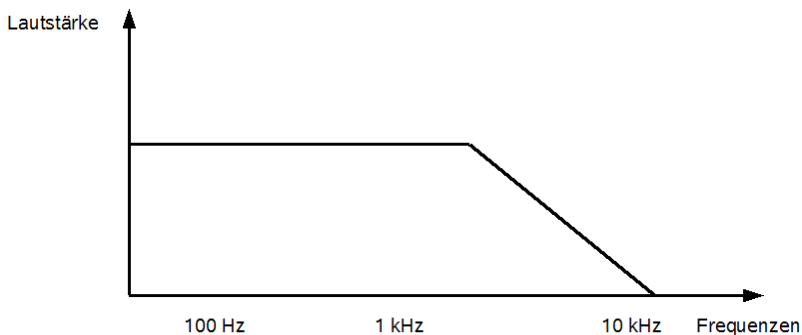


Abbildung 35: Tiefpassfilter

Die genaue Frequenz, ab der die nachfolgenden, höheren Frequenzen gedämpft werden und in welcher Intensität, kann dabei situationsabhängig sein und vom

Tontechniker gewählt werden. Bevor man einen Tiefpassfilter auf ein Instrument anwendet, sollte man aber bedenken, ob das Instrument vielleicht bestimmte hohe Frequenzen für seinen Klang benötigt. Im Falle des Basses könnten dies die weiter oben angesprochenen Quietschgeräusche bei der Bewegung der Finger über die Seiten sein. Bei der Basedrum könnte dies der Anschlag des Schlegels auf das Fell der Basedrum sein. Auch hier entstehen hohe Frequenzanteile, die eventuell charakteristisch sein könnten und durch einen Tiefpassfilter entfernt würden.

Je nach Anwendungsfall kann ein Tiefpassfilter auch als Höhengrenze, Höhenfilter, Treble-Cut-Filter, High-Cut-Filter oder Rauschfilter bezeichnet werden.

Natürlich gibt es auch eine hochfrequente Entsprechung des Tiefpassfilters: den Hochpassfilter. Abbildung 36 zeigt einen solchen.

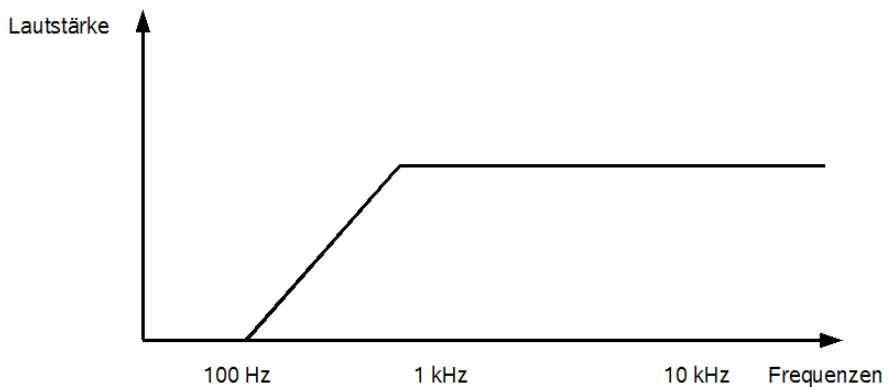


Abbildung 36: Hochpassfilter

Hochpassfilter, anwendungsbezogen auch bezeichnet als Tiefensperre, Bassfilter, Low-Cut-Filter, Bass-Cut-Filter, Trittschallfilter oder sogar Rumpelfilter, entfernen die tiefen Frequenzen aus einem Audiosignal. Bei Instrumenten wie der

E-Gitarre kann ein solcher Filter sehr sinnvoll sein, da viele Gitarristen dazu neigen, einen sehr basslastigen Sound bei ihrer Verzerrung einzustellen. Allerdings „konkurrieren“ diese Frequenzen dann mit denen des Basses, wodurch ein undifferenzierter Sound entsteht. Dies geschieht nicht, wenn man der Gitarrenspur die tiefen Frequenzanteile herausfiltert. Zwar klingt die Gitarre alleingestellt dann relativ dünn, im Gesamtmix dafür umso besser.

Eine letzte Art der Bandfilter ist das Notch, auch genannt Bandsperre oder Bandstoppfiler. Im Gegensatz zum Hoch- oder Tiefpassfilter liegt das blockierte Frequenzband nicht am oberen oder unteren Ende der Frequenzskala, sondern mittendrin. Abbildung 37 zeigt einen Notchfilter.

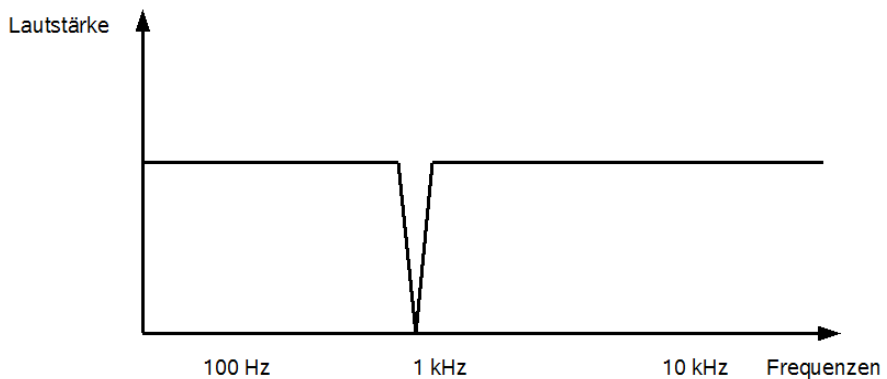


Abbildung 37: Notchfilter

Auch hier kann natürlich die Breite des gedämpften Bereichs und die Härte des Schnittes gewählt werden. Ein solcher Filter kommt in erster Linie bei Live Auftritten zum Einsatz, um bestimmte Frequenzen, die Rückkopplungen verursachen können, unterdrücken zu können. Bei Audioaufnahmen kann der Notchfilter eventuelle Störgeräusche präzise entfernen, sofern diese in einem bestimmten Frequenzband auftreten und in diesem Band keine oder

vernachlässigbar wenig Nutzinformationen liegen.

Wie bereits weiter oben erwähnt, unterliegt es – von diesen speziellen Equalizereinstellungen abgesehen – letztlich dem Geschmack des Hörers, welche Filtereinstellungen gut genug für das zugrundeliegende Audiomaterial klingen. Im Beispielsong wurde vor allem die Rhythmus E-Gitarre in ihren Höhen verstärkt, da der Verstärker in natura recht dumpf klingt. Aber auch fast alle anderen Instrumente wurden geringfügig in ihren Frequenzen manipuliert.

4.3.3 Pitch Shifter / Time Stretcher

Manchmal kann es nötig sein, die Tonhöhe einer Audiodatei nachträglich zu ändern. Dies erreicht man beispielsweise dadurch, dass man sie schneller oder langsamer ablaufen lässt. Im Verhältnis zur Abspielgeschwindigkeit ändert sich dann auch die Tonhöhe. Dies ist ein bekannter Effekt der akustischen Vorspulfunktion an alten Kassettenrekordern oder der Jog Funktion anderer Bandgeräte. Der Nachteil besteht hauptsächlich darin, dass sich sowohl die Geschwindigkeit der Audiodatei ändert, als auch ihr Klang: spielt man beispielsweise die menschliche Stimme schneller ab, entsteht der „Micky Maus Effekt“, der die Stimme klingen lässt, als hätte der Sprecher kurz zuvor Helium eingeatmet. In Hostprogrammen der modernen Audioproduktion arbeiten jedoch mächtige Algorithmen, die diesen Effekten in beeindruckender Qualität entgegenwirken können.

Abbildung 38 zeigt das Manipulationsfenster der Tonhöhe einer Audiodatei in Programm Cubase SX3.

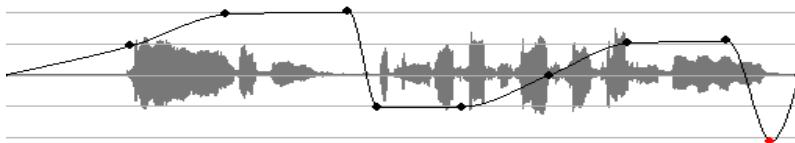


Abbildung 38: Pitch Shifter

Die eigentlichen Audiodaten sind grau dargestellt. Die horizontalen Linien stehen für Halbtonschritte und die frei durch den Tontechniker manipulierbare Linie gibt die Zieltonhöhe an. Die Audiospur in der Abbildung wird also nun zu Beginn einen Halbtonschritt höher wiedergeben, danach zwei Halbtonschritte höher und ab der Hälfte einen Halbtonschritt tiefer als das Original, und so weiter.

