



**Masterarbeit**

im Studiengang Audiovisuelle Medien

mit dem Titel

**Entwicklungen der KI im Bereich der Filmkomposition – Vergleich einer KI-Komposition mit dem Score von Hans Zimmer anhand einer Filmszene**

vorgelegt von

Nina Schaffer

Matrikel-Nr. 44930

an der Hochschule der Medien Stuttgart am 30.09.2024

zur Erlangung des akademischen Grades eines Master of Engineering (M.Eng.)

Erstprüfer: Prof. Oliver Curdt

Zweitprüfer: Prof. Dr. Andreas Koch

## Ehrenwörtliche Erklärung

„Hiermit versichere ich, Nina Schaffer, ehrenwörtlich, dass ich die vorliegende Masterarbeit mit dem Titel: „Entwicklungen der KI im Bereich der Filmkomposition – Vergleich einer KI-Komposition mit dem Score von Hans Zimmer anhand einer Filmszene“ selbstständig und ohne fremde Hilfe verfasst und keine anderen als die angegebenen Hilfsmittel benutzt habe. Die Stellen der Arbeit, die dem Wortlaut oder dem Sinn nach anderen Werken entnommen wurden, sind in jedem Fall unter Angabe der Quelle kenntlich gemacht. Die Arbeit ist noch nicht veröffentlicht oder in anderer Form als Prüfungsleistung vorgelegt worden.

Ich habe die Bedeutung der ehrenwörtlichen Versicherung und die prüfungsrechtlichen Folgen (§26 Abs. 2 Bachelor-SPO (6 Semester), § 24 Abs. 2 Bachelor-SPO (7 Semester), § 23 Abs. 2 Master-SPO (3 Semester) bzw. § 19 Abs. 2 Master-SPO (4 Semester und berufsbegleitend) der HdM) einer unrichtigen oder unvollständigen ehrenwörtlichen Versicherung zur Kenntnis genommen.“

---

Ort, Datum

---

Unterschrift

## Kurzfassung

Die Entwicklungen der generativen künstlichen Intelligenz finden zunehmend Einfluss in der Musikproduktion. Neben diversen Plugins zur Audibearbeitung und Mischung tritt die Künstliche Intelligenz (KI) nun in das Feld der Komposition vor. Die entsprechenden Softwares ermöglichen dabei die Generierung verschiedenster Musiken, darunter auch die Erstellung von Filmmusik. In der vorliegenden Arbeit soll untersucht werden in welchem Umfang schon heute, im Jahr 2024, ein Score zu einem bestehenden Bewegtbild von einer KI erzeugt werden kann.

Dazu werden zunächst ausgewählte Grundlagen der Filmmusik und der generativen künstlichen Intelligenz beleuchtet. Der erste Schwerpunkt liegt auf der anschließenden Analyse des Scores einer ausgewählten Filmszene. Diese Erkenntnisse werden genutzt, um für dieselbe Szene eine neue Filmmusik durch eine KI-Software generieren zu lassen und bildet damit den zweiten zentralen Teil dieser Arbeit. Um die Qualität des generierten Scores einzuordnen und den aktuellen Stellenwert von KI-Kompositionstools zu erfassen, werden Experteninterviews durchgeführt und qualitativ ausgewertet.

Das Ziel dieser Arbeit ist, den aktuellen Stand der Entwicklung sowie die vorhandenen Möglichkeiten und Grenzen der Künstlichen Intelligenz bei der Erstellung eines klassischen Filmscores darzulegen.

## Abstract

Developments in generative artificial intelligence are having an increasing impact on music production. In addition to various plug-ins for audio editing and mixing, artificial intelligence (AI) is now entering the field of composition. The corresponding softwares enable the generation of a wide variety of music, including the creation of film music. This thesis will investigate the extent to which a score for an existing moving image can be generated by an AI today, in the year 2024.

To this end, selected basics of film music and generative artificial intelligence are first examined. The first focus is on the subsequent analysis of the score of a selected film scene. These findings are used to generate a new film score for the same scene using AI software and thus form the second central part of this work. In order to evaluate the quality of the generated score and to determine the current status of AI composition tools, expert interviews are conducted and analysed qualitatively.

The aim of this thesis is to present the current state of development as well as the existing possibilities and limitations of artificial intelligence in the creation of a classic film score.

# Inhaltsverzeichnis

EHRENWÖRTLICHE ERKLÄRUNG .....	II
KURZFASSUNG.....	III
ABSTRACT.....	III
INHALTSVERZEICHNIS .....	IV
ABBILDUNGSVERZEICHNIS .....	VI
ABKÜRZUNGSVERZEICHNIS.....	VII
GENDERHINWEIS .....	VII
DANKSAGUNG .....	VII
1. EINLEITUNG .....	1
1.1 Motivation und Zielsetzung .....	1
1.2 Aufbau der Arbeit.....	3
2. THEORETISCHE GRUNDLAGEN .....	4
2.1 Filmmusik.....	4
2.1.1 Was ist Filmmusik? .....	4
2.1.2 Funktionen der Filmmusik .....	5
2.1.3 Kompositionstechniken .....	7
2.2 Künstliche Intelligenz .....	10
2.2.1 Was ist Künstliche Intelligenz? .....	10
2.2.2 Was ist generative künstliche Intelligenz? .....	11
2.2.3 Verortung KI in der Musikproduktion .....	13
3. DER FILM INCEPTION .....	16
3.1 Handlung .....	16
3.2 Szenenwahl und Begründung.....	17
4. ANALYSE .....	18
4.1 Filmmusik Inception .....	18
4.2 Einführung in die Analyse .....	23
4.3 Analyse der Szene [02:03:58 – 02:15:10].....	25
4.3.1 [02:03:58 – 02:07:20]: Traumbene drei / Kick .....	25
4.3.2 [02:07:21 – 02:09:50]: Limbus / Traumbene eins.....	31
4.3.3 [02:09:51 – 02:15:10]: Limbus / Realität .....	36
5. ERSTELLUNG DER KI-GENERIERTEN KOMPOSITION .....	42
5.1 Vorproduktion .....	42
5.1.1 Vorbereitung.....	42
5.1.2 Konzept .....	43
5.2 Produktion.....	45
5.2.1 Part 1: [00:00:00 – 00:03:31].....	46
5.2.2 Part 2: [00:03:32 – 00:06:08].....	50
5.2.3 Part 3: [00:06:09 – 00:11:39].....	53
5.2.4 Soundalikes.....	55
5.3 Postproduktion.....	57
5.4 Fazit .....	60
5.4.1 Herstellungsprozess.....	60

5.4.2 Vergleich Originalscore .....	62
6. EVALUIERUNG DER KÜNSTLICH GENERIERTEN FILMMUSIK .....	64
6.1 Vorgehen .....	64
6.2 Methodik .....	65
6.3 Durchführung der Experteninterviews .....	67
6.4 Auswertung der Experteninterviews .....	69
7. SCHLUSSBETRACHTUNG .....	78
7.1 Zusammenfassende Beantwortung der Fragen .....	78
7.2 Kritische Würdigung.....	80
7.3 Ausblick .....	82
LITERATURVERZEICHNIS.....	VIII
INTERNETQUELLEN .....	XI
ANLAGEN .....	XII
DIGITALE ANLAGEN .....	XVI

## Abbildungsverzeichnis

Abbildung 2.1 Grundlegende Komponenten eines KI-Systems .....	10
Abbildung 2.2 Maschinelles und Tiefgehendes Lernen als Teilgebiete der Künstlichen Intelligenz.....	11
Abbildung 4.1 Vereinfachte Darstellung des Bläserhauptthemas .....	22
Abbildung 4.2 Einführung des Time-Themas durch Klavier am Filmbeginn.....	22
Abbildung 4.3 Motivvariation der Streicher bei der Erklärung des Totems.....	23
Abbildung 4.4 Rhythmische Melodie der E-Gitarre für das dritte musikalische Thema.....	23
Abbildung 4.5 Zusammenhang zwischen der dramaturgischen und musikalischen Ausarbeitung im ersten Szenenabschnitt (eigene Darstellung) .....	30
Abbildung 4.6 Zusammenhang zwischen der dramaturgischen und musikalischen Ausarbeitung im zweiten Szenenabschnitt (eigene Darstellung) .....	35
Abbildung 4.7 Zusammenhang zwischen der dramaturgischen und musikalischen Ausarbeitung im dritten Szenenabschnitt (eigene Darstellung) .....	41
Abbildung 5.1 Aufbau der Editor-Ansicht von AIVA.....	48
Abbildung 5.2 Zusammenstellung der verfügbaren Nachbearbeitungsmöglichkeiten in der Editor-Ansicht von AIVA.....	58
Abbildung 6.1 Phasen der inhaltlich strukturierenden qualitativen Inhaltsanalyse .....	65
Abbildung 6.2 Leitfaden für die Experteninterviews .....	68
Abbildung 6.3 Experteninterviews .....	69
Abbildung 7.1 Prüfung der internen und externen Gütekriterien .....	81

## Abkürzungsverzeichnis

AI	Artificial Intelligence
AIVA	Artificial Intelligence Virtual Artist
AR	Augmented Reality
BPM	Beats per Minute
DAW	Digital Audio Workstation
dB	Dezibel
DL	Tiefergehendes Lernen/Deep Learning
GANs	Generative Adversarial Networks
kHz	Kilohertz
KI	Künstliche Intelligenz
LUFS	Loudness Units relative to Full Scale
ML	Maschinelles Lernen/Machine Learning
VAEs	Variational Autoencoders
VR	Virtual Reality

## Genderhinweis

In dieser Arbeit wird aus Gründen der besseren Lesbarkeit das generische Maskulinum verwendet. Es sind dabei alle Geschlechteridentitäten ausdrücklich stets eingeschlossen.

## Danksagung

Für die Betreuung meiner Masterarbeit bedanke ich mich recht herzlich bei Herrn Prof. Oliver Curdt und Herrn Prof. Dr. Andreas Koch. Vielen Dank auch an die Experten, die sich bereiterklärt haben, mich mit ihrer Expertise bei dieser Arbeit zu unterstützen.

Zudem möchte ich mich bei meiner Familie und meinen Freunden bedanken, die mich über den gesamten Bearbeitungszeitraum emotional unterstützt und das Korrekturlesen übernommen haben.

## 1. Einleitung

### 1.1 Motivation und Zielsetzung

*All music is based in one way or the other, or influenced through the ages, on technology.*

*- Hans Zimmer, 2014*

Diese Worte stammen von Hans Zimmer, einem weltweit renommierten Filmkomponisten, der durch die Verbindung traditioneller Orchester mit Klängen von Synthesizern einen neuartigen Kompositionsstil in der Filmmusik etabliert hat. Dieses Zitat ist im Zusammenhang eines Interviews aus dem Jahr 2014 entstanden, bei dem er ferner anführt, dass er Instrumente gleichermaßen wie den Computer spiele und, dass der Unterschied zwischen diesen verschwindend gering sei (Zimmer, 2014). Diese Anschauung begründet er durch sein Interesse an Fortschritt und Technologien. Er sieht ein Abenteuer darin, Neues in vielfältiger und durchaus in unkonventioneller Weise auszuprobieren (Zimmer, 2014). Zu dem Zeitpunkt des Interviews konnte Zimmer jedoch nicht ahnen, inwieweit sich die Technologie in den darauffolgenden Jahren entwickeln würde. Heute, zehn Jahre später, sind Maschinen nun in der Lage, verschiedenste Musikstücke mithilfe von generativer künstlicher Intelligenz eigenständig und innerhalb kürzester Zeit zu erstellen.

Der Begriff der Künstlichen Intelligenz (KI) ist keineswegs neu und wurde bereits in den 1950er Jahren etabliert. Die Weiterentwicklung der dahinterstehenden Technologien haben nun generative künstliche Systeme hervorgebracht, die durch maschinelles Lernen zur Erstellung von Inhalten eingesetzt werden können (Feuerriegel, Hartmann, Janiesch, & Zschech, 2023). Diese umfassen vielseitige Anwendungen von der Informationsgewinnung bis hin zur Unterstützung und Erledigung von Aufgaben. Die möglichen Anforderungen und Nutzungen gewinnen zunehmend an Komplexität, sodass KI-Systeme schon heute problemlos Texte, Bilder, Videos und Audio generieren können (Feuerriegel et al., 2023).

Durch die rasanten Entwicklungen ist zu erwarten, dass sich die Qualität der künstlich generierten Kompositionen weiter verbessern wird und die Einsatzmöglichkeiten, folglich die Präsenz der KI im Bereich der Musikproduktion weiter zunehmen werden (Berens & Bolk, 2023). Diese Technologie bringt neben Chancen auch Risiken mit sich. Für diese müssen Wege zu einer verantwortungsvollen Entwicklung und Nutzung gefunden werden, denn insgesamt besitzt die generative KI das Potential, weltweit bis zu 300 Millionen Arbeitsplätze mitunter von Fachkräften zu übernehmen (Feuerriegel et al., 2023).

Mit der Generierung visueller und auditiver Inhalte dringen intelligente Systeme in den Bereich des künstlerischen Schaffens vor. Die Kunst war bisher den Menschen vorbehalten und deren

zugeschriebene kreativen Fähigkeiten haben dazu beigetragen Mensch von Maschine zu separieren (Catani, 2024a). Wenn nun die KI einen künstlerischen Anspruch erhebt, stellt sich die Frage, ob es einer Neudefinierung der über die Jahrhunderte fortgeführten Konzepte von Originalität, Schöpfertum, Kreativität und Autorschaft bedarf (Catani, 2024a).

Dabei ist die Idee, Musik von oder mit einer generativen KI komponieren zu lassen nicht abwegig. Die Musik ist eng mit der Mathematik verflochten und kann formal in weiten Teilen durch sie beschrieben werden. So werden beispielsweise Intervallabstände, Tondauern, Metronomangaben oder auch Tonhöhen mit Hilfe von Zahlen dargelegt (Heintz, 2024). Auch findet sich im digitalen Umfeld der Musikproduktion aktuell eine Varianz an Plugins wieder, die bereits mit intelligenten Algorithmen arbeiten (Frieler, Zaddach, & Meyer, 2023). Daher ist es konsequent, dass die KI nun auch in den Bereich der Komposition vordringt. Erste Anwendungen wurden bereits im Jahr 2016 veröffentlicht, seither nimmt deren Verfügbarkeit stetig weiter zu (Berens & Bolk, 2023).

Ein Musikstück kann demnach mit einem KI-System unkompliziert erstellt werden. Doch inwiefern eignen sich diese Kompositionen im Hinblick auf die Verwendbarkeit für Filmmusik? Kann eine Komposition unter Berücksichtigung eines Handlungsverlaufs generiert werden? Filmmusiken dienen dem gezeigten Bewegtbild und fungieren nur äußerst selten eigenständig. Es ist vorgesehen, dass sie hintergründig agieren und erst durch das audiovisuelle Zusammenspiel die volle Wirkung entfalten (Kloppenburger, 2017a). Weder eine KI noch der Komponist können dabei im Voraus garantieren, dass der vorgesehene Ausdruck und die konzipierte Wirkung beim Publikum erzielt werden. Für jede Filmproduktion bedarf es stets einer individuell zur Dramaturgie und zum Stil angepassten musikalischen Ergänzung. Während es in einem Fall minimalistische und subtile Musiken sind, werden in einem anderen Fall prachtvolle opulente Kompositionen gefordert (Kreuzer, 2009).

Weiterhin stellt sich die Frage, ob auch die klangliche Komponente und die Qualität der verwendeten Instrumente von den intelligenten Systemen überzeugend sind. Über die Jahre hat sich der Klang eines synthetischen Orchesters deutlich verbessert, jedoch kann es nicht die Dynamik, Klangmodulation und Raumakustik eines echten Orchesters simulieren (Lensing, 2018). Arbeitet ein Komponist bei der Aufnahme mit den Musikern und dem Dirigenten zusammen können Feinheiten angepasst und Interpretationen eingespielt werden, die künstlich so nicht erzeugt werden können (Lensing, 2018).

Neben der technischen Umsetzbarkeit implizieren diese Produkte unabhängig ihrer äußeren Form weitere Herausforderungen bezüglich Urheberschaft, Originalität sowie zu ethischen Fragen, welche noch nicht ausreichend geklärt sind und in naher Zukunft aufgearbeitet werden müssen (Heintz, 2024).

Die vorliegende Arbeit untersucht die aktuell verfügbaren Umsetzungsmöglichkeiten sowie die hervorgebrachte Qualität intelligenter Systeme bezüglich der Erstellung einer Filmmusik. Dazu

werden zunächst die Gestaltung und Wirkung eines Scores von Hans Zimmer analysiert und anschließend ein KI-Programm zur Erzeugung einer neuen Komposition herangezogen. Ziel ist, herauszufinden in welchem Umfang generierte Musikstücke für den Einsatz als Filmmusik geeignet sind und welchen Nutzen solche intelligenten Anwendungen mit sich bringen. Es soll geklärt werden, ob die generierten Produkte bereits imstande sind, mit den Scores bekannter Komponisten mithalten und womöglich deren Arbeit in der Zukunft übernehmen könnten. Diesbezüglich wurden drei Teilfragen formuliert, die in den nachfolgenden Kapiteln beantwortet werden:

1. Worin liegen die Unterschiede einer von Menschen gestalteten bzw. durch KI generierten Filmmusik?
2. Wie organisiert sich die Zusammenarbeit zwischen Menschen und KI-Software und welche Auswirkungen ergeben sich auf den Herstellungsprozess?
3. Wie wird die Übernahme kreativer Aufgaben durch die generative KI und die Qualität der Erzeugnisse von Komponisten und Filmmusikern empfunden?

## **1.2 Aufbau der Arbeit**

Diese Arbeit ist in sieben Kapitel untergliedert. Sie beginnt mit der Einleitung, welche die Motivation und Zielsetzung sowie den Aufbau der Arbeit beinhaltet. Im nachfolgenden zweiten Kapitel werden die theoretischen Grundlagen, die für das Verständnis der Arbeit erforderlich sind, beschrieben. Dabei werden zunächst Aspekte der Filmmusik betrachtet und anschließend Begrifflichkeiten der Künstlichen Intelligenz eingeordnet. Zudem wird aufgezeigt inwiefern intelligente Systeme bereits heutzutage einen Bestandteil innerhalb der Musikproduktion bilden.

Das dritte Kapitel befasst sich mit der Vorstellung des Filmes Inception. Neben der Filmhandlung wird die zu analysierende Szene vorgestellt und deren Wahl begründet. Anschließend wird in Kapitel vier die filmmusikalische Analyse durchgeführt. Zunächst wird die Filmmusik in ihrer Gesamtheit betrachtet, um ein Grundverständnis für die Szenenanalyse herzustellen. In Kapitel 4.2 wird der Analyseansatz dargelegt, der nachfolgend für die Sequenz abschnittsweise angewendet wird.

Im fünften Kapitel erfolgt die Erstellung der alternativen Filmmusik mittels einer KI-Software. Dazu werden in der Vorproduktion die Ziele und Voraussetzungen für die Generierung der Komposition geschaffen, welche anschließend in Kapitel 5.2 für die drei Szenenbereiche durchgeführt wird. Nach der Produktion wird die erstellte Musik mit den weiteren Elementen des Soundtracks gemischt und abschließend das Resultat mit dem Originalscore verglichen. Kapitel sechs beschäftigt sich mit der Evaluierung der generierten Filmmusik, die in Form von Experteninterviews umgesetzt wird. Zunächst werden das Vorgehen und die Methodik

erläutert, bevor die Durchführung der Interviews erfolgt. Die Auswertung der Inhalte wird anschließend in Kapitel 6.3 anhand verschiedener Kategorien dargelegt.

Die Schlussbetrachtung mit zusammenfassender Beantwortung der Forschungsfragen, einer kritischen Würdigung sowie einem kurzen Ausblick schließen diese Arbeit ab.

## **2. Theoretische Grundlagen**

Da sich diese Arbeit thematisch auf die Schnittstelle von künstlicher Intelligenz und Filmmusik fokussiert, werden in diesem Kapitel theoretische Grundlagen für beide Bereiche erklärt. Diese sind für das weitere Vorgehen in den folgenden Kapiteln relevant. Zunächst werden filmmusikalische Begriffe eingeordnet und anschließend der Filmmusik zugeschriebene Funktionen und Kompositionstechniken erläutert. Im zweiten Abschnitt wird die Künstliche Intelligenz allgemein sowie unter dem Gesichtspunkt der Generierung von Daten definiert und in den Kontext der (Film-)Musikproduktion gesetzt.

### **2.1 Filmmusik**

#### **2.1.1 Was ist Filmmusik?**

Filmmusik wird als funktionale Musik beschrieben, deren Bedeutung und Wirkung sich in direktem Zusammenhang mit den gezeigten Bewegtbildern entfaltet. Die Kompositionen werden so gestaltet, dass sie durch den Rezipienten eindeutig aufgefasst werden können. Das Filmkonzept bestimmt dabei die Anforderungen der musikalischen Einsätze und definiert die zu erfüllenden Aufgaben der Filmmusik (Kloppenburg, 2017a). So kann diese unter anderem dafür eingesetzt werden, um unterbewusst die Emotionen des Publikums zu beeinflussen und hervorzubringen, oder um die Spannung innerhalb einer Szene zu steigern (Lensing, 2018). Die Funktionen der Filmmusik sind jedoch deutlich weitreichender und werden im folgenden Unterkapitel 2.1.2 ausführlicher dargelegt.

Prinzipiell erleichtert die Musik die Rezeption eines Filmes, ermöglicht dem Publikum eine höhere Immersion und unterstreicht dessen Gesamtbedeutung (Kreuzer, 2009). Dabei kann die Filmmusik sowohl aus eigens für den Film komponierter Musik, als auch aus bereits veröffentlichten Werken oder einer Kombination beider bestehen (Kreuzer, 2009).

Es wird grundlegend zwischen diegetischer und nichtdiegetischer Filmmusik unterschieden. Ist die Musik Teil der filmischen Realität und für die Protagonisten ebenso hörbar wie für das Publikum, wird von diegetischer Filmmusik oder auch source Musik gesprochen (Heldt, 2018; Tieber 2020). Dabei muss die musikalische Quelle für den Zuschauer nicht zwingend sichtbar

sein, lediglich ihr Ursprung liegt eindeutig in der Filmwelt. Dies können beispielsweise Plattenspieler oder ein Radio, ebenso wie mit Instrumenten musizierende Personen sein (Tieber, 2020). Die nichtdiegetische Musik dient der filmischen Narration und wird nachträglich zu dem Bewegtbild hinzugefügt. Sie wird lediglich von den Zuschauern wahrgenommen und kann je nach Einsatz verschiedene Wirkungen erzielen und für objektive bzw. subjektive Perspektiven eingesetzt werden (Heldt, 2018). Ein weiterer Begriff für diesen Bereich der Filmmusik ist der Score, welcher stets durch einen Komponisten realisiert wird (Tieber, 2020). Dieser entwirft ein musikalisches Konzept in Abstimmung mit dem Regisseur und Produzent, das anschließend ausgearbeitet und durch Orchesteraufnahmen, einzelne Instrumente sowie durch elektronische Produktion umgesetzt wird (Lensing, 2018).

Die Zuordnung zwischen diegetischer und nichtdiegetischer Musik ist dabei nicht immer eindeutig, da deren Übergänge ineinanderfließen können. So kann beispielsweise eine zunächst diegetisch erklingende Musik durch eine einsetzende nichtdiegetische Musik übernommen und von ihr abgelöst werden oder umgekehrt wird die Quelle den Zuschauern vorenthalten, bevor sie sich ihnen erschließt (Heldt, 2018).

Unter dem Begriff Soundtrack wird oftmals ausschließlich die Musik des Filmes assoziiert. Dieser beinhaltet jedoch zusätzlich die Dialoge sowie das Sounddesign bestehend aus Geräuschen, Atmosphären und klanglichen Effekten (Filmtone Berufsvereinigung e.V., 2023).

### **2.1.2 Funktionen der Filmmusik**

Zofia Lissa war eine der ersten Musikwissenschaftlerinnen, die sich mit der Funktionalität von Filmmusik intensiv beschäftigt und in einer gegliederten Form niedergeschrieben hat (Kloppenburger, 2017a). Um die umfangreichen Funktionen zu vereinheitlichen hat der Schweizer Musikwissenschaftler Hansjörg Pauli ein Modell entwickelt, bei dem er diese in drei Kategorien unterteilt: in paraphrasierende, polarisierende und in kontrapunktierende Filmmusik (Pauli, 1976). Als paraphrasierend bezeichnet Pauli (1976) „Musik, deren Charakter sich direkt aus dem Charakter der Bilder, aus den Bildinhalten, ableitet“ (S. 104). Sie untermauert das Bewegtbild, indem sie mit Bewegungen mitgeht, die emotionale Stimmung aufgreift und parallel zu den Bildern wiedergibt und verstärkt (Pauli, 1976). Diese Art der Filmmusik kann mit Hilfe der deskriptiven Technik komponiert werden (Kreuzer, 2009).

Die Filmmusik ist polarisierend, wenn sie „inhaltlich neutrale oder ambivalente Bilder in eine eindeutige Ausdrucksrichtung“ (Pauli, 1976, S.104) wertet. Damit werden die Interpretationen der Rezipienten in eine bestimmte Richtung gelenkt, um Fehldeutungen zu vermeiden (Kreuzer, 2009). Wie auch die paraphrasierende Musik dient die Polarisation der Immersion sowie der Identifikation der Zuschauer mit den Filmfiguren. Daher wird polarisierende Musik auch gerne für die Filmeröffnung verwendet, um diese emotional auf die kommende Handlung einzustimmen (Pauli, 1976).

Widerspricht der musikalische Charakter den gezeigten Bildinhalten und erzeugt dadurch eine Diskrepanz zwischen Ton und Bild wird von kontrapunktierender Filmmusik gesprochen (Pauli, 1976). Die gegensätzlichen Inhalte werden mitunter kritisch von den Zuschauern wahrgenommen. Sie induzieren Gedankenprozesse und führen zu einer erschwerten Interpretation, da sie nicht mit der Bild- und Tonebene in Einklang gebracht werden können (Pauli, 1976).

Neben Pauli haben noch weitere Musikwissenschaftler die Funktionalität von Filmmusik untersucht und alternative Kategorisierungen entwickelt. Unter anderem auch Claudia Bullerjahn (2001), deren Funktionsmodell nun grundlegend dargelegt wird. Sie unterscheidet zwischen Metafunktionen und Funktionen im engeren Sinne. Die Metafunktionen sind in rezeptionspsychologische und ökonomische Aspekte unterteilt und müssen stets in den zeitlichen Kontext gesetzt werden. Die rezeptionspsychologischen Funktionen beschreiben die Fähigkeit der Filmmusik, technische als auch von den Zuschauern erzeugte Störgeräusche zu maskieren und dadurch für eine bessere Filmrezeption zu sorgen (Bullerjahn, 2001). Gerade während der Anfänge der Filmwiedergabe half die Musik, Projektorengeräusche und Filmrollenwechsel zu kaschieren. Die ökonomischen Funktionen zeigten sich früher vor allem darin, mittels der Musik das Zielpublikum zu erreichen und zu Kinobesuchen zu bewegen. Heutzutage ist die Vermarktung wesentlich komplexer, bei der zahlreiche weitere Faktoren eine Rolle spielen (Bullerjahn, 2001).

Neben dem übergeordneten Charakter der Metafunktionen sind die Funktionen im engeren Sinne genreabhängig und stehen in direktem Filmbezug. Auch hier wird zwischen vier Kategorisierungen unterschieden: den dramaturgischen, epischen, strukturellen und persuasiven Funktionen.

Die dramaturgischen Funktionen beschreiben die Darstellung der Stimmung sowie die atmosphärischen Elemente einer Szene. Zudem kann die Musik den Spannungsverlauf einer Szene führen und die inneren Gefühlszustände der Protagonisten verdeutlichen (Bullerjahn, 2001). Werden durch die Filmmusik narrative Aufgaben übernommen, wird von den epischen Funktionen gesprochen. Dabei kann die musikalische Ebene Handlungselemente wie den Schauplatz und die Zeit der Handlung verdeutlichen, die Wahrnehmung der Erzählzeit steuern oder auch das Bewegtbild kommentieren (Bullerjahn, 2001; Lensing, 2018).

Den strukturellen Funktionen werden die Hervorhebung und Kaschierung von Schnitten und Bewegungsabläufen sowie die Ankündigung von Szenenwechseln zugeschrieben. Die Filmmusik verleiht eine gewisse Kontinuität, denn die „Geschlossenheit der kompositorischen Konzeption durch homogene Aufführungsmittel und einheitliches thematisches Material trägt dazu bei, einen ganzen Film formal zu integrieren“ (Bullerjahn, 2001, S.72). Die persuasiven Funktionen bestehen darin, sowohl die im Bewegtbild implizierten Emotionen darzustellen, als

auch die beim Zuschauer hervorgerufenen Gefühle und dessen Aufmerksamkeit zu beeinflussen (Bullerjahn, 2001). Dies gelingt dadurch, dass das Publikum in der Regel die Haltung des Regisseurs als die eigene übernimmt und dadurch das Verhältnis von Bild und Musik als stimmig empfindet (Kühn, 1976).

### **2.1.3 Kompositionstechniken**

Mit der Entwicklung der Filmmusik haben sich über die Jahre drei Kompositionstechniken etabliert, die bis heute von Filmkomponisten in vielfältiger Weise verwendet werden. Diese sind die deskriptive Technik, die Mood Technik sowie die Leitmotivtechnik (Bullerjahn, 2001). Sie ermöglichen ein effizientes und strukturiertes Vorgehen und eignen sich für die Arbeit mit Bewegtbild. Dabei werden die Kompositionstechniken sowohl in reiner Form, als auch parallel und in Kombination miteinander verwendet (Bullerjahn, 2001; Kloppenburg 2017a). Nachfolgend werden die drei Techniken kurz beschrieben, da sie im Verlauf der Arbeit erneut aufgegriffen werden.

#### **Deskriptive Technik**

Die deskriptive Technik, auch Underscoring genannt, ist eine musikalische Illustration der Bildebene, indem sie die „sichtbaren Vorkommnisse, Bewegungen und dargestellten Gefühle möglichst synchron mitvollzieht“ (Kloppenburg, 2017a, S.126). Die Musik verdoppelt die Bildaussage und kann die Bedeutung einzelner Bilder beeinflussen (Bullerjahn, 2001). Dazu werden auch Geräusche durch die Musik imitiert. So können rhythmische Bewegungen von Personen, Maschinen oder Transportmitteln durch rhythmische Muster in der Musik aufgegriffen oder kurze prägnante Elemente, wie zum Beispiel ein Schuss, durch musikalische Akzente hervorgehoben werden (Kloppenburg, 2017a; Bullerjahn, 2001).

Diese Verdopplung birgt jedoch auch das Risiko für Klischees. So kann der Einsatz bestimmter Klangfarben durch ausgewählte Instrumentierung Assoziationen zu bestimmten Orten, Zeiten oder Gefühlen hervorrufen. Da diese Mittel jedoch in vielfältiger Weise bereits verwendet wurden, können sie zu einer flachen und klischeehaften Wirkung führen (Bullerjahn, 2001). Folglich erfordert die deskriptive Technik entsprechend des Anwendungszwecks durchdachte und sorgfältig erstellte Kompositionen und sollte nicht nur auf einfachen, plakativen Stücken basieren (Kloppenburg, 2017a).

Bereits während der Stummfilmzeit wurden Bewegungen in Filmen musikalisch unterstrichen, aber erst durch die Entwicklungen von Max Steiner hat sich die deskriptive Technik als Kompositionsstil etabliert (Kloppenburg, 2017a).

Wird diese Technik in extremer Form verwendet, dann wird von Mickeymousing gesprochen. Dabei übernimmt die Musik alle Vorgänge der Bildebene und zeichnet diese exakt synchron

nach (Schramm, Spangardt, & Ruth, 2017). Wie der Begriff bereits verdeutlicht, wird besonders in Zeichentrickfilmen und Cartoons mit Mickeymousing gearbeitet, bei denen die Musik „weniger eine Ergänzung der Handlung als ein integraler Bestandteil“ (Bullerjahn, 2001, S.80) ist.

### **Mood Technik**

Die Mood Technik unterscheidet sich von der deskriptiven Technik dahingehend, dass nicht die einzelnen Details, sondern die Szene im Gesamten musikalisch unterlegt wird (Kloppenburger, 2017a). Die Komposition erfährt durch musikalische Parameter wie Instrumentierung, Tonarten und Tonlagen oder auch Tempi einen Stimmungsgehalt, welcher der Vermittlung von Affekten dient (Bullerjahn, 2001). Es wird demnach über die musikalische Ebene eine Stimmung transportiert, die dem Bewegtbild und der Handlung entspricht und gleichzeitig die gezeigte Emotion innerhalb der Szene intensiviert (Kloppenburger, 2017a). Neben der stimmunggebenden Einfärbung kann der vorgezogene Einsatz dieser Musik auf folgende Handlungsgeschehnisse vorbereiten (Bullerjahn, 2001). Dadurch soll dem Zuschauer eine höhere emotionale Immersion ermöglicht werden (Schramm, Spangardt, & Ruth, 2017).

Bei dieser Technik kann zwischen expressiver und sensorischer Filmmusik unterschieden werden. Während die expressive Filmmusik dazu dient, die Gefühlslage der Protagonisten zu verdeutlichen oder zu offenbaren, wirkt sich die sensorische Filmmusik auf die Rezipienten aus und beeinflusst deren Empfindungen (Schramm, Spangardt, & Ruth, 2017). Wie auch bei der deskriptiven Technik ist die vertikale Beziehung zwischen Bild und Musik für die Erstellung der Komposition ausschlaggebend (Bullerjahn, 2001).

### **Leitmotivtechnik**

Bei der Leitmotivtechnik werden Handlungselemente wie Personen, Gegenstände, Orte oder auch Situationen an eine „charakteristische, wiederkehrende und wiedererkennbare musikalische Figur oder Tonfolge“ (Albrecht, 2021, S.7) gekoppelt. Diese Verbindungen werden vom Zuschauer über den Handlungsverlauf erlernt, wobei mit jedem Auftreten weitere Assoziationen durch die hinzukommenden Bildinhalte entstehen und sich somit die außermusikalische Bedeutung der Leitmotive erweitert (Bullerjahn, 2001). Dadurch entsteht eine wechselseitige Verknüpfung der Bilder und der musikalischen Motive in der filmischen Narration (Kreuzer, 2009). Zudem dient der gezielte Einsatz von Leitmotiven dazu, Gedankenprozesse bei den Zuschauern anzuregen und zu steuern (Bullerjahn, 2001). Laut Tieber (2020) verläuft die Analyse der Figuren bezüglich der musikalischen Gestaltung und des Auftretens in der Filmhandlung zunächst zweigeteilt und wird erst im Anschluss für die Bild-Musikbeziehung zusammengeführt. Leitmotive können in ihrer ursprünglichen Form

wiederkehren und erst mit den dramaturgischen Entwicklungen musikalisch und charakterlich angepasst werden (Albrecht, 2021).

Bei der Verwendung von Leitmotiven differenziert Schneider (1983) zwischen Motivatoren, bei denen die musikalischen Figuren sich über die Handlung nicht verändern. Zum anderen benennt er die sogenannte *idée fixe*, welche die Variationen der musikalischen Motive entsprechend der dramaturgischen Entwicklungen beinhaltet. Mit der dritten Verwendungsweise spricht Schneider (1983) die Leitmotivtechnik in ihrer ausgeprägten Version an, bei denen die Motive tief in die musikalische Gestaltung verflochten sind und diese mit dominieren.

Albrecht (2021) hingegen hat die Funktionen von Leitmotiven untersucht und diese in formale, semantische und dramaturgische Funktionen gegliedert. Die formalen Funktionen des Leitmotivs liegen darin, eine Zusammengehörigkeit der Filmmusik zu erreichen. Dabei ist es nicht von Relevanz, dass die Leitmotive „die Substanz des Tonsatzes konstituieren noch die übergreifende filmische Struktur determinieren“ (Albrecht, 2021, S.10). Die wiederkehrenden und vom Rezipienten erkennbaren Motiveinsätze verhelfen zu einer strukturierten Form und bewirken damit eine homogene Klangwirkung der Filmmusik (Albrecht, 2021).

Die semantischen Funktionen zeichnen sich durch die Wechselwirkung zwischen Leitmotiv und Rezipient aus. Die filmischen Inhalte werden mit den Motiven verknüpft und mit jeder Wiederkehr in ihrer Bedeutung erweitert. Damit spiegelt das Leitmotiv nicht nur die visuelle Ebene wider, sondern erfährt einen außermusikalischen Gehalt. Es fungiert zudem auf diversen zeitlichen Ebenen wie in Erinnerungen, der Gegenwart und in zukünftigen Handlungen (Albrecht, 2021; Kloppenburg, 2017a).