Wohlgemerkt handelt es sich hier um eine im Zeitverlauf frei wählbare Tonhöhenkorrektur, die die Abspielgeschwindigkeit *nicht* beeinflusst. Darüberhinaus kann man den sogenannten *Formant Modus* aktivieren, der zusätzlich bewirkt, dass die Stimmcharakteristik erhalten bleibt. Dies ist zwar rechenintensiver, funktioniert aber in beeindruckender Art und Weise: selbst bei einer ganzen Oktave Unterschied nach oben oder unten ist die Stimme eindeutig wiederzuerkennen, ohne dass ein Micky Maus Effekt o. ä. eintritt. Mit diesem Effekt ist es ohne weiteres möglich, einige verunglückte Töne des Sängers im Nachhinein präzise gerade zu rücken.

Wie weiter oben bereits angemerkt wurde, muß ein Sänger aber nicht nur die richtigen Töne treffen (wobei man ihm auch - wie gezeigt - technisch helfen kann), sondern auch ein gutes Timing besitzen. Mit den Möglichkeiten der digitalen Audioproduktion kann aber auch ein zu unpräzises Timing im Nachhinein korrigiert werden. Ein Time Stretcher ist ein Effekt, der die Abspielgeschwindigkeit ändert, ohne die Tonhöhe (und die Stimmcharakteristik)

zu beeinflussen. Er arbeitet also mit dem oben beschriebenen Pitch Shifter zusammen und kann in entsprechend hochwertigen Programmen ebenso selektiv angewendet werden. Abbildung 39 zeigt die Möglichkeit zur Timingkorrektur in einem Vorher- und Nachherbild:

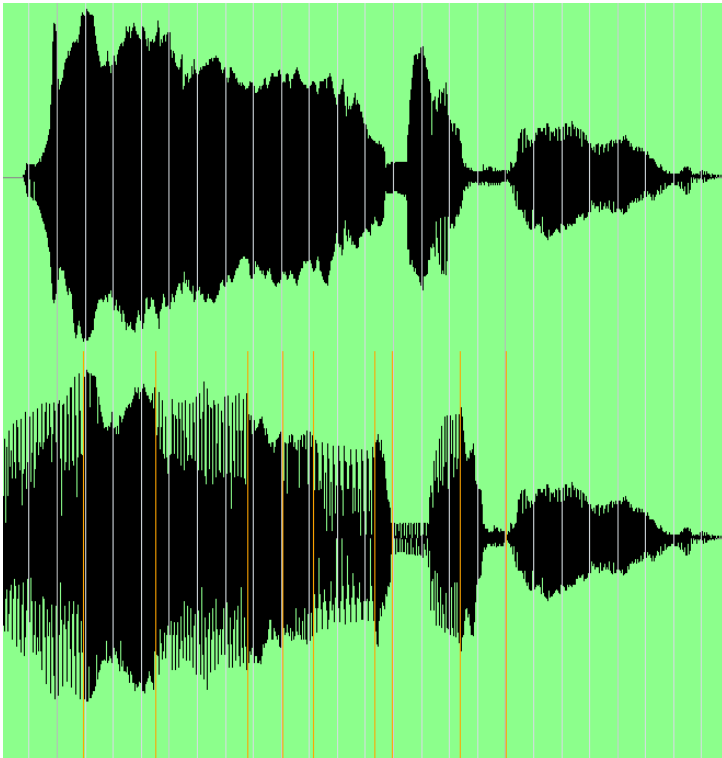


Abbildung 39: Time Stretcher

Oben sind die Original-Audiodaten zu sehen. Unten sind im zeitlichen Verlauf manipulierte Audiodaten zu sehen, die jeweils an einem Ankerpunkt (vertikale Linien) verschoben wurden. Dies geschieht per Drag and Drop. Ein praktischer Anwendungsfall wäre, wenn ein Sänger die entscheidenden Silben nicht genau auf

dem zugrundeliegenden Rhythmus gesungen hat. Diese Silben können akustisch identifiziert und mit einem Anker versehen werden. Danach werden sie schlicht weiter nach vorne oder nach hinten geschoben. Es ist so tatsächlich möglich, ein eher unpräzises Timing exakt auf den Beat zu setzen.

Da es in einem solchen Falle aber passieren kann, dass eine Passage der Audiodaten durch ihre starke Dehnung zu grob aufgelöst ist, ist es grundsätzlich sinnvoll, Audioaufnahmen mit mehr als der Zielsamplerate von 44,1 kHz aufzuzeichnen.

Sinn machen diese Effekte vor allem bei einem zeit- und ortsversetzten Aufnahmeablauf, in dem der Tontechniker nicht ohne weiteres eine zusätzliche Aufnahme vom Sänger verlangen kann. Kleine Fehler können so zuverlässig korrigiert werden.

Tatsächlich kam der Effekt des Time Stretchings mehrfach im Beispielsong zur Anwendung, da das Timing eines Sängers zwar gut, aber nicht perfekt war. Diese Möglichkeit der nachträglichen Korrektur eines bestehenden Audiofiles ist eine wichtige Eigenschaft der digitalen Audioproduktion, die man in der klassischen Audioproduktion wohl vergebens sucht. Die Qualität des Beispielsongs wäre letztlich sehr viel geringer gewesen, wenn das Timing des Sängers nicht präzise gewesen wäre. Ein Pitch Shifter kam nicht zur Anwendung.

Selbstverständlich können Time Stretcher und Pitch Shifter auch bei anderen Instrumenten angewandt werden.

4.3.4 Hall / Roomsimulator / Pre-Delay

Halleffekte verleihen den im Studio aufgenommenen Klängen im Nachhinein die Natürlichkeit, die eine trockene Aufnahme nicht besitzt. Hall oder auch Nachhall (Englisch: Reverb) ist die kontinuierliche Reflektion von Schall in geschlossenen Räumen. Die einfachsten Effekte dieser Art besitzen lediglich einen einzigen Regler, der die zeitliche Länge des Nachhalls regelt (die sogenannte Hallfahne). Die meisten Reverbeffekte besitzen jedoch zusätzlich einen Regler, mit dem man den lautstärkemäßigen Anteil des „trockenen“ und „nassen“ Signals einstellen kann.

Eine weiterentwickelte Form des Halleffektes ist der sogenannten Raumsimulator (Room Simulator), der die Reflektionscharakteristika ganz spezieller Räumlichkeiten akustisch nachzubilden versucht. Dies können Kirchen, kleine oder große Hallen, Konferenzräume oder sogar Garagen sein. Zitat [9]: *„Die grundlegende Idee besteht darin, dass man die Akustik eines schon bestehenden Raumes als Vorlage nimmt, um vorhandenes „trockenes“ Audiomaterial zu verhallen, so als ob es in eben diesem Raum aufgenommen worden wäre.“*

Ein Raumsimulator imitiert, wie auch höherwertige Halleffekte, zusätzlich den sogenannten Pre-Delay. Die Pre-Delay Zeit ist die Zeit, die vergeht, bis ein Klangereignis von den Wänden eines Raumes zum ersten Mal reflektiert wurde. Pre-Delay ist gleichbedeutend mit dem Begriff der frühen Reflektionen, die zwischen einem Klang und dessen Hallfahne auftreten. Durch die Manipulation der Pre-Delay Zeit in einer Aufnahmeumgebung ist es möglich, Schallquellen nachträglich räumlich zu positionieren. Je höher die Pre-Delay Zeit gewählt ist, desto weiter erscheint das Objekt, das das Klangereignis erzeugt hat, für den Zuhörer entfernt. Siehe auch Abbildung 40 auf Seite 98.

Im Beispielsong wurde in erster Linie der Gesang mit Hall versehen, zusätzlich

die Textzeile „Ki you ancient Tiamat...“ mit einer höheren Pre-Delay Zeit. Die akustisch aufgezeichneten E-Gitarren enthielten bereits Hall aus ihrem Effektgerät und das Schlagzeug aus seinem Klangerzeuger. Es wurde grundsätzlich darauf geachtet, ähnlichen Hall für alle beteiligten Instrumente zu vergeben, da diese sonst keine akustische Einheit bilden können.

4.3.5 Stereoeffekte

Nachdem die Akustikquellen nachträglich in ihrer virtuellen Entfernung zum Hörer durch das Pre-Delay positioniert wurden, sollten sie auch im Stereopanorama verteilt werden. Mit dem Pan bezeichnet man in der Tontechnik den Panorama-Regler. Also den Regler, dessen Einstellung darüber entscheidet, wieviel des Signals im linken, bzw. rechten Lautsprecher zu hören ist. Seine Normalposition ist demnach mittig zentriert. Abbildung 40 zeigt das Zusammenspiel von Pan und Pre-Delay:

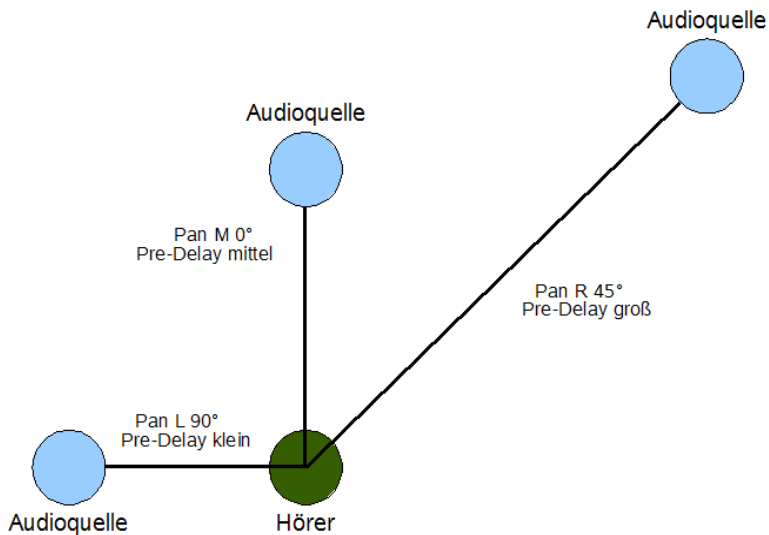


Abbildung 40: Pan und Pre-Delay

Vor der Anwendung des Pan liegen alle Audioquellen direkt frontal vor dem Hörer. Vor der Anwendung des Pre-Delay liegen sie dem Hörer sehr nah.

Alle mit einem Mikrofon aufgezeichneten Audiospuren sind naturgemäß Monokanäle. Durch ihre Positionierung links oder rechts vom Zentrum erreicht man einen differenzierten und breiteren Klang.

Im Beispielsong sind alle Instrumente leicht um die Panoramamitte verteilt. Insbesondere der Eingangsdialog der beiden Cheruben ist leicht links, bzw. leicht rechts gelegt. Enki singt zentriert. Die unterschiedlichen Solo E-Gitarren sowie die Rhythmus E-Gitarren sind komplett jeweils 90° links und 90° rechts angeordnet.

Dieser Effekt kann allerdings auch künstlich mit einem sogenannten Stereo Enhancer hergestellt werden, ohne dass der Musiker zweimal dieselbe Spur

aufzeichnen muß. Ein Stereo Enhancer (auch Stereo Widener genannt) wendet in der Regel folgende Effekte auf eine Audiospur an:

- Er legt zwei Kopien des Audiomaterials an
- Die beiden Kopien werden zeitlich leicht gegeneinander verschoben
- Eine Kopie wird 90° links, die andere 90° rechts ins Panorama gelegt.

Gewöhnlicherweise kann man die Verzögerung der Kopien und die Lautstärke des Originals im Kontrast zu den Kopien mit entsprechenden Reglern variieren. Der Effekt ist in Worten schwer zu beschreiben. Es entsteht ein voluminöserer Klang, der ein viel breiteres Spektrum einzunehmen scheint.

Die Entscheidung, im Beispielsong auf diesen Effekt bei den von Hand gespielten Gitarren zu verzichten und stattdessen zwei Spuren von Hand im Panorama zu verteilen, basiert letztlich auf einer Geschmacksfrage.

4.3.6 Weitere Effekte

An dieser Stelle sei noch kurz auf weitere Effekte eingegangen, die im Rahmen des Beispielsongs zur Anwendung kamen.

Allen voran erwähnenswert ist der De-Esser, der die stimmlosen Zischlaute (s, ss, ß, sch, z, tz) der Sänger dämpfen soll. Diese Zischlaute werden vor allem durch eine starke Kompression hervorgehoben und müssen gedämpft werden. Ein De-Esser senkt die charakteristischen Frequenzen im Pegel ab. Er hat meistens Presets für weibliche und männliche Zischlaute, die sich eindeutig zu

unterscheiden scheinen.

Ein Exciter ist ein Effekt, der im Beispielsong nicht zur Anwendung kam und hauptsächlich bei analogen Aufnahmen Anwendung fand, wenn nach oftmaligen Abspielen die Höhen verloren gingen. Besonders dumpfen Aufnahmen verhilft er zu mehr (empfundener) Klarheit, in dem er sehr hohe Frequenzen verzerrt und lauter macht. Tatsächlich kann es schnell passieren, dass der Klang der so bearbeiteten Aufnahme bei der Anwendung viel besser klingt, aber ein unangenehmes Gefühl in den Ohren hinterlässt. Wenn er dosiert angewendet wird, kann er dem Audiomaterial aber eine gewisse Brillanz verleihen.

Sogenannte Multibanddynamiken sind Kompressoren, die nur spezielle Frequenzbänder betreffen. Meistens sind dies grob die Bereiche Tiefen, Mitten und Höhen, es gibt aber auch Effekte, die viel mehr Frequenzbänder zulassen. Im Gegensatz zum normalen Equalizer werden also hier Frequenzbänder komprimiert, statt nur im Pegel abgesenkt. Das Ergebnis ist, wie bei Kompressoren üblich, dass vormals zu leise Signale in den entsprechenden Bändern lauter werden, was zu einem sehr differenzierten Sound führen kann. Im Beispielsong wurden Multibanddynamiken erst im Mastering angewendet.

4.4 Mix

Die in Kapitel 4.3 beschriebenen Effekte behandelten 4 Dimensionen einer Audioaufnahme:

- das Panorama (Pan/Stereoeffekte)

- den Frequenzbereich (Equalizer)
- den Raum (Hall und Stereo Enhancer)
- die Dynamik (Kompressoren)

Bobby Owsinski identifiziert in seinem Buch *Mischen wie die Profis* [10] noch zwei weitere Dimensionen:

- die Ausgewogenheiten (Balance der Pegel)
- das Interesse

dass die Instrumente untereinander den richtigen Lautstärkepegel haben sollten, versteht sich von selbst. Es macht keinen Sinn, einzelne Instrumente unter anderen gewissermaßen zu begraben. Dies wird nachfolgend noch genauer beschrieben.

Zur sechsten Dimension schreibt Owsinski auf Seite 99:

Wenn Sie die ersten fünf Elemente beherrschen, dann mag das für die meisten Audiojobs völlig genügen. Es wird auch ausreichen, um eine recht ordentliche Mischung abzuliefern. Aber in den meisten heutigen Musik-Genres wird erwartet, dass man dieses Level noch einmal deutlich überschreitet. Natürlich ist das immer leichter, wenn man mit fantastischen Tracks, einem soliden Arrangement und spektakulären musikalischen Darbietungen arbeitet. Aber ein hervorragender Mix kann durchschnittliche Tracks in Hitmaterial verwandeln, das die Leute so mitreißt, dass sie gar nicht genug davon bekommen können. Das ist mit Songs gemacht worden, die wahrscheinlich zu Ihren Lieblings-Hits zählen.

Diese Beschreibung Owsinskis sechster Dimension ähnelt dem Grundproblem einer guten Audioaufnahme: sie ist unkonkret. Das Mischen eines Songs ist keine

exakte Wissenschaft und neben dem Know How braucht ein Tontechniker (oder auch Toningenieur – wie auch immer man den Menschen bezeichnet, der das Rohmaterial vereint und ein einzelnes, gut klingendes Audiostück daraus produziert) sowohl eine Begabung, als auch das durch Owsinski vorgeschlagene Interesse.