Die dramaturgischen Funktionen beinhalten die Verknüpfung der Protagonisten mit den Leitmotiven sowie deren Variationen und Abspaltungen. Entsprechend der Entwicklung der Protagonisten, deren Verhältnissen zueinander sowie der Dramaturgie werden die musikalischen Figuren verändert, um diese auch auf auditiver Ebene wiederzugeben. Das Erklingen der Tonfolgen wertet auf subtile Weise die Bildebene und kann, nachdem es von dem Rezipienten erlernt wurde, für weitere Funktionen, wie zum Beispiel Spannungserzeugung oder auch Täuschung verwendet werden (Albrecht, 2021). Zudem werden durch Leitmotive „unausgesprochene Gedanken und verborgene Gefühle einzelner Personen enthüll[t] oder geheime Wünsche musikalisch artikulier[t], die mit sprachlichen und filmischen Mitteln nicht ausgedrückt oder absichtsvoll verschwiegen werden“ (Albrecht, 2021, S.13-14).

## 2.2 Künstliche Intelligenz

Die Künstliche Intelligenz ist ein Teilbereich der Informatik, deren Entwicklungen in verschiedensten Branchen zunehmend an Bedeutung gewinnen (Frierer, Zaddach, & Meyer, 2023). Es werden zunächst grundlegende Begrifflichkeiten der KI erläutert, bevor der Teilbereich der generativen künstlichen Intelligenz, welche im Rahmen dieser Arbeit untersucht wird, separiert hervorgehoben wird.

### 2.2.1 Was ist Künstliche Intelligenz?

Eine präzise Definition für den Begriff Künstliche Intelligenz zu formulieren ist kaum möglich, da bereits die Intelligenz nur unscharf beschrieben ist und sich zudem „das Verständnis für KI stetig und dynamisch weiterentwickelt“ (Mockenhaupt & Schlagenhauf, 2024). Auch die fachspezifische Betrachtungsweise beeinflusst den Bedeutungsraum Künstlicher Intelligenz maßgeblich (Catani, 2024a). Prinzipiell kann bei der KI von der „Fähigkeit eines Systems, externe Daten korrekt zu interpretieren, aus diesen Daten zu lernen und diese Erkenntnisse zur flexiblen Anpassung zu nutzen, um bestimmte Ziele und Aufgaben zu erreichen“ (Kaplan & Haenlein, 2019, S.15) gesprochen werden. Demnach wird versucht, die menschliche Intelligenz durch eine Maschine nachzubilden, wobei das erlernte Wissen durch algorithmische Analyse und Verarbeitung enormer Datenmengen erfolgt (Mockenhaupt & Schlagenhauf, 2024). Ein KI-System besteht dabei aus einer Vielzahl an Komponenten. Laut Mockenhaupt und Schlagenhauf (2024) sind dies die grundlegenden Bestandteile:

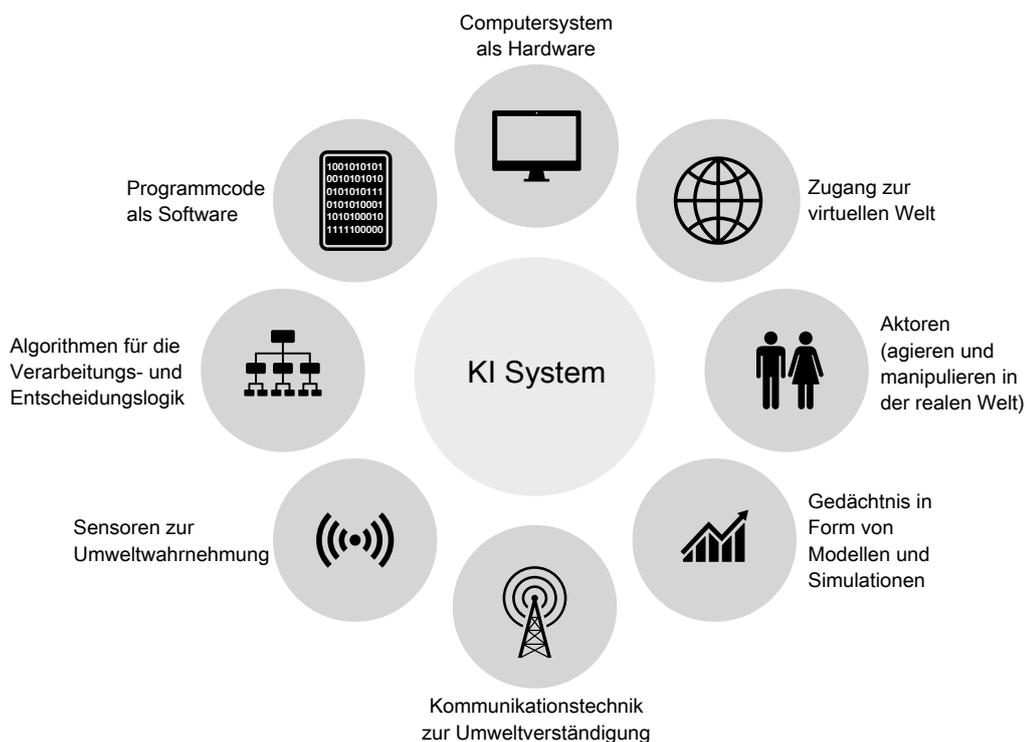


Abbildung 2.1 Grundlegende Komponenten eines KI-Systems  
In Anlehnung an (Mockenhaupt & Schlagenhauf, 2024, S.58)

Zentrale Teilgebiete der KI sind Maschinelles Lernen (Machine Learning) sowie Tiefergehendes Lernen (Deep Learning). Der Lernprozess erfolgt dabei jeweils durch Datenanalysen mittels Algorithmen. Während beim Maschinellen Lernen Anpassungen durch manuelle Eingriffe erfolgen, übernehmen diese beim Tiefergehenden Lernen die Algorithmen selbst (Mockenhaupt & Schlagenhaut, 2024). Ermöglicht wird dies durch künstliche neuronale Netze, die aus einer Vielzahl an miteinander verbundenen künstlichen Neuronen bestehen, welche in Schichten angeordnet sind (Frieler, Zaddach, & Meyer, 2023).

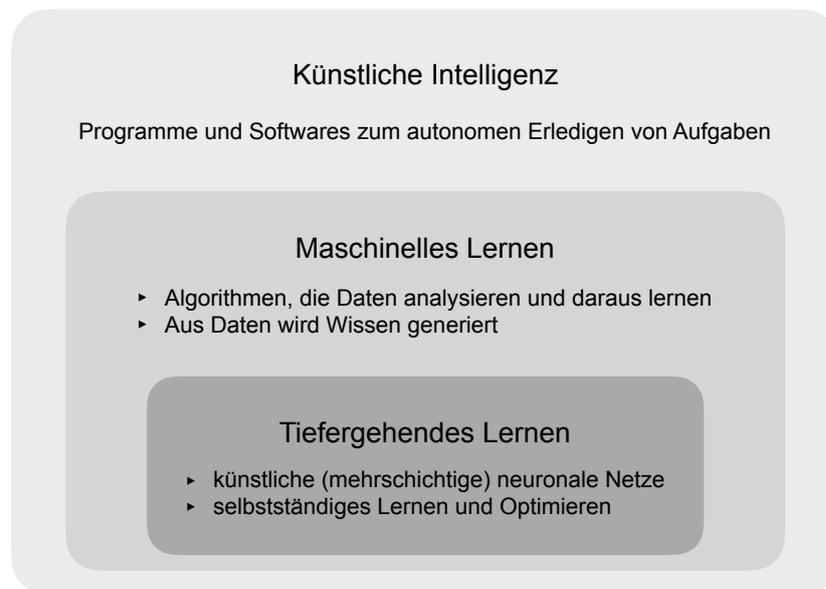


Abbildung 2.2 Maschinelles und Tiefergehendes Lernen als Teilgebiete der Künstlichen Intelligenz  
In Anlehnung an (Mockenhaupt & Schlagenhaut, 2024, S.170)

Es wird zwischen starker und schwacher KI unterschieden. Ein schwaches KI-System ist für wenige bis hin zu einer Anwendung optimiert, deren Ausführung durchaus der des Menschen überlegen ist. Für abweichende Aufgabenbereiche erweisen sich diese Systeme jedoch als ungeeignet (Koehler, 2024). Folglich repräsentiert die schwache KI einzelne Aspekte der menschlichen Intelligenz und ist in den derzeitigen KI-Anwendungen wiederzufinden (Frieler, Zaddach, & Meyer, 2023). Die starke KI, auch Künstliche Allgemeine Intelligenz (Artificial General Intelligence), hingegen begrenzt sich nicht auf einzelne Aufgaben, sondern simuliert die menschliche Intelligenz in umfassendem Maße als mindestens ebenbürtig (Koehler, 2024). Die Forschung starker KI-Systeme wird noch mehrere Jahre andauern, sodass derzeit keine Anwendungen mit dieser Technologie verfügbar sind (Mockenhaupt & Schlagenhaut, 2024).

### 2.2.2 Was ist generative künstliche Intelligenz?

Eine Subkategorie der Künstlichen Intelligenz bildet die generative künstliche Intelligenz. Diese generativen KI-Systeme zeichnen sich durch die Generierung von beispielsweise Bild-, Text- oder Audioinhalten mithilfe von erlernten Trainingsdaten aus (Feuerriegel et al., 2023).

Derzeit gesellschaftlich populäre Anwendungen sind das Sprachmodell GPT-4 und der Bildgenerator Dall-E, welche von Open AI entwickelt wurden. Demnach wird der Anschein geweckt, dass Maschinen kreative Prozesse durchlaufen und folglich etwas Neues erschaffen. Die Realisierung erfolgt jedoch mittels verschiedener generativer Modelle, die über Architekturen für Maschinelles und Tiefergehendes Lernen verfügen, wodurch neue Dateninstanzen durch die erlernten Muster der Trainingsdaten erstellt werden (Feuerriegel et al., 2023; Catani, 2024a). Laut D’Onofrio (2024) kann zwischen drei häufig verwendeten Modellarten differenziert werden. Zum einen gibt es Generative Adversarial Networks (GANs). Diese bestehen aus einem Datengenerator und einem Diskriminator, der diese Daten auf deren Echtheit prüft. Dabei wird das Training so gestaltet, dass die Komponenten gegeneinander konkurrieren, mit dem Ziel, hochwertige Daten zu erzeugen. Zum anderen gibt es Transformer-Modelle, die mit Attention Mechanismen ausgestattet sind. Durch diese fokussiert sich die Verarbeitung der Eingabe auf ein spezielles Token, das sich auf die Modellierung der Ausgabe auswirkt. Das dritte Modell sind Variational Autoencoders (VAEs), bei denen die Daten durch eine Wahrscheinlichkeitsverteilung generiert werden. Damit soll eine sich den Eingabedaten annähernde Datenrekonstruktion erlernt werden (D’Onofrio, 2024). Zudem können diese Modelle in unimodale und multimodale Modelle unterteilt werden. Der Unterschied ist, dass bei unimodalen Modellen die Anweisungsart der Eingabe gleich der generierten Ausgabeform ist, während bei multimodalen Modellen verschiedene Eingaben angenommen und ebenso ausgegeben werden können (Feuerriegel et al., 2023).

Durch das autonome Erstellen von Inhalten eröffnet die generative KI neue Anwendungsmöglichkeiten und bringt zugleich Herausforderungen und Grenzen mit sich. Zum einen kann die Richtigkeit der generierten Informationen aufgrund der maschinellen Lernmodelle nicht gewährleistet werden. Dadurch können Falschinformationen verbreitet werden mit dem Potenzial Benachteiligungen zu verursachen (D’Onofrio, 2024). Zum anderen können Bias, die durch das Lernen mit den Trainingsdaten bedingt wurden, auftreten und zu verzerrten bzw. unpassenden Ergebnissen führen. Daher müssen diese Daten sorgfältig ausgewählt werden, um die auftretenden Bias zu mindern (Feuerriegel et al., 2023). Auch die noch unzureichenden Richtlinien in Bezug auf den Umgang mit dem Urheberrecht stellen eine weitere Schwierigkeit dar. Fragen zur Zuschreibung des Urheberrechts und damit des Gedankenguts eines künstlich generierten Inhalts sind nicht final geklärt, ebenso wie der Umgang mit generierten Kopien, welche ohne Einholung einer Erlaubnis erstellt wurden (Feuerriegel et al., 2023). Weiterhin bedarf es an Regelungen sowie Maßnahmen zum Schutz der Privatsphäre und zur Aufklärung bestehender Risiken (D’Onofrio, 2024). Eine ökologische Herausforderung ist der hohe Energiebedarf und der damit verbundene CO<sub>2</sub> Ausstoß von KI-Systemen, der sich umweltbelastend auswirkt (Feuerriegel et al., 2023).

Gerade in den letzten Jahren ist die Entwicklung der generativen KI vorangeschritten, sodass sie zukünftig nicht nur im kreativen, künstlerischen Bereich Transformationen bewirken wird. Vielmehr wird sie sich auch vermehrt im Alltag der Menschen, wie derzeit als Assistenzsysteme zur Informationsbereitstellung, etablieren (D'Onofrio, 2024).

### **2.2.3 Verortung KI in der Musikproduktion**

Um den aktuellen Einsatz und die Verwendung von Künstlicher Intelligenz in der Musikindustrie vor allem in Bezug auf Musikproduktionen nachzuvollziehen, ist ein Blick in die Vergangenheit nützlich. Denn die bereits im Jahr 1956 für Streichquartett veröffentlichte Illiac Suite wird als einer der ersten computergenerierten Kompositionen angesehen (Zickgraf, 2024). Sieben Jahre später, 1963, haben die Mathematiker und Kybernetiker Helmar Frank und Manfred Berger untersucht, in welcher Weise und in welchem Umfang Musikstücke im Stil eines Komponisten wie beispielsweise von Johann Sebastian Bach maschinell generiert, demnach komponiert werden können (Catani, 2024a). In diesem Zusammenhang sollte auch das Projekt Experiments in Musical Intelligence von David Cope der 1980er Jahre erwähnt werden. Es beschreibt ein dreiteiliges Verfahren, das auf Basis analysierter Werke dazu dient, neue Kompositionen im Stil bekannter klassischer Künstler zu erzeugen (Heintz, 2024).

Durch die gesteigerte Rechenleistung, Datenverfügbarkeit sowie durch die Entwicklungen der maschinellen Lernmechanismen in den letzten Jahren wurde ein großer Fortschritt in musikbezogenen KI-Anwendungen ermöglicht. Hinzukommt, dass nun Unternehmen wie Google, Sony und Spotify in eigene Forschungen investieren und somit die Technologien weiter vorantreiben (Zickgraf, 2024).

Folglich haben die KI-gestützten Technologien heutzutage in verschiedenen Bereichen der Musikindustrie Einzug gefunden. Sie reichen von analytischen und kompositionellen Anwendungen in der Musikproduktion und qualitativen Aufwertungen von Audioaufnahmen über intelligente performative Systeme bis hin zur Umwandlung von Texten in Musik und Hörspiele oder - umgekehrt - Audioinhalte in Textform (Brown, 2021).

Gerade bei den beiden Kernelementen Mischung und Mastering von Musikproduktionen „wurden in den vergangenen Jahren zahlreiche marktaugliche Tools entwickelt, die automatisierte und am konkreten Material ausgerichtete Prozesse umsetzen“ (Frieler, Zaddach, & Meyer, 2023, S.22). Sie werden derzeit noch als unterstützende Technologien eingesetzt, da sie vor allem bei großen Mehrspurprojekten noch keine, dem professionellen Standard entsprechenden Gesamtmischungen erzeugen können. Neben der Mischung selbst dienen diese Tools auch zur Soundoptimierung, was durch fortlaufende Analyse des Audiomaterials ermöglicht wird. Ein Beispiel dafür stellen unter anderem die verschiedenen Produkte des Herstellers iZotope dar (Frieler, Zaddach, & Meyer, 2023).

Auch im Reverse-Engineering wird vermehrt mit Künstlicher Intelligenz gearbeitet. Intelligente Systeme werden dazu eingesetzt, die Einzelspuren und Instrumenten-Stems qualitativ hochwertig aus der Mischung zu extrahieren, um diese anschließend zum Beispiel in einem Re-Mastering weiterzuverarbeiten (Frieler, Zaddach, & Meyer, 2023).

Aber auch bei kompositorischen Aufgaben, der Musikgenerierung und Erzeugung von Instrumenten und Klängen wird die KI immer präsenter. Dabei gibt es zum einen den Ansatz, dass die KI den Menschen in seinem kreativen Schaffen unterstützt und dessen Fähigkeiten dadurch erweitern kann. „Der Computer kann hier als Tool fungieren, das einen kreativen Raum der Interaktion eröffnet und somit individuelle kreative Entdeckungen ermöglichen kann“ (Frieler, Zaddach, & Meyer, 2023, S.20). Zum anderen kann die KI bereits schon heute Musikstücke eigenständig erzeugen und mithilfe von neurologischen und biologischen Fachkenntnissen individuell an den Hörer anpassen. Ein Beispiel dafür ist die App Endel, die Aspekte wie die Tageszeit, Standort und auch biologische Faktoren für die Wiedergabe der Instrumentalstücke miteinbezieht (Liegmal, 2020).

Ein weiterer Aspekt stellt die adaptive Musikkomposition dar, welche im Bereich Games sowie für AR- und VR-Formate eingesetzt wird (Brown, 2021). Im Hinblick auf die funktionale Musik hält Heintz (2024) es für möglich, dass diese auch durch KI-Systeme realisiert werden kann. Sie vermögen es, Symboliken wie zum Beispiel Freude, Trauer oder Angst klanglich zu illustrieren, um bestimmte Wirkungen hervorzurufen. Jedoch stößt die KI an ihre Grenzen, sobald „der Kontext [eine] Komponistenpersönlichkeit erfordert“ (Heintz, 2024, S. 249).

Laut Gioti (2021) sind die KI-Systeme für Künstler nur bedingt verwendbar, da sie mit Stilimitationen erlernter Inhalte arbeiten und folglich künstlerische Neuschöpfungen limitieren. Zudem kritisiert sie, dass die Möglichkeiten zur Interaktion und Kollaboration zwischen Nutzer und dem KI-System noch nicht ausreichend gegeben sind.

Mittlerweile existiert eine Reihe an generativen Programmen und Anwendungen, die bei Kompositionen unterstützen oder diese durchweg selbst übernehmen. Sie liefern in wenigen Sekunden bis Minuten musikalische Entwürfe, die den Eingaben des Nutzers entsprechen. Dies beinhaltet prinzipiell „die Erstellung von Melodie-, Akkord- oder Schlagzeugmustern [sowie] Arrangement, Orchestrierung, Harmonisierung [und die] Auswahl von Klangfarben“ (Brown, 2021, S.7). Zumeist kann ebenfalls die gewünschte Stilrichtung gewählt werden. Nachfolgend wird eine Auswahl dieser vorgestellt.

MusicLM ist eine KI-Anwendung aus dem Jahr 2023, bei welcher textuelle Eingaben in Musik transferiert werden. Durch das Training mit einem großem Datensatz können bei der Generierung verschiedene Musikstile bedient werden. Zudem ist MusicLM mit der Funktion *Story Mode* versehen. Bei dieser können innerhalb der Eingabe verschiedene musikalische Anforderungen mit Zeitstempeln formuliert werden, die bei der Generierung berücksichtigt

werden (Berens & Bolk, 2023). Gerade hinsichtlich der Erstellung von Filmmusiken kann diese Funktion herangezogen werden.

Ecret Music ist für die Erstellung von Musik für Videoinhalte und Spiele spezialisiert. Bei dieser Anwendung erfolgt die Eingabe des Nutzers durch die Wahl von mindestens einer Szene, einer Stimmung und einem Genre, woraufhin ein Musiktrack erstellt wird. Dabei können im Anschluss die Instrumentengruppen und der strukturelle Aufbau in gewissem Umfang angepasst werden (Loth, 2024). Ecret Music wird von Soundraw betrieben, eine Anwendung, die ebenfalls zahlreiche Funktionen für Kreativschaffende sowie für Künstler zur Verfügung stellt (Soundraw, 2023).

Besonders umfangreich ist die Onlineplattform Melobytes, die eine Vielzahl an Anwendungen zur Erstellung von Musikinhalten anbietet. Neben Text können auch Bilder oder Videos als gültige Eingaben zu Musik transformiert werden. Zudem werden weitere Funktionen wie die Stimmenbearbeitung, die Generierung von Kurzfilmen sowie die visuelle und auditive Animation von Selfies angeboten (Melobytes, 2022).

Ein weiteres Tool ist AIVA, das bereits 2016 veröffentlicht wurde. Es ermöglicht den Nutzern, sich ebenfalls Musikstücke in verschiedenen Stilrichtungen generieren zu lassen (Catani, 2024b). Zunächst wurden die Kompositionen auf klassische Musik spezialisiert, bevor sich die Produktpalette stetig erweiterte. AIVA bietet die Möglichkeit sowohl Variationen bestehender Werke als auch neue Kompositionen zu erstellen und wird heutzutage kommerziell in diversen Medienformaten eingesetzt (Loth, 2024).

Auch Open AI verfügt über mehrere musikgenerierende Programme wie unter anderem MuseNet und Jukebox (Berens & Bolk, 2023). Weitere Tools sind AmperMusic (Loth, 2024), Suno AI (Suno, 2024) sowie die im April 2024 veröffentlichte Betaversion des Musikgenerators Udio (Udio, 2024).

### 3. Der Film Inception

Inception ist eine Science-Fiction-Heist-Thriller-Produktion von Christopher Nolan aus dem Jahr 2010. Der zweieinhalbstündige Film wurde durch die Produktionsfirmen Warner Bros., Legendary Entertainment sowie Syncopy realisiert. Neben der Regie hat Christopher Nolan auch das Drehbuch übernommen. Die Filmmusik wurde von Hans Zimmer komponiert. Bekannte Schauspieler wie Leonardo DiCaprio, Joseph Gordon-Levitt, Tom Hardy, Elliot Page, Ken Watanabe und Cillian Murphy agieren als Hauptbesetzung. Inception wurde mit zahlreichen Preisen ausgezeichnet, unter anderem mit vier Oscars im Jahr 2011 für die Kategorien beste Kamera, Ton, Tonschnitt und visuellen Effekte. Die Premiere des Filmes wurde am 8. Juli 2010 in London gefeiert. Drei Wochen später folgte in Deutschland der Kinostart. Insgesamt hat Inception weltweit einen Betrag von 839.030.630\$ eingespielt (IMDb, kein Datum). Christopher Nolans Ziel von Inception war es, einen Heist Film zu produzieren, in dem Personen Träume miteinander teilen, sodass eine neue Realitätsebene mit validen zwischenmenschlichen Interaktionen kreiert wird. Dabei sollten der Heist selbst sowie Cobbs Geschichte auf emotionalen Konzepten aufbauen (Nolan, 2010b). Die Genese und Inspiration entsprangen seiner eigenen Erfahrung mit Lucid Dreaming, welches das bewusste Befinden im Traum und dessen Versuch diesen zu kontrollieren, beschreibt (*James Cameron's Story of Science Fiction*, 2018).

#### 3.1 Handlung

Dom(inik) Cobb ist Experte in seinem Fachgebiet: das Stehlen von Informationen in Träumen, der *Extraction*. Dazu nutzt er die Militärtechnologie des Traum-Sharings, bei der im gemeinsamen Erleben von Träumen in das Unterbewusstsein einer Person eingedrungen wird, um wichtige Informationen zu extrahieren. Im Schlaf eröffnen sich tiefere Bewusstseins-ebenen, wodurch die Zeit intensiver erlebt wird und folglich langsamer als in der Realität vergeht. Von diesen Gesetzmäßigkeiten machen sich Cobb und sein Team bei der Durchführung von Extractions Gebrauch.

Ein Auftrag scheitert, da die Zielperson Saito realisiert, dass er sich in einem Traumkonstrukt befindet. Daraufhin wollen Cobb und sein Geschäftspartner Arthur die Stadt verlassen, werden jedoch von Saito mit einem Gegenangebot abgefangen. Er ist ein Geschäftsmann und möchte verhindern, dass ein Konkurrenzunternehmen die Vormachtstellung im Energiemarkt erlangt. Dies sieht er nur möglich, wenn dem Erben des Unternehmens, Robert Fischer, der Gedanke zur Auflösung des Imperiums seines Vaters in den Verstand eingepflanzt wird. Gelingt Cobb diese Mission, wird Saito ihm im Gegenzug die Rückkehr in sein altes Leben in den Vereinigten Staaten mit seinen Kindern ermöglichen. Das Säen einer Idee in den Verstand, einer sogenannten Inception, gilt jedoch als unmöglich. Cobb hingegen weiß, dass eine Inception

realisierbar ist, da er selbst eine solche bei seiner Frau Mal bereits durchgeführt hat. Diese war auch der Auslöser dafür, dass er die USA verlassen musste, denn infolge der Inception hat Mal den Bezug zur Realität verloren und sich das Leben genommen. Da sie den Selbstmord ihrem Mann angehängt hat, steht ihm bei Verbleib in den USA ein strafrechtlicher Prozess bevor.

Cobb nimmt den Auftrag an und stellt mit Arthur ein Team zusammen. Dieses besteht aus Eames, dem Identitätsfälscher, Ariadne, der Architektin der Traumebenen und Yusuf, der für die Sedierung und das Traumserum zuständig ist. Auch Saito wird zum Teil des Teams, um den Erfolg der Inception zu überprüfen. Um zwischen Traum und Realität unterscheiden zu können, besitzt jedes Teammitglied ein spezielles Totem. Bei Cobb ist es ein Kreisel, der sich in der Realität ausdreht und umkippt, im Traum hingegen sich unendlich weiterdreht. Die Inception soll während eines Fluges von Sydney nach Los Angeles stattfinden. Für die Mission entwirft das Team drei Traumebenen, um den Gedanken glaubwürdig in Fischer zu platzieren. Während dieser kommt es zu Komplikationen, sodass Saito und Fischer in den Limbus fallen. Der Limbus ist ein nicht entworfener Traumraum, eine unendliche Weite des rohen Unterbewusstseins. Der Film endet damit, dass Fischer, Cobb und sein Team im Flugzeug aufwachen und in die USA einreisen. Cobb kehrt zu seinen Kindern nach Hause zurück und dreht seinen Kreisel auf einem Tisch an. Als dieser zu zucken beginnt und umfallen könnte wird das Bild schwarz, sodass es nicht eindeutig ist, ob die Szene nun real oder geträumt ist.

### **3.2 Szenenwahl und Begründung**

Für die Analyse wird die Sequenz ab der Vollendung der Inception bis zur Heimkehr Cobbs zu seinen Kindern [02:03:58-02:15:10] betrachtet. Ausschlaggebend für die Wahl dieses Filmabschnitts ist die hohe Ereignisdichte, ebenso wie die mehrmals wechselnden Traumebenen und die mutmaßliche Rückkehr in die Realität. Die Musik ist in dieser Phase des Films sehr präsent und untermalt sowohl die ruhigeren, emotionalen als auch die spannungsgeladenen und actionreichen Momente. Damit eignet sich diese Sequenz, um die Wirkungen und Varianz der Filmmusik zu untersuchen. Die Analyse soll dabei zeigen, wie die Traumebenen dargestellt sind und ob es Zusammenhänge zwischen diesen gibt. Zudem soll untersucht werden, wie die Nebenhandlungen, die Beziehung von Fischer zu seinem Vater sowie Cobbs Familiendrama musikalisch umgesetzt werden. Wird durch die Musik eventuell ein Hinweis auf den tatsächlichen Ausgang des Filmes gegeben?

Die Szenenwahl begründet sich auch im Hinblick auf die Erstellung der künstlich generierten Komposition. Es bedarf an unterschiedlichen Musikstücken, die dennoch im Zusammenhang miteinander funktionieren müssen. Eine intensive Auseinandersetzung mit den Möglichkeiten des gewählten KI-Programms ist demnach für die musikalische Neuvertonung notwendig.

## 4. Analyse

Filme bringen Menschen zum Lachen, zum Weinen oder auch zum Fürchten. Das liegt mitunter daran, dass das Zusammenspiel der Bild- und Tonebene Affekte bei den Zuschauern in Form von unmittelbaren Reaktionen und vornehmlich auch unterbewusst hervorruft (Kreuzer, 2009). Um diese Wirkung nachzuvollziehen, empfiehlt es sich eine filmmusikalische Analyse durchzuführen, die zugleich den durch die Musik aufgespannten erweiterten Bedeutungsraum aufzeigt. Für eine umfassende Analyse sollten zusätzlich die bestehenden Produktionsumstände sowie die verwendeten Techniken betrachtet werden (Tieber, 2020).

### 4.1 Filmmusik Inception

Zunächst werden grundlegende Aspekte der Filmmusik bezüglich der Instrumentierung, der verwendeten Themen und des Musikeinsatzes herausgearbeitet. Auf diese Kenntnisse wird in der anschließenden tiefergreifenden Szenenanalyse aufgebaut.

Eine Besonderheit der Filmmusik in Inception ist, dass diese nicht anhand des geschnittenen Bewegtbilds, sondern bereits vor dem Dreh basierend auf dem Skript von Nolan entwickelt wurde. Der Regisseur begründet diese Entscheidung damit, dass er Zimmers Imagination und Interpretation der vorgestellten Idee frei entfalten lassen wollte, ohne dass er durch vorhandene Bilder eingeschränkt wird (Hans Zimmer - Making Of Inception Soundtrack, 2011). Durch diese Vorgehensweise trägt die Musik die Emotionen im Film und wird nachträglich für die Bildfolge angepasst und variiert. Insgesamt ist die Handlung in weiten Teilen mit Musik unterlegt, wobei die non-diegetische Musik deutlich überwiegt. Es gibt nur ein einzelnes Musikstück, das wiederholt in diegetischer Form für die Handlung verwendet wird.

Auch die Abwesenheit von Musik wird in mehreren Szenen bewusst eingesetzt. Sie dient dazu, die Aufmerksamkeit auf wichtige Handlungspunkte sowie Wendungen zu lenken. Momente der Erkenntnis werden durch Stille festgehalten und anschließend durch einsetzende Musik aufgelöst. Dabei fällt das Aussetzen der Musik nur bedingt auf, da sie entweder schrittweise reduziert oder durch einsetzende Soundeffekte maskiert und ausgeblendet wird. Dies kann zum Beispiel beobachtet werden, als Cobb und Ariadne sich in einem Café in Paris unterhalten. Zu Beginn wird die Szene musikalisch begleitet, als er ihr das Konzept des gemeinsamen Träumens erklärt. Mit Cobbs Andeutungen, dass sie sich gerade nicht mehr in der Realität befinden, endet die Musik. In dem Moment, als Ariadne dies realisiert setzen die Streicher erneut ein. Die musikalische Gestaltung führt dazu, dass sich der Zuschauer in der Perspektive der Architektin befindet und vorerst nicht weiß, dass es sich um einen Traum handelt. Ein weiteres Beispiel ist das Gespräch zwischen Cobb und seinem Schwiegervater in einem Hörsaal, bei dem neben dem Dialog nur einzelne Elemente des Sounddesigns hörbar sind. Auf Nachfrage erzählt Cobb von dem Auftrag und seiner Suche nach einem neuen

Architekten. Die Musik beginnt, als Cobb von der Möglichkeit spricht, durch diese Mission wieder nach Hause zu seinen Kindern zurückkehren zu können. Die Musik suggeriert die in Cobb erweckte Hoffnung, welche durch die Hilfe seines Schwiegervaters näher in die Realität rückt.

Die einzige Sequenz im Film, in welcher die Musik über mehrere Minuten aussetzt ist, als Cobb und sein Team die erste Traumebene für die Durchführung der Inception betreten. Parallel zur Freisetzung des Traumserums wird das Verlassen der Realität durch das Innehalten der Musik illustriert. Mit einem einsetzenden Soundeffekt wird der Zuschauer mit in die Traumwelt hinabgezogen. Diese wird in der folgenden Actionszene mit einer Verfolgungsjagd und einem Schusswechsel zunächst ohne musikalische Mittel dargestellt, sondern ausschließlich durch das Sounddesign auditiv etabliert. Erst bei der Erkenntnis, dass Saito angeschossen wurde, setzen die atmosphärischen Klänge wieder ein und die musikalische Untermalung der Traumebene beginnt.

### **Diegetische Musik**

Wie bereits erwähnt, besteht die diegetische Musik in Inception lediglich aus einem wiederkehrenden Liedabschnitt. Bei diesem handelt es sich um das Chanson *Non, je ne regrette rien* von Édith Piaf, das sowohl für die Zuschauer als auch für die Protagonisten in der filmischen Realität hörbar ist. Es wird in der Filmhandlung als musikalischer Countdown für den Kick verwendet, um die Traumwelt zu verlassen bzw. in eine höhere Ebene zu wechseln. Dabei wird die Originalaufnahme von 1960 verwendet, wobei immer der Beginn des Chansons angespielt wird. Die Wiedergabe startet mit dem Einsetzen des Gesangs und den Worten ‚non, rien de rien‘. Neben der Stimme ertönen die markanten Blechbläser, deren Bedeutung für die non-diegetische Musik im Verlauf des Kapitels erläutert wird.

Diese musikalische Ankündigung des Kicks wird als eine gängige Methode des Teams um Cobb vorgestellt und wird im Handlungsverlauf insgesamt drei Mal eingesetzt. Das erste Auftreten erfolgt im Anschluss an die erste Szene, als in einem Zug einer schlafenden Person Kopfhörer aufgesetzt werden. Da sich das Lied auch auf die parallel geschnittene Szene auswirkt, wird ein Bewusstsein für die Funktion dieser Musik entwickelt. Zudem wird deutlich, dass die vorherigen Ereignisse nicht in der Realität, sondern in Traumebenen stattgefunden haben. Der musikalische Countdown überträgt sich über die verschiedenen Traumebenen und den darin enthaltenen Personen und wird dabei mit zunehmender Traumtiefe entsprechend in seiner Wiedergabegeschwindigkeit verlangsamt. Dies zeigt sich, als Yusuf in der ersten Traumebene Arthur die Kopfhörer im Van aufsetzt und die Wiedergabe startet. Arthur nimmt das Lied eine Ebene tiefer verlangsamt in einem Treppenhaus wahr, während das restliche Team in der dritten Traumschicht den Countdown aufgrund der sehr hohen Temporeduktion kaum noch als Lied wahrnehmen kann.

## **Instrumentierung**

Für den Score von Inception hat sich Zimmer noch mehr als in seinen vorherigen Produktionen den elektronischen Klangelementen zugewandt und kombiniert diese mit akustischen Instrumenten. So hat er ausschließlich am Computer erstellte Konzepte durch das Orchester nachimitieren lassen, wodurch viele der Atmosphären und Umgebungssounds entstanden sind (Hans Zimmer - Making Of Inception Soundtrack, 2011).

Die Kombination des orchestralen und elektronischen Klangcharakters erforderte eine Reihe an Instrumenten, die für die Aufnahmen herangezogen wurden. Zum einen besteht die Bläserbesetzung aus sechs Bassposaunen, sechs Tenorposaunen, vier Tuben und sechs Hörnern (Hans Zimmer - Making Of Inception Soundtrack, 2011). Diese Konstellation ermöglicht es, den Bläsern eine sehr druck- und kraftvolle Wirkung zu verleihen, sodass sie bei voller Besetzung einen lauten bis dröhnenden Charakter ausüben. Zum anderen werden auch Holzbläser, beispielsweise ein Fagott, verwendet. Neben den Bläsern liegt eine Streichersektion bestehend aus ersten und zweiten Geigen, Bratschen, Celli und Kontrabässen vor, die besonders für die emotionalen Passagen zum Einsatz kommen. Weiterhin wurden verschiedene Schlaginstrumente, eine E-Gitarre, ein Flügel und diverse Synthesizer verwendet (Hans Zimmer - Making Of Inception Soundtrack, 2011).

## **Non-diegetische Musik**

Die non-diegetische Filmmusik in Inception zeichnet sich durch die Verwendung zweier unterschiedlicher Klangcharakteristiken aus. Zum einen wird in den actionreichen und spannungsgeladenen Szenen vorrangig mit den Bläsern und überlagerten Rhythmen gearbeitet. Erstere erinnern teilweise an Nebelhörner und werden zur Beschreibung der Traumebenen verwendet, die neben den synthetischen Klangflächen zu einer düsteren Atmosphäre beitragen. Zum anderen werden für die ruhigeren und entschleunigten Sequenzen mit langen schwebenden Akkorden, emotionsführenden Flächen und subtilen Tonhöenschwankungen gearbeitet. Zimmer war es wichtig, Emotionen über die Musik zu transportieren. Er wollte dadurch auch den Schmerz und die Schuldgefühle Cobbs zum Vorschein bringen. Dafür nutzt er die Klangfarbe einer E-Gitarre und hat sich mit dem Gitarristen Johnny Marr zusammengeschlossen (Itzkoff, 2010). Sie haben eine repetitive, melancholische E-Gitarrenmelodie entwickelt, die mit dem Orchester verwoben und zugleich an die Figur Cobb geknüpft wird.