Die Effekteinstellungen der einzelnen Spuren, die im vorigen Kapitel beschrieben wurde, stellt zu keinem Zeitpunkt des Mischprozesses einen endgültigen Stand dar. Im jetzt folgenden (vielleicht wichtigsten) Arbeitsschritt wird es oft nötig sein, Einstellungen an den Einzelspuren feiner vorzunehmen oder sogar ganz zu verwerfen. Schlußendlich ist es keine Frage, wie ein Instrument alleine klingt, sondern wie es im Kontext der anderen Instrumente klingt.

Es gibt viele verschiedene Vorgehensweisen, die alle zu einem Erfolg führen können, aber möglicherweise ist es sinnvoll, nachdem die Einzelspuren in ihrer leicht bearbeiteten Form vorliegen, sich zunächst auf das Fundament des Songs zu konzentrieren, das aus Schlagzeug und Bass besteht. Gemäß der sechs Dimensionen muß zuerst der Lautstärkepegel der beiden Instrumente angeglichen werden. Die Anwendung von Panoramafunktionen sind für das Fundament an dieser Stelle vernachlässigbar, denn das Schlagzeug selbst bringt bereits Panoramaeinstellungen mit, die durch den Klangerzeuger produziert wurden. So kann es zum Beispiel sein, dass die verschiedenen Toms unterschiedliche Positionen im Stereopanorama besetzen, um bei Fills einen räumlichen Effekt zu erzielen. Die Snare- und die Basedrum jedoch, als wichtigste Trommeln des Schlagzeuges, sollten im Center oder zumindest sehr nahe beim Center verblieben sein. dasselbe gilt für den Bass: auch dieser befindet sich fast immer im Center

einer Audioaufnahme. Das liegt vor allem daran, dass Menschen tiefe Frequenzen sowieso nicht lokal zuordnen können. Alle Schallquellen, die sich unterhalb von 80 - 90 Hertz abspielen, kann der Mensch nicht lokalisieren. Dies ist auch der Grund, wieso Subwoofer Boxen, also Lautsprecher für besonders tiefe Frequenzen an einer Stereoanlage, an einem beliebigen Ort im Raum aufgestellt werden können. Die dritte Dimension gibt vor, die Frequenzen der beiden Instrumente zu betrachten und gegebenenfalls voneinander zu trennen, wenn ihre Hauptausschläge sich im Spektrum zu nah beieinander befinden. Dies kann in der Regel nur ein Konflikt zwischen Basedrum und Bass sein. Owsinski schreibt:

Um die Wirkung und den Biss zu erzielen, den moderne Mischungen ausstrahlen, muss man im Mix Platz machen für diese beiden Instrumente, so dass sie nicht gegeneinander ankämpfen und der Song zu einem undefinierten Brei verkommt. Man könnte einfach die höheren Frequenzen des Basses und die tieferen Frequenzen der Bassdrum betonen (oder anders herum). Diese simple Strategie kann zwar funktionieren, doch ist es sicherlich besser, eine etwas ausgefeiltere Herangehensweise in der Hinterhand zu haben.

Danach beschreibt Owsinski grundlegend ein solches Vorgehen: die Basedrum soll im Bereich von 60 bis 120 Hertz angehoben werden, um sie auch auf kleineren Lautsprechern besser hörbar zu machen. Außerdem soll der Schlegelanschlag im Bereich von 1 und 4 kHz hörbarer gemacht werden. Frequenzen unterhalb von 30 Hertz (Basedrum) und unterhalb von 50 Hertz sollten komplett ausgeblendet werden. Betont werden muß, dass diese Empfehlungen nur grobe Daumenregeln darstellen können, da der Klang je nach verwendetem Instrument, Klangerzeuger oder Aufnahmeequipment variieren kann.

Halleffekte sind gegebenenfalls schon leicht auf der Schlagzeugspur vorhanden, nachdem sie, wie in Kapitel 4.2.1 beschrieben, aufgezeichnet wurde. Um den Bass und das Schlagzeug aber wie eine Einheit klingen zu lassen, sollten beide denselben (leichten) Halleffekt bekommen. Insbesondere werden hierbei die Geräusche charakterisiert, die nicht direkt mit dem gespielten Ton zusammenhängen. Also hochfrequente Anteile wie die Bewegung der Finger über die Saiten oder der Anschlag des Schlegels auf der Trommel.

Die Dimension Dynamik wird für das Schlagzeug mit einem entsprechenden Klangerzeuger ebenfalls bereits vor der Aufnahme erledigt. Eine weitere Kompression des Audiosignals kann aber dennoch sinnvoll sein. Der Bass mußte im Beispielsong kaum komprimiert werden, da seine akustischen Signale von einem MIDI Klangerzeuger stammten. Ein natürlicher Bass muß in der Regel stark komprimiert werden, da die verschiedenen Töne des Basses unterschiedliche Lautstärken erzeugen.

Am Ende dieses ersten Mixschrittes sollte man einen Drum und Basstrack haben, der keine Wünsche übriglässt. Nun kann man sich den anderen Instrumenten widmen.

Bei Gitarren ist es wichtig, die niedrigen Frequenzen, die einen mächtigen Sound ausmachen, per Equalizer herauszufiltern, da dies der Bereich des Basses ist. Wenn beide Instrumente im selben Frequenzspektrum agieren, führt das zu einem undifferenzierten Sound. dasselbe gilt auch für den hohen Bereich. Owsinski schreibt auf Seite 66 seines Buches[10]:

Wenn Sie zu viele Bässe in einem verzerrten Gitarrensound lassen, konkurrieren

diese mit der Rhythmussektion, und zu viele Frequenzen über 8 kHz geraten sich mit den Becken in die Haare.

Im Panorama wurden die Gitarren des Beispielsongs folgendermaßen verteilt: von Hand gespielte Gitarre 1 45 Grad links, Gitarre 2 45 Grad rechts. MIDI Gitarre 1 und 2 verblieben im Center.

Allen Rhythmusgitarrenspuren wurde derselbe Halleffekt zugeteilt, wie dem Bass, um eine Einheit zu erreichen.

Die Kompression einer verzerrten Rhythmusgitarre ist nicht immer nötig, da die Dynamik einer Verzerrung schon eine fest definiertes Maximum besitzt. Meistens reicht es stattdessen, die Lautstärke entsprechend dem Rest anzupassen. Natürlich kann es aber auch Sinn machen, Gitarrenspuren zu komprimieren. Owsinski, Seite 98:

Wenn Sie bemerken, dass die Gitarre hier und da im Mix untergeht, versuchen Sie, sie mit einer Ratio von 2:1 oder 4:1 zu komprimieren und stellen Sie mittlere Attack- und Release-Zeiten ein. Dies bewirkt, dass das rhythmische Einschwingen unbearbeitet durchkommt, während die Sustain-Resonanzen der Gitarren angehoben werden.

Das Ziel ist, dass Rock-Gitarren im Mix groß und „in your face“ erscheinen. Sie können das mit einer Ratio von 10:1 oder noch höher erreichen, was die Einschwingphase des Signals eliminiert. Stellen Sie sicher, dass das Sustain des Signals wieder auf den Ursprungspegel (1:1) zurückkehrt.

Nun sollte ein grober Mix der Rhythmusinstrumente vorliegen. An dieser Stelle sei angemerkt, dass der Mischvorgang sich nicht zufällig an den

Ablaufdiagrammen der Aufnahme orientiert, wie sie auf Seite 11 ff. zu sehen sind. Es ist grundsätzlich sinnvoll, sich im Frequenzspektrum nach oben durchzuarbeiten. Schlagzeug und Bass bilden das Fundament, die Rhythmusgitarre unterstützt dies später. Erst dann kommen Sologitarren, Ambient und Sänger hinzu.

Wie bereits weiter oben beschrieben, wurden die Sologitarren, die ihre eigenen Effekte mitbrachten, im Beispielsong hart links und hart rechts ins Panorama gelegt und sowohl von der Lautstärke, als auch von ihrer Gewichtung im Equalizer angepasst. Bei einer Mono Version des Songs führte die Panoramaeinstellung allerdings dazu, dass die Sologitarren kaum zu hören waren, wodurch ihr Winkel von -90 und 90 auf -44 und 41 verringert wurden.