Durch die Nutzung dieser gegensätzlich wirkenden Stile ergibt sich eine hohe Dynamikspanne innerhalb der Musik. Auffällig dabei ist, dass gewisse Score-Tracks auf Wiederholungen basierenden Strukturen aufgebaut sind. So beginnt ein Stück zunächst mit nur einem oder wenigen Instrumenten und baut sich zunehmend durch Hinzufügen weiterer Instrumente

während den Wiederholungen der vorherigen Passagen auf. Gerade bei den Titeln *Dream is collapsing*, *Mombasa* und *Time* kann dies beobachtet werden.

Der Score besitzt einen überwiegend paraphrasierenden Charakter, der durch die Verwendung der deskriptiven Technik erreicht wird. Ein Beispiel für die bildunterstützende Wirkung der Musik ist die Szene in *Mombasa*, als Cobb vor Agenten seines alten Arbeitgebers flieht. Sie begleitet die Flucht sowohl in ihrer Dynamik als auch instrumental mit. Dabei treiben die tiefen Bässe mit den überlagerten Rhythmen die Verfolgungsphasen aktiv mit an. Bei dem Versuch sich zu verstecken wird die Musik leiser und auf wenige Instrumente reduziert, bis sie sich für den zweiten Part der Verfolgung erneut aufbaut. In dem Moment, als Cobb in das Auto von Saito einsteigt und die Gefahrensituation geklärt ist, wird die gewonnene Sicherheit durch die abrupte Entspannung auf der musikalischen Ebene dargestellt.

Für die Ausarbeitung des Scores wurde zudem viel mit der Mood Technik gearbeitet, da die musikalischen Konzepte vor dem Bewegtbild erstellt wurden und die Musik dadurch einen stimmungstragenden Charakter erfährt. Besonders bei den Sequenzen, welche Einsicht in Cobbs Vergangenheit und in seine Erinnerungen an Mal geben, werden über die Musik seine Schuldgefühle und sein innerlicher Schmerz offenbart.

In vereinzelt Sequenzen wird der Score aber auch polarisierend eingesetzt. Dies kann bei der Vorstellung von Robert Fischer und der Beziehung zu seinem Vater beobachtet werden. Die Szene beginnt in einem Büro, in dem Geschäftsmänner über ein Anliegen der Firma diskutieren. Musikalisch wird jedoch eine düstere bis bedrohliche Stimmung ausgestrahlt. Anschließend hält die Musik kurz inne, als die Tür zum Nebenzimmer geöffnet und sowohl Robert Fischer als auch der kritische Gesundheitszustand seines Vaters gezeigt werden. Subtile Glockenschläge weisen dabei auf die ihm wenig verbleibende Zeit hin, während die Bilder auf das angespannte Verhältnis zwischen Fischer und seinem Vater hindeuten. Die Musik hingegen signalisiert, dass die Beziehung zwischen Vater und Sohn irreparabel ist und, dass Robert Fischer darunter leidet.

### **Musikalische Themen**

Für *Inception* wurde mit mehreren musikalischen Themen gearbeitet, die wiederholt für die Handlung aufgegriffen und variiert werden. Das Hauptthema resultiert aus einer Note des diegetisch verwendeten Liedes von Édith Piaf und wird durch die Blechbläserbesetzung verkörpert. Entsprechend der Dramaturgie wird das Tempo und der Rhythmus für die jeweiligen Actionszenen variiert, wodurch sich die Wirkung leicht ändert, die Zuordnung zum ursprünglichen Motiv jedoch bestehen bleibt. Das dominante Bläserthema wird bereits im Vorspann eingeführt und endet mit dem ersten Bild der Filmhandlung.



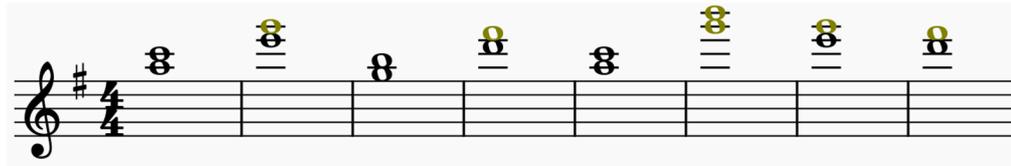


Abbildung 4.3 Motivvariation der Streicher bei der Erklärung des Totems

Erstmalig erklingt diese Melodie mit dem Beginn des Vorspanns durch ein Klavier und wird durch das Bläserhauptthema abgelöst. Weitere Passagen sind, als Arthur die Funktion des Totems erklärt oder als Mal aus dem Limbus in die Realität erwacht. Eine Variation dieses Themas kann beobachtet werden, als Ariadne während Cobbs Traum in seine Erinnerung der Todesnacht von Mal eindringt. Das Motiv wird über den Hall eines zerbrechenden Glases eingeleitet und dissonant weitergeführt. Für das Ende des Films wird die Melodie erneut aufgegriffen. Sie beginnt als Cobb und Saito im Limbus miteinander sprechen und nimmt anschließend im finalen Musikstück *Time* ihre ausgeprägteste Form an.

Ein drittes Thema tritt bei der Planung für und während der Traummissionen auf, bei denen Informationen extrahiert werden. Dieses kennzeichnet sich vor allem durch die durchgehenden Sechzehntelrhythmen, die von einer E-Gitarre gespielt werden. Es wird für eine geheimnisvolle Klangkomponente sowie zur Entwicklung des Spannungsverlaufs eingesetzt.



Abbildung 4.4 Rhythmische Melodie der E-Gitarre für das dritte musikalische Thema

## 4.2 Einführung in die Analyse

Die nachfolgende Analyse der filmmusikalischen Gestaltung basiert auf einem Verfahren von Chion und Neumeyer, welches zunächst vorgestellt wird.

In seinem Werk *Audio-Vision: Ton und Bild im Kino* beschreibt Chion (2012) die sogenannte Abdeckmethode. Bei dieser erfolgt die Analyse durch mehrfaches Sichten der Sequenz. Dabei werden die Bild- und Tonspur sowohl zusammen als auch isoliert voneinander betrachtet. Eine vorgeschriebene Reihenfolge gibt es nicht, es wird jedoch empfohlen zuerst die einzelnen Elemente vor deren Zusammenwirken zu untersuchen. Der Mehrwert dieser Methode liegt darin, dass Bild und Ton in ihrer reinen Form analysiert werden können, ohne dass sie von ihrem Gegenpart beeinflusst werden. Eine Schwierigkeit stellt die Projektion von bestehenden Kenntnissen auf die eigene Wahrnehmung dar. So sollte das bereits Gesehene nicht auf das Gehörte abgebildet werden oder umgekehrt. Die Abdeckmethode funktioniert nur, wenn die

Bild- und Tonebene als einzelne Elemente nebeneinander weiterbestehen und nicht gänzlich ineinander verschmelzen (Chion, 2012).

Daneben gibt es die Zwangsehe, bei welcher der Originalton der Sequenz zunächst verborgen bleibt. Stattdessen werden in mehreren Durchläufen verschiedene Musikstücke unter das Bewegtbild gelegt, wovon sich zumeist in einzelnen Stücken Synchronisationspunkte ergeben und als passend empfunden werden. Durch die Musikwechsel wird deutlich, welche Varianz in der Bilderzählung möglich ist. Erst danach wird der Originalton gezeigt, wodurch die Bild-Tonbeziehung wesentlich bewusster wahrgenommen wird (Chion, 2012). Diese Methode wird nachfolgend nicht berücksichtigt, da die vorliegende Filmmusik bereits bekannt ist und in Kapitel fünf umfassend mit alternativen Musikansätzen gearbeitet wird.

Für ein strukturelles Vorgehen hat Chion (2012) einen Entwurf für einen Fragebogen entwickelt. Da sich dieser auf die Analyse der gesamten Tonspur bezieht, werden die von Neumeyer angefügten Erweiterungen für die Filmmusik berücksichtigt. Die Analyse wird in vier Teilaspekte untergliedert (Tieber, 2020). Zunächst werden in der Auflistung die vorkommenden musikalischen und weiteren akustischen Elemente benannt sowie die Musikeinsätze dokumentiert. Bei der *Charakterisierung* wird anschließend die Balance und die Interaktion der Elemente des Soundtracks zueinander untersucht. Zudem wird die Qualität und die Konsistenz beleuchtet. Unter letzterer wird die Positionierung der klanglichen Elemente verstanden. Sind diese „in eine[n] globalen Teig, in eine Textur eingearbeitet, oder [...] für sich getrennt deutlich hörbar“ (Chion, 2012, S.152). Im dritten Schritt erfolgt die Lokalisierung der Synchronisationspunkte in der Musik, die überdies in deren Bedeutung eingeordnet werden. Für eine tiefergehende Analyse wird ein Vergleich für das Zusammenwirken der Bild- und Tonebene unter vorgegebenen Gesichtspunkten angestellt. Bezüglich technischer Aspekte wird beispielsweise das Verhalten der Tonebene bei der Änderung der Cadrage durch eine Bewegung der Kamera untersucht. Begleitet sie den Wechsel auditiv mit, wird er ignoriert oder eventuell stärker betont?

Werden sowohl Bild und Ton in ihrer Repräsentation bezüglich eines Kriteriums analysiert, wird dies den formalen Vergleichsaspekten zugeordnet. So kann herausgefunden werden, ob die Ebenen verschiedene Tempi aufweisen und wie sie in ihrer „Substanz und Definition“ (Chion, 2012, S.153) zueinander stehen. Die narrative und gestalterische Analyse dient dazu, die Funktionen der beiden Ebenen für die Erzählung nachzuvollziehen. Es wird untersucht, ob sich Bild und Ton „ergänzen, widersprechen oder sich verstärken“ (Chion, 2012, S.153). Dazu kann mit den Fragen „Was sehe ich von dem was ich höre?“ und „Was höre ich von dem was ich sehe?“ (Chion, 2012, S.154) gearbeitet werden. Sie verhelfen zudem auf negative Bilder und Töne einer Einstellung aufmerksam zu werden. Ein negativer Ton ist nicht hörbar, wird

aber durch die Bildebene suggeriert, wodurch sich dessen Bedeutung verstärken kann. Dies gilt äquivalent auch für negative Bilder (Chion, 2012).

### **4.3 Analyse der Szene [02:03:58 – 02:15:10]**

Nach dem beschriebenen Prinzip wird nun zusammen mit der Abdeckmethode die im vorherigen Kapitel aufgeführte Sequenz abschnittsweise analysiert. Dabei konzentriert sich die Analyse auf die verwendete Filmmusik, sodass der Dialog und das Sounddesign nur an den erforderlichen Stellen miteinbezogen werden. Die Sequenz wurde entsprechend des dramaturgischen Verlaufs dreigeteilt. Die Verweise auf den Film werden in [Stunde:Minute:Sekunde] angegeben, wobei sich die Zeitangaben auf die Wiedergabe der DVD-Version (Nolan, 2010a) beziehen. Es empfiehlt sich, die Szene zum besseren Verständnis im Voraus anzusehen.

#### **4.3.1 [02:03:58 – 02:07:20]: Traumebene drei / Kick**

##### **Bild ohne Ton**

Die Szene beginnt mit einer Kamerafahrt in der dritten Traumebene, bei der Fischer eine große Tresortür öffnet und anschließend die Totale eines fensterlosen Krankenzimmers gezeigt wird. Fischer läuft auf seinen Vater zu, der angeschlossen an Geräte in einem Krankenbett liegt. Dieser beginnt zu reden, sodass Fischer sich zu ihm kniet und sein Gesichtsausdruck zunehmend emotionaler wird. Eames öffnet unbemerkt die Tresortür und aktiviert Sprengsätze, die in einer Detailaufnahme gezeigt werden. Der Vater gestikuliert erst in die Richtung seines Sohnes, dann auf einen Safe, der neben seinem Bett steht. Erneut wird Eames an derselben Position gezeigt, der den Moment beobachtet und etwas spricht. Fischer öffnet den Safe und findet das Testament sowie ein Papierwindrad und wird von seinen Emotionen übermannt. Mit einem Schnitt auf die Halbtotalen wird ersichtlich, dass Fischers Vater in der Zwischenzeit verstorben ist. Eames löst daraufhin die Sprengsätze aus, woraufhin Gebäudeteile explodieren und die dritte Traumebene zusammenstürzt.

Parallel zu dieser Sequenz wird mehrmals in die anderen Traumebenen geschnitten, in denen der abstürzende Van sowie der sich schnell bewegende Aufzug inklusive Arthur und das schlafende Team zu sehen sind. Auch der zusammenstürzende Limbus wird gezeigt, bevor Ariadne Cobb dort zurücklässt und aus dem Gebäude springt. Die Kameraeinstellung wechselt von einer Totalen in Nahaufnahmen, als sie und die anderen Teammitglieder über schnell geschnittene Bildfolgen in die erste Traumebene zurückgeführt werden. Die Sequenz endet mit dem Aufprall des Vans, welcher samt aller Insassen versinkt.

## **Ton ohne Bild: Auflistung und Charakterisierung**

Die non-diegetische Musik ist das am stärksten vertretene auditive Element dieser Sequenz. Sie zeichnet sich durch drei unterschiedliche Musikeinsätze aus, welche die Dramaturgie und die Dynamik der Szene auf der auditiven Ebene steuern. Lediglich die ersten und letzten Sekunden sind durch die Abwesenheit der Musik geprägt. Neben dieser sind kurze Dialogeinschübe sowie verschiedene Geräusche und Effekte des Sounddesigns zu hören.

Die Sequenz beginnt ruhig, es sind ausschließlich der Atem einer Person sowie ein metallisches Einrasten zu vernehmen. Ein schleifendes Türengeräusch folgt parallel zu dem Einsatz langsamer Musik. Diese beginnt tief und baut sich durch weiter einsetzende Streicher auf. Sie wirkt atmosphärisch und verleiht dem Moment einen düsteren Charakter. Es findet ein kurzer Dialog zwischen Fischer und seinem Vater statt. Wortfetzen des Vaters werden durch Fischer ergänzt, bevor mit der Aussage „weil du es versucht hast“ abrupt eine neue Musik einsetzt. Diese wird durch schnelle, sich ständig wiederholende Rhythmen angeführt und nach kurzer Zeit durch eine in der Tonhöhe ansteigende Melodie der Streicher ergänzt. Vereinzelt sind kurze Geräusche neben der sich stetig steigenden Musik zu hören, die mit lauten Explosionsgeräuschen endet und nahtlos in den dritten Musikeinsatz übergeht. Dieser beginnt mit einer stark verlangsamten und beinahe dröhnenden Variation des Bläserthemas, welches von einer zweiten höheren und im Tempo etwas schnelleren Variante überlagert wird. Erneute Explosionen leiten eine weitere musikalische Änderung ein, mit welcher Instrumente hinzukommen, die das Bläserthema ebenfalls unterstützen. Durch den Dialog von Cobb und Ariadne wird deutlich, dass es sich in diesem Moment um den Kick handelt. Mit dem Einsetzen eines sehr langgezogenen Gesangs werden die tiefen Bässe in das Bläserhauptthema zurücktransformiert. Es folgt ein kurzer Moment der Stille, bevor lautes Wasserrauschen zu hören ist und der Klang dumpf wird.

Über den gesamten Szenenabschnitt hinweg dominiert die Musik und ist durchgehend zu hören, während das Sounddesign natürlich und ergänzend wirkt. Sie suggeriert Stimmungen und treibt die Szene voran, wobei ausreichend Raum für die punktuellen Einsätze des Sounddesigns gelassen wird, welche separiert von der Musik wahrgenommen werden können. Mit dem Einsetzen des Bläserthemas findet ein deutlicher Lautstärkesprung statt und auch das Sounddesign wird zunehmend präsenter. Die Musik behält die Kontrolle, gleichzeitig treten für einige Sekunden viele Geräusche hervor, die ohne das Bild nicht mehr einzeln identifizierbar sind. Der Dialog ist reduziert und agiert hintergründig, liefert jedoch wichtige Erkenntnisse, die durch den Wechsel der Musik aufgenommen und interpretiert werden.

## **Ton und Bild zusammen**

### Maskierung

Bei der gemeinsamen Betrachtung der visuellen und auditiven Ebene wird zunächst untersucht, ob Maskierungen durch die Musik vorliegen. Bei dem Aufprall des Aufzugs [02:07:04] ist dies der Fall, da weder die Kollision noch der Funkenflug durch Geräusche begleitet werden. Zudem sind weder die Tastengeräusche des Safes [02:05:18 und 02:05:25], noch die lebensüberwachenden Geräte [02:04:19] im Krankenzimmer auditiv unterlegt. Dies kann jedoch auch durch eine bewusste gestalterische Entscheidung begründet werden. Im Gegenzug dazu wird dem Sounddesign Vorrang gegeben und die Musik in der Lautstärke zurückgenommen, als der nach oben fahrende Aufzug erstmalig in der Szene gezeigt wird [02:04:49].

### Synchronisationspunkte

In dieser Szene liegen mehrere zentrale Synchronisationspunkte zwischen der musikalischen und der visuellen Ebene vor. Sie dienen zum einen der Demonstration der erfolgreich durchgeführten Teilpunkte des Plans, Fischer die Idee zur Auflösung des Unternehmens in den Verstand zu pflanzen. Gleichzeitig leiten sie die weiteren Schritte ein und treiben damit die Handlung zu diesem Zeitpunkt maßgeblich voran. Der Dialog findet stets lippen synchron statt und wird daher nachfolgend nicht berücksichtigt.

Bild und Ton fallen erstmals mit der Einführung in diese Sequenz zusammen, als Fischer die Tresortür entriegelt und in das Krankenzimmer seines Vaters eintritt [02:03:59]. Die visuelle Komponente arbeitet hier mit der Musik und dem Sounddesign zusammen, da neben den Öffnungsgeräuschen der Tür auch die Musik einsetzt und die Szene dadurch in eine ernste und bedrückte Stimmung gelenkt wird. In diesem Moment wird zugleich die finale Phase zur Durchführung der Inception eingeleitet.

Kurze Zeit später zeigt das Bewegtbild Fischers Reaktion, nachdem er den wahren Grund der Enttäuschung seines Vaters erfahren hat [02:04:40]. Synchron zum Bildschnitt auf Fischer ändert sich der musikalische Charakter drastisch und die nebeneinander laufenden Handlungsstränge außerhalb des Krankenzimmers werden durch eine Parallelmontage dargestellt. Mit diesem Synchronisationspunkt realisiert Fischer seine vorherige Fehleinschätzung und glaubt nun daran, nicht in die Fußstapfen seines Vaters treten zu müssen. Es wird deutlich, dass ein weiterer Schritt des Plans erfolgt ist und nun die nächste Phase eingeleitet wird.

Ein markanter Zusammenfluss der gesamten Ton- und Bildebene erfolgt, als Eames die Sprengsätze auslöst [02:05:56]. Die Explosionen werden synchron mit Geräuschen und Effekten unterlegt, während parallel der neue Musikeinsatz beginnt. Dieser Moment signalisiert den Start des Kicks in der dritten Traumebene, der alle Teammitglieder in die erste Schicht

zurückführen soll, um die Inception zu vollenden. Die musikalische Akzentuierung verdeutlicht, dass sich der Kick auch auf den Limbus übertragen hat. Die Geräusche begleiten die Ereignisse im Bewegtbild, so wird Ariadnes Fall aus dem Limbus in Ebene drei durch einen Impact-Sound betont [02:06:51].

Der Aufprall des Vans in Traumebene eins [02:07:08] ist weniger eindeutig als die vorangegangenen Synchronisationspunkte. Er wirkt verzögert, da Musik und Sounddesign unabhängig voneinander positioniert sind. Mit der Rückkehr in die erste Traumebene klingt die Musik aus, es folgt eine minimale Pause und anschließend isoliert das Sounddesign. Diese Kombination illustriert den Wechsel der Ebene und verdeutlicht, dass auch das Bewusstsein der Protagonisten wieder in diese zurückgekehrt ist.

### Technischer und formaler Vergleich

Durch den technischen Vergleich wird deutlich, dass weder die Kamerafahrten noch die Kameranews durch die Musik begleitet werden. Stattdessen bleibt sie über die gesamte Sequenz auf einer übergeordneten Ebene und unterstützt die dramaturgische Entwicklung. Durch die Musik werden die Handlungsstränge der Parallelmontage verbunden sowie die Atmosphäre und die Spannung in großem Maße transportiert (Kloppenburg, 2017b).

Formal fällt auf, dass Bild und Ton in ihrer Substanz durchweg zueinander passen. Zu Beginn dominieren bedrückende und dunkle Bilder, welche durch die Streicher musikalisch verstärkt werden. Der Fokus liegt auf der angespannten und zugleich fragilen Vater-Sohn-Beziehung, die in dieser Szene zu einem versöhnlicheren Ende gebracht wird. Sobald sich die Aufmerksamkeit in der Bildebene auf mehrere zeitgleiche Handlungen verteilt, wird dies ebenso auf die detailreichere musikalische Struktur übertragen. Die visuellen Eindrücke des brennenden Aufzugs, der aktivierten Sprengsätze und des Krankenzimmers werden mittels der durchgehenden Musik dramaturgisch miteinander verbunden. Für den Höhepunkt werden ausdrucksstarke Bilder mit einer ebenso monumentalen und lauten Musik zusammengeführt.

Die Substanz korrespondiert mit den Tempi der Ebenen, die mit dem Verlauf der Sequenz mehrmals variieren. Die zuvor beschriebenen Musikeinsätze geben dabei jeweils das Tempo vor, an das sich das Bewegtbild stets anpasst. Während die Szene im Krankenzimmer auditiv und visuell langsam beginnt und eine gewisse Intimität ausstrahlt, erfahren beide Ebenen eine Beschleunigung mit dem zweiten Musikeinsatz [02:04:46]. Dies wird einerseits durch die schnellen und sich wiederholenden Rhythmen, andererseits durch die zügigen Bildschnitte innerhalb der Traumebene sowie zwischen den Ebenen realisiert. Durch die darüberliegende höherwertige Streichermelodie wirkt die Musik zunehmend schneller und verstärkt die entstehende Spannung. Diese audiovisuelle Kombination bringt die Handlung in kurzer Zeit voran und arbeitet auf den Höhepunkt der Szene hin [02:05:56], der sich durch eine abrupte Entschleunigung und einen Tonlagenwechsel in der Musik abzeichnet. Die rhythmisch

überlagerten Bassschläge dominieren, bis sie vereinheitlicht durch das Orchester übernommen werden. Zusammen mit den schnellen Bildschnitten führen sie dazu, dass Ton und Bild einen gemeinsamen Rhythmus annehmen. Die Kamera verfolgt die einzelnen Protagonisten, während die Musik die Gesamtsituation zusammenhält. Die visuelle Darstellung der Schichten berücksichtigt dabei die Zeitverhältnisse zwischen den Traumebenen. Die Haupterzählung findet zu diesem Zeitpunkt noch in der dritten Ebene statt, sodass diese in Normalgeschwindigkeit abgespielt wird. Bei den Schnitten in den höheren Ebenen wird das Bewegtbild in deutlich reduzierten Geschwindigkeiten wiedergegeben. Besonders bei der einsetzenden Schwerelosigkeit im Aufzug kann dies beobachtet werden. Auch auditiv wird dies ansatzweise übernommen, als der Van auf der Wasseroberfläche aufschlägt [02:06:10] und durch einen dumpfen und zeitlich gedehnten (time-stretched) Impact unterlegt wird.

Eine weitere Tempoveränderung geht mit dem Verlassen der zweiten Traumebene einher [02:06:59]. Zusammen mit dem Bild verlangsamt sich auch die Musik, bremst regelrecht ab und klingt mit dem Schnitt auf den ins Wasser stürzenden Van aus.

### Narrativer Vergleich

In Bezug auf die Narration spielt die musikalische Ausarbeitung eine wichtige Rolle. Sie spiegelt die einzelnen Phasen des Plans wider und ändert sich dabei stilistisch von atmosphärischen und flächigen Klängen hin zu lauten und rhythmischen Kompositionen. Die Musik ist paraphrasierend eingesetzt und bildet einen Spannungsbogen über die gesamte Sequenz.

Als Fischer auf seinen Vater im Krankenzimmer trifft, erklingt dieselbe Musik wie bei deren Vorstellung [00:44:20] im ersten Drittel des Filmes. Dadurch wird erneut das geschädigte Vater-Sohn-Verhältnis impliziert. Ebenso wird deutlich, dass das Gelingen des Plans von diesem kritischen Moment abhängt. Die Musik greift der Handlung nicht voraus und wartet mit der musikalischen Reaktion, bis die Information über den Dialog vermittelt wurde.

Kurze Zeit später [02:04:46] wird das Voranschreiten und die entstehende Eigendynamik der Mission durch die Musik übernommen. Die durchgehenden Rhythmen wirken wie Zahnräder, die ineinanderlaufen. Sie zeigen zusätzlich zu der visuellen Ebene an, dass zentrale Punkte des Plans auf allen Traumebenen im vollen Gange sind und auf den Kick zusteuern. Die kontinuierliche Steigerung der überlagerten Streicher führt sowohl zu dem emotionalsten Moment von Fischer als auch zur Einleitung des Kicks in der dritten Traumebene, um den Plan zu vollenden. Dabei widmet sich die auditive Ebene primär der Durchführung der Inception, während Fischers Emotionen größtenteils über das Bild suggeriert werden. Die Musik verbleibt in ihrem kontinuierlichen Rhythmus und kehrt nicht mehr zu den zuvor verwendeten atmosphärischen Klängen zurück. Dies wird deutlich als Fischer eine Bewegung seines Vaters

als körperliche Zuneigung missinterpretiert [02:05:05] und stattdessen das Testament sowie sein Kinderspielzeug von hohem emotionalem Wert in dem Safe findet. Als er sieht, dass sein Vater tot ist, wird die Szene visuell von Trauer und Schmerz dominiert. Auf der auditiven Ebene unterstützt nur die tonal ansteigende Streichermelodie diese Emotionalität.

Zudem steuert die Musik die Rückkehr aus den Tiefen des Unterbewusstseins. Nach der ersten Explosion werden über die geschichteten Bläserhythmen, welche aus dem verlangsamten Bläserhauptthema resultieren, die parallelen Ereignisse der Traumebenen vereint. Nach der Verbindung der Ebenen [02:06:18] formt sich die Musik zu einer antriebsvollen Kraft aus Bläsern und Orchester zusammen, die den Kick in unmittelbare Nähe bringt. Der Dynamiksprung in der Musik unterstreicht die Wichtigkeit des Moments, den die Figuren nicht verpassen dürfen, da zeitgleich auch die Traumebenen mit zunehmender Geschwindigkeit zusammenbrechen. Folglich ist dies die einzige Möglichkeit, in die Realität zurückzukehren. Kurz vor der Vollendung des Kicks wird für den Übergang der Ebenen wenige Sekunden lang der stark verlangsamte Gesang von *Non, je ne regrette rien* durch die non-diegetische Musik aufgegriffen. Theoretisch wird das Lied noch immer über die Kopfhörer an Arthur wiedergegeben, was jedoch nicht im Bild gezeigt wird. Stattdessen reduzieren sich die Instrumente und das aufkommende Bläserhauptthema fokussiert sich auf das Verlassen der zweiten Ebene und beendet damit die Realitäten der tieferliegenden Traumschichten.

### Negative Bilder und Töne

In dieser Sequenz werden keine negativen Bilder oder Töne verwendet. Die auditive Ebene zeigt das gesamte Bewegtbild und umgekehrt. Lediglich die in der Maskierung beschriebenen Stellen könnten als negative Töne aufgefasst werden. In diesen Momenten wird jedoch der Musik der Vorrang gegenüber den Geräuschen gelassen, da sie die Handlung dirigiert und die Geräusche nicht als fehlend empfunden werden.

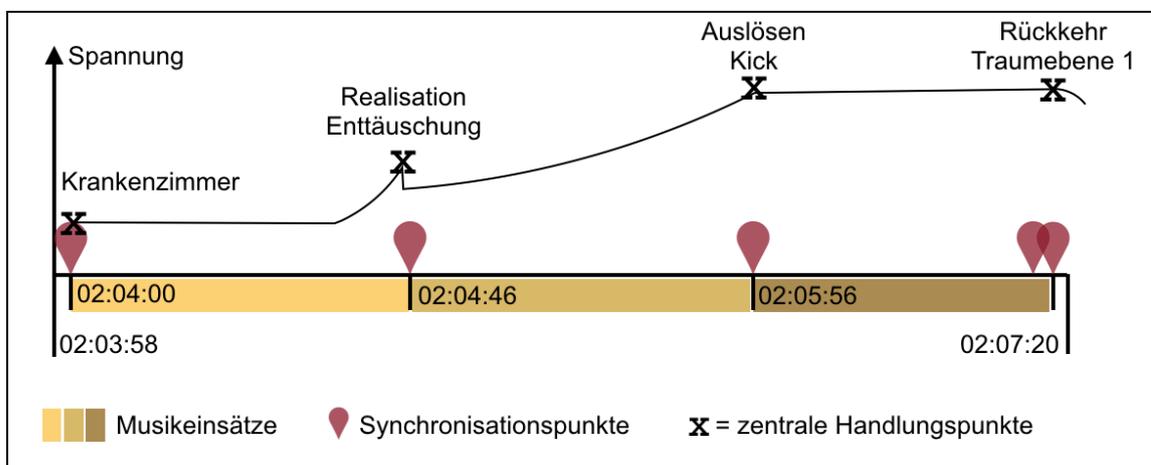


Abbildung 4.5 Zusammenhang zwischen der dramaturgischen und musikalischen Ausarbeitung im ersten Szenenabschnitt (eigene Darstellung)

#### **4.3.2 [02:07:21 – 02:09:50]: Limbus / Traumebene eins**

##### **Bild ohne Ton**

Die Weiterführung der Szene beginnt damit, wie Cobb im Limbus Mal in seinen Armen hält und sie miteinander reden. Die wechselnden Großaufnahmen zeigen einen emotionalen Moment, da Mal Tränen in den Augen hat und Cobb sie liebevoll anschaut und berührt. Es folgt die Totale eines modernen Stadtviertels, in der ein altes Paar sich an den Händen haltend von der Kamera wegschneidet. Die Detailaufnahme der Hände wird über den Schnitt weitergeführt, woraufhin sich diese nahe des Bodens fest aneinanderklammern. Danach wird wieder auf Cobb und Mal zurückgeschnitten. Sie schließt ihre Augen und lässt ihren Kopf bewusstlos zur Seite kippen. Im Gegenschnitt berührt Cobb sie behutsam an der Wange, wendet seinen Blick von ihr ab und atmet mit geschlossenen Augen tief aus.

Anschließend springt die Kameraperspektive in die erste Traumebene zurück und zeigt den sinkenden Van. Zuerst befreien sich Fischer und sein Onkel aus dem Auto und retten sich an das Ufer. Es regnet, während die Kamera um den sprechenden Fischer schwenkt und eben nicht sein Onkel, sondern Eames neben ihm auf dem Boden sitzt. Parallel dazu befreien sich auch Ariadne, Yusuf und Arthur aus dem Van. Arthur sieht den bewusstlosen Cobb und realisiert, dass er nicht mit in die erste Traumebene zurückgekehrt ist. Am Steinufer sprechen Arthur und Ariadne miteinander und es wird kurz zu Cobb im Van zurückgeschnitten. Es lässt sich vermuten, dass sie sich über ihn unterhalten.

##### **Ton ohne Bild: Auflistung und Charakterisierung**

Die Gestaltung der auditiven Ebene wird wie zuvor mit non-diegetischer Musik, Soundeffekten und Dialog fortgeführt. Gleichmaßen wird dieser Szenenabschnitt durch die Abwesenheit von Musik eingeleitet und anschließend durchgehend musikalisch unterlegt. Die Balance der einzelnen Komponenten ist dabei wesentlich ausgewogener, sie erfüllen jeweils eine eigene Funktion und treten an den entsprechenden Stellen in den Vordergrund. Während das Sounddesign die Umgebung klanglich illustriert, trägt die Musik die Stimmung der Szene. Insgesamt wirkt diese Sequenz ruhig, beginnt deutlich leiser und wird auch über ihren Verlauf in der Lautstärke kaum angehoben. Die Konsistenz ist so gestaltet, dass Musik und Geräusche einzeln verfolgt werden können, sich im Zusammenspiel dennoch gegenseitig ergänzen. Den Dialogpassagen wird ausreichend Platz gegeben, sodass sie immer gut verständlich sind. Dies zeigt die Wichtigkeit der darin geteilten Informationen, die unbedingt vom Zuschauer verstanden werden müssen.

Zunächst sind ausschließlich Windgeräusche zu hören, bevor der Dialog zwischen Cobb und Mal in den Vordergrund rückt. Sie weint, während das Paar über deren gemeinsame Vergangenheit spricht und die Musik nach Mals Aussage „du hast gesagt, [...] dass wir

zusammen alt werden“ einsetzt. Die Violinen-Flageoletts leiten das in Kapitel 4.1 beschriebene Time-Thema ein, welches wiederholt und variiert wird. Parallel dazu erklärt Cobb, dass sie diese Zeit zusammen erlebt haben und offenbart ihr seine Gefühle. Mit der dritten Wiederholung endet der Dialog und das Thema wird durch das einsetzende Orchester unterstützt und dissonant weitergeführt. In dieser Phase dominiert die Musik, neben ihr sind lediglich einzelne Atemgeräusche zu hören.

Die Umgebung ändert sich und wird durch dumpfe Unterwassergeräusche sowie Atemzüge einer Sauerstoffflasche ersetzt. Auch das musikalische Thema klingt durch ein langgehaltenes dreigestrichenes e aus und leitet zugleich in den neuen Musikeinsatz über. Dieser wird ebenfalls durch die Streicher angeführt, wobei besonders die tieferen Streichinstrumente wie Celli hervortreten und einen melancholischen bis tragischen Charakter erzeugen. Die Musik klingt weiter, während die Unterwasserumgebung mit einer zweiten an Land mehrmals wechselt. Im Rauschen des Regens verkündet Fischer seinem Onkel, dass er nun auf eigenen Beinen stehen werde. Daraufhin baut sich die Musik kurzzeitig auf und wirkt aufmunternder, bevor ein weiterer Dialog zwischen Arthur und Ariadne beginnt. Sie sprechen über Cobb und Arthur zeigt sich misstrauisch zu Cobbs Entscheidung, im Limbus zu bleiben und nach Saito zu suchen. Die Worte „er wird sich verlieren“ werden von einem nachfolgenden Impact unterlegt, mit dem sich die Lautstärke des Sounddesigns deutlich reduziert. Erneut beenden gehaltene Streicher die Szene, die erst im nächsten Szenenabschnitt musikalisch aufgelöst werden.

## **Ton und Bild zusammen**

### Maskierungen

Maskierungen liegen in dieser Sequenz nicht vor. Trotz dessen, dass sich Musik und Sounddesign regelmäßig überlagern, sind die Frequenzbereiche ausreichend voneinander getrennt. So kann die dumpfe Unterwasserumgebung eindeutig von der klaren Musik unterschieden werden, ebenso wie die leisen Seufzer oder auch der Kuss an Mal.

### Synchronisationspunkte

In dieser Passage sind die Synchronisationspunkte der visuellen und auditiven Ebene wesentlich subtiler ausgearbeitet. Die Geräusche der ersten Traumebene sind stets passend dem Bewegtbild unterlegt und lassen die Umgebung realitätsnah wirken. Sie stellen jedoch in diesem Zusammenhang keine relevanten Synchronisationspunkte dar. Vielmehr zeigen sich diese in den musikalischen Akzenten, die zum einen Handlungselemente hervorheben und zum anderen die dargestellte Emotion verstärken. Dadurch, dass der Zusammenfluss von Bild und Ton in dieser Sequenz diskret und nicht plakativ gestaltet wird, zeigt sich die Ernsthaftigkeit und Wichtigkeit der Momente in beiden Ebenen.

Dies kann in drei Fällen beobachtet werden. Der markanteste Moment der Synchronisation findet statt, als Mal das Bewusstsein verliert und stirbt [02:08:06]. Während sie in Cobbs Armen liegt und ihr Kopf gegen seinen Körper kippt, wird die Variation des Time-Themas mit einem dissonanten Akkord weitergeführt. In diesem Moment schließt Cobb innerlich mit der Beziehung zu Mal ab und signalisiert damit eine wichtige und langerwartete Entwicklung seiner Filmfigur.

Auch die Überleitung in die Geschehnisse der ersten Traumebene [02:08:24] können als eine Art der Synchronisation aufgefasst werden. Die Streicher ziehen durch das gehaltene e3 den Moment in die Länge, sie halten beinahe wie die Protagonisten die Luft an und führen anschließend mit der einsetzenden Melodie die Handlung in der ersten Traumebene fort.

Zudem bewirkt der Dialog zwischen Mal und Cobb den Einsatz des Time-Themas [02:07:33], auf welchen die Erinnerungen des gemeinsamen Lebens im Limbus folgen. Die sich aufbauende Musik lädt die Szene dabei zunehmend emotional auf.

#### Technischer und formaler Vergleich

In technischer Hinsicht wirken die visuelle und auditive Ebene homogen, es gibt weder schnell geschnittene Bildfolgen noch auffällige Kamerafahrten, die durch musikalische Verzierungen begleitet werden. Stattdessen bleibt die Kadrierung in der ersten Hälfte der Szene im Limbus relativ konstant. Durch die nahe Kameraperspektive sind die Gesichter von Cobb und Mal angeschnitten, wodurch Intimität entsteht (Kuchenbuch, 2005). Gleichzeitig wird dies auditiv durch den überbetonten Wind und durch das von den Violinen gespielte Thema bestätigt.

Im sinkenden Van [02:08:30] wirkt das Bild etwas instabil, als würde die Kamera im Wasser treiben. In der Musik wird dies jedoch nicht aufgegriffen, sie hält durch die gehaltenen Streicher lediglich kurz inne, wie bereits im vorangegangenen Abschnitt dargelegt wurde.

Die letzten Sekunden des Kameraschwenks während Fischers Monolog [02:09:28] werden von einer musikalischen Steigerung begleitet. Diese beruht allerdings nicht auf der Kamerabewegung selbst, sondern auf der sich daraus ergebenden Information, welche für die Narration von Bedeutung ist.

Die vorangegangenen Erkenntnisse geben zugleich Hinweise auf die Substanz der beiden Ebenen, die auch in dieser Szene ähnlich gestaltet ist. Im ersten Abschnitt liegt die detaillierte Darstellung der Emotionen von Cobb und Mal sowohl im Bild als auch im Ton im Fokus. Dazu dominiert die Mimik der beiden den Bildausschnitt, während die restlichen Bildanteile im Schatten liegen und musikalisch motivisch gearbeitet wird. Mit dem Schnitt in die erste Traumebene wird in eine einfache Streichermelodie übergeleitet und mit dem bläulich wirkenden Bewegtbild kombiniert. Obwohl die auditive Ebene in dieser Phase lediglich die Erzählung widerspiegelt, wirkt sie wesentlich voller als das gezeigte Bewegtbild.

Das Tempo bleibt über den Verlauf dieser Sequenz weitestgehend konstant und ist vergleichsweise langsam. Visuell wird zu Beginn wiederholt zwischen den zwei Großaufnahmen geschnitten, um die Reaktionen von Mal und Cobb während des Gesprächs zu zeigen. Trotz dieser Bildwechsel findet keine Beschleunigung statt, da die intime Gesprächssituation bestehen bleibt. Im Ton kann zunächst aufgrund der vorherrschenden Stille kein Verhältnis bezüglich der Geschwindigkeit hergestellt werden. Erst mit dem Einsatz der Musik [02:07:33] etabliert sich ein konstanter Rhythmus, der bis zum Wechsel in die erste Traumebene bestehen bleibt. Die Rückblende in den Limbus ist passend auf die Worte Cobbs geschnitten, sodass sich diese adäquat in den Rhythmus einfügt.

Der musikalische Wechsel in Traumebene eins [02:08:24] bewirkt eine dezente Entschleunigung, die sich auch in den Bildern in Form von längeren, ungeschnittenen Passagen wiederfindet. Zwar werden die Handlungsereignisse von Fischer und dem restlichen Team parallel gezeigt, jedoch bringen die Schnitte innerhalb und zwischen den Schauplätzen nur kleinschrittige Veränderungen mit sich. So retten sich beispielsweise Fischer und sein Onkelin einer Einstellung ans Ufer und bleiben anschließend an derselben Stelle sitzen. Auffällig sind die kontrolliert wirkenden Bewegungen der Figuren unter Wasser, die sich an das wahrgenommene Gesamttempo anpassen.

### Narrativer Vergleich

Auch mit Hinblick auf die Erzählung ist das audiovisuelle Zusammenwirken von Bedeutung. Der Musikstil ist in dieser Szene einheitlich gehalten, wobei sich für dessen Umsetzung an der Leitmotiv- sowie der Mood Technik bedient wurde. Ton und Bild ergänzen sich, während die Musik eine zusätzliche Ebene für die Emotionalität aufspannt. Sobald sie einsetzt, ist sie durchgehend präsent, allerdings vereinnahmt die Musik das Bild nicht und erzeugt dadurch einen großen Mehrwert in dieser Szene.

Die Kombination aus minimalistischem Sounddesign und der nahen Kameraperspektive [02:07:21] lenkt den Fokus auf die Interaktion zwischen Cobb und Mal. Es wird deutlich, dass es sich um einen eminenten Moment der Zweisamkeit handelt. Die Bedeutung der Szene steigert sich mit dem Einsatz des Time-Motivs, dessen Wiederholungen die Vergangenheit sowie die jetzige Realität zusammenführen. Cobb erkennt nun, dass er sich von seinen Schuldgefühlen lösen und ihren Tod akzeptieren muss. Er kann nicht weiterhin in seine Traumerinnerungen flüchten, sondern muss in die Realität zu seinen wartenden Kindern zurückkehren. Als Mal während der Motivvariation stirbt, wird deutlich, dass er sie nun wirklich loslässt und schweren Herzens mit der Situation Frieden schließt.

Des Weiteren verhilft die musikalische Ausarbeitung dabei, die erste Traumebene aus der Perspektive von Fischer als Realität darzustellen. Seine Worte verdeutlichen, dass durch die vorherige Konfrontation die Vater-Sohn-Beziehung zu einem gewissen Umfang aufgearbeitet

wurde. Die stimmungsgeladene Musik impliziert zugleich sein Bedauern, das Zerwürfnis nicht zu Lebzeiten seines Vaters geklärt zu haben. Mit der Verkündung, das Unternehmen nicht fortzuführen, wird die Täuschung durch Eames offenbart und der Erfolg der eingepflanzten Idee und der damit abgeschlossenen Inception musikalisch bestätigt [02:09:31].

Die Musik fungiert demnach als verbindendes, narratives Element, das den Limbus mit der ersten Traumschicht und der Realität verknüpft. Sie ist ein immerwährender Begleiter der Handlung und passt sich dieser unauffällig an. Darüber hinaus endet die Musik in dieser Sequenz nicht, sondern leitet über den Bildschnitt hinweg die folgende Szene ein. Auch das Sounddesign wird in ähnlicher Weise eingesetzt. Das Bild von Cobb im sinkenden Van wird von einem dumpfen Impact begleitet [02:09:44], welcher erneut auf den Schnitt in die folgende Sequenz im Limbus ertönt. Diese bereiten den Szenenübergang vor und implizieren die Distanz seines Bewusstseins zu der ersten Traumebene.

### Negative Bilder und Töne

An einer Stelle der untersuchten Szene wurden negative Töne verwendet. Die Rückblende zu der gealterten Mal und Cobb im Limbus [02:07:37] wird ausschließlich durch die Musik und Cobbs Stimme begleitet. Die fehlende Stadtatmosphäre und Schrittgeräusche fallen zunächst nicht auf, was sich jedoch bei der Detailaufnahme auf die sich fest umklammernden Hände ändert. In dieser Einstellung wackelt das Kamerabild und impliziert die Erschütterung des näherkommenden Zuges, der das Paar aus dem Limbus zurück in die Realität bringen soll. Die Tonebene weist jedoch in keiner Weise auf den bevorstehenden Suizid hin. Ohne das Vorwissen der über den Handlungsverlauf gezeigten Rückblenden sowie den musikalischen Hinweis wäre die Szene in dieser Form für den Zuschauer nicht interpretierbar (Kloppenburg, 2017b). Dadurch dass die Situation im Limbus weiterhin mit derselben Musik und Cobbs Stimme fortgeführt wird, bleibt die Perspektive bestehen. Die Rückblende zeigt seine positiven Erinnerungen an die gemeinsame Zeit mit Mal im Limbus, die sie zusammen bis zum Ende ausgekostet haben.

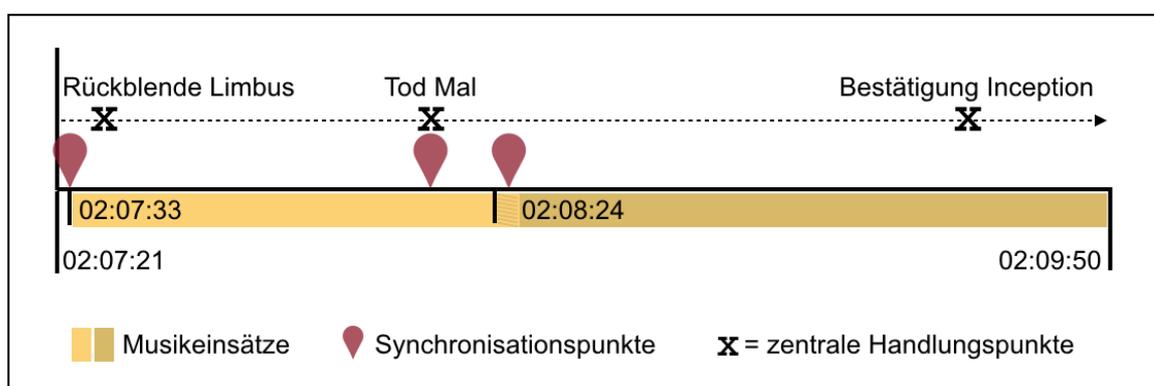


Abbildung 4.6 Zusammenhang zwischen der dramaturgischen und musikalischen Ausarbeitung im zweiten Szenenabschnitt (eigene Darstellung)

#### **4.3.3 [02:09:51 – 02:15:10]: Limbus / Realität**

##### **Bild ohne Ton**

Der dritte Szenenabschnitt beginnt mit einer verkürzten Variante des Filmanfangs. Nach der Gischt einer großen Welle wird auf den in flachem Wasser liegenden Cobb geschnitten, der langsam zu Bewusstsein kommt. Er wird von uniformierten Soldaten gefunden, die ihn in den Saal eines japanischen Tempels führen. Die Totale zeigt einen Raum sowie einen alten Mann, der an einem langen Tisch sitzt. Cobb sitzt ihm gegenüber und starrt verwirrt und desorientiert in seine Richtung. Es wird deutlich, dass es sich bei dem Mann um den gealterten Saito handelt. Beide wirken lädiert und beginnen miteinander zu reden, wobei sich die Kamera kontinuierlich Cobbs Gesicht annähert. Es wird eine Detailaufnahme des drehenden Kreisels eingeschoben und kurze Zeit später erneut aufgegriffen. In dieser zweiten Einstellung liegt neben dem Kreisel eine Pistole mit der Schaftöffnung in Richtung Cobb. Im unscharfen Hintergrund bewegt sich Saitos Hand zu der Pistole hin.

Es folgt ein harter Schnitt, Cobbs Augen öffnen sich in einer neuen Umgebung. Der Bildausschnitt vergrößert sich, sodass die Flugzeugkabine sichtbar wird. Cobb wirkt irritiert, als er seine Teamkollegen anschaut, ebenso wie Saito, der zu einem Telefon greift und eine Nummer wählt. Fischer ist ebenfalls wach und schaut nachdenklich aus dem Flugzeugfenster. Bei der Einreisekontrolle am Flughafen werden Cobbs Papiere von einem Polizeibeamten geprüft. Er darf passieren, woraufhin sich seine Mimik und Körperhaltung aufhellen. Die Kamera begleitet ihn durch die Gepäckhalle, während er seinen Koffer holt und an seinen Teammitgliedern vorbei zum Ausgang geht. Sein Schwiegervater empfängt ihn und bringt ihn nach Hause, wo er mit misstrauischem Blick seinen Kreisel andreht. Er wird durch seine Kinder abgelenkt, die wie in seiner Erinnerung im Garten spielen, sich nun aber zu ihm drehen und auf ihn zulaufen. Die Kamera schwenkt von der Totalen des Wohnzimmers mit Cobb und seinen Kindern zu dem Kreisel zurück, der sich noch immer dreht, zu zucken beginnt und abrupt in einem Schwarzbild endet.

##### **Ton ohne Bild: Auflistung und Charakterisierung**

Für die klangliche Ausarbeitung der letzten Minuten von Inception wurden erneut non-diegetische Musik sowie Dialog und Sounddesign verwendet. Diese Szene wird durchgängig von Musik begleitet, welche entschieden das präsenteste auditive Element darstellt. Während die Balance zwischen Dialog und Musik zu Beginn noch ausgeglichen ist, nimmt die Musik über den Szenenverlauf stetig mehr Platz ein und führt die Handlung im letzten Drittel alleinig an. Dabei steuert sie den Dynamikverlauf der Szene maßgeblich mit. Das Sounddesign hingegen ist den anderen Elementen untergeordnet. Es gibt lediglich zwei Stellen, an denen die Umgebung durch eine Geräuschkulisse illustriert wird sowie ein wiederkehrendes,

metallenes Geräusch. Die Konsistenz in dieser Szene ist zweigeteilt, so sind zu Beginn und am Ende Dialog und Musik klar voneinander getrennt. Das metallische Kreiselgeräusch des Totems hingegen ist in die Textur der Musik eingearbeitet und tritt erst mit zunehmender Lautstärke aus dieser hervor. Auch weitere unauffälligere Elemente des Sounddesigns fügen sich in die übergeordnete musikalische Struktur ein.

Die Sequenz wird durch den zuvor erwähnten dumpfen Impact eingeleitet, auf den diverses Wasserrauschen sowie tiefe, ausklingende Bässe folgen. Einige schwer zu deutende Geräusche sind vernehmbar, bevor über die atmosphärische Musik eine bedrohliche Stimmung suggeriert wird. Diese wird durch tiefe Bässe zusammen mit Streichern umgesetzt, welche in einem wiederholten Muster in ihrer Lautstärke anschwellen und kurz darauf wieder abfallen. Parallel dazu beginnt ein Dialog zwischen Cobb und Saito, bei dem sich die beiden in ihren Aussagen gegenseitig ergänzen. In dieser Phase erscheinen Musik und Dialog ebenbürtig, da sie unterschiedliche Funktionen übernehmen, wodurch neben der Stimmung auch Informationen vermittelt werden. Ein stehender synthetischer Klang setzt ein und geht in ein metallenes Kreiselgeräusch über, welches hintergründig bestehen bleibt.

Die Musik erfährt eine erste fließende Änderung mit der ausgesprochenen Erinnerung von Cobb, dass es sich in diesem Moment nicht um die Realität handelt. Die bedrohlichen Klänge werden von den Violinen abgelöst und das Time-Motiv beginnt. Wie auch in der Szene zuvor, wird das Motiv mehrfach wiederholt und variiert. Nach den Worten „kommen Sie zurück“, die Cobb mehrmals an Saito richtet, setzt das Orchester in Form von weiteren Streichern, Klavier und Schlagwerk ein. Ein kurzer Dialogeinschub folgt, bevor das Motiv weiter ausgearbeitet wird und die Musik an Dominanz gewinnt. Das Sounddesign ist weiterhin unauffällig, lediglich Kleidungsgeräusche sowie die Tasten eines Telefons sind hintergründig hörbar.

Die Musik wird leiser, statt des Time-Motivs erklingt eine statisch wirkende Melodie der hohen Streicher und auch die Geräuschebene ändert sich hin zu einer geschäftigen Umgebung. Mit der Aussage „Willkommen zuhause Mr. Cobb“ baut sich die musikalische Figur erneut auf. Eine E-Gitarre tritt nun in den Vordergrund, die durch das Streicherensemble und die sukzessiv einsetzenden Bläser ergänzt wird. Mit jeder Motivwiederholung steigert sich die Musik, sie wird lauter und nimmt einen monumentalen Charakter an. Sie vereinnahmt zu diesem Zeitpunkt die gesamte Tonspur und erreicht ihren Höhepunkt. Dies ändert sich erst, als Cobbs Schwiegervater zu seinen Enkelkindern spricht. Musikalisch findet eine unvermittelte Rückkehr auf das ursprüngliche Time-Motiv statt, welche durch die instrumentale Reduktion auf Klavier und Streicher unterstrichen wird. Es entsteht eine gewisse Intimität, in welcher die Stimmen der Kinder in den Vordergrund treten. Sie begrüßen euphorisch ihren Vater, während erneut das metallische Kreiseln einsetzt und der Dialog ausgeblendet wird. Das Time-Motiv endet mit dem Klavier, die Streicher führen eine Melodie weiter, welche auf einem dreigestrichenen Ges in mitten eines Crescendo endet.

## **Ton und Bild zusammen**

### Maskierungen

Auch in den Schlussminuten von Inception sind Maskierungen innerhalb der Tonebene zu erkennen. Diese unterscheiden sich subtiler von den negativen Tönen als zuvor, da abgesehen von dem Beginn der Sequenz das Sounddesign weniger explizit raushörbar ist. Es findet eine Maskierung innerhalb des Sounddesigns statt, als der im Wasser liegende Cobb nach Luft schnappt [02:09:58]. Statt seiner Atmung sind ausschließlich die nahen, kleinen Wellen zu hören. Wenige Sekunden danach verdeckt die Musik im japanischen Tempel die Raumumgebung [02:10:07]. Eine wesentlich auffälliger Maskierung kann bei dem Auftreten des Totems beobachtet werden [02:10:56], als sich das korrespondierende Geräusch aus der Musik löst, bevor der Kreisel im Bild sichtbar wird. Es bleibt über den weiteren Dialogverlauf bestehen und wird stellenweise durch die darüberliegende Musik erneut maskiert. Ein weiteres Beispiel findet sich in der Flughafenhalle zur Einreisekontrolle [02:12:53] wieder. Die Umgebung wird kaum über eine Geräuschkulisse illustriert, erst durch die Lautstärkereduktion in der Musik gewinnt sie an Präsenz.

### Synchronisationspunkte

Momente der Synchronisation von Bild und Musik werden auch in diesem Szenenabschnitt mehrmals verwendet. Sie dienen der Verdeutlichung zentraler Handlungsereignisse und geben Hinweise auf den Fortgang der Erzählung. Dabei werden die im Bild gezeigten Entscheidungsmomente besonders durch den musikalischen Verlauf hervorgehoben. Erfolgt die Synchronisation auf den Bildschnitt wird eine neue Umgebung etabliert bzw. der Wechsel zwischen den Bewusstseisebenen signalisiert. Wie zuvor sind Dialog und Geräusche stets synchron dem Bild unterlegt und spielen daher eine untergeordnete Rolle. Insgesamt wurden bei der Analyse fünf zentrale Synchronisationspunkte festgestellt.

Mit dem Schnitt in diese Sequenz [02:09:51] kommen Bild und Ton erstmalig zusammen. Die brechende Welle wird auditiv durch das Sounddesign, aber auch durch die dröhnenden Bässe unterstützt und leitet die visuelle Wiederholung des Filmbeginns ein. Die alternative Vertonung weist jedoch darauf hin, dass der Wiederaufgriff dieser Szene mehr als die bereits gezeigten Bilder mit sich bringen wird.

Anschließend erfolgt eine weitere Synchronisation, als Cobb den drehenden Kreisel sieht und realisiert, dass Saito und er sich noch immer im Limbus befinden [02:11:04]. Diese Erkenntnis hat sich bereits im vorherigen Austausch angedeutet, wird nun aber durch die einsetzende motivische Figur in der Musik bestätigt und bildet damit den zentralen Punkt des Gesprächs. Bei der Rückkehr in die Realität [02:11:40] werden Bild und Ton erneut markant synchronisiert. In dem Moment, als Cobb seine Augen öffnet, setzen weitere Instrumente des Orchesters ein, welche das Time-Motiv ebenfalls unterstützen. Dem Zuschauer wird dadurch der Eindruck

vermittelt, dass sowohl Cobb als auch Saito durch den Gebrauch der Pistole den Limbus verlassen konnten.

Während der Einreisekontrolle im Flughafen folgen zwei Synchronisationspunkte zeitnah aufeinander. Zunächst halten Bild und Musik inne, während Cobbs Dokumente kontrolliert werden [02:12:56 - 02:13:13]. Dadurch entsteht auf beiden Ebenen Spannung, die erst mit dem zweiten Synchronisationspunkt aufgelöst wird. Dieser ergibt sich, als der Stempel zur Einreise in Cobbs Reisepass eingetragen wird [02:13:12]. Auditiv wird dies von einem passenden Geräusch sowie einer einsetzenden E-Gitarrenmelodie unterlegt. In diesem Moment wird deutlich, dass Cobb keine Strafverfolgung mehr in den USA droht und Saito sein Wort gehalten hat.

### Technischer und formaler Vergleich

Bei der weiteren Analyse der Bild-Ton-Gestaltung wird ersichtlich, dass die technischen Bildveränderungen bis auf wenige Ausnahmen musikalisch nicht mitberücksichtigt werden. So wird beispielsweise die visuelle Annäherung an Cobb mit der gegenläufigen Entfernung von Saito zu Beginn des Gesprächs im Limbus von der Musik ignoriert. Mit dem Einsatz des Time-Motivs [02:11:04] wird der Bildausschnitt verkleinert und zugleich die Lautstärke angehoben. Im Flugzeug [02:12:17] ergibt sich erneut ein technischer audiovisueller Zusammenhang. In dieser Sequenz passt sich die Geschwindigkeit der Kamerabewegung dem Tempo der Musik an, wodurch eine stimmige Einheit zwischen Bild und Ton entsteht.

Der finale Kameranachschwenk [02:14:45] wird ebenfalls subtil von der gesamten Tonspur begleitet. Durch diesen wechselt der Fokus von Cobb und seinen Kindern auf sein Totem, indem die Stimmen ausgeblendet werden und das Kreiselgeräusch zusammen mit der sich lösenden Streichermelodie präsenter wird.

Die Substanz von Bild und Ton ergänzen sich im ersten Part dieser Szene. Im Mittelpunkt steht der Austausch zwischen Saito und Cobb, deren Mimik durch den Bildausschnitt detailliert dargestellt und durch repetitive musikalische Muster versehen werden. Nach dem Musikwechsel in der Hälfte des Gesprächs ändert sich der raue Charakter und wird sanfter. Letztlich wird die Situation durch klare Musik und einem fast gänzlich unscharfen Bild abgeschlossen. Ein Gegensatz in der Textur ergibt sich im letzten Teil der Szene. Die üppige und detailreiche Musik dominiert klar über die relativ einfachen Bilder, welche die Flughafenhalle und Cobbs Zuhause zeigen. Sie erfüllt in diesem Moment eine wichtige narrative Funktion, die durch die unvermittelte Sanftheit bei dem Schnitt auf die Kinder deutlich wird.

Das Tempo variiert mehrere Male in dieser Sequenz, wobei Bild und Ton durchweg einen ähnlichen Rhythmus annehmen. Der Abschnitt im Limbus beginnt langsam, was durch die langen tiefen Bässe und die in Zeitlupe spritzende Wellengischt realisiert wird. Der visuelle

Wechsel in das Gebäude erfolgt zügig, wobei der Dialogbeginn den Bildern vorgezogen wird. Mit dem Schnitt in die Realität findet eine leichte Beschleunigung statt, als die Streichergruppen sich gegenseitig ablösen und die Musik mehr an Fahrt aufnimmt. Die Bildschnitte dienen den zügigen Umgebungswechseln in Saitos Tempel, zum Flughafen sowie in Cobbs Zuhause und führen damit zu einer Raffung der erzählten Zeit.

Bei der Einreisekontrolle [02:12:57] kommt es sowohl im Bild als auch im Ton beinahe zu einem Stillstand. In der Musik sind kaum Bewegungen vorhanden und auch das Bewegtbild fokussiert sich auf die Situation um Cobb. Dadurch wird seine Anspannung zusätzlich zu seiner Körperhaltung und dem Zucken seiner Kiefermuskulatur offengelegt. Nach der Auflösung dieser Situation entwickelt die Musik eine Eigendynamik und wird von der visuellen Ebene übernommen. Diese wird erst unterbrochen, als Cobb seine Kinder wiedersieht. Anschließend wird der Film in dem entschleunigten Tempo kurz darauf zum Ende geführt.

### Narrativer Vergleich

Die gestalterische Analyse zeigt, dass die Musik auch in dieser Sequenz bildunterstützend fungiert. Dabei wechselt der Stil von atmosphärischen Klängen hin zu einer musikalisch ausgearbeiteten Komposition, welche auf dem Time-Motiv basiert. Mittels der Abdeckmethode wurde die beträchtliche Wirkung der Musik für die Narration ersichtlich. Sie dominiert das Geschehen und führt den Film ab der Rückkehr in die Flugzeugkabine dramaturgisch zu Ende. Zudem fällt auf, dass Inception durch das Time-Thema sowohl eröffnet als auch beendet wird. Die motivischen Aufgriffe über den Handlungsverlauf führen zu dem finalen Scoretrack, bei dem die musikalische Figur entfaltet wird.

Durch den Einsatz dieses Themas im Limbus [02:11:04] wird erneut der Bezug zur Realität hergestellt. Wie auch bei dem zu vorigen Gespräch zwischen Cobb und Mal gibt die dissonante Motivvariation einen Hinweis darauf, dass sich die Figuren nicht in der realen Welt befinden und dies auch erkennen. Nach der Rückkehr aus dem Limbus wird von weiteren Dissonanzen für die musikalische Gestaltung abgesehen. Die Musik steigert sich nicht kontinuierlich, sondern reagiert auf das Bewegtbild und baut sich entsprechend auf bzw. fällt ab, um Handlungseignisse nicht vorwegzunehmen. Dadurch wird bei dem Gespräch im Limbus durch den fließenden musikalischen Übergang subtil die Stimmung verändert und bei der Einreisekontrolle über die Musik ein Spannungsbogen erzeugt.

Der musikalische Höhepunkt wird durch den Einsatz der druckvollen Blechbläser erreicht, welche die positive, fast schon ungläubige Mimik von Cobb unterstreichen [02:14:02]. Die Musik verleiht den Bildern eine höhere Bedeutung, denn sie spiegelt den Missionserfolg des Teams und dessen Auswirkungen für Cobbs Privatleben wider. Als er sein Haus betritt und nach seinem Totem greift, werden seine Zweifel musikalisch ignoriert [02:14:17]. Ob es sich nun um die reale Welt oder einen Traum handelt, scheint zweitrangig. Cobb ist wieder mit

seinen Kindern vereint und damit in seiner eigenen Realität angekommen. Dies wird durch das ursprüngliche Motiv der Streicher und des Klaviers pointiert [02:14:30].

In den letzten Filmsekunden verstärken sich Bild und Ton nochmals gegenseitig, als die Streicher auf einer einzelnen Note stehen bleiben und der Kreisel zu zucken beginnt [02:15:05]. Durch das offene musikalische Ende und das abgeschnittene Drehgeräusch werden auch im Ton keine eindeutigen Hinweise auf den Fall des Kreisels gegeben, sodass es dem Zuschauer selbst überlassen wird, sich mit dem Handlungsausgang auseinanderzusetzen und diesen zu interpretieren.

### Negative Bilder und Töne

Während es auch in diesem Szenenabschnitt keine negativen Bilder gibt, wird erneut mit negativen Tönen gearbeitet. Dies zeigt sich zum einen durch das Außenvorlassen nebensächlicher Sounddesignelemente. So werden beispielsweise die Schritte im japanischen Tempel [02:10:04] oder die des Schwiegervaters [02:14:03] nicht vertont. Diese Geräusche sind nicht für das Verständnis der Erzählung notwendig, weswegen der verfügbare Platz an die wirkungsvollere Musik übergeben wird. Mehrere markante negative Töne ergeben sich kurz bevor Cobb den Flughafen verlässt [02:13:20]. Visuell wird eine Halle mit Gepäckbändern, Koffern und vielen Menschen gezeigt, die jedoch in keiner Weise auf der Tonspur verortet werden können. Auch naheliegende Bildelemente werden nicht vertont, die Musik ist das einzig vertretene Element. Lediglich die Stimme des Stiefvaters kann diese Barriere durchbrechen und bringt nachfolgend Geräusche einer häuslichen Umgebung zurück. Durch diese Positionierung wird die Vormachtstellung der Musik in diesem Abschnitt deutlich. Das fehlende Schussgeräusch [02:11:39], nachdem sich Saitos Hand der Pistole nähert, trägt einen wesentlichen Beitrag dazu bei, dass die anschließende Handlung nicht eindeutig als Realität eingeordnet werden kann. Über das Bild wird ein Vorgang angedeutet, aber nicht zu Ende erzählt. Der Ton könnte die Situation eindeutig auflösen, wovon jedoch entschieden abgesehen wird, um das Filmende nicht vorwegzunehmen.

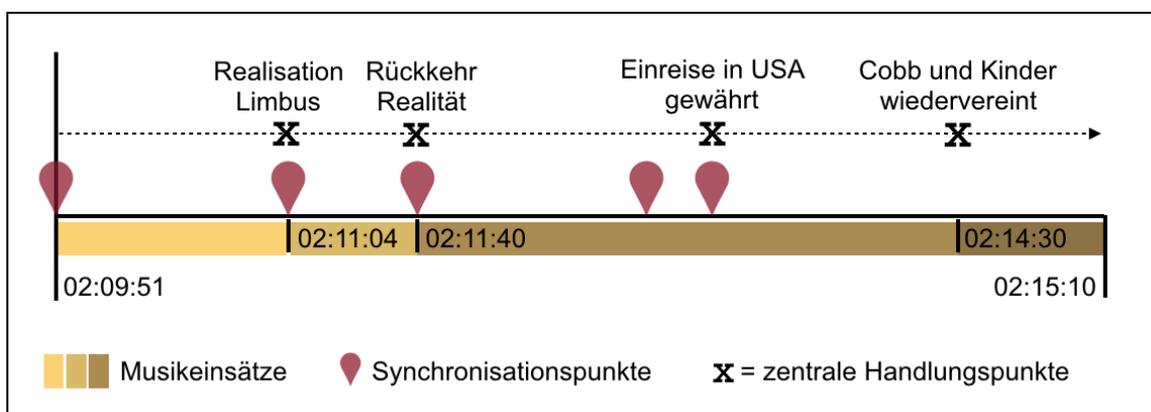


Abbildung 4.7 Zusammenhang zwischen der dramaturgischen und musikalischen Ausarbeitung im dritten Szenenabschnitt (eigene Darstellung)

## **5. Erstellung der KI-generierten Komposition**

Im folgenden Kapitel wird eine alternative Filmmusik für die letzten zwölf Minuten von Inception entwickelt. Dazu werden die Ergebnisse der Szenenanalyse verwendet, um eine vergleichbar wirkungsvolle Musik zu generieren. Der Herstellungsprozess wird in Vorproduktion, Produktion und Postproduktion gegliedert, wobei die Sequenz für die Produktion erneut in dieselben Abschnitte unterteilt und bearbeitet wird.

### **5.1 Vorproduktion**

#### **5.1.1 Vorbereitung**

Bevor mit der Generierung des neuen Scores begonnen werden kann, wird zunächst der Film entsprechend der analysierten Szene zugeschnitten und in eine separate Videodatei exportiert. Die vorliegende Stereotonspur wird isoliert bearbeitet, da die Stems – bestehend aus Dialog, Musik und Effekten – einzeln aus der Gesamtmischung extrahiert werden müssen. Dieser Schritt ist notwendig, um in der Postproduktion den Dialog und das Sounddesign mit der neu erstellten Filmmusik zusammenzuführen. Dadurch wird eine Vergleichbarkeit zwischen den Soundtracks der originalen sowie der veränderten Szene hergestellt.

Für die Extraktion der einzelnen Komponenten wird der iZotope RX 11 Audio Editor verwendet. Die Wahl für dieses Programm begründet sich in der intuitiven Handhabung sowie in den umfangreichen Bearbeitungsmöglichkeiten sowohl innerhalb der Wellenform als auch im Spektrogramm. Die Stems werden abschnittsweise entsprechend des musikalischen Verlaufs voneinander getrennt. Dabei wird für jeden Bereich die Bearbeitung einzeln angepasst. Liegt neben dem Dialog ausschließlich Musik vor, funktioniert dessen Isolation gut. Gegensätzlich verhält es sich in den actionreichen Passagen, die über viel Musik und Sounddesign verfügen. Die Elemente sind ineinander verflochten, sodass deren Trennung an Präzision verliert und die klanglichen Artefakte zunehmen. Aus diesem Grund wird das Sounddesign während des Kicks sowie in der ersten Traumbene möglichst originalgetreu nachgebaut, um weiterhin eine klangliche Similarität beider Szenen zu gewährleisten. Zudem wird die maskierende Funktion der Musik bei der alternativen Vertonung an Verwendung finden, um die bei der Extraktion hervorgerufenen klanglichen Defizite zu verdecken.

Insgesamt wird deutlich, dass bei der Trennung der Einzelspuren ein Kompromiss zwischen der Klangqualität und der Isolationsschärfe eingegangen werden muss. Besonders bei den auditiv dichten Passagen kommt der Audio Editor an seine Grenzen, sodass manuell nachgearbeitet werden muss.

Ein weiterer Teil der Vorproduktion liegt in der Wahl der Software, welche für die Generierung der Filmmusik genutzt wird. Dabei fiel die Entscheidung auf die kostenpflichtige Pro Version des Programms AIVA (Artificial Intelligence Virtual Artist), welches bereits in Kapitel 2.2.3 erwähnt wurde. Ein Grund dafür ist die Spezialisierung von AIVA auf klassische Musik, da auch bei dem neuen Score vorrangig mit Orchesterklängen gearbeitet werden soll. Es wird prinzipiell ein ähnlicher Ansatz zum Originalscore von Zimmer verfolgt, welcher über weite Teile des Films von den akustischen Streichern und Blechbläsern dominiert wird. Um analog zu den synthetischen Klängen eine gewisse Besonderheit innerhalb der filmmusikalischen Gestaltung zu erreichen, können in AIVA zusätzliche Instrumente zu der bestehenden Komposition hinzugefügt werden. Zudem können Midi- und Audiodateien, welche als Inspirationen dienen, durch die Algorithmen verarbeitet und für die Musikerstellung mitberücksichtigt werden (AIVA, kein Datum).

Zum anderen eignet sich die Software, da vor der Generierung eingegriffen und auch im Anschluss die Komposition nachbearbeitet werden kann. Es können unter anderem die verwendeten Instrumente und deren Midi-Spuren verändert, klangliche Anpassungen vorgenommen oder auch das Tempo justiert werden (AIVA, kein Datum). Der Nutzer hat so die Möglichkeit mit der Software zusammenzuarbeiten. Die klangliche Qualität der verfügbaren Referenzen überzeugte, sodass im nachfolgenden Kapitel die alternative Filmmusik mit AIVA erstellt wird.

### **5.1.2 Konzept**

Das Ziel des neu generierten Scores ist, eine wirkungsvolle und ausdrucksstarke Filmmusik zu kreieren, welche das Bewegtbild unterstützt und zum dramaturgischen Verlauf beiträgt. Dabei soll kein Soundalike erstellt werden, welches dem Original gleicht, sondern eine neue und unabhängige Komposition. Grundsätzlich werden für die drei Szenenabschnitte non-diegetische Musikstücke generiert, die sowohl einzeln als auch im Zusammenspiel miteinander funktionieren sollen. Nach der Trennung der Stems wurde nochmals deutlich, wie viel Platz die Musik in der Tonebene von Inception eingenommen hat. Diesen gilt es nun adäquat zu füllen, wobei über die knapp zwölf Minuten ein musikalischer Verlauf erkennbar werden soll. Dazu kann laut Chion (2012) mit der Klangperspektive sowie variierenden musikalischen Ausdrücken gearbeitet werden, denn die Musik vermag die gezeigten Emotionen im Bewegtbild zu untermalen und folglich empathisch zu wirken. Agiert sie gegensätzlich, etwa indifferent zu den Bildern, können dadurch die Emotionen noch intensiviert werden (Chion, 2012). Bezüglich der Instrumentierung wird sich primär an den von AIVA zur Verfügung gestellten Orchesterinstrumenten bedient.