Die Refrainstreicher wurden ohne Nachbearbeitung lediglich per Stereo Enhancer verbreitert.

An dieser Stelle liegt eine instrumentale Version des Songs vor, in die jetzt der Gesang integriert werden muß.

Gesang muß zunächst komprimiert werden, um eine gleichmäßige Lautheit der einzelnen Worte zu erreichen. Danach wird er gewöhnlich anhand der Instrumentallautstärke angepasst und mit einem Hall versehen. Diese Kombination an Effekten führt zum Zischen der stimmlosen Sprachanteile, was durch einen De-Esser beseitigt werden muß. Gesangsstimmen verbleiben meist im Center. Im Beispielsong jedoch belegt lediglich der Protagonist das Panoramazentrum, die beiden Nebencharaktere Cherub 1 und Cherub 2 wurden

leicht nach links bzw. rechts gelegt. Darüberhinaus muß, je nach Aufnahme und Stimme, der Equalizer angepasst werden, bis sich die Stimme deutlich durchsetzt, ohne zu laut zu sein. Owsinski gibt als grundlegende Tipps auf Seite 66 seines Buches[10]:

Heben Sie 125 bis 250 Hz ein wenig an, um das Fundament der Stimme zu betonen. Zudem klingt es dann mehr „aus dem Brustkorb heraus“. Der Bereich von 2 bis 4 kHz betont die Konsonanten, und die Stimme erscheint näher am Zuhörer.

Die in diesem Kapitel beschriebenen Vorgehensweisen spiegeln die grundsätzlichen Handlungen wieder, die vorgenommen wurden oder vorgenommen werden sollten. Selbstverständlich ist ein bestimmtes Instrument im Rahmen des Mischvorganges nie fertig und unantastbar. Im Gegenteil: weitere Instrument verändern den Klang und das in Gebrauch befindliche Frequenzspektrum, was zu unzähligen nachträglichen Anpassungen führt. Die leichte Änderung an einem einzigen Instrument hat zwangsläufig kleine bis große Auswirkungen auf alle anderen Instrumente. Dadurch müssen auch diese wieder angepasst werden.

Abbildung 41 zeigt alle Mixerspuren im Beispielsong. Zu erkennen sind neben den Einzelspuren auch Gruppenspuren und VST Ausgänge, sowie ein Effektkanal für Hall. Man sieht die Lautstärkeeinstellungen der Fader, die Panoramaverteilungen und das Signalarouting. Kurze Erläuterung der Knöpfe eine Signalzuges: M/S = Spur ist gemutet oder solo. R/W = Spur liest oder schreibt Automationsdaten (Die Spuren mit einem aktiven R blenden sich zum Endes des Songs hin aus). e = öffnet weitere Einstellungen für Equalizer, Insert- und

Sendeffekte, etc. Danach folgen Inserteffektstatus, Equalizerstatus und Sendeffektstatus. Diese leuchten jeweils, wenn Einstellungen vorhanden und aktiv sind. Gruppen- und Effektspuren enthalten keine weiteren Anzeigen. MIDI Spuren enthalten einen zusätzlichen Knopf zum Aufruf des VST UI und Audiospuren einen Knopf zur Aufnahme- und Monitoraktivierung.



Abbildung 41: Mixereinstellungen des Beispielsongs

Besonders wichtig hervorzuheben ist, dass der Song selbst in dieser Form extrem manipuliert werden kann. Jede Bassnote liegt nach wie vor als MIDI Datum vor, ebenso wie zwei Gitarren und die Streicher im Refrain. Lediglich die Audiospuren sind im Rahmen ihres Inhaltes fix. Sie können zwar in ihrem Klang, ihrer Tonhöhe und ihrer Geschwindigkeit und Länge bearbeitet werden (siehe Kapitel Effekte), nicht jedoch nachträglich andere Melodien spielen/singen. Während also MIDI Spuren von der Aufzeichnung bis zum Mastering absolut flexibel sind, sind Audiospuren vor ihrer Aufnahme gemäß ihres Inhaltes flexibel und nach ihrer Aufnahme gemäß begrenzter Eigenschaften, die durch Effekte hinzugefügt werden können. Die Flexibilität von Audiospuren ist demnach lediglich additiv bedingt (d.h. Effekte können nur hinzugefügt werden). Im Rahmen der klassischen Audioproduktion wären hinzugefügte Effekte nicht mehr entfernbar. Stattdessen müsste man im Notfall auf eine Sicherheitskopie des trockenen Originals zurückgreifen. Die digitale Audioproduktion erlaubt es, die Effekte, die in Echtzeit dem trockenen Signal hinzugefügt werden, wieder zu entfernen oder zu manipulieren.

4.5 Mastering

Das Mastering bezeichnet im klassischen Sinne eigentlich lediglich die Übertragung des fertigen Songs auf ein Medium. Wenn man aber heutzutage von Mastering spricht, dann meint man meistens den Prozess des Pre-Masterings, der den fertigen Downmix eines Songs noch einmal in seinen Klangeigenschaften perfektioniert, bevor er auf CD (oder früher LP) gepresst wird. Owsinski schreibt auf Seite 116:

Mastering ist der Prozess, der eine Kollektion von Songs in eine professionelle Platte, bzw. CD verwandelt. Danach klingen sie in Bezug auf Ton, Lautstärke und Timing (den Pausen zwischen den Songs) so, als gehörten sie zusammen. [Mastering] ...ist eine Kunstform, deren Erfolg, wenn sie gewissenhaft durchgeführt wird, von dem Können, der Erfahrung in verschiedenen Musik-Genres und dem guten Geschmack des Mastering Engineers abhängt.

Das Mastering stellt also sozusagen den letzten Schliff eines Musikstückes dar und ist besonders wichtig, da es sicherstellen muß, dass der erste Eindruck beim Hörer (der ja bekanntlich der wichtigste ist) sehr gut ist. Gewissermaßen macht ein Mastering den Unterschied zwischen einem kommerziellen Musikstück und einer guten Demoversion aus. Owsinski schreibt:

Ein Projekt, das gemastert wurde (besonders von einem angesehenen erstklassigen Mastering-Studio), klingt einfach besser. Es klingt komplett, brillant und abgeschlossen. [...] Der Grund liegt darin, dass der Mastering-Engineer hochwertige Equalizer und Kompressoren eingesetzt hat, um Ihr Projekt größer, fetter, wertiger und lauter klingen zu lassen.

Das Mastering ist etwas, das in aller Regel von absoluten Profis vorgenommen wird. Owsinski führt dies aus:

Warum können Sie also nicht einfach die Mastering-Plug-Ins Ihrer digitalen Workstation aufrufen, statt zu einem teuren, professionellen Mastering-Studio zu gehen? [...] Das Mastering-Studio ist besser ausgerüstet. Es kann auf Geräte zurückgreifen, die Ihnen eine typische Workstation nicht bietet. [...] Tatsächlich kostet das Monitorsystem solcher Profistudios wahrscheinlich mehr als ein gesamtes Heimstudio. [...] Aber der wahre Schlüssel zum Erfolg ist die Person des Mastering Engineers. Diese Tätigkeit ist alles, was er tagein und tagaus

macht. Er hat „große Ohren“, weil er mindestens acht Stunden am Tag mastert, und er kennt seine Monitore so gut, wie Sie Ihr Lieblingspaar Schuhe. Außerdem ist seine Referenz – nämlich was eine gute, ausgewogene Produktion darstellt – sehr fein ausgeprägt, dank der tausenden und abertausenden von Stunden, die er mit den besten und den schlechtesten Mixes jedes Musik-Genres zugebracht hat.

Es ist zu erkennen, dass ein guter Mastermix keine triviale Angelegenheit darstellt, die wohl auch im Rahmen dieser Arbeit nicht erschöpfend behandelt werden kann. Es gibt viele Plug-Ins auf VST Basis für Hostprogramme wie Cubase, von denen behauptet wird, dass sie einen guten Mastermix ermöglichen. Allerdings ist der Unterschied zwischen einem kommerziellen Musikstück und einem selbst gemasterten Song sogar für ungeübte Ohren immernoch eklatant.