Da über die KI-Software nur bedingt in die musikalische Gestaltung eingegriffen werden kann, wird überwiegend mit der Mood Technik gearbeitet. Eventuell sind auch paraphrasierende

Passagen möglich, die Leitmotivtechnik hingegen kann nicht eingesetzt werden. Gründe dafür sind zum einen die technische Umsetzbarkeit, da keine musikalische Figur bei der Generierung als ein solches Motiv deklariert und wiederverwendet werden kann. Zum anderen entfalten Leitmotive erst durch ihr wiederkehrendes Auftreten in verschiedenen Handlungsebenen ihre volle Wirkung (Rabenalt, 2020). Für diese Szenen ist die verfügbare Zeitspanne der gezeigten Handlungseignisse nicht ausreichend, um ein solches Motiv einzuführen.

Die aus der Szenenanalyse hervorgebrachten Erkenntnisse werden ebenfalls für die Generierung mitberücksichtigt, um eine ebenbürtige musikalische Wirkung zu erzielen. Es wird sich an dem musikalischen Verlauf inklusive der Musikeinsätze sowie den Momenten der Abwesenheit von Musik orientiert. Auch die zentralen Synchronisationspunkte, beispielsweise während der Traumebenen-Wechsel oder des Kicks, sollen übernommen werden. Durch den Generierungsvorgang werden auch abweichende und neue Synchronisationspunkte erwartet. In diesem Zusammenhang ist es von Vorteil, dass sich Bild und Ton in vielfältiger Weise verbinden können. So können diese als „unerwarteter doppelter Bruch im audio-visuellen Fluss, [...] in Form von Interpunktion am Ende einer Sequenz“ (Chion, 2012, S.55) oder mit dem Wechsel der Bildeinstellung sowie in semantischer Art zusammenfinden (Chion, 2012). Bezüglich des finalen Scoretracks wird die neue Filmmusik vor eine große Aufgabe gestellt. Der Regisseur Nolan äußerte sich in einem Interview zu dem musikalischen Stellenwert dieser Szene mit „we just let the music take over everything [...] because you realize that the momentum of the film is entirely defined by the structure of the music“ (Hans Zimmer - Making Of Inception Soundtrack, 2011). Es wird sich zeigen, ob und inwiefern die künstlich generierte Musik mit dem Scoretrack *Time* mithalten kann.

In technischer Hinsicht liegt die Filmdatei in stereo vor, weswegen die neue Filmmusik ebenfalls in einer Stereomischung angefertigt wird. Dies ermöglicht im Anschluss einen einfacheren Vergleich beider Varianten über eine Stereoabhöre sowie über Kopfhörer. Im professionellen Kontext müssten die Audiokomponenten für ein mehrkanaliges System aufbereitet werden, da Filmproduktionen heutzutage in den Formaten 5.1 bis hin zu Dolby Atmos erstellt werden (Filmtone Berufsvereinigung e.V., 2023).

Zusammengefasst liegen die Herausforderungen darin, einen Score aus mehreren Musikstücken zu entwickeln, welcher die Bilder verstärkt und die Erzählung unterstützt. Das Zusammenspiel der verschiedenen Tracks muss in sich stimmig sein und auch die Übergänge müssen fließend gestaltet werden. Über die Musik sollen Emotionen aufgebaut und transportiert werden, sodass die gezeigte Handlung authentisch wirkt (Rabenalt, 2020).

## 5.2 Produktion

Zunächst wird ein Überblick über die Software gegeben, da AIVA verschiedene Ansätze zur Erstellung einer Komposition anbietet. Die Generierung kann über die Auswahl eines Musikstils, einer Akkordfolge sowie Schritt für Schritt erfolgen. Zudem kann anhand einer importierten Midi-Datei oder eines bestehenden Referenztracks eine neue Musik erzeugt werden. Der Generierungsmodus bestimmt dabei den Umfang der Einflussnahme auf die musikalischen Parameter. Vor jeder Generierung werden die Dauer und die Anzahl der zu erstellenden Kompositionen abgefragt bzw. bei fehlender Angabe von AIVA festgelegt.

Bei der zuerst genannten Methode können lediglich der Musikstil und optional eine Tonart vorgegeben werden. In Bezug auf die Erstellung von klassischen Filmmusiken stehen zwölf verschiedene Cinematic-Stile zur Verfügung. Neben der Wahl des Musikstils ermöglicht die zweite Variante die Festlegung einer Akkordfolge. Diese kann manuell beschrieben und generiert, selbstständig erstellt oder von einer bereits bestehenden Akkordfolge wiederverwendet werden. Es werden acht Takte definiert, welche anschließend die Songstruktur bilden. Zusätzlich können Tonart, Taktart und Tempo beeinflusst werden.

Die Schritt-für-Schritt-Methode stellt die zuvor genannten Angabeoptionen ebenfalls zur Verfügung. Weiterhin ermöglicht diese Generierungsform eine Vorschau der Komposition. Bei dieser repräsentieren mehrere Midi-Spuren die verschiedenen musikalischen Schichten, wie beispielsweise Melodie, Akkorde und Bass. Die Noten einer einzelnen Midi-Spur können in dieser Ansicht neu generiert, die Instrumente verändert und die Lautstärke eingestellt werden. Wird ein Referenztrack als Eingabe geladen, können daraufhin optional die Instrumentierung, die Tonart und die gewünschte Emotion beeinflusst werden. Bei der Verwendung einer Midi-Datei ist der Ablauf vergleichbar.

Die verschiedenen Methodiken werden getestet, um eine passende Vorgehensweise für die Erstellung des neuen Scores zu finden. Wird vor der Generierung lediglich der Musikstil gewählt, wirkt das Ergebnis zufällig. Die audiovisuelle Verbindung funktioniert nur bedingt, zudem scheint es über weite Teile keinen Zusammenhang zwischen Ton und Bild zu geben. Durch Festlegen der Akkorde kann die Musik präziser an die vorherrschende Stimmung angepasst werden. Darüber hinaus eignet sich diese Funktion, um in kurzer Zeit bei vorliegender Akkordfolge eine klangliche Nachbildung der originalen Scoretracks zu entwerfen. Für das Filmende wird mit der Eingabe einer Referenz experimentiert. Dazu wird der Track *Time* verwendet und zwei neue Kompositionen werden generiert. Beide Ergebnisse können weder in ihrem musikalischen Aufbau noch klanglich dem Original zugeordnet werden. Auch der zweite Versuch, bei dem Tonart, Stimmung und Instrumentierung festgelegt werden, kann in keiner Weise überzeugen (siehe digitale Anlagen 6). Durch den mehrstufigen Erstellungsprozess ermöglicht die Schritt-für-Schritt-Methode die größte Einflussnahme auf die Komposition und die erfolgreichste Integration der musikalischen Vorstellungen. Daher

wird diese Generierungsform gewählt, um in den folgenden Kapiteln die neue Filmmusik zu erstellen.

In AIVA können bis zu fünf Variationen bei der Generierung einer Komposition erzeugt werden. Die verfügbaren Zeitspannen sind in 30-Sekundenschritten geordnet und reichen bis zu 5:30 Minuten. Werden keine Angaben zu Dauer und Trackanzahl getroffen, werden diese zufällig festgelegt. Die zeitliche Limitierung führt dazu, dass bei der neuen Filmmusik für jeden Musikeinsatz separate Kompositionen erstellt werden müssen. Es besteht keine Möglichkeit, musikalische Änderungen über Timecode-Angaben zu integrieren. Diese werden durch manuelle Eingriffe vorgenommen. Der Generierungsvorgang selbst dauert zwischen wenigen Sekunden bis Minuten, wobei ein vollständiges Musikstück erzeugt wird. Die Komposition verfügt über ein Intro, ein Outro sowie entsprechende Zwischenparts, welche durch Buchstaben und Ziffern gekennzeichnet sind.

Alle nachfolgenden Zeitreferenzen werden erneut in [Stunde:Minute:Sekunde] angegeben. Jedoch beziehen sich diese nun auf die zugeschnittene Originalfilmdatei, welche in der digitalen Anlage 1 zu finden ist. Neben dieser sind alle Links zu den einzelnen Kompositions-Dateien von AIVA in der zweiten digitalen Anlage aufgelistet. Es wird nachfolgend der Herstellungsprozess der neuen Filmmusik beschrieben und deren Zusammenwirken mit dem Bewegtbild beleuchtet. Die audiovisuelle Analyse ist lediglich unter strukturellen und formalen Gesichtspunkten möglich, da sich die Musikerstellung unabhängig von der Narration vollzieht. Um eine Struktur innerhalb der Generierungsumgebung von AIVA zu wahren, werden mehrere Ordner angelegt, in welchen die Scores entsprechend der Verwendung gespeichert werden.

### **5.2.1 Part 1: [00:00:00 – 00:03:31]**

Für den ersten Part werden insgesamt zwei Kompositionen erstellt. Zunächst wird die musikalische Untermalung des Gesprächs zwischen Fischer und seinem Vater bis zu dem Beginn des Kicks angefertigt. Dabei werden die Musikeinsätze aus dem Original übernommen und sich an dessen musikalischen Verlauf orientiert. Die Herausforderung liegt darin, ein Musikstück zu generieren, welches ruhig beginnt, sich anschließend aufbaut und den bevorstehenden Kick ankündigt. Die ersten Generierungen liefern keine verwertbaren Ergebnisse. Entweder eignen sich die Kompositionen für nur einen Teilbereich der Szene oder passen in ihrer Grundstimmung nicht zu dem Bewegtbild. So liefert AIVA beispielsweise eine Komposition im Stil Cinematic Solo String, welche trotz der beigefügten Attribute dark (dunkel) und serious (ernst) einen lebendigen und freundlichen Klangcharakter besitzt. Auch bei den folgenden Generierungen werden die Attributeingaben nur bedingt berücksichtigt, sodass mehr Einfluss auf die Tonart und die Akkordreihenfolge genommen wird. Mit dieser Vorgehensweise wird die gewünschte Stimmung wesentlich zielgerichteter als über die

Attributeingaben erreicht. Weiterhin wird deutlich, dass sich die finalen Kompositionen von der vorangegangenen Vorschau unterscheiden. So werden während des Generierungsvorgangs unter anderem weitere Instrumente hinzugefügt. Die Verbindung von zwei unabhängigen Kompositionen erweist sich als kompliziert, da diese über verschiedene Instrumentierungen, Akkordfolgen und Spieltechniken verfügen und nur durch manuelles Eingreifen einander angepasst werden können.

Diese Erkenntnisse führen zu der Entscheidung, dass über beide Szenen ein umfangreiches Musikstück generiert werden muss, damit sich dieses in das Bewegtbild einfügt. Es wird eine 2:00-2:30 minütige Komposition eines Streicher Ensembles in der Tonart e-Moll erstellt, welche in der grundlegenden Stimmung zu dem Szenenabschnitt passt. Sie enthält sowohl ruhigere als auch rhythmische Passagen und wird daher weiterverwendet. Die acht vordefinierten Akkorde gestalten sich wie folgt:

I Em<sup>add2</sup> | C<sup>maj9</sup> | D<sup>maj9</sup> | C D | Em<sup>add2</sup> | G<sup>add2</sup> | C<sup>maj7</sup> | B |

Die Komposition wird daraufhin im Editor (siehe Abbildung 5.1) geöffnet, da dieser eine nachträgliche Bearbeitung ermöglicht. Der Editor enthält unter anderem eine Ansicht der verschiedenfarbigen Midi-Spuren, welche wie in der Generierungsvorschau die musikalischen Schichten repräsentieren. Jede Midi-Spur kann zudem solo geschaltet, gemutet oder aufgeklappt werden, um die darin enthaltenen Instrumente sichtbar zu machen. Über veränderbare Farbbalken wird für jedes Instrument angezeigt, wann dieses aktiv ist. In dieser Ansicht können auch weitere Instrumente innerhalb einer Spur hinzugefügt werden. Darüber hinaus besteht die Möglichkeit, eine neue, benutzerdefinierte Midi-Spur als Pitched Layer oder Percussion Layer zu erstellen. Das Tempo der Komposition wird in Beats per Minute (BPM) angezeigt und kann bei Bedarf über eine Edit-Option angepasst werden. Für eine zeitliche und strukturelle Orientierung werden oberhalb des Arrangierfensters die Taktzahlen sowie die Akkordfolgen mit den entsprechenden Musikabschnitten dargestellt.

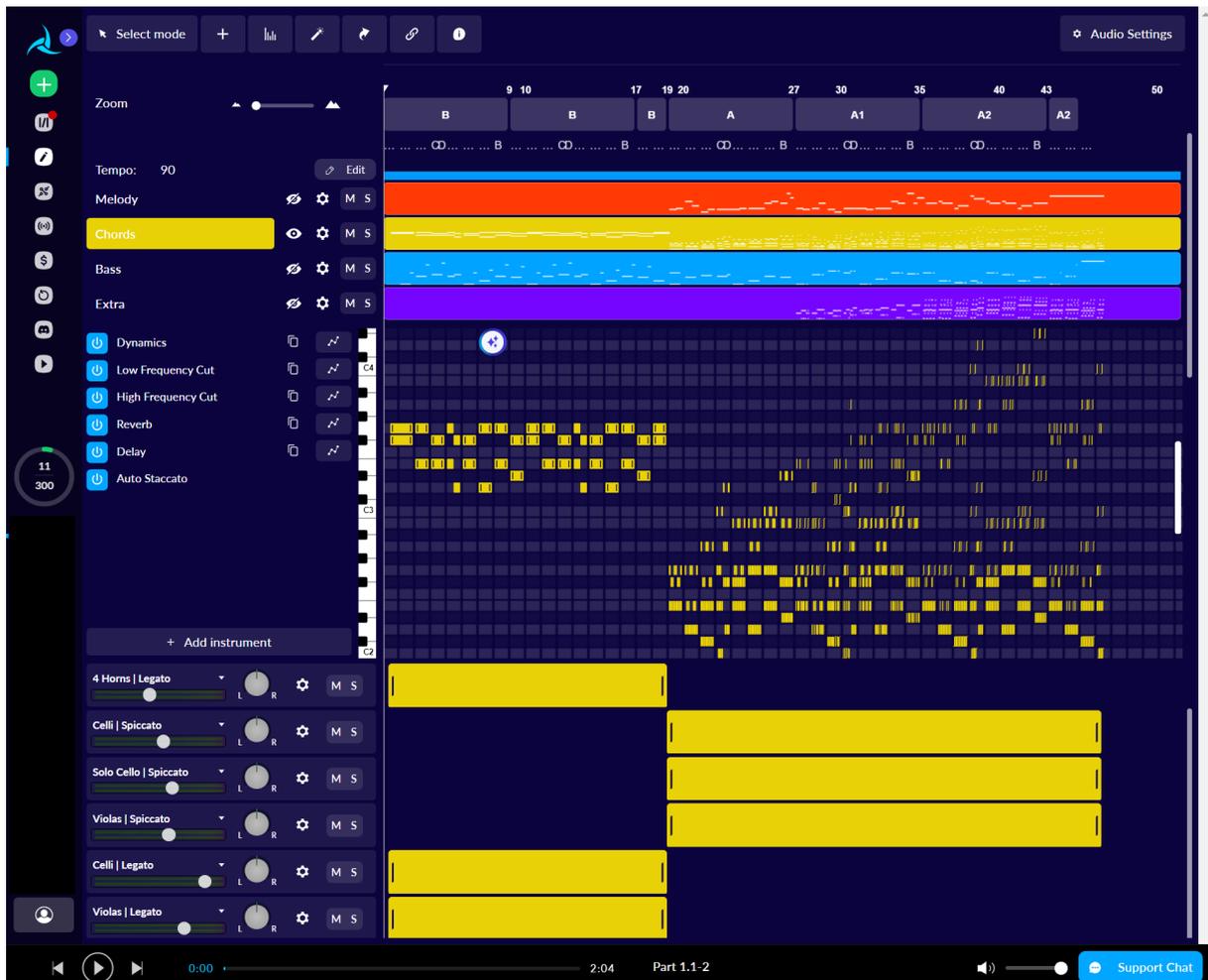


Abbildung 5.1 Aufbau der Editor-Ansicht von AIVA  
(AIVA, kein Datum)

Zunächst werden bei der generierten Komposition einzelne Musikabschnitte umgestellt. Die anfängliche Gesprächssituation soll von der ruhigen Passage mit wenig Bewegung unterlegt werden. Die rhythmischen Parts hingegen begleiten die parallelen Handlungsereignisse in den verschiedenen Traumebenen. Dabei müssen beide Bereiche verlängert werden, um nicht bereits vor den entsprechenden Bildschnitten der Szene zu enden. AIVA ermöglicht es, im Editor weitere Musikabschnitte zu generieren. Diese können sich an bestehenden Parts orientieren, diese kopieren oder ohne Vorgabe erstellt werden. Die verfügbaren Taktlängen sind mit zwei, vier oder acht Takten vordefiniert. Für den erstellten Score werden jeweils zwei Takte zu den bestehenden Musikabschnitten hinzugefügt. Dadurch, dass die rhythmisch verschiedenen Passagen lediglich aneinandergesetzt sind, wirkt der Übergang wenig fließend. Zudem endet die Komposition durch die Verschiebung der Musikabschnitte etwas abrupt. Dies stellt allerdings kein Problem dar, da die Musik unvermittelt durch die Explosionen unterbrochen werden soll und anschließend ein neuer Musikeinsatz bevorsteht. Die Instrumentierung wird von den Streichern dominiert, wobei die Akkordspur zu Beginn durch zusätzliche Hörner verstärkt wird. Celli und Bratschen führen den Score neben einem

Bass an und wechseln dabei in der Vortragsweise zwischen legato und spiccato. Die Melodie setzt erst im rhythmischen Part ein und wird von ersten Geigen besetzt.

Wird die Musik nun im Zusammenhang mit dem Bewegtbild betrachtet, wird eine Ähnlichkeit des neuen Scores zum Original ersichtlich. Die zentralen Synchronisationspunkte wurden übertragen, sodass direkt mit dem Öffnen der Tresortür [00:00:04] Bild und Musik erstmalig zusammenkommen. Auch in dem Moment, als Fischer den Grund der Enttäuschung seines Vaters realisiert [00:00:51], wechselt die musikalische Gestaltung. Zu Beginn liegt ein langsames Tempo vor, während die Akkordfolgen mehrmals identisch wiederholt werden. Durch die wechselnden Streicherrhythmen im zweiten Part findet eine Beschleunigung statt, mit welcher sich die Musik zunehmend aufbaut. Insgesamt wird die Dramaturgie in dieser Szene von der Musik unterlegt, sie vermag jedoch nicht, den Handlungsverlauf anzuführen.

Die Generierung der zweiten Komposition für die folgende Sequenz während des Kicks erweist sich als eine Herausforderung. Die Tonart e-Moll wird beibehalten und erste Kompositionen werden in den verschiedenen filmischen Musikstilen generiert. Der Score soll in diesem Abschnitt kraftvoll und dominant wirken und zugleich die Handlung mitvorantreiben. Mit der zuerst erstellten Komposition im Stil des Epic Cinematic Thriller ist dies jedoch nicht möglich. Weder die vielen synthetischen Klänge noch der musikalische Aufbau passen zu den gezeigten Filmbildern. Ähnlich verhält es sich mit den nachfolgenden Kompositionen, welche im Stil Epic Orchestra generiert werden. Diese können aufgrund fehlender Wirkungen, unpassender Stimmungen oder wegen des Klangcharakters nicht überzeugen. Es wird mit zusätzlichen Attributeingaben für die Erstellung der Akkordfolgen gearbeitet, die nur kleine Verbesserungen bringen. Hinzu kommt, dass während der Generierungen wiederholt sehr ähnliche Varianten erstellt werden, welche wenig Varianz aufweisen und oftmals einem undurchsichtigen Streicherteppich gleichen. Erst nach mehreren Durchläufen werden verwertbare Ergebnisse von AIVA geliefert. Bei dieser wird erneut das Epic Orchestra verwendet und eine aus nur zwei Akkorden bestehende Akkordfolge gewählt:

I Em I Em I Bm I Bm I Em I Em I Bm I Bm I

Mit diesen Einstellungen werden zwei Generierungen mit einer Länge von 3:00-3:30 Minuten angefordert. Die Eingabe der mehr als doppelt benötigten Dauer wurde bewusst getroffen, um eine höhere Wahrscheinlichkeit an verschiedenen, passenden Musikabschnitten zu erhalten. Die Entscheidung fällt auf die erste Variante, da diese im Vergleich zu allen vorherigen Kompositionen am kraftvollsten wirkt. Im Editor müssen allerdings mehrere Veränderungen vorgenommen werden, um die Musik passend unter das Bild zu legen. Zunächst wird das Tempo von 68 BPM auf 120 BPM erhöht und die Komposition auf die Szenenlänge verkürzt. Das Intro wird manuell bearbeitet und die Posaune durch eine Bassposaune ersetzt, da diese sonst von den parallelen Explosionen verdeckt wird. Zudem werden die Akkorde in Takt zwei

und vier entfernt, um dem Sounddesign an dieser Stelle mehr Platz zu geben. Das Hinzufügen oder Entfernen von Midi-Noten ist mittels des Pencil-Mode möglich. Dieser kann durch Umschalten des voreingestellten Select-Mode aktiviert werden und erlaubt eine beliebige Bearbeitung aller musikalischen Schichten. Der Select-Mode hingegen ermöglicht nur die Verschiebung bereits bestehender Midi-Noten einer einzelnen Spur.

Die Komposition besteht aus einer Hörner-Melodie, welche ab Takt zehn einsetzt. Die Akkorde werden wie erwähnt zunächst von Bassposaunen gespielt und anschließend durch zwei von AIVA vorgefertigte Orchestral Layers übernommen. Die umfangreiche Bass-Sektion besteht aus Tuben, Bässen, Celli, einem Flügel und einer Pauke und wechselt konstant zwischen den Grundtönen E und B. Eine Extra-Akkordspur wird von einem Streicherensemble und weiteren Hörnern besetzt. Diese wird viele Male neu generiert, da sie sehr inkonsistent wirkt und von unmusikalischen Wechseln durchsetzt ist. Zusätzlich ist eine Percussion-Spur vorhanden.

Für die Ankunft in der ersten Traumbene wird ein alternatives Outro angefordert. Nach mehreren ergebnislosen Durchläufen erfolgt die Generierung ohne Einfluss eines vorgegebenen Musikabschnitts und liefert vier Takte mit den abweichenden Akkorden D-Dur und C-Dur. Der Schlussakkord e-Moll wird abschließend manuell hinzugefügt.

Daraus ergibt sich ein Score, welcher die Bilder oberflächlich begleitet. Der Kick wird als eine Actionszene dargestellt, dessen narrative Bedeutung allerdings nicht über die Musik vermittelt wird. Wichtige Synchronisationspunkte werden erhalten oder leicht verändert. So beginnt die Musik nach der ersten Explosion [00:02:04] und erst mit dem Schnitt in den Aufzug [00:02:12] tritt sie zunehmend in den Vordergrund. Die Rhythmuswechsel der Streicher während den Schnitten auf Cobb im Limbus [00:02:29] sowie auf Ariadne bei der Rückkehr in Traumbene drei [00:03:01] fungieren ebenfalls als eine Art der Synchronisation. Allgemein kennzeichnen die vermehrten Rhythmuswechsel diese Komposition. Die abweichenden D-Dur- und C-Dur-Akkorde begleiten den Aufprall des Aufzugs, während mit dem Schnitt auf den abstürzenden Van der Schlussakkord folgt [00:03:18].

Das Tempo bleibt über die Szene konstant, wobei durch die Streicherrhythmen eine Beschleunigung suggeriert wird. Die generierte Musik ist insgesamt weniger kraftvoll und dominiert das Handlungsgeschehen nicht. Es findet keine auditive Verbindung der Traumbenen statt und auch die musikalische Ausarbeitung ist deutlich weniger ausgefeilt als im Original. Die fehlende narrative Bedeutung der Musik wirkt sich besonders in diesem Szenenabschnitt auf dessen Qualität aus.

### **5.2.2 Part 2: [00:03:32 – 00:06:08]**

Für den zweiten Szenenabschnitt werden zwei verschiedene Musikstücke benötigt, die fließend ineinander übergehen. Es wird erneut mit der Schritt-für-Schritt-Methode gearbeitet, wobei chronologisch mit der Sequenz im Limbus begonnen wird. Auch die neue Filmmusik soll

die intime und emotionale Handlung wiedergeben. Dazu werden der Musikstil Cinematic Solo Piano gewählt und die Tonart F-Dur vorgegeben. Dadurch dass keine Taktangabe vorliegt, wird diese Entscheidung von AIVA übernommen und ein 4/4-Takt festgelegt. Die Akkordfolge wird mehrmals unter der Eingabe des Attributs sad (traurig) generiert, bis sich für diese entschieden wird:

I Gm I Dm I F I C I Gm I Bb<sup>maj7</sup> I F I C I

Die Vorschau liefert eine Komposition, bei welcher sowohl die Melodie- als auch die Akkordstimme von einem Soft Piano übernommen wird sowie einen zusätzlichen Bass. Die Melodiespur wird mehrmals neu angefordert, da sie zunächst zu freudig und lebendig wirkt. Anschließend werden mit diesen Einstellungen drei Variationen von 1:00-1:30 Minuten Länge generiert, wobei die Wahl auf den dritten Vorschlag fällt. Dieser wirkt am stimmigsten und wird daher im Editor weiterverarbeitet. Die Trackdauer wird an das Bewegtbild angepasst und anschließend die Melodie für den ersten Abschnitt neu generiert, um etwas mehr Verzierungen zu erhalten. Zudem werden ab Takt elf die Akkorde durch Bratschen verstärkt, um diese klanglich aufzufüllen und eine leichte Steigerung in dem musikalischen Aufbau zu erreichen. Weiterhin ist nur die zweite Hälfte der Komposition mit Midi-Noten für den Bass versehen. Bei der erneuten Durchsicht von Bewegtbild und Musik wird deutlich, dass das Outro zu lang ist und bis nach den Schnitt auf die Insassen des sinkenden Vans reicht. Daraufhin entsteht die Idee, die Komposition auf denselben Tönen enden zu lassen wie die nachfolgende beginnt, um einen nahtlosen Übergang zu kreieren. Der Generierungsprozess hat für diese Sequenz zügig und gut funktioniert. Bereits die ersten Musikstücke konnten verwertet und für das Bewegtbild weiterbearbeitet werden. Im Zusammenhang mit dem Bild ergibt sich eine ähnliche, wenn auch nicht so ausdrucksstarke Verbindung wie im Original. Die Musik versucht primär die Stimmung einzufangen und widerzuspiegeln. Dafür liegt ein langsames, konstantes Tempo vor, das die Bildfolgen um Mal und Cobb begleitet. Auch die weiche und klare Substanz der Musik soll helfen eine gewisse Emotionalität herzustellen. Die Synchronisationspunkte der Bild- und Tonebene sind in der neuen Filmmusik reduziert. So wird der Moment von Mals Tod nicht mehr explizit in der Musik herausgearbeitet. Stattdessen setzen mit Cobbs Erkenntnis und Aussage „ich muss dich loslassen“ [00:04:09] die Streicher ein und unterstreichen diesen Moment. Der Musikeinsatz ist ähnlich wie im Original gestaltet und findet erst während des Dialogs, als Cobb auf Mals Aussagen reagiert, statt.

Für die anschließende Sequenz in der ersten Traumebene wird in der parallelen Molltonart d-Moll gearbeitet. Da sich in diesem zweiten Abschnitt die Musik hintergründig verhalten soll, werden die Solo Strings als Musikstil gewählt und anschließend drei Vorschläge mit einer Dauer von 2:00-2:30 Minuten generiert. Jedoch kann keine dieser allein stehend noch zusammen mit den Filmbildern überzeugen und werden daher aussortiert. Stattdessen fügt

sich eine Komposition, welche ursprünglich für eine andere Szene gedacht war, deutlich besser ein und wird folglich für diesen Abschnitt weiterverwendet. Diese steht in der Tonart a-Moll und wurde mittels des Cinematic String Ensembles generiert. Die gegebene Akkordfolge ist:

I F<sup>maj7</sup> | G | I Am | I Am | I F<sup>maj7</sup> | G | I Am | I Am | I

Die Komposition besteht aus einer Geigen- und Celli-Melodie, Bratschen-Akkorden, einem Bass sowie einer Extra-Akkordspur, welche neben ersten und zweiten Geigen ebenfalls von Bratschen besetzt ist. Zunächst wird die Musik wieder in ihrer Länge an das Bild angepasst und zugleich die unpassenden bzw. überflüssigen Musikabschnitte entfernt. Um den Übergang zur vorherigen Musik zu gewährleisten, muss das Intro manuell angepasst und verkürzt werden. Der erste Track klingt auf einem Zweiklang der Noten G und D aus, sodass der folgende Musikeinsatz erst auf dem zweiten Akkord, G-Dur, mit denselben Tönen beginnen darf. Durch diese Anpassung verkürzt sich die Gesamtlänge der Komposition, weswegen die Musik bereits vor dem Szenenwechsel endet. Aus diesem Grund werden 2-taktige Abschnitte mit den Akkorden F<sup>maj7</sup> und G im Stil des musikalischen Hauptteils hinzugefügt, um die Musik ausreichend zu verlängern. Die Generierung dieser Takte gestaltet sich zeitintensiv, da erst nach mehreren Versuchen verwertbare und mit dem angegebenen Stil vergleichbare Ergebnisse geliefert werden.

Bei der Zusammenführung von Musik und Bild wird deutlich, dass die Streicher dominieren und beinahe erdrückend wirken. Um dem entgegenzuwirken, wird die zuvor negativ aufgefallene Streichermelodie entfernt. Sie erklingt nun erst ab Takt 21 und wird durch vier weiche, in legato spielende Hörner ersetzt. Zudem wird die Lautstärke der Extra-Akkordspur zurückgenommen und die Anzahl der Streicher reduziert, sodass abgesehen von den letzten Takten lediglich zweite Geigen vorhanden sind. Dadurch geben sich die einzelnen Instrumente mehr Platz und auch das Sounddesign kommt wieder zur Geltung. In dieser Komposition kann ebenfalls eine gestalterische Nähe zum Original erkannt werden. So wird durch die musikalische Veränderung nach Fischers Monolog versucht [00:05:48], den Erfolg der Ideeneinpflanzung in der Tonebene zu verankern. Synchronisationspunkte sind in diesem Abschnitt hingegen kaum vertreten. Lediglich der Schnitt auf den gesunkenen Van [00:05:05] parallel zu den in der Lautstärke anschwellenden Streichern kann als ein solcher in Erwägung gezogen werden. Die weiterhin bestehende, niedrige Geschwindigkeit der Musik führt zu einer Entschleunigung des Moments, während diese fließend und weich bleibt. Mit dem Ende des Gesprächs von Arthur und Ariadne klingt die Musik aus und die Handlung der ersten Traumbene wird abgeschlossen.

### 5.2.3 Part 3: [00:06:09 – 00:11:39]

Im dritten Part werden das Gespräch von Cobb und Saito im Limbus sowie die mutmaßliche Rückkehr in die Realität bis zum Filmende musikalisch unterlegt. Dies wird durch zwei Kompositionen realisiert, wobei sich die Herangehensweise analog zu den vorherigen Szenen gestaltet. Zunächst wird die Komposition für den ersten Teil der Szene erstellt. Ziel ist, den Dialog im Limbus musikalisch zu unterlegen, sodass eine gewisse Bedrohung ausgestrahlt wird. Dazu wird eine erste Komposition in d-Moll unter der Attributeingabe danger (Gefahr) erstellt. Diese klingt jedoch nach einem traurigen Klavierstück und wird folglich nicht weiter berücksichtigt. Weitere Generierungen im Stil des Epic Orchestra werden mit zusätzlichen Attributen versehen, aber auch diese Kompositionen liefern keine zu dem Bewegtbild passenden Ergebnisse.

Daraufhin wird das bestehende Material verworfen und erneut begonnen. Der Musikstil Epic Orchestra wird gewählt, die Tonart bleibt in d-Moll. Es wird eine Akkordfolge verwendet, bei der lediglich vier Akkorde je einen Takt lang erklingen und anschließend in derselben Reihenfolge wiederholt werden. Diese sind:

I Dm            I Am            I F            I Gm            I

Die Vorschau liefert eine Gitarrenmelodie, flächige Streicherakkorde, ein Tuba-Ensemble als Bass-Sektion, weitere Celli für die Extra-Akkorde sowie ein Drumkit. Aus dieser werden drei Kompositionen mit einer Dauer von 2:00-2:30 Minuten generiert, wobei die Entscheidung für die erste Variante getroffen wird. Bei der Durchsicht mit dem Bewegtbild wird deutlich, dass die Musik zu viel Platz einnimmt und den Dialog beinahe verdeckt. Aus diesem Grund werden sowohl die Extra- als auch die Percussion-Spur gemutet und das Melodieinstrument durch ein Klavier ersetzt. Die Bass-Sektion wird insgesamt in der Lautstärke verringert, da neben den Tuben noch weitere Instrumente wie beispielsweise eine Pauke und Celli während des Generierungsvorgangs von AIVA hinzugefügt wurden.

Anschließend werden die Musikabschnitte für die Szenenlänge angepasst und vier neue Takte für das Trackende generiert, bei denen die Akkorde zwischen F und Gm wechseln. Zudem wird der Einsatz der Klaviermelodie auf die zweite Dialoghälfte verzögert, sodass diese erst mit der zurückkehrenden Erinnerung Cobbs beginnt. In den ersten beiden Takten werden die Bässe durch oktavierte Hörner ergänzt, um ein Gegengewicht zu den dominierenden tiefen Frequenzen zu bilden. Die Komposition endet auf einem Zweiklang der Noten B und D, was für den nachfolgenden Musikeinsatz mitberücksichtigt werden muss.

Zusammen mit dem Bild ergibt sich dadurch eine Filmmusik, welche die gezeigte Stimmung der Bilder wiederzugeben versucht. Sie agiert hintergründig und wirkt entgegen der Vorgabe wenig bedrohlich, sondern durch das hinzukommende Klavier beinahe sanft und demütig. Die Verkleinerung des Bildausschnitts hin zu den Gesichtern von Saito und Cobb wird von der

Musik ignoriert und auch die Synchronisationspunkte zwischen Ton und Bild haben sich verändert. Lediglich das einsetzende Klavier kurz vor der Einblendung des Kreisels kann als ein solcher gewertet werden. Zu Beginn der Szene kommen die Bild- und Tonebene nun versetzt zusammen, da die Musik erst im Anschluss an die Wellengicht einsetzt. Das Tempo ist konstant langsam und passt sich der Geschwindigkeit des Dialogs an. Die Komposition findet kein wirkliches Ende und bleibt auf einem Zweiklang stehen, welcher in den nachfolgenden Musikeinsatz überleiten wird.

Für das Filmfinale soll der Score eine deutliche Steigerung erfahren. Es wird das Epic Orchestra verwendet und die Tonart auf F-Dur festgelegt. Damit wird in der parallelen Dur-Tonart des vorherigen Musikstücks weitergearbeitet, welche bereits zu Beginn des zweiten Parts erstmalig verwendet wurde. Die vorgeschlagenen Akkordfolgen wirken platt und können nicht überzeugen, sodass die Progression von dem Originalscoretrack *Time* übernommen wird. Damit gestalten sich die Akkorde wie folgt:

I Gm I Dm I F I C I Gm I B<sup>maj7</sup> I F I C I

Die ersten generierten Kompositionen enthalten bereits geeignete Passagen, sind jedoch von mehreren unmusikalischen Wechseln und Gesangseinsätzen gezeichnet. Aus diesem Grund wird bei einem zweiten Versuch eine präzisere Instrumentenauswahl in der Vorschau-Ansicht getroffen. Dies liefert drei Kompositionen mit einer vornehmlich aus Hörnern bestehenden Melodie, Streicher- und Posaunenakkorden, einem Tuba-Ensemble und Celli für den Bass sowie einer Percussion-Sektion. Die generierten Vorschläge von 4:00-4:30 Minuten Länge werden verglichen und anschließend die zweite Variante für die weitere Verwendung gewählt. Ausschlaggebend dafür sind der passende musikalische Verlauf sowie die in den anderen Varianten erneut hinzugefügten dominierenden Gesangsstimmen.

Die Komposition wird im Editor weiterbearbeitet und an das Bewegtbild angepasst. Es besteht bereits ein musikalischer B-Teil, welcher von dem rhythmischen Charakter der Streicherakkorde geprägt ist. Diese Passage wird verlängert, sodass sie über die gesamte Flughafensequenz reicht. Zudem wird in diesem Abschnitt der hinzugekommene Gesang durch eine gezupfte E-Gitarrenmelodie ersetzt und in der Lautstärke reduziert. Die Reihenfolge der weiteren Musikabschnitte wurde nur leicht verändert, sodass sich die Musik kontinuierlich mit der Handlung aufbaut und kraftvoller wird. Für den musikalischen Höhepunkt, als Cobb auf seinen Stiefvater trifft und den Flughafen verlässt [00:10:15-00:10:55], werden die Melodiespur gedoppelt und zusätzliche Celli hinzugefügt. Zugleich werden einzelne Lautstärkeverläufe angepasst, um diese Steigerung zu präzisieren.

Das vorgefertigte Outro ist deutlich zu kurz und wird manuell überarbeitet. Dazu wird das Intro in dreimaliger Ausführung vor dem Outro positioniert und erklingt ab dem Moment, als Cobb seine Kinder im Garten spielen sieht. Es entsteht ein Einbruch in der Musik, welcher zusätzlich

durch den Instrumentenwechsel gestaltet wird. So reduzieren sich die mitwirkenden Instrumente mit jeder Wiederholung der Akkordfolge bis nur noch ein Flügel, eine Geige und eine Bratsche zu hören sind. Der Schlussakkord wurde entfernt, um ebenfalls musikalisch offen zu enden.

Bei dieser Komposition wird bereits während der Erstellung in gewisse Lautstärkeverhältnisse eingegriffen, um das Zusammenwirken von Bild und Musik einzuschätzen. In mehreren Iterationen werden kleinere Anpassungen an das Bewegtbild vorgenommen und ein Paukenschlag zu der ersten Note hinzugefügt. Dadurch sollen der Augenaufschlag Cobbs sowie der Schnitt zwischen Limbus und Realität [00:08:02] unterstrichen werden und zugleich eine Art der Synchronisation zwischen Ton und Bild bilden. Weitere Synchronisationspunkte aus dem Original werden ebenfalls übernommen. Bei der Einreisekontrolle [00:09:14] ändert sich die Musik hin zu dem erwähnten B-Teil, wobei statt des Innehaltens nun die antreibenden Streicherrhythmen dominieren. Nach dem Stempfeleintrag des Polizeibeamten [00:09:44] kehrt die Musik wieder zu ihrer vorherigen Form zurück und baut sich zunehmend auf. Dadurch dass das Sounddesign an dieser Stelle schlicht ist, muss die Musik vordergründig agieren und die Wirkung transportieren. Neben den zusätzlichen Instrumenten wird dies auch über eine Steigerung in der Dynamik versucht. In dieser Sequenz ist das Tempo höher als zuvor, wobei für die letzten Filmbilder eine Entschleunigung durch den Wegfall der Melodiespur und durch die Reduktion der Instrumente angestrebt wird.

#### 5.2.4 Soundalikes

Für den Beginn des ersten Parts sowie für das Filmende werden mit der Erstellung von Soundalikes experimentiert, welche letztendlich jedoch keine Verwendung für die alternative Filmmusik finden. Die generierten Kompositionen werden zusammen mit den Videosequenzen in zwei neue Filmdateien exportiert und sind in den digitalen Anlagen 3 bis 4 unter den Bezeichnungen „Soundalike Part 1“ und „Soundalike Part 3“ vorzufinden.

Für die Erstellung des ersten Soundalikes wird sich stark an den Akkorden des Originalscoretracks *Radical Notion* orientiert und über weite Passagen die folgende Akkordprogression verwendet:

I Gm I Gm I Gb I Gm I Eb I Eb I B I Bm I

Der Musikstil Modern Cinematic Epic Orchestra wird verwendet und drei 2:00-2:30 minütige Kompositionen werden angefordert. Darunter liefert AIVA eine rhythmische Variante, welche sich in die Richtung des Originals bewegt und für die weitere Ausarbeitung gewählt wird. Der ruhigere Anfangspart der Komposition wird über die gesamte Gesprächssituation verlängert und anschließend die Trackdauer an die Filmszene angepasst. Die Akkorde werden von mehreren Streichern und Hörnern übernommen und bilden die Basis des Scores. Die Bass-

Sektion wird von einer Pauke, einem Flügel, einer Tuba und weiteren Celli besetzt. Die Geigenmelodie sowie die Percussion-Elemente sind erst in der zweiten Hälfte der Komposition zu hören.

Im Vergleich zum Original ist eine klangliche Parallelität aufgrund der kopierten Akkordfolgen erkennbar. Die musikalische Entwicklung ist neben dem Stimmungswechsel innerhalb der generierten Komposition weniger ausgeprägt als im Original. Auch das Anschwellen der Emotionalität wird in dem Score von Zimmer subtiler und wirkungsvoller über die Streicher umgesetzt. Die generierte Musik versucht dies über die offensichtlicheren Rhythmenwechsel. Allerdings werden auch in der Version von AIVA die Handlungsereignisse von musikalischen Wechseln unterlegt, welche ohne nachträgliches Eingreifen entstanden sind. Für die kurze Erstellungsdauer ist das Ergebnis akzeptabel.

Für das zweite Soundalike wird versucht, eine Annäherung an den Klangcharakter des Scoretracks *Time* zu erreichen. Dazu wird erneut der Musikstil Modern Cinematic Epic Orchestra verwendet. Die Tonart, G-Dur, wird ebenso wie die Akkordfolge vom Original übernommen. Diese ist:

I Am I Em I G I D I Am I C<sup>maj7</sup> I G I D I

Die Vorschau der Komposition setzt aus einem Streicherensemble, flächigen Pads, Hörnern, mehrere Tuben und einem Piano zusammen. Neben diesen Instrumenten wird die Melodie durch eine zusätzliche Gesangsstimme ergänzt, welche entfernt wird. Die Generierung zweier Variationen mit einer Länge von 4:30-5:00 Minuten liefert verwertbare Ergebnisse. Auch wenn die Kompositionen über die vier Minuten etwas eintönig wirken, besitzt die zweite Variante einen geeigneten musikalischen Aufbau und wird daher weiterverwendet. Wie zuvor werden einzelne Musikabschnitte entsprechend des Handlungsverlaufs angepasst und ein statischer Zwischenpart für die Einreisekontrolle am Flughafen generiert. Diese Passage erweist sich als zeitaufwändig, da die gelieferten Generierungen klanglich wenig ansprechend sind. Nach der Durchsicht mit dem Bild werden weitere Kürzungen vorgenommen und das Tempo von 78 BPM auf 82 BPM erhöht, wodurch es über dem des Originals liegt. Für den musikalischen Höhepunkt wird auch hier die Melodiestimme gedoppelt und durch Hörner verstärkt. Eine weitere Bearbeitung erfolgt im Outro, welches ebenfalls durch eine Reduktion der Instrumente gekennzeichnet ist. Ein Flügel und mehrere Geigen führen das Musikstück zu Ende und bleiben bis zur Schwarzblende auf den Noten A und D stehen.

Die Wirkung des Soundalikes ist verglichen zum Original weniger monumental. Auch die musikalische Ausarbeitung ist weniger detailreich. Der Aufbau hingegen wird weitestgehend nachgebildet und auch die Stimmung nähert sich aufgrund des vollen Orchesters und derselben Akkorde dem Original hörbar an. Hinzu kommt, dass diese Komposition in nur wenigen Stunden vollständig erstellt wurde.

### 5.3 Postproduktion

Nach der Fertigstellung der Kompositionen werden die Möglichkeiten der Klanganpassung in der Editor-Umgebung von AIVA näher betrachtet. Die einzelnen Komponenten werden bereits während des Generierungsvorgangs automatisch vorgemischt, sodass diese ohne weitere Bearbeitung vom Nutzer übernommen werden können. Es werden aber auch diverse Einstellungsoptionen bereitgestellt, um in die Mischung einzugreifen und diese nach der gewünschten Klangvorstellung auszuarbeiten.

Zum einen können in einem separaten Mixing-Fenster die aktuellen Pegeleinstellungen aller musikalischen Schichten einer Komposition eingesehen werden. In 1dB Schritten können diese je Schicht erhöht oder verringert werden. Dieses Tool wird bei der generierten Filmmusik für die Angleichung der Lautstärkeverhältnisse genutzt, wenn eine Instrumentengruppe insgesamt zu leise bzw. zu laut ist. Dies ist beispielsweise während der Sequenz in der ersten Traumbene der Fall [00:04:38 - 00:06:08], da die melodieführenden Streicher alle weiteren Instrumente verdecken und folglich deutlich zurückgenommen werden.

Um präzisere Lautstärkeanpassungen sowie Einfluss auf einzelne Instrumente vorzunehmen, wird eine Midi-Spur ausgewählt und aufgeklappt. Dadurch werden die enthaltenen Instrumente sichtbar und können individuell mit einem Fader angepasst werden. Diese Anpassungen erfolgen meist nur geringfügig, da von AIVA bereits stimmige Instrumentenverhältnisse vorgemischt wurden. In dem Musikstück zu Beginn der Filmszene [00:00:04-00:02:04] werden beispielsweise kaum Veränderungen vorgenommen. In dem darauffolgenden Track [00:02:05-00:03:20] wird aufgrund des lauten Sounddesigns inklusive der Explosionen hingegen mehr eingegriffen. In dieser Ansicht besteht zudem die Möglichkeit, weitere Einstellungen für die gesamte Instrumentengruppe ein- oder auszuschalten. Es können ein Hochpass- und ein Tiefpassfilter gesetzt, sowie die Ausprägungen von Hall und Delay bestimmt werden. Zudem kann ein Auto Staccato für die automatische Anpassung der Artikulation (de)aktiviert werden. Die für die Filmmusik am meisten verwendete Funktion ist allerdings die Lautstärkeautomation. Bereits während der Generierung werden von AIVA Automationen geschrieben, welche den musikalischen Verlauf unterstützen. Diese werden teilweise fast gänzlich für die erstellten Kompositionen übernommen, wie zum Beispiel im finalen Scoretrack, bei dem die kontinuierliche Lautstärkesteigerung mit dem kurzzeitigen Einbruch des Zwischenparts als passend empfunden wird. In anderen Passagen werden die Lautstärkekurven etwas abgeändert oder neu gezeichnet, um die Musik an das gezeigte Bewegtbild anzupassen. So werden unter anderem einzelne Instrumentengruppen wie die Melodiestimmen stärker in ihrer Dynamik bearbeitet als zum Beispiel die Bass-Sektion. Für die Streicher in der letzten Bildeinstellung [00:11:35] oder auch für wichtige Synchronisationspunkte, wie die Rückkehr in die erste Traumbene [00:03:18], werden kurze Crescendi eingefügt, um musikalische Akzente zu setzen. Ebenso werden Instrumentengruppen zurückgenommen, um Einbrüche

zu erzeugen. Interessant ist, dass AIVA bei längeren Sequenzen gleicher Lautstärke eigenständig minimale Lautstärkeautomatationen generiert, wodurch gerade die Streicher lebendiger und organischer wirken. Insgesamt führen die Dynamikverläufe dazu, dass die Kompositionen den Handlungsverlauf glaubwürdiger untermalen.

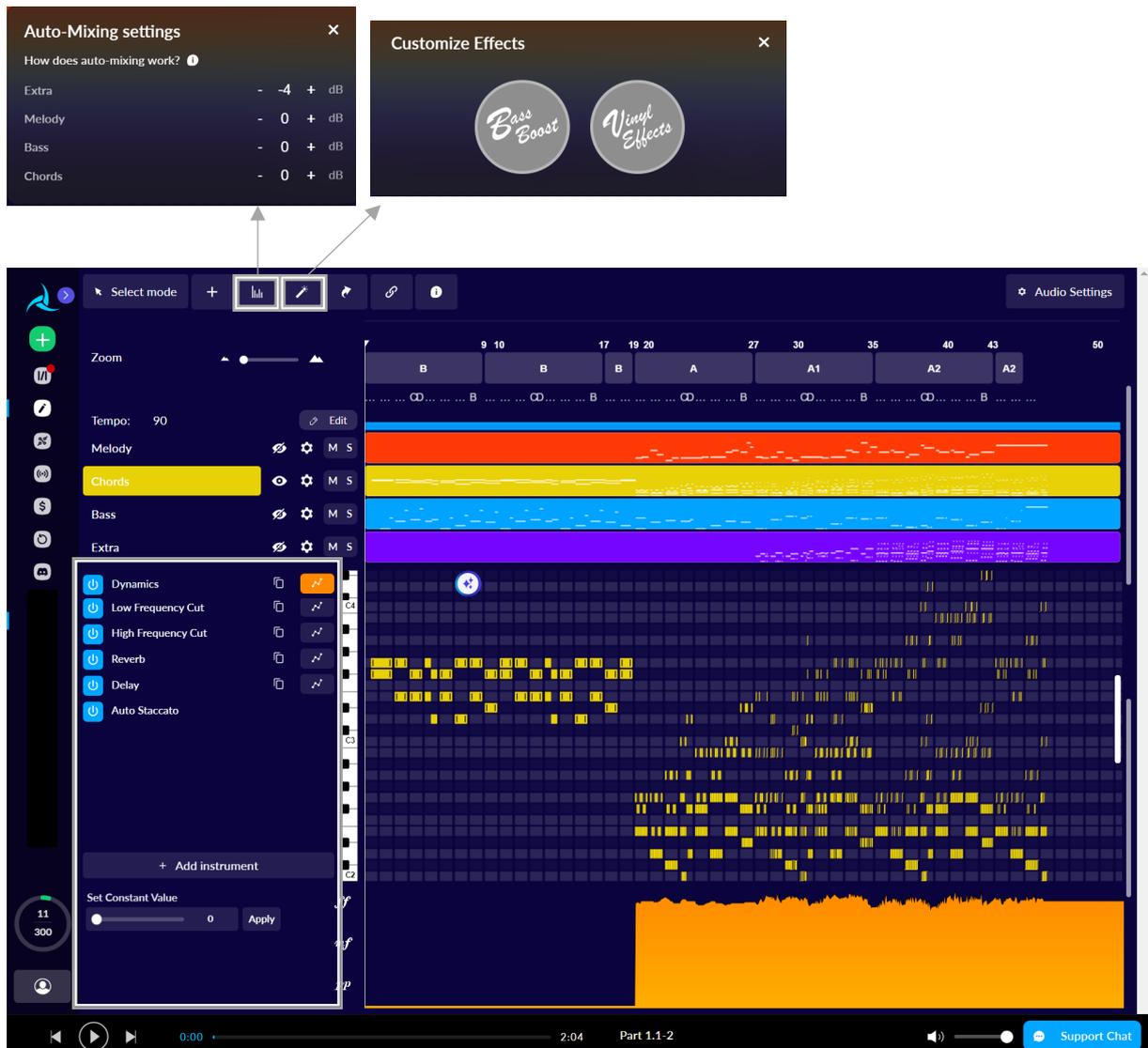


Abbildung 5.2 Zusammenstellung der verfügbaren Nachbearbeitungsmöglichkeiten in der Editor-Ansicht von AIVA (AIVA, kein Datum)

Ein Nachteil ist, dass die Automationskurven nur global auf die gesamte Instrumentengruppe angewendet werden können. Ein einzelnes Instrument zu automatisieren ist demnach nur möglich, wenn eine neue Schicht mit diesem erstellt wird und die entsprechenden Midi-Noten übertragen werden. Zudem erfolgt die Automation entweder durch Festlegen eines konstanten Wertes oder durch manuelles Einzeichnen der Kurve. Es können keine zentralen Punkte gesetzt werden, anhand derer die Automationen verschoben bzw. durch lineare Steigung miteinander verbunden werden können. Soll eine Automation hingegen auf alle Schichten übertragen werden, ist dies über eine Kopierfunktion möglich. Des Weiteren besteht kein

Einfluss auf die Parameter von Hall und Delay, lediglich deren Anteil in der Komposition kann gesteuert werden. Dabei stehen für jede musikalische Instrumentenschicht dieselben Bearbeitungsmöglichkeiten zur Verfügung. Eine klangliche Bearbeitung der verwendeten Instrumente ist in AIVA hingegen nicht möglich. Aus Gründen der Vollständigkeit wird noch eine globale Effekt-Option genannt, welche jedoch nicht auf die generierten Musikstücke angewendet wird. Bei dieser können die Audioeffekte Bass Boost und Vinyl Effects einzeln oder zusammen über die gesamte Komposition gelegt werden. Ein genauer Einblick in diese Effekte sowie eine weitere Bearbeitung dieser wird jedoch nicht gestattet.

Um die Bild- und Tonebene final aneinander anzupassen, muss das erstellte Material auf eine andere DAW ausgelagert werden. In AIVA besteht weder die Möglichkeit eine Videodatei zu importieren und zusammen mit der Musik in einer Session zu öffnen, noch können mehrere Kompositionen in einer gemeinsamen Ansicht miteinander verbunden werden. Daher wurde jeder Track einzeln im Editor angepasst und die Verhältnisse zwischen den Instrumenten festgelegt sowie zentrale Lautstärkeverläufe hinzugefügt. Somit ist die Musik in AIVA weitestgehend fertiggestellt, sodass diese nachfolgend dem Bewegtbild unterlegt werden kann. Die fertigen Audiodateien können nur als 44,1 kHz Stereo-Wave-Dateien aus AIVA exportiert werden, mehrkanalige Formate werden von der Software nicht unterstützt. Es besteht jedoch die Möglichkeit, einzelne Instrumente, Instrumentengruppen oder gesamte Midi-Spuren als Stems zu exportieren und mit diesen anschließend weiterzuarbeiten, um beispielsweise Surroundmischungen zu erstellen.

Nach dem Export aus AIVA werden die Stems in eine Reaper-Projektdatei geladen, in welcher auch das Sounddesign nachgebaut wurde. Dazu werden die Dateien auf 48 kHz konvertiert. Für die Mischung wird wie zuvor bei der Produktion mit den Kopfhörern Beyerdynamic DT 770 PRO gearbeitet und auf diversen Geräten gegengehört. Insgesamt orientiert sich die Mischung an dem Original, wobei zugleich die neue Musik zur Geltung gebracht werden soll. In der Regel werden Filmproduktionen über die integrierte Lautheit auf -23 LUFS gemischt (Filmtone Berufsvereinigung e.V., 2023). In diesem Fall handelt es sich nur um eine Filmsequenz, sodass sich an der gemessenen Lautheit der originalen Filmdatei orientiert wird, welche bei -19,5 LUFS liegt. Zunächst werden die extrahierten Dialogpassagen und Effekte sowie die nachträglich hinzugefügten Elemente in Lautstärke und Panning einander angepasst. Erst im Anschluss wird die Musik abschnittsweise hinzugefügt und die fertigen Kompositionen in ihrer Gesamtlautstärke und in ihrem musikalischen Verlauf entsprechend des Bewegtbilds automatisiert. Zudem werden Fades für die Übergänge zwischen den Tracks gesetzt.

So kennzeichnet sich der erste Part durch einen Lautstärkesprung nach dem Gespräch von Fischer mit seinem Vater [00:00:51]. Anschließend steigert sich die Musik bis zu den Explosionen [00:02:04] und bleibt während des Kicks die treibende auditive Kraft. Eingeleitet

wird der Score durch einen langen Fade-In, welcher parallel zu der sich öffnenden Tresortür und Fischers Betreten des Raumes positioniert wird.

Im zweiten Part soll die Musik zu Beginn dem Dialog ausreichend Platz lassen und sich mit dem Wechsel in die erste Traumbene [00:04:38] etwas steigern. Mittels eines Crossfades werden die beiden Musiken miteinander verbunden, sodass deren Übergang weicher wird. Ein kurzer Akzent wird zudem nach dem Monolog Fischers [00:05:48] zur Bestätigung des Erfolgs der Inception eingesetzt.

Der Score des dritten Parts enthält zu viele Bassanteile und konkurriert dadurch mit den weiteren Elementen des Soundtracks. Daher werden diese reduziert und nur kurzzeitig bei dem Schnitt in die Sequenz hervorgehoben [00:06:10]. Während des Gesprächs im Limbus bleibt die Musik hintergründig und verstärkt erst mit dem Einsatz des Klaviers [00:07:18] ihre Präsenz. Nach der Rückkehr in die Realität [00:08:02] soll die Musik zunehmend an Dominanz gewinnen und wie das Original die gezeigten Bilder anführen.

Durch das Hinzufügen der Musik müssen einzelne Elemente des Sounddesigns in ihrer Lautstärke erneut angepasst werden, um im Gesamtkontext ausreichend hörbar zu bleiben. Die Mischung wird iterativ bearbeitet und dem Bewegtbild angepasst. Zuletzt wird die fertige Filmszene neben dem Original exportiert und ist in der digitalen Anlage 5 unter dem Dateinamen „Inception-Szene mit generierter Filmmusik“ zu finden.

## **5.4 Fazit**

### **5.4.1 Herstellungsprozess**

AIVA ist eine facettenreiche KI-Software, um Kompositionen verschiedenster Art zu generieren. Die Herstellung kann gänzlich durch AIVA übernommen werden, wobei auch eine gewisse Zusammenarbeit mit dem Nutzer angeboten wird. Die Stärke der möglichen Einflussnahme wird entsprechend der Generierungsmethode gewählt und geht zugleich mit dem individuellen Wissen zur Musiktheorie einher. Allerdings kann durch die verschiedenen Generierungsmodi unabhängig des Wissenstands mit der Software gearbeitet werden. AIVA verfügt zudem über weitere Hilfestellungen, sodass beispielsweise die gewählte Tonart eingehalten wird. Musiktheoretische Kenntnisse sind hilfreich, um kreativ und experimentell zu arbeiten und die gegebenen Möglichkeiten bei der Gestaltung auszuschöpfen. Zu Beginn bedarf es einer kurzen Einarbeitung in die Software, um zielgerichtet und effektiv vorgehen zu können. Zumeist wird in unter zwei Minuten ein vollständiger Track generiert. Die weitere Ausarbeitung bis zur Fertigstellung dauert ebenfalls nur wenige Stunden, wobei die erstellten Musiken von einfachen bis hin zu komplexeren Kompositionen reichen. Die sich daraus ergebende Zeit- und Kostenersparnis im Vergleich zur Beauftragung eines Komponisten ist

erheblich, denn in AIVA können entsprechend des gewählten Abonnements monatlich bis zu 300 Kompositionen für einen geringen Preis exportiert werden (AIVA, kein Datum).

Ein Nachteil von AIVA ist, dass für die Kompositionen nur wenige nachträgliche Bearbeitungsmöglichkeiten zur Verfügung stehen. So sind keine Tonartenwechsel innerhalb eines Musikstücks möglich. Die zu Beginn festgelegten acht Takte werden regelmäßig wiederholt und können nach der Generierung nicht mehr beeinflusst werden. Lediglich durch manuelles Eingreifen oder durch Neuerstellung eines Musikabschnitts kann eine abweichende Akkordfolge derselben Tonart erreicht werden. Eine höhere Flexibilität ist ebenso im Umgang mit der Taktaufteilung bei den generierten musikalischen Abschnitten erwünscht. Diese können weder in ihrer Länge beeinflusst noch in ihrer Positionierung verschoben werden. Einerseits unterbindet dies unmusikalisches Arbeiten, andererseits ist es arbeitserschwerend, da sowohl bei Verlängerungen, Verkürzungen als auch bei Verschiebungen neue Musikabschnitte generiert werden müssen. AIVA liefert Kompositionen, welche stets dem durch die Akkordfolge festgelegten Muster folgen. Dadurch fehlt eine gewisse Varianz, welche für die musikalische Untermalung der Filmbilder nötig ist. Zudem wird dadurch die Arbeit mit AIVA phasenweise eintönig, da sich die Kompositionen trotz vieler Generierungen nur geringfügig unterscheiden und teilweise unmusikalische Ergebnisse entstehen.

Bei der Erstellung von Soundalikes mittels der KI-Software ergeben sich sowohl Möglichkeiten als auch Risiken. Grundsätzlich können diese durch die Verwendung derselben Akkorde und Instrumente sowie durch die Kopie des strukturellen Aufbaus ohne Weiteres umgesetzt werden. Diese klangcharakterlichen Replikat implizieren erneut die Problematik bezüglich des Umgangs mit dem Urheberrecht. Zudem stellt sich die Frage, ob die KI zur Kreation neuer Werke oder als Tool zur Herstellung von Kopien eingesetzt werden soll.

In Bezug auf die erstellte Filmmusik steht und fällt die Wirkung des generierten Tracks mit der Auswahl des Musikstils und den gewählten Akkordfolgen. Synchronisationspunkte zwischen Ton und Bild sind auch in den generierten Kompositionen gut erreichbar. Es können neue Punkte entstehen oder vorhandene durch Eingreifen in die einzelnen Musikabschnitte genauer positioniert werden. Die Erzeugnisse der KI wirken jedoch unpräziser und bedeutungsärmer, da sie nicht direkt auf die Handlung zugeschnitten sind. Die fehlende musikalische Filigranität führt dazu, dass die im Bewegtbild vorherrschenden Stimmungen und Emotionen nicht im Umfang eines beispielsweise von Hans Zimmer komponierten Scores erreicht und transportiert werden. Die unbearbeitete Verwendung einer Komposition liefert zumeist nur unzufriedenstellende Ergebnisse. Das Setzen einzelner musikalischer Akzente über AIVA ist ebenfalls schwierig und erfordert eine manuelle Ausarbeitung. Dadurch dass die Taktlängen vordefiniert sind, mussten teilweise die Intros und Outros der Kompositionen verlängert bzw. verkürzt werden, um die Bildsequenz passend zu unterlegen oder, um in den neuen Musikeinsatz überzuleiten. Für manche Parts verlief die Erstellung zügig und effektiv, sodass

bereits die ersten Vorschläge weiterverwertet werden konnten. Für andere Sequenzen waren weitaus mehr Generierungen nötig, wie beispielsweise für das Gespräch von Cobb und Saito im Limbus oder während des Kicks. Insgesamt gestaltet sich die Generierung als ein iterativer Prozess. Die Kompositionen werden zusammen mit dem Bild in mehreren Durchläufen gesichtet und daraufhin angepasst. Dabei besteht die Gefahr sich an Kleinigkeiten aufzuhalten und fortlaufend neue Kompositionen zu generieren.

Eine Herausforderung bei der Arbeit mit AIVA ist, eine Konstanz in Klang, Stil, und Instrumenten über einen längeren Zeitraum und verschiedene Musikstücke zu erreichen, da jede Komposition separat erstellt wird. Zudem ist die Verwendung von Leitmotiven kaum möglich und wird aufgrund der unzureichenden Ergebnisse bei der Musikgenerierung über Referenzen weiter erschwert. Dennoch wird eine Fülle an Möglichkeiten geboten, sodass in kurzer Zeit vieles ausprobiert und mit verschiedenen Musikstilen experimentiert werden kann. Bei einem klaren Konzept ist ein effektives und zielgerichtetes Arbeiten möglich, welches schnell zu Ergebnissen führt. Allerdings entsteht während des Herstellungsprozesses auch eine große Menge an Ausschuss. So wurden für die Filmmusik insgesamt über 80 Kompositionen erstellt, von denen sechs verwendet und weiter ausgearbeitet wurden.

Bei den Generierungen muss zumeist ein Kompromiss aus der eigenen Vorstellung und dem tatsächlich gelieferten Musikstück eingegangen werden. Die musikalischen Rahmenbedingungen können mitgestaltet werden, die weitere Entscheidungsfreiheit ist allerdings eingeschränkt und wird von der Software übernommen. AIVA ermöglicht die Generierung von Filmmusik, ist jedoch nicht auf diesen Bereich spezialisiert. Die Implementierung weiterer Funktionen wäre daher erstrebenswert. Zum einen umfasst dies die Arbeit mit verschiedenen Abstraten sowie eine mehrkanalige Arbeitsumgebung. Zum anderen würde eine höhere Flexibilität innerhalb der Session sowie die Klangbearbeitung der verwendeten Instrumente weitere Vorteile mit sich bringen. Die parallele Bearbeitung und Verbindung mehrerer Kompositionen in einer Session würde die Gesamtqualität des Scores steigern, ebenso wie die Möglichkeit längere, zusammenhängende Kompositionen zu generieren. Zusätzlich wären die Optionen, eine Videosequenz als Eingabe für die Generierung zu verwenden oder auch eine generierte Komposition in eine Partitur umzuwandeln, um diese anschließend von einem echten Orchester einspielen zu lassen, von Nutzen.

#### **5.4.2 Vergleich Originalscore**

Wie bereits in den vorigen Kapitelabschnitten angedeutet, fällt die Komposition der KI-Software im Vergleich zum Originalscore ab. Zimmers Filmmusik ist etabliert und wurde millionenfach zusammen mit dem Bewegtbild gesehen. Der Score wird durch einen speziellen

Sound geprägt und ist über den gesamten Handlungsverlauf durchdacht. Dadurch hält dieser eine Bedeutungsebene inne, welche die generierte Komposition nicht wiedergeben kann.

Zudem wird deutlich, dass das Original die Generierung der neuen Filmmusik beeinflusst hat. Die verschiedenen Stimmungen orientieren sich ebenso wie der musikalische Verlauf an dem Score von Zimmer, wobei die klanglichen Defizite der KI-Produktion hörbar sind. Besonders die Streicher wirken weniger natürlich, sondern artifiziell. Auch die Frequenzverteilungen von Sounddesign, Dialog und Musik liegen enger beieinander als im Original, sodass mitunter Verdeckungen stattfinden und die Elemente zeitweise gegeneinander konkurrieren.

Zu Beginn der Arbeit wurde erwähnt, dass die Musik für Inception bereits vor dem Bewegtbild angefertigt und als dessen Erstellungsgrundlage verwendet wurde. Somit ist es fast unmöglich, diese audiovisuelle Verbindung zu übertreffen. Zudem fehlen bei den generierten Kompositionen eine gewisse Raffinesse und Varianz, welche die Filmmusik besonders machen. Dies zeigt sich beispielsweise während der Kicksequenz, als im Original die Traumschichten in Bild und Ton zusammengeführt werden. Die Komposition von AIVA bringt eine gewisse Dynamik mit sich, jedoch greift sie die verschiedenen Handlungsstränge der Traumebenen nicht auf und enthält keine narrative Bedeutung. Auch in den letzten Takten der finalen Filmszene verhält es sich ähnlich. In dem Score von Zimmer wird die Melodie in den Streichern weitergeführt und damit das offene Handlungsende musikalisch unterstrichen. Bei der generierten Musik hingegen werden lediglich die Akkorde wiederholt und reduziert bis sie auf einem Schlussakkord enden.

Trotz dessen, dass die generierte Filmmusik im direkten Vergleich in ihrer musikalischen Ausarbeitung nicht mit der Qualität des Originals mithalten kann, können mit AIVA solide Musikstücke erzeugt werden. Das Tool kann dabei unterstützend, als Inspiration bzw. Ideengeber oder als Kreativeur eingesetzt werden. Es werden vielseitige Ansätze geboten, um eine Filmmusik zu erstellen, auch wenn diese derzeit noch nicht die Präzision und den Bedeutungsgehalt eines frei von der Hand komponierten Scores erreicht. Nach den gesammelten Erfahrungen wird empfohlen, die KI als Hilfsmittel für Bewegtbildprojekte sowie für erste musikalische Entwürfe einzusetzen oder wenn die Kompositionen alleinstehend unter wenigen Vorgaben erstellt werden.

## 6. Evaluierung der künstlich generierten Filmmusik

Um das generierte Material zu evaluieren und um einen Einblick in den derzeitigen Einfluss von künstlicher Intelligenz in der Filmmusikbranche zu erhalten, werden Experteninterviews durchgeführt. Zunächst werden das Vorgehen und die angewandte Methodik dargelegt, bevor die herausgearbeiteten Ergebnisse in der Auswertung geschildert werden.

### 6.1 Vorgehen

In diesem Kapitel soll die dritte der zu Beginn aufgestellten Teilfragen geklärt werden.

- ◆ Wie wird die Übernahme kreativer Aufgaben durch die generative KI und die Qualität der Erzeugnisse von Komponisten und Filmmusikern empfunden?

Dafür wird die generierte Filmmusik, verglichen mit dem Originalscore, von Komponisten und Filmmusikern in ihrer Qualität eingeordnet und bewertet. Zudem werden Erfahrungen aus dem professionellen Bereich bezüglich der KI gestützten Filmmusikkomposition eingeholt und rekonstruiert. Auch der derzeitige Einfluss von künstlicher Intelligenz innerhalb der Branche soll erfasst werden. Umgesetzt wird dies mithilfe von Experteninterviews. Laut Brink (2013) ist ein Experte eine Person, „die sich über mehrere Jahre auf dem zu untersuchenden Gebiet eine besondere Kompetenz oder ein spezifisches Wissen angeeignet hat, das nicht allgemein zugänglich ist“ (S.131). Da während des Interviews auch individuelle Erfahrungen und die persönliche Haltung der Experten miteinfließen, sind die Antworten zu einem gewissen Grad subjektiv. Aus diesem Grund werden mehrere Experten befragt, um die Subjektivität zu reduzieren (Lehnen, 2017). Experteninterviews werden zumeist leitfadengestützt durchgeführt. Der Leitfaden wird dabei im Vorfeld erstellt und strukturiert den inhaltlichen Rahmen des Gesprächs (Liebhold & Trinczek, 2009). Zudem kann dieser bereits erste zu prüfende Hypothesen enthalten. Spontane Nachfragen und eine wechselnde Reihenfolge einzelner Fragen bleiben dennoch möglich (Liebhold & Trinczek, 2009). Die Nutzung eines einheitlichen Leitfadens hilft weiterhin dabei, das Gespräch ausschließlich auf das Wissen der Experten einzugrenzen, um anschließend die gegebenen Antworten leichter vergleichen zu können (Berger-Grabner, 2022). Jedes Interview sollte als Audiodatei aufgezeichnet und möglichst zeitnah transkribiert werden, da anhand der Transkripte die anschließende Auswertung der Daten erfolgt. Im Forschungskontext wird bei Interviews häufig die qualitative Inhaltsanalyse verwendet, wobei das Vorgehen systematisch und regelgeleitet ist (Kuckartz & Rädiker, 2022). Im nachfolgenden Abschnitt wird die gewählte Methode näher beschrieben.

## 6.2 Methodik

Udo Kuckartz und Stefan Rädiker haben drei verschiedene Verfahren für die qualitative Inhaltsanalyse entwickelt: die inhaltlich strukturierende, die evaluative und die typenbildende qualitative Inhaltsanalyse. Für diese Arbeit wird mit der zuerst genannten, der inhaltlich strukturierenden qualitativen Inhaltsanalyse gearbeitet. Diese beinhaltet ein mehrstufiges Verfahren, welches in der Abbildung 6.1 dargestellt wird.

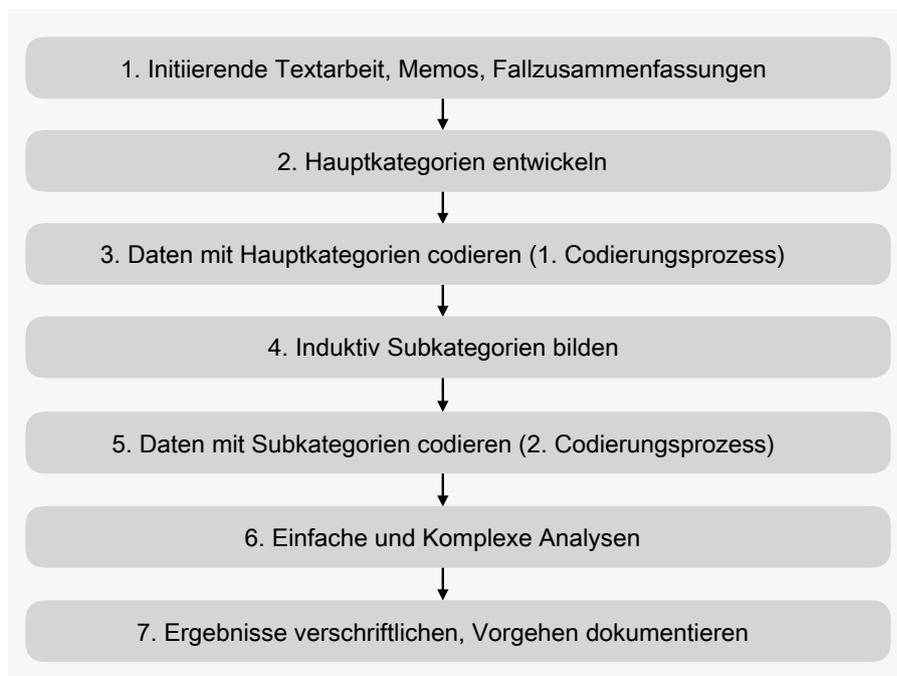


Abbildung 6.1 Phasen der inhaltlich strukturierenden qualitativen Inhaltsanalyse  
In Anlehnung an (Kuckartz & Rädiker, 2022, S.132)

Zunächst wird in der ersten Phase das erhobene Material gesichtet und zentrale Textpassagen durch Markierungen hervorgehoben. Zusätzlich können Anmerkungen und Kommentare in Memos dokumentiert werden. Abgeschlossen wird diese erste Bearbeitung durch das Erstellen von Fallzusammenfassungen (Kuckartz & Rädiker, 2022).