Eine der Aufgaben, die gemeinhin als besonders wichtig angesehen wird, ist die Anhebung der Lautheit eines Songs. Im Rahmen seiner Produktion wird dafür gesorgt, dass die Pegel der Einzelspuren stets weit unter 0 dB liegen (siehe Abbildung 41), was das absolute Maximum für ein digitales Audiosignal darstellt. So kann das Instrument sich innerhalb seiner Dynamik entfalten und es kommt zu keinem Zeitpunkt zu einer digitalen Übersteuerung (auch Clipping genannt). Nach einem Downmix besitzt der Song dann zwar eine Dynamik, die durchaus für den Song adäquat ist, seine Gesamtlautstärke hängt allerdings von dem höchsten Ausschlag eines bestimmten Instrumentes ab und diese ist in aller Regel klein. Menschen empfinden aber Songs, die lauter sind, als der Song, der vorher oder nachher gespielt wird, subjektiv als besser. Zumindest wird dies seit den 50er Jahren von Menschen aus der Musikbranche behauptet und daher wird seitdem versucht, die Lautheit (also die subjektiv empfundene Lautstärke) der Aufnahmen zu erhöhen. Abbildung 42 verdeutlicht den Effekt:

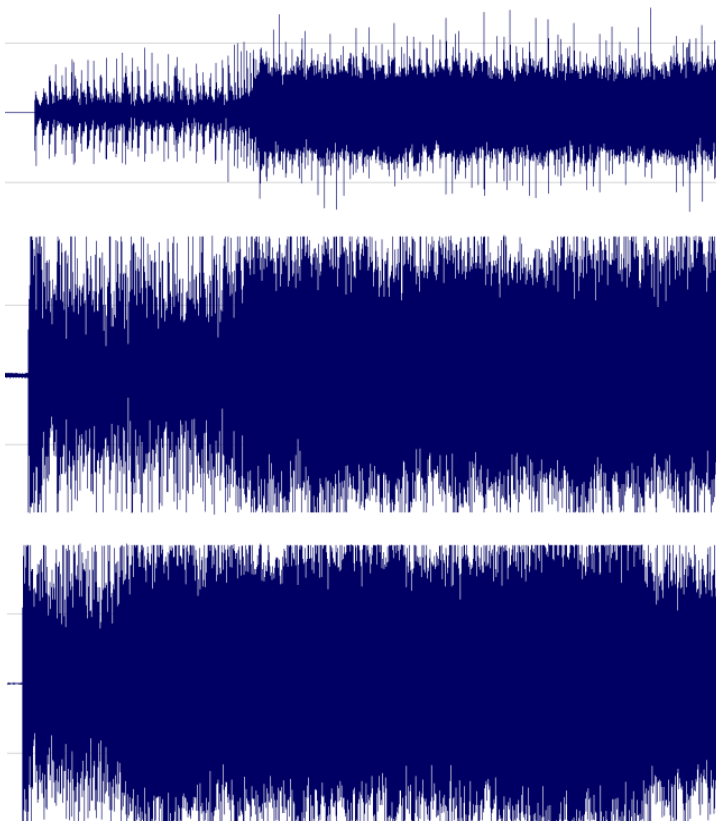


Abbildung 42: Dynamikschwund bei Überkompression

Oben in Abbildung 42 ist das Audiosignal zu sehen, wie es aus dem Hostprogramm nach einem Downmix herauskommt. Erkennbar sind die einzelnen Peaks und die relativ geringe Lautheit in Bezug auf den möglichen Umfang (die durchgezogenen Linien jenseits der Mittellinie liegt bei -6 dB). Die mittlere und die untere Darstellung zeigen dasselbe Signal, nachdem es durch einen Kompressor oder Limiter plattgedrückt wurde und durch eine Normalisierung

aufgezogen. Unter Toningenieuren ist diese Methode eigentlich eher unbeliebt, allerdings verlangt sie der Markt. Der Wettbewerb, Songs immer lauter und lauter zu machen, ohne eine Übersteuerung herbeizuführen, wird in der Branche bereits als „Loudness War“ bezeichnet. In einer Quelle [11] wird Mastering Ingenieur Bob Katz zitiert:

In den Achtzigern, als die CD sich durchsetzte, hatten wir bis zum Lautstärkelimit der CD bei 0 dB noch einen Headroom von 6 bis 8 dB. Dem Originalmix während des Masterings Lautstärke hinzuzugeben war deshalb ungefährlich. Viele Künstler wollen ihre Produkte lauter haben als die der Konkurrenz, weil sie sich davon mehr Aufmerksamkeit versprechen. [...] Ab 2000 war es nicht mehr möglich, den Pegel weiter anzuheben: Clipping wurde bewusst herbeigeführt. Das sind ernsthafte Verzerrungen, die mitunter die Anlage beschädigen können.

Ob und in wie weit man sich als Musikschafter diesem Trend der maximalen Aussteuerung unterwerfen muß, bleibt jedem selbst überlassen. Der Beispielsong wurde bewußt mit sanfteren Kompressoren weniger extrem in seiner Lautheit erhöht. Das resultiert darin, dass er nicht so laut ist, wie kommerzielle Songs.

Zusätzlich zur Erhöhung der Lautheit modifiziert ein Mastering aber auch den Klang des Songs. Dies wird zumeist unter dem Einsatz von Multibandkompressoren erreicht, die bereits weiter oben beschrieben wurden. Ziel ist es, den Lautheitsanteil der verschiedenen Frequenzbereiche in etwa an die auf dem Markt befindlichen Songs anzugleichen. Denn zu wenige Höhen klingen alleinstehend akzeptabel, aber im Wechsel mit anderen Werken mit mehr Höhen zu dumpf. Umgekehrt gilt dasselbe für einen zu wenig ausgeprägten Bassbereich, der eher einen schwachbrüstigen Eindruck beim Hörer hinterlässt oder einem zu stark ausgeprägten Mittenbereich, der schlicht ungewohnt und metallisch klingt.

Weiterhin kann ein Mastering die Breite des Songs beeinflussen. Dies kann unter dem Einsatz von Stereo Enhancern passieren, da das Panning in einer Stereosumme nicht mehr beeinflusst werden kann.

Der Beispielsong wurde aufgrund von mangelndem Equipment und unzureichender Erfahrung auf dem Gebiet des Masterings an ein professionelles Mastering-Studio übergeben.

5. Bestandsaufnahme und Qualitätsbewertung

Chapter 2 – The Quest for Gold, wie es in seiner finalen Version auf Enuma-Elish.de vorliegt, muß den Vergleich zu professionellen kommerziellen Aufnahmen nicht scheuen. Dies hat insbesondere folgende Gründe:

Die Musiker hatten eine sehr hohe Flexibilität in Bezug auf die Aufnahme ihrer Instrumente. Sie konnten einzelne Passagen unabhängig voneinander einspielen und im Nachhinein modifizieren, indem Sie Noten im MIDI Format geraderückten oder ins richtige Tempo setzten. Das Instrumentenspiel ist daher auch bei der hohen Geschwindigkeit des Beispielsongs fehlerfrei. Instrumente, die ohne MIDI Technologie aufgezeichnet wurden, konnten ebenfalls in kleine Passagen heruntergebrochen werden und theoretisch in gewissen Grenzen im Nachhinein modifiziert werden (Timestretching und Pitchshifting), auch wenn dies im Beispielsong nicht nötig war. Darüberhinaus konnten die Musiker ihr eigenes Aufnahmeequipment und ihre eigenen Räumlichkeiten nutzen und unterlagen so keinem unmittelbaren Zeitdruck. Die Überstellung des Materials geschah per Internet oder, bei entsprechender räumlicher Nähe, auf Medien wie

CDs oder USB Sticks. Im Rahmen der Instrumentierung handelt es sich also – vor allem dank der digitalen Audiotechnik und ihrer Flexibilität – um ein Werk, das die ursprünglichen Vorstellungen des Komponisten erfüllt oder sogar übertrifft.

Für den Gesang, mit dem die Qualität eines Songs im Ohr des Hörers meist steht und fällt, konnten zwei hervorragende Musiker gewonnen werden, die weitgehend autonom auf Grundlage der bereitgestellten Demoversion arbeiten konnten. Auch in diesem Aspekt wurden die Vorstellungen des Komponisten weit übertroffen.