Im zweiten Arbeitsschritt werden die Hauptkategorien entwickelt. Die Definition von Kategorien bildet die Grundlage für die Auswertung der Daten und dient dazu die Inhalte zu strukturieren (Kuckartz & Rädiker, 2022). Häufig enthalten diese die zentralen Themen, die bereits anhand der Forschungsfrage und des Leitfadens bei der Datenerhebung festgelegt werden. Jedoch ist es möglich, dass sich bei der Sichtung noch weitere wichtige Aspekte aus dem Textmaterial herauskristallisieren. Prinzipiell sollte die Kategorienbildung den Gütekriterien folgen und möglichst präzise, eindeutig und ergebnisorientiert beschrieben sein (Kuckartz & Rädiker, 2022).

Phase drei umfasst den ersten Codierungsprozess. Bei diesem wird das gesamte Transkript zeilenweise analysiert und den definierten Kategorien zugewiesen. Dabei können einzelne Textpassagen mehrfach codiert werden oder auch bei fehlender Relevanz uncodiert bleiben.

Zu beachten ist, dass die Codierung in „Sinneinheiten“ (Kuckartz & Rädiker, 2022, S.136) erfolgt, sodass diese für sich stehend nachvollziehbar sind. Der Umfang reicht demnach von einzelnen Sätzen bis zu längeren Textpassagen mit eventuellen Zwischenfragen (Kuckartz & Rädiker, 2022).

Im Anschluss daran werden im vierten Schritt aus den bestehenden Hauptkategorien inhaltlich differenzierte Subkategorien herausgearbeitet. Dazu werden alle codierten Textstellen einer Kategorie aufgelistet und anhand dessen mögliche Subkategorien formuliert. Diese werden systematisiert und in einem Kodierleitfaden mit einer Definition und einem Ankerbeispiel aus dem Transkript versehen (Kuckartz & Rädiker, 2022).

Die fünfte Phase beinhaltet den zweiten Codierungsprozess. Wie zuvor wird das Transkript erneut feingliedrig durchlaufen und die Textstellen werden nun mit den neuen Subkategorien codiert (Kuckartz & Rädiker, 2022).

Auf diesen Arbeitsschritt folgen in der sechsten Phase die einfachen und komplexen Analysen basierend auf den verwendeten Kategorien. Dabei stehen verschiedene Methoden zur Auswertung des Materials zur Verfügung. Im Rahmen dieser Arbeit wird mit der „kategorienbasierte[n] Analyse entlang der Hauptkategorien“ (Kuckartz & Rädiker, 2022, S.148) gearbeitet. Die Inhalte der Kategorien werden dabei fallorientiert und aufbauend strukturiert und die Gewichtungen der genannten Themen berücksichtigt. Auch fehlende Aspekte können hervorgehoben werden. Das Ziel ist, die Transkripte qualitativ auszuwerten und dadurch gehaltvolle Ergebnisse zu erarbeiten. Vermutungen und Interpretationen dürfen in diesem Fall mit in die Analyse einfließen. Zudem sollen auch die Zusammenhänge zwischen den Hauptkategorien sowie den Sub- und Hauptkategorien beleuchtet werden.

Die letzte Phase der inhaltlich strukturierenden qualitativen Inhaltsanalyse beinhaltet die Verschriftlichung der herausgearbeiteten Erkenntnisse und die Beantwortung der zu Beginn gestellten Forschungsfragen. Die Ergebnisse der Analyseschritte werden zusammengeführt und können beispielsweise durch Hinzufügen von Tabellen und Grafiken zusätzlich veranschaulicht werden (Kuckartz & Rädiker, 2022).

Um die gewonnenen Erkenntnisse hinsichtlich ihrer wissenschaftlichen Güte zu beurteilen, muss sich der gesamte Prozess von der Datenerhebung bis zur Auswertung gewissen Gütekriterien stellen. Kuckartz und Rädiker (2022) unterscheiden dabei zwischen der internen und externen Studiengüte. Die interne Studiengüte beinhaltet Aspekte wie „Verlässlichkeit, Auditierbarkeit, Regelgeleitetheit, intersubjektive Nachvollziehbarkeit [und] Glaubwürdigkeit“ (Kuckartz & Rädiker, 2022, S.236) während der Datenerfassung und der Durchführung der qualitativen Inhaltsanalyse. Die externe Studiengüte dagegen fokussiert sich auf die Generalisierung und allgemeine Übertragbarkeit der ausgewerteten Daten (Kuckartz & Rädiker, 2022).

### 6.3 Durchführung der Experteninterviews

Aufgrund der in Kapitel 6.1 dargelegten Rahmendaten wird für diese Arbeit mit dem leitfadengestützten Experteninterview gearbeitet. Der Gesprächsaufbau des Leitfadens orientiert sich dabei an der Fachliteratur und beginnt mit der Begrüßung und Vorstellung des Experten (Berger-Grabner, 2022). Die anschließende Befragung umfasst in diesem Fall sechs zentrale Fragen. Eine Danksagung an den Experten rundet das Interview ab (Berger-Grabner, 2022). Der entwickelte Leitfaden enthält neben den Kernfragen auch mögliche Rückfragen, die kursiv aufgelistet sind. Diese werden lediglich dann gestellt, wenn sie in den Antworten des Experten nicht aufgegriffen werden. Dadurch soll den Befragten zunächst ein offenes und uneingeschränktes Beantworten der Themen ermöglicht werden, bevor bei Bedarf auf spezifische Aspekte eingegangen wird (Berger-Grabner, 2022).

Abschnitt	Inhalt	
<b>Vorarbeit</b>		
E-Mail an Experten	<ul style="list-style-type: none"> <li>- Vorstellung der eigenen Person (bei erstem Kontakt mit Experten)</li> <li>- Vorstellung des Interviewthemas</li> <li>- Einholen des Einverständnisses als Interviewpartner</li> </ul>	
<b>Eröffnung</b>		
Begrüßung	<ul style="list-style-type: none"> <li>- Bedanken für die Unterstützung und die Zeit des Experten</li> <li>- Einholen des Einverständnisses zur Audioaufzeichnung</li> <li>- Kurze Vorstellung des Experten</li> </ul>	
<b>Interview</b>		
Frage 1	Konnten Sie in die vorab geschickten Hörbeispiele Reinhören?	Gestellte Frage
Frage 2	Wie beurteilen Sie die Varianten eins und zwei und wie wirken diese auf Sie?  - <i>Was ordnen Sie der KI und was dem Original zu und warum?</i>	Gestellte Frage
		Mögliche Rückfrage
Frage 3	Worin sehen Sie die Stärken und Schwächen dieser, aber auch allgemein von generierten Filmmusikkompositionen?  - <i>Wie ordnen Sie im Gegensatz dazu eine von Menschen angefertigte Filmmusikkomposition ein?</i>	Gestellte Frage
		Mögliche Rückfrage
Frage 4	Wie beurteilen Sie das Verhältnis von Qualität und Nutzen von KI generierten (Filmmusik-)Kompositionen?  - <i>Kann eine KI vergleichbare Ergebnisse wie ein Komponist erzielen?</i> - <i>Zeitkomponente: Ist die Einbuße in der Qualität durch den Zeitgewinn vertretbar?</i>	Gestellte Frage
		Mögliche Rückfragen

	<ul style="list-style-type: none"> <li>- <i>Wie lange bräuchten Sie als Komponist, um einen Score für die 12 Minuten des Hörbeispiels zu erstellen?</i></li> </ul>	
Frage 5	Sehen Sie die KI als Risiko der als Chance für Ihre Branche und auch Ihre Berufsgruppe?	Gestellte Frage
	<ul style="list-style-type: none"> <li>- <i>Ist eine Zusammenarbeit mit den intelligenten Systemen möglich oder besteht ein Konkurrenzverhältnis?</i></li> <li>- <i>Haben Sie selbst schon mit etwaigen Softwares gearbeitet?</i></li> <li>- <i>Wenn ja, mit welchen und in welchem Kontext?</i></li> <li>- <i>Ändert die KI Ihre Rolle als Komponist/Mischer und welche Rolle spielt diese in ihrem derzeitigen Arbeitsalltag?</i></li> </ul>	Mögliche Rückfragen
Frage 6	<p>Komponist: Werden KI-generierte Kompositionen in der Zukunft die Arbeit des Filmmusikkomponisten ersetzen können?</p> <ul style="list-style-type: none"> <li>- <i>Wie begründen Sie diese Aussage?</i></li> </ul> <p>Filmmischtonmeister: Macht es für Sie einen Unterschied, ob das Musikmaterial von einer Person oder einer Maschine angeliefert wird?</p> <ul style="list-style-type: none"> <li>- <i>Wie begründen Sie diese Aussage?</i></li> </ul>	Gestellte Frage
<b>Abschluss</b>		
Verabschiedung	<ul style="list-style-type: none"> <li>- Bedanken für das Interview</li> <li>- Ausblick für das weitere Vorgehen</li> </ul>	

Abbildung 6.2 Leitfaden für die Experteninterviews

Es fanden sechs Experteninterviews statt, die alle mittels einer Videotelefonie über Zoom realisiert wurden. Die Befragten unterteilen sich in Komponisten und Filmmischtonmeister, wobei die nachfolgende Tabelle eine Übersicht zu deren Qualifikation als Experten und der Durchführung der Interviews gibt.

Experte	Abkürzung	Expertenqualifikation	Durchführung	Dauer
Filmmusikkomponist	FK1	25 Jahre Berufserfahrung als Filmmusikkomponist, 15 Jahre Berufserfahrung als Jazzmusiker	Zoom	34:29min
Komponist	FK2	Filmmusikkomponist für unterschiedliche Bewegtbildformate und Game-Musik, Entwicklung einer Orchestrationssoftware	Zoom	28:24min

Komponist	FK3	Mehr als 20 Jahre Berufserfahrung als Komponist für verschiedene Bewegtbildformate, Musiker	Zoom	27:18min
Mischttonmeister	FM1	Mehrjährige Berufserfahrung als Mischttonmeister sowie in der gesamten Audiopostproduktion verschiedener Bewegtbildformate	Zoom	23:07min
Mischttonmeister, Sound-designer, Supervising Sound-Editor	FM2	Mehr als 25 Jahre Berufserfahrung in der Tonpostproduktion, Mischung vieler (Kinospiegel-)Filme angefertigt	Zoom	33:09min
Sound-Editor, Sound-designer, Re-recording Mixer	FM3	Mehrjährige Berufserfahrung in verschiedenen Bereichen der Audiopostproduktion inklusive Filmmischungen, Tätigkeit als Head of Sound	Zoom	21:46min

Abbildung 6.3 Experteninterviews

Im Anschluss an die Datenerhebung werden die Interviews mittels des Tools TurboScribe transkribiert. Die Transkripte sind in der digitalen Anlage sieben zu finden. Die Auswertung der Daten erfolgt nun im nächsten Kapitel mittels der strukturierenden qualitativen Inhaltsanalyse nach Kuckartz und Rädiker.

#### 6.4 Auswertung der Experteninterviews

Die ersten sechs Schritte der Analyse wurden durchgeführt und sind in den Anlagen eins und zwei in Form von Fallzusammenfassungen sowie des Kodierleitfadens zusammengefasst. Die ausgewerteten Ergebnisse werden nachfolgend anhand von Kategorien dargelegt.

##### Evaluierung der generierten Filmmusik

Alle Experten konnten die beiden Hörbeispiele dem Original und der KI-Software zuordnen, wobei der Originalscore durchweg als qualitativ hochwertiger empfunden wurde. Die Evaluierungen der Komponisten fallen im Vergleich zu den Filmmischttonmeistern kritischer aus. Die KI misst sich mit Inception und dem Score von Hans Zimmer mit einer der ikonischsten Filmmusiken des vergangenen Jahrzehnts, was sowohl kompositorisch konzeptionelle Aspekte als auch den Klang angeht (FM3, FK2). Die generierte Filmmusik ist in ihren Ansätzen dem Original ähnlich, erreicht jedoch nicht dieselbe Wirkung. Für einen weniger epischen Film, beispielsweise eine romantische Komödie oder auch im TV-Umfeld wäre eine solche Musik jedoch denkbar (FK3). Der generierte Score verfügt über eine

Grundstruktur, die etablierte Hörkonventionen erkennbar macht und diese triggert (FK1). Dadurch werden mitunter Emotionen hervorgerufen (FK3), sodass vor allem die Action-Szenen funktionieren (FM1).

Jedoch reichen passende Harmoniefolgen nicht aus, um eine ebenbürtige Filmmusik zu erstellen (FK3). Es fehlt die musikalische Idee und durch die sich ständig wiederholenden Muster wirkt der Score repetitiv und teilweise maschinell (FM3, FK1). Trotz dessen, dass in der KI-Variante mehr Instrumente für die Kompositionsverläufe verwendet werden, nutzt diese für den Aufbau von Emotionen immer dieselben Mittel (FM2). Dadurch lässt die Wirkung nach und wird flach und eintönig. Es fehlt der Detailgrad sowie die kleinen organischen Variationen, die den Score lebendiger machen (FM3, FM1). Diese Aspekte führen dazu, dass die Musik die dramaturgischen Elemente nur bedingt bedient und einen generischen Anklang erhält. Auch durch die musikalischen Übergänge wird erkenntlich, dass die KI diesen Score nicht zusammenhängend erstellt hat (FK2). Es wird kritisiert, dass die generierte Filmmusik keine Verbindung zum Zuschauer herstellt und es folglich zu einer Distanzierung durch die Wahrnehmung der versuchten Intention des Scores kommt (FK1). Zwei Experten setzen die Qualität der generierten Kompositionen auch mit den Leistungen von weniger guten Komponisten gleich (FK1, FK2). Im Kontrast dazu wird angeführt, dass dieser Score das Bewegtbild eben anders bedient und entsprechend Emotionen entwickelt (FM1). So greift die Musik der Handlung weniger voraus als im Originalscore (FM3). Ein Experte sieht bei den KI-Systemen mehr Potenzial für Filmmusiken aus populärer Musik anstelle des klassischen orchestralen Scores, da diese klanglich einfacher und zielgerichteter umgesetzt werden können (FM3).

Ein weiterer wichtiger Faktor für die Beurteilung liegt in der Klangqualität der generierten Musik. Diese wird aufgrund der hörbaren virtuellen Instrumente verglichen zum Original als schwächer eingeordnet (FM3). Würden die Kompositionen von einem echten Orchester eingespielt werden, müssten Teile neu orchestriert werden, da in der generierten Version beispielsweise große Sprünge in den Streichern vorhanden sind, die so nicht spielbar sind (FK2). Durch den Verlust des speziellen Originalscore-Klangs gehen in der neuen Variante wichtige Emotionen verloren, die nicht anderweitig kompensiert werden können (FK3).

Bezüglich der Mischung wird deutlich, dass die Musik über deutlich weniger Dynamik verfügt und die Verhältnisse von Dialog und Musik weniger filigran sind. Zusätzlich führt die geringere Frequenzrange dazu, dass die Musik weniger Platz für die übrigen Soundelemente im Vergleich zum Original lässt (FM1, FM2). Daher funktioniert die generierte Mischung gut im TV, im Kino hingegen weniger, beim Originalfilm Inception ist dies genau umgekehrt (FM1). Die von AIVA generierte Mischung lässt vermuten, dass die Software mehr für Consumer-Devices und nicht für Kinoproduktionen und Filme ausgelegt ist.

Anzumerken ist, dass die Experten die Qualität der Komposition subjektiv nach dem bestehenden Wissenstand einordnen (FK3, FM3). Die Ergebnisse eines Hörversuchs mit durchschnittlichen, weniger audiophilen Personen, welche nicht primär auf die Musik achten, könnte daher aufschlussreiche Ergebnisse liefern.

Aktuell sind noch zentrale Unterschiede zwischen der Arbeit einer KI und der eines Komponisten(-teams) bei der Erstellung einer Filmmusik vorhanden. Ein Alleinstellungsmerkmal des Menschen ist, dass konzeptionelle, tief mit der Handlung verbundene Entscheidungen getroffen werden können, welche die Musik und die Dramaturgie miteinander verbinden (FK2). Bezüglich der Genauigkeit der Positionierung der Musik und der Ausarbeitung des Musikstils ist der Mensch der KI momentan noch deutlich überlegen (FM1). Dies überträgt sich auch auf die Mischung, welche in der verwendeten Software ausschließlich für die Musik und ohne Berücksichtigung weiterer Tonelemente angefertigt wurde (FM3). Ein Mensch hingegen stellt ein ausgewogenes und der Dramaturgie entsprechendes Verhältnis zwischen Dialog, Musik, Geräuschen und Bild her, was in den beiden Hörbeispielen deutlich wird (FM3). Zudem werden noch keine softwareinternen Surroundmischungen angeboten, die für Filmproduktionen jedoch notwendig sind (FM2).

Bei der Erstellung einer Filmmusik sind mehr als nur strukturelle Vorgaben zu erfüllen, eine Mehrdeutigkeit der Komposition bestimmt die Qualität maßgeblich mit (FK1). Diese entsteht während der Rezeption durch den Zuschauer und basiert auf dessen persönlichen, sozialen und kulturellen Erfahrungen. Aktuell werden den KI-Softwares gewisse Kompositionsweisen antrainiert, es fehlt jedoch eine Grundlage, um diese Bedeutungsebene aufzuspannen und um bedeutungstragende musikalische Filmthemen zu etablieren (FK1).

Durch den Einsatz einer generativen KI ändert sich neben der Arbeitsweise auch die Intention des Projekts (FM2). Wer hält die Entscheidungskontrolle inne und mit wem soll interagiert werden sind dabei zentrale Fragen. Oftmals ergibt sich ein Mehrwert aus der Kommunikation und der Zusammenarbeit zwischen Menschen und auch die Beschränkung der verfügbaren Ressourcen wie Zeit und Produktionsmittel können die Kreativität fördern, sodass Kunst entsteht (FM2). Einer KI hingegen stehen immer dieselben Mittel zur Verfügung. Die wahrgenommene Musik ruft bei den Zuschauern Emotionen hervor und provoziert Gedankenprozesse. Diese wurden während des Herstellungsprozesses ebenfalls von dem Komponisten selbst durchlebt, sodass eine Verbindung zwischen diesem und dem Zuschauer entsteht (FK2). Eine KI empfindet keinerlei Emotionen und kann daher keinen Bezug zum Zuschauer aufbauen, da das entstandene Produkt nicht mit dem menschlich komponierten vergleichbar ist (FK2). Die Kommunikation stellt insgesamt einen entscheidenden Unterschied für die gesamten Produktionsprozess dar. Wie werden beispielsweise Revisionsschleifen von einer KI umgesetzt und welche Begrifflichkeiten können dazu genutzt werden? Menschen

können auf geäußerte Kritiken aufgrund ihrer langjährigen Erfahrungen passend reagieren und die notwendigen Änderungen vornehmen, bei einer KI funktioniert dies nicht (FM2).

Allerdings wird der Mensch auch mit einer humanen KI verglichen, denn Menschen nutzen Erlerntes und Gehörtes, um anschließend dieses Wissen zu verknüpfen und daraus etwas Neues und Eigenes zu erschaffen (FK3). Das wahrgenommene und verarbeitete Material dient im Grunde als Trainingsdaten, welche die weitere Arbeit beeinflussen. Vereinfacht sind dies elektronische Vorgänge, welche das Gehirn steuern und somit ebenfalls einer KI antrainiert werden können (FK1). Zudem spielen auch zufällige Kombinationen während des Komponierens eine Rolle, besonders wenn verschiedene Perspektiven aufeinandertreffen (FK3, FM2).

Der Begriff Perfektionismus wird oftmals mit der KI in Verbindung gebracht und steht im Kontrast zur Kreativität. Laut einer weitergegebenen Erfahrung sind es die individuellen Fehler, welche die Handschrift der Regisseure und auch der Komponisten letztendlich ausmachen (FM2).

#### Möglicher Nutzen sowie Chancen und Risiken durch KI (Kompositions-) Tools

Der Nutzen einer KI-Software für die Erstellung einer Filmmusik wird als gering eingeordnet. Vielmehr kann diese als Inspirationsquelle eingesetzt werden (FK1, FK3, FM1). Bei der Erstellung von Temp-Musiken hingegen, wird eine Möglichkeit zur Nutzung dieser Kompositionstools gesehen. Durch die Generierung eines neuen Tracks müssen auf keine bekannten Filmmusiken zurückgegriffen werden, die wiederum Assoziationen hervorrufen und Erwartungen schüren (FM1). Zudem wird dem Konflikt entgangen, Material aus Libraries zu nutzen, welches final nicht verwendet werden darf (FM3). Durch die vielseitigen Generierungen der KI können auch Lücken in den Archiven abgedeckt werden, um passendes Material für das bestehende Projekt zu finden (FM2). Der Zeitgewinn durch die Nutzung einer KI bringt aus Sicht der Komponisten keinen Vorteil, da das Ergebnis qualitativ nicht dasselbe bleibt. Eine Beschleunigung des Herstellungsprozesses ist wünschenswert, jedoch können die menschlichen Gedankengänge und Gefühle nicht beliebig beschleunigt werden (FK2). Die Erstellung der generierten Filmmusik hat mit AIVA insgesamt etwa eine Woche beansprucht. Im Vergleich dazu bräuchten die befragten Komponisten zwischen zwei bis drei Tagen und einer Woche für einen ersten musikalischen Entwurf der Filmsequenz (FK3, FK2).

Aktuell liegt der Mehrwert von KI-gestützten Programmen in anderen Bereichen der Audiopostproduktion. Besonders die verbesserten und weitreichenden Möglichkeiten der O-Tonbearbeitung wurden mehrfach hervorgehoben (FM3, FM2). Dazu gehören Aspekte wie die Trennung von Sprache und Noise oder auch die Reduktion von Hall-Anteilen. Durch die KI wird heutzutage Material mischbar, das vor ein paar Jahren noch komplett durch Vollsynchron hätte ersetzt werden müssen (FM2).

Durch die KI eröffnen sich auch Chancen für Muskschaffende. So wird Komponisten dadurch ermöglicht, sich aus den eigenen Gedankenschleifen zu befreien und neue Impulse durch die KI zu erhalten, die anschließend eigenständig weiterentwickelt werden (FK1, FM1). Die Technologie steht in diesem Bereich noch in den Anfängen. Zukünftige Erweiterungen der Funktionalität der bestehenden Softwares und explizite Spezialisierungen auf die Filmmusik, zum Beispiel durch Timecode-Angaben, könnten nicht nur die Zusammenarbeit verbessern, sondern auch zu hochwertigeren Ergebnissen führen (FM3). Ein Experte sieht gerade bei der filmmusikalischen Untermalung von Genrefilmen die Chance, einer KI die entsprechenden Hör- und Erzählkonventionen anzutrainieren (FK1). Dadurch, dass Musikrechte sehr teuer geworden sind und bei den Produktionen nicht ausreichend Budget für bestimmte Tracks vorliegt, können dafür künstlich generierte, aber ähnlich klingende Musiken als kostengünstige Alternative verwendet werden (FM3). AIVA bietet diese Funktionalität bereits an, aufgrund mangelnder Ergebnisse wurden diese jedoch im Rahmen der Arbeit nicht verwendet.

Die Entwicklungen der Branche selbst stellen auch vielseitige Chancen für die KI dar. Im Zusammenhang mit den steigenden Streamingangeboten der letzten Jahre wird zunehmend Content produziert, wodurch die Nachfrage an originell produzierter Musik steigt (FM1). Dadurch dass die herausragenden, zeitlich überdauernden Musikthemen an Seltenheit gewinnen (FK1), eröffnet sich eine Chance für die künstlichen Generierungen. Die Nachfrage und der Spielraum für weitere Entwicklungen sind aktuell gegeben, wobei seitens der Komponisten Interesse an Notations-Softwares oder an der klanglichen Aufwertung von eigens erstellten Entwürfen besteht (FK2). Die tatsächlichen Large Language Models, welche unveränderliche Audiofiles basierend auf Orchesteraufnahmen generieren, können bereits heute klanglich mehr überzeugen als die von AIVA genutzte Sample Engine (FK2). Die Bereitstellung von kreativer Zusammenarbeit mit diesen Large Language Models könnte daher einen tatsächlichen Mehrwert für die Komponisten mit sich bringen. Ein von den Experten hervorgehobener Aspekt steht nicht in direktem Zusammenhang mit der Musik, sondern betrifft den Bereich der Synchronisation. Durch die Entwicklungen in der Sprach- und Stimmengenerierung könnten zukünftig Filme unabhängig der Produktionssprache mit den Originalstimmen versehen werden (FM2). Dies würde jedoch zu einer ernsthaften Bedrohung der Synchronbranche führen, wodurch eine direkte Verbindung zu den einhergehenden Risiken durch die KI-Systeme hergestellt wird.

Die Frage inwiefern die KI nützlich ist und zugleich Arbeit wegnimmt, stellt einen zentralen Risikofaktor dar. Aktuell verfolgen die KI-Kompositionstools nicht den Ansatz Komponisten bei ihrer Arbeit zu unterstützen, sondern sie versuchen diese zu ersetzen (FK2). Die befragten Experten halten dies für problematisch, da es keine Bestrebungen geben sollte, gesamte Berufsfelder an die KI zu übergeben und sehen dadurch auch den künstlerischen Mehrwert

von Kompositionen gefährdet (FM2, FM3). Sobald sich die Qualität der Generierungen insoweit verbessert, dass sie im professionellen Kontext eingesetzt werden können, wird der Druck auf die Komponisten steigen. Sie müssen in kürzerer Zeit mehr Musik kreieren, wobei sich auch die Gagen verändern werden (FM1). Für Komponisten ist es demnach wichtig die Entwicklungen im Blick zu behalten und diesen offen zu begegnen.

Die Entwicklung praktischer Anwendungen, wie beispielsweise die automatisierte Umwandlung einer Midi-Session in eine Partitur, würden gewissen Personengruppen während der Arbeit einen Nutzen bringen. Gleichzeitig würden diese Implementierungen die Aufgaben und damit geldbringende Jobs von Personen übernehmen und diese überflüssig machen (FM3, FK2). Daher stellt sich die Frage, ob diese Entwicklungen weiter vorangetrieben und unterstützt werden sollten. Eine Abwägung zwischen wirtschaftlichen Geldeinsparungen und dem möglichen Verlust von künstlerisch handwerklichen Berufsgruppen muss vollzogen werden (FK1).

Neben diesen Aspekten führen die noch ungeklärten rechtlichen Grundlagen bezüglich des Umgangs mit dem Urheberrecht zu weiteren Herausforderungen (FM1). Zudem wird sich das Hörverhalten der Menschen durch das Vordringen von generierter Musik in deren Alltag verändern. Es besteht das Risiko, dass viele Menschen zukünftig nicht mehr zwischen künstlich generierten und echten Kompositionen unterscheiden können (FM3, FK1). Diese Entwicklungen werden kommen, sodass bereits heute ein Umgang zur Eindämmung der Risiken und einer nachhaltigen Nutzung entscheidend ist (FM2).

#### Erfahrungen mit Kompositionssoftwares und deren Relevanz im Arbeitsalltag

Die Relevanz und der Gebrauch von KI-Kompositionssoftwares hält sich im Arbeitsalltag der Experten aktuell in Grenzen. Ein ausschlaggebender Grund dafür ist die unzureichende Qualität der Generierungen, sodass diese nicht auf professioneller Ebene genutzt und weiterverwendet werden können (FK2, FK3). Dadurch besteht noch keine zwingende Notwendigkeit, sich intensiv mit diesen Tools auseinanderzusetzen (FK3). Zudem bringt der Einsatz von KI für die Komponisten aktuell keinen Mehrwert während des Arbeitsprozesses mit sich (FK1). Die generativen Tools scheinen sich noch nicht im professionellen Kontext etabliert zu haben, denn beispielsweise werden die klassischen Musikarchive den Softwares noch vorgezogen (FM3). Jedoch wird mit einem Vordringen der KI in diesem Feld gerechnet. In anderen Bereichen hingegen, wie in der Sprachbearbeitung, im Sounddesign oder auch bei der Entwicklung von virtuellen Instrumenten spielen die KI-gestützten Programme bereits eine wichtige Rolle im Arbeitsalltag und werden regelmäßig eingesetzt (FK2, FM1, FM3, FM2).

Die persönliche Haltung gegenüber diesen Technologieentwicklungen wirkt sich ebenso auf die Nutzung und den Umgang mit diesen Tools aus. In der Praxis interessieren sich einige Komponisten für die Arbeit mit KI, während andere das Konzept der künstlichen

Generierungen kategorisch ablehnen (FK3). Diese Beobachtungen stehen in engem Zusammenhang mit den Erfahrungen der Experten bezüglich der Kompositionssoftwares. Teilweise liegen noch keine Berührungspunkte vor oder es besteht nur eine ungefähre Übersicht über die verfügbaren Programme. Dies wird jedoch nicht anhand mangelnden Interesses begründet, sondern aufgrund fehlender zeitlicher Kapazitäten (FM1, FM3). Aus Sicht der Mischtonmeister geben weiterhin die Zusammenarbeit und die Verfügbarkeit von Komponisten keinen Anlass dafür, sich mit den Kompositionssoftwares tiefergehend zu beschäftigen. Bei den Komponisten werden die Entwicklungen verfolgt und untereinander diskutiert (FK2). Diverse Programme wurden bereits getestet, darunter Udio und Suno, die mehrfach genannt wurden. Mit Udio wurde beispielsweise eine Idee weitergeführt, wobei das Verhältnis von interessanten Entwürfen und entstandenem Ausschuss nicht überzeugen konnte. Die Anzahl der verwertbaren Ergebnisse ist noch zu gering, weswegen sich diese Tools nur zum Experimentieren eigenen oder als Inspirationsquelle eingesetzt werden (FK3). Auch wenn die Nutzung aktuell noch begrenzt ist, wird erwartet, dass die Generierungstools in der Zukunft an Einfluss gewinnen und die Zusammenarbeit mit diesen selbstverständlich werden wird (FK1). Daher sollte bereits heute eine Auseinandersetzung mit diesen stattfinden, um einen adäquaten Umgang zu finden (FM1).

#### Zukünftige Entwicklungen und die mögliche Ersetzung des Komponisten durch die KI

Die Entwicklungen der letzten Jahre haben die Implementierung generativer Musiksoftwares ermöglicht, die bereits heute beachtliche Ergebnisse erzielen können (FK3). Die Experten sind sich einig, dass die Entwicklungen mindestens so rasant fortschreiten werden, sodass sich die Qualität der Generierungen zunehmend steigern wird. Dazu gehören die handwerklichen Aspekte, wie zum Beispiel musikalische Übergänge zwischen zwei Tracks, aber auch die klanglichen Komponenten (FK1, FK2). Eine gewisse Limitierung bleibt jedoch durch die begrenzt verfügbaren Trainingsdaten (FK2). Der Musikstil wirkt sich ebenfalls auf die hervorgebrachte Klangqualität der Generierungen aus. Bei elektronischer Musik wird die KI aufgrund der künstlichen Herstellung einfacher überzeugende Ergebnisse erzielen, während sich die Umsetzung akustischer und organischer Musik deutlich aufwändiger und schwieriger gestalten wird (FM3). Um schon heute einen hochwertigeren Klang einer KI-Komposition zu erhalten wird empfohlen, diese von echten Instrumenten einspielen zu lassen (FK3). Neben den technischen und kompositorischen Aspekten wirken noch eine Reihe weiterer außermusikalischer Faktoren mit, welche den zukünftigen Einsatz von KI und die Stellung der Komponisten beeinflussen werden.

Besonders die Filmmusik zeichnet sich durch deutlich mehr als die reine Notenfolge aus. Dadurch, dass die Musik nicht eigenständig ist und einem definierten Zweck unterliegt, setzt sich diese nur zur Hälfte aus musikalischen Kriterien zusammen. Die anderen 50% ergeben

sich aus der Kommunikation zwischen Komponisten und Regisseur, dem Filmverständnis und dem Zusammenwirken von Ton und Bild (FK1). Dieser ineinander verflochtene Herstellungsprozess wird durch viele Entscheidungen gesteuert, welche durch die menschlichen Erfahrungen, Assoziationen und Sinne geleitet werden (FK1). Es stellt sich die Frage, auf welcher Basis die KI diese Aspekte nachbilden bzw. integrieren kann. Ist es möglich, einer KI dieses Wissen zukünftig anzutrainieren? Die bisherigen Kommunikationsmöglichkeiten reichen bei Weitem nicht, um dies ebenbürtig umsetzen zu können, wie durch die generierte Filmmusik deutlich wurde (FK1). Aus diesem Grund sind sich die Experten sicher, dass diese menschliche Verbindung maschinell nicht ersetzt werden kann. Zudem steht die KI vor der Herausforderung einen ca. 60 bis 70-minütigen zusammenhängenden Score zu erstellen. Die dazu erforderliche Rechenleistung ist erheblich und erscheint unrealistisch (FK2), denn die maximale Trackdauer der genutzten Software liegt derzeit bei 5:30 Minuten. In der nahen Zukunft wird die KI demnach nicht in hochqualitativen Film-, Image- oder Werbefilmbereichen mitwirken können (FK2, FM1).

Aus Sicht der Mischtonmeister macht es unter bestimmten Voraussetzungen keinen Unterschied, ob das Material von einer Person oder einer KI angeliefert wird (FM1, FM3). Sofern die musikalische Qualität vorhanden ist und das Material in Stems vorliegt, kann in derselben Weise gearbeitet werden. Der Wegfall der vorherigen Absprachen bezüglich des Zusammenspiels von Musik und Sounddesign, wird jedoch als negativ betrachtet (FM3). Die pünktliche Anlieferung des Materials wiederum, kann gegenüber den Verzögerungen bei der Zusammenarbeit mit Personen als Vorteil gewertet werden (FM3). Da aus Budgetgründen bereits heute die Komponisten immer seltener bei der Endmischung mit dabei sind, stellt dies keinen Unterschied dar (FM1). Von Bedeutung ist nur, dass es sich weiterhin um Arbeitsmaterial handelt, sodass während der Mischung Cues verändert und angepasst werden können (FM2, FM3). Sobald dies nicht mehr gegeben ist, sollte die Mischung ebenfalls von einer KI übernommen werden. Damit beginnt jedoch die Grundsatzdiskussion, was erreicht werden soll und wer die Entscheidungen trifft, von Neuem (FM3).

Wird nun isoliert die Situation der Komponisten betrachtet, dann stellen die KI-Systeme durchaus eine Bedrohung dar. Gerade in Library Musik, in weniger qualitativen Formaten wie Scripted Reality Shows und Telenovelas oder auch in Online-Content, Werbung und Marketing wird damit gerechnet, dass die KI vordringen und in weiten Teilen Aufgaben übernehmen wird (FK2). Auch die Synchronbranche ist potenziell durch die KI gefährdet (FM2). Durch die mögliche Reduktion bzw. den Wegfall dieser Berufsfelder wird folglich die Konkurrenz in anderen Bereichen aufgrund der Umorientierung der Komponisten steigen (FK2).

Es werden jedoch auch wirtschaftliche Gründe genannt, die das Weiterbestehen von Komponisten unterstützen. Sender und Verlage generieren Einnahmen, indem sie den Komponisten die Verlagsrechte für die erstellten Musiken abkaufen und anschließend die entstandenen Kosten über Verwertungsgesellschaften wie GEMA wieder einspielen. Ein eventueller Wegfall dieser Einnahmen durch die Nutzung von generierter Musik ist nicht vorgesehen (FK2). Zudem wird die Bedeutung der Mensch zu Mensch Interaktion hervorgehoben. Der Bezug zu Interpreten eröffnet eine außermusikalische Verbindung, die als Alleinstellungsmerkmal gegenüber der KI eingeordnet wird (FK3). Damit einher geht die Entwicklung des Hörverhaltens und der Hörästhetik der Menschen. Analog zu den Innovationen der Vergangenheit wird sich dieses den neuen Bedingungen anpassen, wobei die Haltung der Menschen gegenüber den künstlich generierten Musiken über den zukünftigen Einfluss der KI mitentschieden wird (FM2).