Die klangliche Modifikation der Einzelspuren sowie der Gesamtmix ist in digitalen Hostprogrammen wie Cubase, Samplitude oder Logic mit den entsprechenden Plug-Ins und dem nötigen Know-How zu bewerkstelligen. Aufgrund der immensen Anzahl von Optionen (und Fehlern) ist es aber unmöglich, ähnlich im Rahmen der Instrumentierung oder des Gesangs, von „Perfektion“ zu reden. Stattdessen ist ein solcher Mix, der immer extrem zeitaufwendig ist (nicht nur aus dem Grunde, weil die Ohren nach einiger Zeit nicht mehr objektiv hören, sondern auch, weil Zwischenversionen immer wieder auf anderen Wiedergabegeräten wie Autoradios, großen Anlagen oder Laptoplautsprechern gegengehört werden müssen) stets nur ein Kompromiss zwischen dem, was der Komponist eines Rocksongs sich wünscht und dem, was möglich ist. So gibt es auch kein methodisches Vorgehen, sondern viel eher eine Art Herantastung an einen Mixsound, der dem (vermeintlichen) Optimum am nächsten kommt.

Das Mastering potenziert die beschriebene Problematik noch, weswegen nach einigen unzufriedenstellenden Versuchen professionelle Hilfe für diesen finalen Schliff in Anspruch genommen wurde.

Ein Direktvergleich mit aktueller Musik aus derselben Stilrichtung offenbart nach

wie vor den Unterschied zwischen einer industriell gefertigten Aufnahme von hauptberuflichen Musikern, Ton- und Masteringingenieuren und dem semiprofessionellen, aber ambitionierten Laien. Dieser Unterschied ist allerdings aufgrund der Mächtigkeit der digitalen Audioproduktion, mit digitalen Hostprogrammen, VST Instrumenten, MIDI-Technologie und Plug-Ins extrem geschrumpft. Vor 20 Jahren wäre dieselbe Aufnahme in dieser Form nur mit einem extremen Aufwand möglich gewesen: die Musiker hätten möglicherweise Jahre der Übung aufwenden müssen, um so zu spielen, wie auf der Aufnahme zu hören, der Sänger hätte von Kanada eingeflogen werden müssen und der Preis für das Equipment wäre statt im vierstelligen Bereich im sechststelligen gewesen. Chöre hätten ebenso engagiert werden müssen, wie Streicher und die Tonbandaufnahme wäre bei jedem Abspielen schlechter geworden, wodurch die Anschaffung weiterer Effekte vonnöten gewesen wäre..

Damit ist *Quest for Gold* möglicherweise ein Song, der einen hohen Qualitätsmaßstab in den frühen 90er Jahren gesetzt hätte. 2008 erkennt man als Musikliebhaber jedoch nach wie Unterschiede, die von Laien beim derzeitigen Stand der digitalen Technik nicht beseitigt werden können.

6. Musiker- und Zeitmanagement

Bei der Aufzeichnung eines Songs sind Menschen voneinander abhängig. Sie benötigen die Zuarbeit der anderen und müssen dementsprechend auch bei ihrer eigenen Arbeit Deadlines einhalten, um den Gesamterfolg des Projektes nicht zu gefährden. Insofern unterscheidet sich die Aufnahme eines Songs kaum von

anderen menschlichen Vorhaben, in denen Teamarbeit gefragt ist.

Im Falle des Beispielsongs bildeten die beteiligten Musiker darüberhinaus keine Band. Stattdessen handelte es sich bei den Mitarbeitern ausschließlich um Gastmusiker, die vorher in keiner Art zusammen musiziert haben. Insofern war das Projekt *The Quest for Gold* im Hinblick auf seinen Erfolg als Projekt ein interessantes Experiment.

Weiterhin ist zum Charakter der Zusammenarbeit der einzelnen Personen zu sagen, dass niemand (abgesehen vom Mastering-Studio) für seine Arbeit Geld oder anderweitige Zuwendungen bekommen hat. Die Mitarbeit fand auf rein freiwilliger Basis statt.

Der Zeitrahmen war auf 4 Monate festgesetzt, von Ende April bis zum 31. August.

Soweit also die nicht ganz opportunen Rahmenbedingungen, in denen der Beispielsong aufgezeichnet wurde. Abbildung 43 zeigt einen groben zeitlichen Ablauf des Projektes:

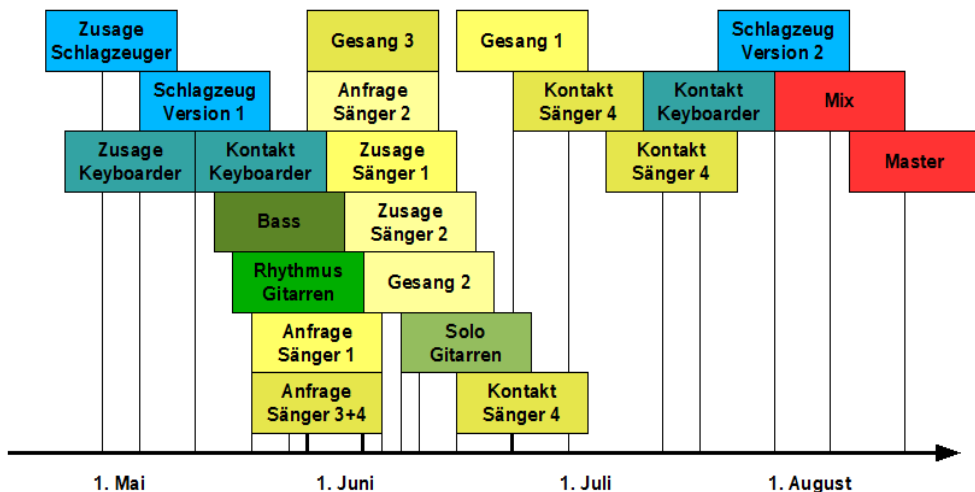


Abbildung 43: Zeitlicher Ablauf des Projektes

Dieser Ablaufplan enthält die ersten Anfragen an die Musiker, sowie deren Zusagen. Einzelne Instrumente sind farblich markiert. Wird nur der Name des Instrumentes angegeben, handelt es sich um die tatsächliche Aufnahme.

Gut zu erkennen ist, dass die Schlagzeugaufnahme Mitte Mai die Voraussetzung für die folgenden Arbeiten der Musiker war, wie auch in Abbildung 5 auf Seite 14 vorgeschlagen. Eine zweite, überarbeitete Version folgte ziemlich gegen Ende des Projektes im August.

Ein zu einem sehr frühen Zeitpunkt angefragter Musiker, der Keyboarder, wurde Ende Mai und Mitte Juli erneut kontaktiert, allerdings blieben Ergebnisse trotz fester Zusagen bis zuletzt aus.

Der Bass und die Rhythmusgitarren wurden auf Grundlage des Schlagzeuges zeitnah der Aufnahme hinzugefügt. Die nun Ende Mai vorliegende Rhythmusversion konnte für Gesang und Sologitarren benutzt werden.

Drei mögliche Sänger wurden daraufhin kontaktiert (vgl. Kapitel 4.2.7), von denen einer den Gesang sofort lieferte, durch die falsche Tonlage schaffte es diese Version jedoch nicht ins Endprodukt. Ein anderer wurde noch dreimal nach seiner Zusage erfolglos kontaktiert. Der dritte Sänger lieferte seine Version (Cherub 1 & 2) Ende Juni ab. Der kanadische Sänger der Hauptrolle (Enki) wurde Anfang Juni kontaktiert und erbrachte wenige Tage später alle geforderten Teile.

Die Sologitarren wurden Mitte Juni aufgezeichnet. Im gesamten Juli wurde erfolglos versucht, die Musiker, die zugesagt hatten, aber keine Ergebnisse lieferten, zu kontaktieren. Einhellige Antwort war stets eine sichere Zusage, die dann aber niemals eingehalten wurde. Außerdem wurden andere Musiker gesucht (nicht eingezeichnet), leider ebenfalls erfolglos.

Ende August wurden die Aufnahmen dann für abgeschlossen erklärt. Der Mix wurde am 16. August und das Mastering am 27. August vorgenommen.

Ohne die erfolglosen Versuche, angeworbene Musiker zu motivieren oder neue Musiker zu finden, hätte der Song schätzungsweise auch in der Hälfte der Zeit produziert werden können, vorausgesetzt, alle Musiker hätten effektiv zusammengearbeitet. Diese effektive Zusammenarbeit wäre durch eine entsprechende Bezahlung wahrscheinlich noch besser geworden.

Dem Autor dieser Arbeit und Komponisten des Beispielsongs standen jedoch keinerlei Druckmittel zur Verfügung, um die beteiligten Personen bei eventuellem Ausbleiben von Ergebnissen hinreichend zu motivieren. Da das Endergebnis aber dennoch mehr als zufriedenstellend vorliegt, kann es wohl auch als höchst positives Beispiel für menschliche Hilfsbereitschaft verstanden werden. Natürlich – wie immer – von ein paar wenigen Individuen abgesehen.

7. Persönliches Fazit und Schlußwort

Es wurde weiter oben schon erwähnt, dass das in dieser Arbeit behandelte Musikstück *The Quest for Gold* aus der Metal Oper *Enuma Elish* objektiv betrachtet nicht auf demselben Qualitäts- und Soundniveau gesehen werden kann, wie moderne kommerzielle Aufnahmen, die von Profimusikern gespielt und von Profi-Toningenieuren gemischt wurden. Stattdessen sind feine, aber nichtsdestotrotz vorhandene Unterschiede im Sound zu erkennen, die möglicherweise erst bei einer speziellen Konfiguration der Abhöranlage auffallen; dies kann eine Mono Wiedergabe sein, ein Autoradio oder die Wiedergabe in einem Raum, in dem es einen unvorteilhaften Abstand zwischen den Stereolautsprechern gibt. Ein Laie mit mäßigem Anspruch an die Soundqualität, sowohl die der Musik als auch die seiner Wiedergabegerätschaften, wird diesen Unterschied entweder nicht wahrnehmen oder sich davon nicht weiter gestört fühlen. Tatsächlich kann wohl angenommen werden, dass ein Großteil der Musik heutzutage im Auto, mit dem MP3-Player oder mit einem Computer wiedergegeben wird, der an kleine Stereolautsprecher aus dem Billigangebot eines Baumarktes angeschlossen ist. In einem solchen Rahmen ist die Qualität des Beispielsongs hervorragend, vor allem in Hinblick auf das technisch einwandfreie Spiel, das die MIDI Technik einigen Instrumenten eröffnet hat.

Auf einen audiophilen Menschen jedoch wird die im Rahmen dieser Arbeit erreichte Qualitätsstufe unzureichend wirken, besonders, wenn er ansonsten nur Musik aus sehr teuren Produktionen gewohnt ist, die er in verlustfreien Formaten auf seiner sehr teuren Anlage in einem akustischen vorteilhaft gestalteten Raum

hört. Die Frage mag allerdings augenzwinkernd erlaubt sein, wieviele audiophile Menschen Fans von Heavy Metal sind, der seinerseits zwar spielerisch durchaus virtuos sein kann, aber eher ein Sinnbild für rauhe Töne ist. Und rauhe Töne klingen nunmal selten butterweich.

Davon abgesehen stellt der Beispielsong aber für den Autoren dieser Arbeit einen Meilenstein dar, denn es handelt sich in Sachen Audioqualität um das mit einigem Abstand beste Werk, das er je produziert hat – und dieser Aussage liegen 10 Jahre letztlich unbefriedigender Audioaufnahmen zugrunde.

Eine wichtige Regel hat sich im Verlauf dieser Produktion aber als wahr erwiesen: wenn die Tonquelle schlecht klingt, dann wird das Endergebnis nicht viel besser klingen. In diesem Sinne bin ich überaus stolz und glücklich, soviele wunderbare Musiker für diese Arbeit gewonnen zu haben. Hervorzuheben sind hier vor allem der Schlagzeuger Jan Häger, der die Schlagzeugspur filigran und immer rechtzeitig perfektioniert hat, der Sänger Georg Laudenberg, der den Song schnell und unkompliziert durch seinen klaren Gesang veredelt hat, sowie Calvin Turko, den ich im Rahmen dieser Arbeit kennenlernen durfte und dessen Gesang mich stets auf's Neue beeindruckt. Die Hilfe solcher Leute in Anspruch nehmen zu dürfen, ist für mich als Komponisten eine große Ehre, denn sie haben es geschafft, meine Vorstellung des Songs in die Realität umzusetzen.

Aber auch dem Bassisten Daniel Gambal und dem Schlagzeuger Daniel Büttner bin ich zu Dank verpflichtet, denn sie haben mit mir zusammen das Arrangement des Songs erarbeitet.

Ohne diese Menschen wäre die Fertigung des Songs nicht möglich gewesen. Genauso wäre es aber unmöglich gewesen, ohne die Fortschritte der digitalen Technik, insbesondere auf dem Gebiet der Audioproduktion, ein Endergebnis zu erzielen, das dem Erreichten gleichkommt, ohne, vom Mastering abgesehen, auf die Hilfe von Profis angewiesen zu sein. *Das ist der Verdienst der Digitaltechnik.*

Die Formalisierung der Produktionsprozesse, die in dieser Arbeit vorgenommen wurde, bestätigt die neue konzeptionelle Flexibilität, die dem Musikschaaffenden im 21. Jahrhundert an die Hand gegeben wurde. Diese neuen Freiheiten eröffnen dem kreativen Kopf die Möglichkeiten, die er braucht, um sich frei entfalten zu können.

*Die Musik drückt aus, was nicht gesagt werden kann
und worüber es unmöglich ist zu schweigen.*

Victor Hugo (1802 - 1885)

Der Song „*Enuma Elish – Kapitel 02 – The Quest for Gold*“ steht zum kostenlosen Download in allen üblichen Audioformaten auf <http://www.enuma-elish.de/> zur Verfügung.

A. Glossar

BPM	Beats per Minute, Maß für die Geschwindigkeit eines Songs
FLAC	Free Lossless Audio Codec
Headroom	siehe <i>Übersteuerungsreserve</i>
LP	Langspielplatte
PCM	Puls-Code-Modulation; die Digitalisierung eines analogen Audiosignals.
SNR	Signal-to-Noise Ratio / Signal-Rausch Abstand
Übersteuerungsreserve	Ausgepegelter Abstand zur Volllaussteuerung.
e	Gewöhnlicherweise bei etwa -5 dB, um Übersteuerungen zu vermeiden.
VST	Virtual Studio Technology
VSTi	VST Instrument. Klangerzeuger auf Softwarebasis.

B. Literaturverzeichnis

- 1 Fielding DSP, o. J., **Fielding DSP**, <http://www.fieldingdsp.com/fieldingdsp/index.php?do=alias>, download am 10.07.2008
- 2 Mitra, S. K., Kaiser, J. F., 1993, **Handbook for Digital Signal Processing**, Wiley & Sons, ISBN: 0-471-61995-7
- 3 Nave, C. R., o. J., **Sensitivity of Human Ear**, <http://hyperphysics.phy-astr.gsu.edu/hbase/sound/earsens.html>, download am 14.07.2008
- 4 Resnick, P., 2001, **RFC 2822 Internet Message Format**, <http://tools.ietf.org/html/rfc2822#section-2.3>, download am 15.07.2008
- 5 Häger J., 2008, **Das Wichtigste auf einen Blick - Edrums**, http://www.ionstorm.de/Edrums/Das_Wichtigste_auf_einen_Blick, download am 28.07.2008
- 6 Steinberg GmbH, 2007, **History :: Steinberg Media Technologies GmbH**, http://www.steinberg.net/de/company/steinberg_company_history.html, download am 30.07.2008
- 7 Ögren, J, o.J., **DIN Audio – HwB**, http://www.hardwarebook.info/DIN_Audio, download am 31.07.2008
- 8 Telefunken GmbH, **Telefunken Laborbuch für Entwicklung Werkstatt und Service**, Band 1., 5. Ausgabe, Franzis-Verlag Münschen, 1962
- 9 Haefeli, R., 2003, **Raumsimulation durch Faltung**, http://romanhaefeli.net/Diplomarbeit_SAE_2003.pdf, download am 15.8.2008
- 10 Owsinski, B., 2007, **Mischen wie die Profis**, GC Carstensen Verlag, ISBN: 3910098363
- 11 Fust, B., 2008, **Der Loudness War - Kollateralschäden ohne Kriegsoffer**, http://www.intro.de/magazin/technik/23047104/der_loudness_war_kollateralschaeden_ohne_kriegsoffer, download am 20.8.2008