Eine persönliche Gefahr empfinden die befragten Experten derzeit nicht, da sie sich mit ihren Kompositionen in der Branche etabliert haben und über ein ausgedehntes Wissen verfügen (FK2, FK3). Gestützt wird diese Aussage durch das Qualitätsverhältnis zwischen den Generierungen und den menschlich erbrachten Leistungen (FK3). Zusammenfassend wird die KI den Filmmusikkomponisten vorerst nicht ersetzen können, jedoch wird sie in anderen Bereichen und Formaten in der Zukunft die Aufgaben von Personen übernehmen (FK2). Daher ist es von Bedeutung mit den Entwicklungen mitzugehen, um nicht abgehängt zu werden (FK1). Es bedarf zudem an Möglichkeiten, wie diese Tools sinnvoll als Werkzeug und nicht als Methode zur Gewinnsteigerung oder Ersetzung von Jobs eingesetzt werden können (FM2).

## 7. Schlussbetrachtung

### 7.1 Zusammenfassende Beantwortung der Fragen

Das Ziel dieser Arbeit war es, eine Filmmusik für eine Szene aus Inception mittels einer KI-Software zu erstellen und dessen Qualität mit dem Originalscore zu vergleichen. Damit sollen der aktuelle Entwicklungsstand sowie die Möglichkeiten und Grenzen der KI-Systeme im Bereich der Filmmusikkomposition aufgezeigt werden. Für dessen Umsetzung wurden zu Beginn drei Teilfragen formuliert, die in diesem Kapitel abschließend aufgegriffen und zusammenfassend beantwortet werden.

- ◆ „Worin liegen die Unterschiede einer von Menschen gestalteten bzw. durch KI generierten Filmmusik?“

Für die Beantwortung der ersten Teilfrage wurde zunächst der Score von Hans Zimmer analysiert. Dazu wurde die zwölfminütige Szene in drei dramaturgische Abschnitte unterteilt und anschließend die Bild- und Tongestaltung sowohl einzeln als auch in deren Zusammenwirken untersucht. Auch die zentralen filmmusikalischen Themen und deren Bedeutungsumfang wurden herausgearbeitet. Während der Analyse wurde die enge Verbindung des musikalischen Konzepts mit der Dramaturgie deutlich. Die Musik gestaltet den Handlungsverlauf mit und spannt eine zusätzliche, rein auditive Bedeutungsebene auf. So werden Musikpassagen aus vorherigen Ereignissen der Filmhandlung erneut aufgegriffen und die Leitmotivtechnik in ihrer ausgeprägten Form eingesetzt. Die Filmmusik von Hans Zimmer interagiert mit dem Sounddesign, sodass ein dynamisches Konstrukt entsteht. Sie gibt den weiteren Tonelementen ausreichend Platz, nimmt aber auch an anderen Stellen fast den gesamten Raum ein. Die klangliche Ausarbeitung mit den dominierenden Blechbläsern, den weichen Streicher-Flageolets und den synthetischen Elementen führen zu einem besonderen Klangcharakter, der diesen Film definiert.

Die aus der Analyse gewonnenen Ergebnisse wurden als Anhaltspunkte verwendet, um im nächsten Schritt mittels der KI-Software AIVA einen ebenfalls klassischen orchestralen Score für dieselbe Szene zu erstellen. Die Gestaltung wurde dabei gänzlich von der KI übernommen, lediglich über die Wahl der Generierungsmethode und die entsprechenden Eingabeoptionen konnte Einfluss auf die Komposition genommen werden. Wie genau die KI arbeitet und woran sich diese bei den jeweiligen Generierungen orientiert wird jedoch nicht offengelegt. Während der Erstellung der Filmmusik wurde deutlich, dass in AIVA noch keine Möglichkeit besteht eine bestimmte Kompositionstechnik zu wählen, sodass die Arbeit mit Leitmotiven ausgeschlossen ist. Stattdessen wird versucht, mit der Musik die vorliegende Stimmung einzufangen, wobei die Kompositionen durchweg auf repetitiven Mustern basieren. Interessanterweise nutzt

Zimmer für den Aufbau bestimmter Scoretracks in Inception ebenfalls sich wiederholende Figuren. Der direkte Vergleich dieser Musiken zeigt jedoch die Überlegenheit und die kompositorische Stärke von Zimmer, welcher die KI noch deutlich unterlegen ist. Bezüglich der Klanggestaltung verfügt AIVA über einen gewissen Instrumentenvorrat, der implementiert wurde. Bei diesem handelt es sich um virtuelle Instrumente, die nicht weiter modifizierbar sind und in ihrer Klangqualität gegenüber echten Instrumenten abfallen. Ein großer Unterschied bei der Gestaltung in AIVA ist, dass kein Bildmaterial als Eingabe verwendet werden kann und dadurch die KI unabhängig vom Bewegtbild arbeitet. Folglich fehlt die Präzision bei der musikalischen Untermalung und auch die übergeordnete Bedeutungsebene kann nur in geringem Umfang kreiert werden. Letztendlich ändert sich durch die Arbeit mit einer KI der gesamte Herstellungsprozess, was zu der zweiten Teilfrage überleitet.

- ◆ „Wie organisiert sich die Zusammenarbeit zwischen Menschen und KI-Software und welche Auswirkungen ergeben sich auf den Herstellungsprozess?“

Durch die Arbeit mit einer KI-Software, in diesem Fall AIVA, ändert sich der Kompositionsprozess grundlegend. Die Entscheidungsfreiheit und Kontrolle werden an die Software übertragen, wobei lediglich die Rahmenbedingungen festgelegt werden können. Somit findet keine Unterstützung bei der Komposition durch die KI statt, sondern wird durch diese gänzlich übernommen. Um möglichst viele Angaben bezüglich der musikalischen Gestaltung zu treffen, wurden die sechs verschiedenen Musikstücke in dieser Arbeit mittels der Schritt-für-Schritt-Methode generiert. Es bestand keine Möglichkeit, den Score zusammenhängend anzufertigen, da AIVA eine maximale Musiklänge von 5:30 Minuten anbietet und die Kompositionen selbst in ihrem Verlauf nur geringfügig variierbar sind. Die Verbindung mehrerer Kompositionen in einer Session ist nicht möglich und gestaltet sich folglich als schwierig. Der in wenigen Sekunden bis Minuten generierte Vorschlag kann entweder angenommen oder verworfen werden. Etwaige Nachbearbeitungen durch den Nutzer sind im Editor möglich, um beispielsweise die musikalische Struktur an den Handlungsverlauf anzupassen. Allerdings besteht keine Möglichkeit, Einfluss auf den Klang oder bestimmte Passagen durch Timecode-Angaben zu nehmen. Daher gestaltet sich die musikalische Untermalung in der bearbeiteten Filmszene grobflächiger und steht nicht immer im Bezug zum Bewegtbild. Die Zusammenarbeit mit der Software ist demnach begrenzt möglich. Es können Anforderungen vorgegeben werden, jedoch zeigt sich erst nach der Generierung ob und inwiefern diese eingehalten und umgesetzt wurden. AIVA ermöglicht eine vergleichsweise hohe Einflussnahme zu anderen KI-gestützten Softwares, verfügt dafür aber nicht über dasselbe Klangniveau wie die Large Language Models. Für die Erstellung einer gehaltvollen Filmmusik sind daher noch weitere Implementierungen in den Softwares nötig,

wie unter anderem eine Erweiterung der Eingabemöglichkeiten sowie ausgedehnte Bearbeitungswerkzeuge, die eine tatsächliche Zusammenarbeit ermöglichen.

- ◆ „Wie wird die Übernahme kreativer Aufgaben durch die generative KI und die Qualität der Erzeugnisse von Komponisten und Filmmusikern empfunden?“

Für die Beantwortung der dritten Teilforschungsfrage wurden sechs Experteninterviews mit jeweils drei Komponisten und Filmmusikern durchgeführt. Um die Qualität der generierten Filmmusik einzuordnen, wurden die Originalfilmszene sowie die alternative Vertonung bereits im Voraus den Experten zugeschickt. Aus diversen kompositorischen und klanglichen Gründen, die in den vorherigen Abschnitten erwähnt wurden, wird das generierte Material verglichen mit dem von Hans Zimmer komponierten Score als schwächer bewertet. Jedoch werden die Erzeugnisse der KI-Systeme durch die rasanten Fortschritte zunehmend besser, sodass diese zukünftig in der Lage sein werden, qualifizierte Jobs beispielsweise in der Library Musik zu übernehmen. Im hochqualitativen Bereich sehen die Experten für die KI allerdings keine Chance sich zu etablieren und die Arbeit von Komponisten zu ersetzen. Aus deren Sicht entsteht durch diese Softwares kein kreativer Mehrwert, sondern vielmehr geht das künstlerische Handwerk durch den Einsatz von KI verloren.

Insgesamt ist bereits heute im Jahr 2024 vieles mit künstlich generierten Kompositionen möglich, auch wenn diese im Bereich der Filmmusik derzeit auf ihre Grenzen treffen. Dies liegt mitunter daran, dass die Musik in Filmen nicht als eigenständiges Element fungiert, sondern primär dem Bewegtbild dient und durch viele außermusikalische Faktoren beeinflusst wird. Aktuell können diese komplexen Zusammenhänge der KI noch nicht antrainiert werden, sodass keine enge Verflechtung der Bild- und Tonebene erreicht werden kann. Die Entwicklungen stehen in diesem Bereich jedoch noch am Anfang, sodass in den nächsten Jahren richtungsweisende Fortschritte zu erwarten sind.

## **7.2 Kritische Würdigung**

Im Rahmen dieser Arbeit wurden mehrere Methodiken angewandt, um die untersuchte Thematik unter einer forschenden und wissenschaftlichen Perspektive zu beleuchten. So wurde die filmmusikalische Analyse mittels der erweiterten Abdeckmethode von Chion und Neumeyer durchgeführt, welche als eine anerkannte Verfahrensweise diesbezüglich gilt (Chion, 2012). Bei der Erstellung der künstlich generierten Filmmusik wurde das Vorgehen deskriptiv dargelegt, wobei das experimentelle und forschende Umfeld im Fokus stand. Der Autorin ist bewusst, dass es sich dabei um eine aktuelle Wiedergabe des Entwicklungsstandes handelt. Voraussichtlich werden die beobachteten Erkenntnisse bereits in den nächsten sechs

bis zwölf Monaten nur noch in Teilen verwertbar sein, da sich dieser Bereich der generativen KI ständig weiterentwickelt.

Die durchgeführten Experteninterviews sollen wissenschaftlich verwertbar sein. Dazu wurde in Kapitel 6.2 die Methode der qualitativen Inhaltsanalyse erläutert und die entsprechenden Gütekriterien vorgestellt. Abschließend kann nun die Bewertung dieser erfolgen. Die von Kuckartz und Rädiker erhobenen internen und externen Kriterien sind in der nachfolgenden Tabelle zusammenfassend aufgeführt und werden auf Erfüllung geprüft. Die internen Gütekriterien wurden umfassend erfüllt, das externe Kriterium hingegen nicht.

Gütekriterium	Argumentation bei Prüfung auf Erfüllung	Erfüllung
Intern: Verfahrens- dokumentation	<ul style="list-style-type: none"> <li>- Erläuterung der Vorgehensweise (Kap. 6.1 und 6.2)</li> <li>- Darstellung der befragten Experten, des Leitfadens und der Durchführung der Interviews (Kap. 6.3)</li> <li>- Auswertung der Interviews (Kap. 6.4)</li> <li>- Verwendung von Originalzitat, Verfassen von Memos</li> </ul>	<ul style="list-style-type: none"> <li>✓</li> <li>✓</li> <li>✓</li> <li>✗</li> </ul>
Intern: Auditierbarkeit	<ul style="list-style-type: none"> <li>- Fixierung der Daten mittels einer Audioaufnahme</li> <li>- Durchführung vollständiger Transkription und Offenlegung des Transkriptionsprozess (Kap.6.3), Anonymisierung der Daten</li> <li>- Weitere interviewbegleitende Dokumentation</li> </ul>	<ul style="list-style-type: none"> <li>✓</li> <li>✓</li> <li>✗</li> </ul>
Intern: Regelgeleitetheit	<ul style="list-style-type: none"> <li>- Leitfaden gestützte Experteninterviews</li> <li>- Anwendung einer definierten Auswertungsmethode nach Kuckartz und Rädiker</li> </ul>	<ul style="list-style-type: none"> <li>✓</li> <li>✓</li> </ul>
Intern: Zuverlässigkeit und intersub- jektive Nach- vollziehbarkeit	<ul style="list-style-type: none"> <li>- Kategorienbildung basierend auf dem Ansatz zur inhaltlich strukturierenden qualitativen Inhaltsanalyse nach Kuckartz und Rädiker, Ausarbeitung eines Kodierleitfadens mit Definitionen und Beispielen</li> <li>- Auswertung der Ergebnisse auf Basis der vorgenommenen Codierungen der Haupt- und Subkategorien</li> </ul>	<ul style="list-style-type: none"> <li>✓</li> <li>✓</li> </ul>
Intern: Glaubwürdig- keit	<ul style="list-style-type: none"> <li>- Unabhängige Bearbeitung des Materials durch mehrere Codierende</li> <li>- Berücksichtigung aller erhobener Daten (inklusive Extremfälle und Abweichungen)</li> <li>- Ausschließliche Befragung von Experten, die aufgrund ihrer professionellen Tätigkeit über Wissen in diesem Feld verfügen</li> </ul>	<ul style="list-style-type: none"> <li>✗</li> <li>✓</li> <li>✓</li> </ul>
Extern: Verallgemeinerung und Übertragung	<ul style="list-style-type: none"> <li>- Einfangen des aktuellen Stimmungsbildes, spiegelt eine Momentaufnahme wider, keine Übertragung vorgesehen</li> </ul>	<ul style="list-style-type: none"> <li>✗</li> </ul>

Abbildung 7.1 Prüfung der internen und externen Gütekriterien  
In Anlehnung an (Kuckartz & Rädiker, 2022, S.237-238)

Im Hinblick auf den fachlich und zeitlich begrenzten Umfang können die erfüllten Gütekriterien als angemessen bewertet werden. Eine Prüfung der Ergebnisse durch weitere Experten wäre grundsätzlich sinnvoll, um eine wissenschaftliche Relevanz zu erreichen. Die fehlende Übertragbarkeit und Verallgemeinerung mindern in diesem Fall die Qualität der Analyse nicht,

da sie laut Kuckartz und Rädiker (2022) ein optionales Kriterium ist. Im Rahmen dieser Arbeit würde eine Generalisierung der Ergebnisse keinen Nutzen bringen, da es sich bei den Experteninterviews um eine subjektive Momentaufnahme der aktuellen Situation handelt, die sich aufgrund der schnelllebigen Entwicklungen in den nächsten Monaten und Jahren mehrmals verändern wird.

### **7.3 Ausblick**

In den nächsten Jahren werden die technologischen Entwicklungen weiter voranschreiten, sodass sich die KI wie bereits in anderen Fachbereichen auch nachhaltig auf die Musikproduktion und -industrie auswirken wird (Berens & Bolk, 2023). Sie besitzt das Potenzial Branchen zu verändern, die auf Kreativität, Innovation und Wissensverarbeitung angewiesen sind (Feuerriegel et al., 2023). Bereits heute können durch KI-gestützte Softwares verschiedene Musikstücke komponiert werden. Es sollte dabei lediglich definiert werden, welche Art der Musik dies impliziert und wie der Begriff Komposition definiert wird. Denn laut Heintz (2024) gibt es „nicht eine Musik, sondern viele Musiken; sowohl bezogen auf ihr Material, also ihre Klänge und Verbindungsmöglichkeiten, als auch bezogen auf den Kontext, in dem sie komponiert und rezipiert wird.“ (S.246).

Im Bereich der Filmmusik wird aufgrund vorgegebener Parameter und den vordefinierten Wirkungen weiterhin eine Chance für KI-generierte Filmmusik gesehen (Heintz, 2024). Unterstützend kommt hinzu, dass diese Musik nicht isoliert wiedergegeben und während des Films im Idealfall nur unbewusst wahrgenommen wird (Bullerjahn, 2001). Aktuell können die KI-Systeme Tempo-, Instrumenten-, und Stimmungsangaben verarbeiten und gewisse Stilimitationen erzeugen. Wie jedoch in dieser Arbeit deutlich wurde, reichen diese Komponenten noch nicht aus, um eine überzeugende Filmmusik zu generieren. Um einen zu dem Bewegtbild passenden und über die Klischees hinausreichenden Score zu entwickeln, bedarf es an weiteren spezifischen Anpassungen. Diese könnten zeitnah zu beträchtlichen Fortschritten führen, sodass die derzeit auftretenden Irritationen während des expliziten Hinhörens verschwinden würden (Heintz, 2024). Es ist daher wichtig, stets die Entwicklungen im Blick zu halten und sich mit den aufkommenden Technologien auseinanderzusetzen (Berens & Bolk, 2023).

Zudem stellt sich die Frage, ob aktuell ein adäquater Ansatz bei der Entwicklung der Kompositionssoftwares verfolgt wird. Anstelle der Aufgabenübernahme und Automatisierung durch die KI sollte die Integration und Förderung menschlicher Kreativität angestrebt werden, sodass eine weitreichende Zusammenarbeit zwischen Mensch und Maschine ermöglicht wird (Gioti, 2021). Die KI sollte durch ihren Input demnach die menschliche Kreativität erweitern und nicht versuchen, diese nachzuahmen und dadurch die Menschen zu ersetzen. Fest steht, dass die generative KI zunehmend komplexere Aufträge übernehmen kann und wird. Daher

sollte eindeutig separiert werden, welche dieser Aufgaben an eine KI übergeben werden und welche von Personen ausgeführt werden. Weiterhin sollen neue gesetzliche Regelungen dabei helfen, Menschen vor „risikobehafteten KI-Anwendungen zu schützen“ (D’Onofrio, 2024, S.341).

Insgesamt reicht die Erstellung von Musikkompositionen durch künstliche, inhumane Komponenten weit in die Vergangenheit zurück. Es verändern sich lediglich die Ausprägungen der technologischen Entwicklungen, die aktuell in Form tiefer neuronaler Netze vorliegen (Heintz, 2024). Noch erreicht der Einsatz der KI-Systeme den angestrebten Anspruch nicht, jedoch wird sich der Trend der letzten Jahre fortsetzen und die generative KI sich weiter verbessern. Das Niveau der Generierungen wird weiter steigen, sodass die zukünftigen Entwicklungen und entstehenden Möglichkeiten aufmerksam beobachtet werden müssen.

## Literaturverzeichnis

- Albrecht, H. (2021). *Leitmotivik in der Filmmusik: Einflüsse auf die visuelle Aufmerksamkeit und emotionale Wirkungen während der Filmrezeption*. Baden-Baden: Tectum Verlag.
- Berens, A., & Bolk, C. (2023). *Content Creation mit KI*. Bonn: Rheinwerk Verlag.
- Berger-Grabner, D. (2022). Angewandte qualitative Sozialforschung. In D. Berger-Grabner, *Wissenschaftliches Arbeiten in den Wirtschafts- und Sozialwissenschaften* (4. Ausg., S. 137-168). Wiesbaden: Springer Gabler.
- Brink, A. (2013). *Anfertigung wissenschaftlicher Arbeiten* (5. Ausg.). Wiesbaden : Springer Gabler.
- Brown, O. (2021). SocioculturalandDesignPerspectivesonAI-BasedMusic Production: Why Do We Make Music and What Changes if AI Makes It for Us? In E. R. Miranda, *Handbook of Artificial Intelligence for Music* (S. 1-20). Springer.
- Bullerjahn, C. (2001). *Grundlagen der Wirkung von Filmmusik*. (R.-D. Kraemer, Hrsg.) Augsburg: Wißner.
- Catani, S. (2024b). Künstliche Intelligenz in der Filmindustrie und im Filmbetrieb. In S. Catani (Hrsg.), *Handbuch Künstliche Intelligenz und die Künste* (S. 191-201). De Gruyter.
- Catani, S. (2024a). Künstliche Intelligenz und die Künste: Einleitung. In S. Catani (Hrsg.), *Handbuch Künstliche Intelligenz und die Künste* (S. 1-8). De Gruyter.
- Chion, M. (2012). *Audio-Vision: Ton und Bild im Kino*. (J. Lensing, Hrsg.) Berlin: Schiele & Schön.
- D'Onofrio, S. (28.. 03. 2024). Generative Künstliche Intelligenz – die neue Ära der kreativen Maschinen. *HMD Praxis der Wirtschaftsinformatik*, S. 331-343.
- Feuerriegel, S., Hartmann, J., Janiesch, C., & Zschech, P. (12. 09 2023). Generative AI. *Business & Information Systems Engineering*, S. 111-126.
- Filmton Berufsvereinigung e.V. (2023). *Jahrbuch der Berufsvereinigung Filmton e.V.* Abgerufen am 19. 04 2024 von bvft: <https://www.bvft.de/unsere-service/filmton-guide/>
- Frieler, K., Zaddach, W.-G., & Meyer, S. (2023). Künstliche Intelligenz in der Musikproduktion. In P. Moormann, & N. Ruth (Hrsg.), *Musik und Internet: Aktuelle Phänomene populärer Kulturen* (S. 3-28). Springer VS.
- Gioti, A.-M. (2021). Artificial Intelligence for Music Composition. In Eduardo Reck Miranda , *Handbook of Artificial Intelligence for Music* (S. 53-74). Springer.
- Heintz, J. (2024). Künstliche Intelligenz und Musik. In S. Catani (Hrsg.), *Handbuch Künstliche Intelligenz und die Künste* (S. 227-252). De Gruyter.

- Heldt, G. (2018). Was uns die Töne erzählen: Narratologie und Filmmusik. In F. Hentschel, & P. Moormann (Hrsg.), *Filmmusik Ein alternatives Kompendium* (S. 123-146). Wiesbaden: Springer VS.
- Kühn, H. (1976). Hinweise auf eine Analyse von Filmmusik. In H.-C. Schmidt (Hrsg.), *Musik in den Massenmedien Rundfunk und Fernsehen: Perspektiven und Materialien* (S. 120-125). Mainz: Schott.
- Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, 62, S. 15-25.
- Kloppenburger, J. (2017a). *Das Handbuch der Filmmusik* (2. Ausg.). (J. Kloppenburger, Hrsg.) Laaber.
- Kloppenburger, J. (2017b). Musik im Film. In G. Rötter (Hrsg.), *Handbuch Funktionale Musik: Psychologie - Technik - Anwendungsgebiete* (S. 429-456). Wiesbaden: Springer.
- Koehler, J. (2024). Zum Begriff der 'künstlichen Intelligenz'. In S. Catani (Hrsg.), *Handbuch Künstliche Intelligenz und die Künste* (S. 9-26). De Gruyter.
- Kreuzer, A. C. (2009). *Filmmusik in Theorie und Praxis*. Konstanz: UKV Verlagsgesellschaft mbH.
- Kuchenbuch, T. (2005). *Filmanalyse: Theorien - Methoden - Kritik* (2. Ausg.). Böhlau Verlag Wien Köln Weimar.
- Kuckartz, U., & Rädiker, S. (2022). *Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung* (5. Ausg.). Weinheim: Beltz Juventa.
- Kungel, R. (2008). *Filmmusik für Filmemacher: Die richtige Musik zum besseren Film*. Heidelberg: dpunkt.verlag GmbH & mediabook Verlag.
- Lehnen, J. (2017). Forschungsdesign: Kombination quantitativer/qualitativer Forschung. In J. Lehnen, *Integration von Lead Users in die Innovationspraxis* (S. 73-84). Wiesbaden: Springer Gabler.
- Lensing, J. U. (2018). *Sound-Design Sound-Montage Soundtrack-Komposition* (3. Ausg.). Berlin: Fachverlag Schiele & Schön .
- Liebhold, R., & Trinczek, R. (2009). Experteninterview. In P. S. Stefan Kühl, & A. Taffertshofer (Hrsg.), *Handbuch Methoden der Organisationsforschung* (S. 32-56). Wiesbaden: VS Verlag.
- Liegmal, M. (2020). *Wenn der Computer zum Künstler wird: Wie Big Data und KI die Musik-, Literatur-, Kunst- und Entertainmentbranche revolutionieren*. München: Redline Verlag.
- Loth, A. (2024). *KI für Content Creation: Texte, Bilder, Audio und Video erstellen mit ChatGPT & Co.* mitp Verlag.

- Mockenhaupt, A., & Schlagenhaut, T. (2024). *Digitalisierung und Künstliche Intelligenz in der Produktion: Grundlagen und Anwendung* (2. Ausg.). Wiesbaden: Springer Vieweg.
- Nolan, C. (Regisseur). (2010a). *Inception* [Kinofilm].
- Pauli, H. (1976). Filmmusik: ein historisch-kritischer Abriß. In H.-C. Schmidt (Hrsg.), *Musik in den Massenmedien Rundfunk und Fernsehen: Perspektiven und Materialien* (S. 91-119). Mainz: Schott.
- Rabenalt, R. (2020). *Musik Dramaturgie im Film - Wie Filmmusik Erzählformen und Filmmwirkung beeinflusst*. München: edition text + kritik.
- Schramm, H., Spangardt, B., & Ruth, N. (2017). *Medien und Musik*. Wiesbaden: Springer VS.
- Tieber, C. (2020). Musik im Film, Musik für den Film: Analysefelder und Methoden. In M. Hagener, & V. Pantenburg (Hrsg.), *Handbuch Filmanalyse* (S. 97-110). Wiesbaden: Springer VS.
- Zickgraf, L. (2024). Künstliche Intelligenz als Werkzeug in der Musik – Versuch einer historischen Einordnung. In S. Catani (Hrsg.), *Handbuch Künstliche Intelligenz und die Künste* (S. 217-226). De Gruyter.

## Internetquellen

AIVA. (kein Datum). (Aiva Technologies SARL) Abgerufen am 09. 07. 2024 von AIVA:

<https://www.aiva.ai>

*Hans Zimmer - Making Of Inception Soundtrack*. (30.03.2011). [Video]. *YouTube*. Abgerufen

am 03.06.2024 von: <https://www.youtube.com/watch?v=W1Flv7rFbv4>

*IMDb*. (kein Datum). Abgerufen am 23. 04. 2024 von IMDb Inception:

<https://www.imdb.com/title/tt1375666/>

Itzkoff, D. (28. 07. 2010). *The New York Times*. Abgerufen am 01. 06. 2024 von Hans

Zimmer Extracts the Secrets of the 'Inception' Score:

<https://archive.nytimes.com/artsbeat.blogs.nytimes.com/2010/07/28/hans-zimmer-extracts-the-secrets-of-the-inception-score/>

*James Cameron's Story of Science Fiction: Christopher Nolan Clip (AMC)*. (2018). [Video].

*YouTube*. Abgerufen am 25.05.2024

von: <https://www.youtube.com/watch?v=qyjybrtSoFU>

Martens, T. (20. 07. 2010). *Los Angeles Times*. Abgerufen am 01. 06. 2024 von Hans

Zimmer and Johnny Marr talk about the sad romance of 'Inception':

<https://www.latimes.com/archives/blogs/hero-complex-blog/story/2010-07-20/hans-zimmer-and-johnny-marr-talk-about-the-sad-romance-of-inception>

*Melobytes*. (2022). Abgerufen am 18. 05. 2024 von <https://melobytes.com/en/>

Nolan, C. (29. 11. 2010b). Q&A: Christopher Nolan on Dreams, Architecture, and Ambiguity.

(Wired, Interviewer, & Wired, Herausgeber) Abgerufen am 25.05.2024 von:

<https://www.wired.com/2010/11/pl-inception-nolan/>

*Soundraw*. (2023). Abgerufen am 18. 05. 2024 von <https://soundraw.io>

*Suno*. (2024). Abgerufen am 18. 05. 2024 von <https://suno.com/about>

*Udio*. (2024). Abgerufen am 18. 05. 2024 von <https://www.udio.com/about-us>

Zimmer, H. (03. 11. 2014). Composer Hans Zimmer Talks 'Interstellar' Origin, Punk Influence

on 'Dark Knight' (Exclusive Video). (S. Galloway, Interviewer, & T. H. Repoter,

Herausgeber) Abgerufen am 21.05.2024 von:

<https://www.hollywoodreporter.com/news/general-news/composer-hans-zimmer-talks-interstellar-745891/>

# Anlagen

## Anlage 1: Kodierleitfaden

Kategorie	Definition	Ankerbeispiele
K1: Evaluierung der generierten Filmmusik	Zusammenführung verschiedener Kriterien anhand welchen die Qualität der generierten Filmmusik eingeordnet wird.	<p>„Zum Beispiel fand ich gerade so dieses Anfangsthema sehr repetitiv, also von der Instrumentalisierung nicht sehr ausgefallen. Es war sehr streicherlastig und bei dem anderen hatte man halt immer noch so eine andere Basis quasi drunter. Man hatte ein bisschen eine höhere Frequenzrange, die da mitgespielt hat.“ (FM1, Pos. 58-62)</p> <p>„[...] es war schon alles da, aber es war ein gutes Stück, beim Hans Zimmer ist es eben filigraner und emotionaler, also es reicht offensichtlich nicht nur quasi die Harmonik richtig zu haben, sondern natürlich auch dann die Instrumente, wie sie spielen [...]“ (FK3, Pos.71-73)</p>
K2: Möglicher Nutzen sowie Chancen und Risiken durch KI (Kompositions-) Tools	Aspekte, welche Aufschluss über die Vor-/Nachteile der KI-Softwares geben und welche Chancen und Risiken sich dadurch für die Filmmusikbranche ergeben.	<p>„Die [Programme], die ich für mich interessant fände, sind noch nicht gekommen. Und ich weiß auch nicht, ob sie jemals kommen werden.“ (FK2, Pos. 209-210)</p> <p>„Natürlich, man macht sich die Gedanken und ja, es kann eine Gefahr sein.“ (FK3, Pos. 63-64)</p>
K3: Erfahrungen mit Kompositionssoftwares und deren Relevanz im Arbeitsalltag	Aspekte, welche die Erfahrungen sowie den derzeitigen Einsatz und die Auswirkungen von KI im Arbeitsalltag thematisieren.	<p>„[...] das Interesse ist da, aber ich hatte einfach noch nicht die Zeit, es auszuprobieren“ (FM3, Pos. 238-239)</p> <p>„es waren viele, viele Versuche, die so völlig im Nichts gelandet sind, [...] es war mir noch zu wenig Outcome, als dass ich jetzt sagen würde, wow, ich lasse mich jetzt immer davon inspirieren [...]“ (FK3, Pos.153-155)</p>
K4: Zukünftige Entwicklungen und die mögliche Ersetzung des Komponisten durch die KI	Erwartungen und Prognosen bezüglich der Entwicklung generierter Filmmusik und der Stellung des	<p>„Also ich glaube schon, dass man da technologiemäßig als Firma [...] Ahnung von [haben sollte], was da passiert in dem Bereich, weil sonst wird man irgendwann abgehängt.“ (FM1, Pos. 243-245)</p>

	Komponisten in den nächsten Jahren.	<p>„[...] solange es elektronische Musik ist, ist es schwierig das zu erkennen, weil die ja [...] auch künstlich erzeugt wird. Wohingegen alles, was eher Akustik oder Singer-Songwriter ist oder eher organisch ist, wird es die KI [...] noch ein bisschen schwerer haben“ (FM3, Pos. 211-214)</p> <p>„[...] da kommt noch was, wenn [das Orchester] auch noch wertig klingt [...] hast du immer noch das Problem, dass du halt einen 60, 70 Minuten Filmscore generieren musst und die Rechenleistung exponentiell steigt, wenn längere Sachen sich aufeinander beziehen müssen“ (FK2, Pos.197-201)</p>
--	-------------------------------------	--

Anlage 2: Stichwortartige Fallzusammenfassungen (nach Kuckartz & Rädiker, 2022, S.124-125)

Interview FK1

- Deutlicher Qualitätsunterschied zwischen dem Original und der generierten Version
- Generierte Musik erfüllt strukturelle Grundkriterien, Kompositionen sind jedoch sehr platt in Bezug auf das Bewegtbild
- KI ist nicht in der Lage hochwertige Filmmusikthemen zu kreieren, kann keine Mehrdeutigkeit erzeugen
- Wirkung der Filmmusik ergibt sich durch ein komplexes Zusammenspiel vieler Komponenten, auf welcher Grundlage soll die KI dies nachbilden
- Filmmusik unterliegt einem definierten Zweck, jede Note muss daran begründet werden können (macht es für die KI schwieriger als in anderen Formen der Kunst)
- Sieht keinen Nutzen und keine Vorteile in generierten Musiken
- Als Inspirationsquelle durchaus denkbar
- KI-Tools werden im Arbeitsalltag nicht genutzt
- Eine KI wird den Filmmusikkomponisten nicht ersetzen können, da sie nicht über die außermusikalischen Faktoren wie Kommunikation zwischen Filmmemacher und Komponist sowie über die menschlichen Sinne, Assoziationen und das menschliche Bewusstsein verfügt

### Interview FK2

- Geringere Klangqualität durch virtuelle Instrumente
- Musik erfasst die Dramaturgie nicht ausreichend, wirkt generisch
- Unmusikalische Sprünge in einzelnen Instrumenten, Artefakte hörbar (daran KI erkannt)
- Deutlich geringere Dynamik als beim Original
- Konzeptionelle Arbeit des Menschen derzeit noch nicht durch eine KI ersetzbar
- Verbesserungen und Erweiterungen der Tools werden in naher Zukunft erwartet
- Wirtschaftlicher Aspekt: Umgang KI und Musikrechte, Royalties
- KI wird voraussichtlich in anderen Bereichen (als Filmmusik) mehr Einfluss nehmen
- KI-Kompositionstools bringen keinen Nutzen, andere Tools hingegen schon – diese werden aber nicht entwickelt
- KI übernimmt die Komposition anstelle dabei zu unterstützen
- KI spielt keine Rolle im Arbeitsalltag, Tools sind jedoch bekannt
- Zeitersparnis bringt keinen Vorteil
- Im High-End Bereich wird die KI keinen Einzug finden

### Interview FK3

- Musik ist flacher, schlichter, überzeugt nicht ganz, teilweise werden ähnliche Stimmungen erzielt
- Nur eine passende Harmonik reicht nicht für eine überzeugende Filmmusik aus, es fehlen die dramaturgischen Feinheiten und auch die Klangqualität ist schwächer
- Vergleich Mensch mit humaner KI
- KI-Kompositionstools werden als Inspiration genutzt, persönliches Interesse vorhanden, sieht jedoch keinen Zwang, sich damit auseinander zu setzen
- Zeitersparnis kann wirtschaftlichen Vorteil bringen
- KI ist derzeit noch keine Gefahr für Komponisten, Mensch zu Mensch Kommunikation ist sehr wichtig, Menschen wollen Bezug zu Interpreten
- KI spielt keine Rolle im Arbeitsalltag
- Die Qualität der Komposition und des Sounds einer KI kann derzeit noch nicht mit dem Score eines Menschen mithalten und damit auch den Menschen nicht ersetzen

### Interview FM1

- Ähnlichkeit zwischen Original und neuer Filmmusik besteht, verschiedene Emotionen werden transportiert
- Unterscheidung der Varianten anhand der Mischungen (Kino- vs. Fernsehmischung)
- Stärken der KI liegen im Kompositionsprozess selbst: Zeitersparnis, Ideengeber
- Schwächen innerhalb der Komposition: repetitiv, fehlende Varianz in den Instrumenten, Frequenzverteilung der Musik
- Selbst keine Erfahrungen mit KI-Musiktools, sieht jedoch Potenzial für Tempomusik
- KI in mehr Bereichen als nur in der Musik vertreten, nutzbare Tools
- Offenheit gegenüber den Entwicklungen: muss sich mit diesen auseinandersetzen, eigenen bestmöglichen Umgang finden, auf aktuellem Stand bleiben
- Komponist wird nicht obsolet werden, da die KI derzeit noch nicht im hochqualitativen, künstlerischen Bereich mitwirken kann
- Für die Mischung macht es keinen wirklichen Unterschied, ob das Material von Komponisten oder einer KI geliefert wird

## Interview FM2

- Hinterfragen des Produktionsprozesses, da ähnliche Grundmuster bei der KI-Variante und dem Original
- Originalscore lässt dem Dialog mehr Platz als die KI-Musik
- Problematik, dass die KI die fertigen Files nur in stereo exportieren kann
- Prinzipielle Fragen nach der Anwendbarkeit: zwischenmenschliche Interaktionen oder Illusion totaler Kontrolle
- Die Veränderung des Arbeitsprozesses durch den Einsatz der KI ist entscheidender, als die derzeitigen erzielbaren Ergebnisse (werden sich verbessern)
- Durch die technischen Entwicklungen ändert sich auch das Hörverhalten der Konsumenten
- KI bringt großen Nutzen in der O-Ton Bearbeitung und Archivmaterial
- KI birgt Gefahr für gewisse Berufsfelder - daher Nutzen als Werkzeug, aber nicht um Berufe zu ersetzen
- Entstehung von Kunst ist ein Zusammenspiel mehrerer Komponenten, noch kann die KI nicht wirklich kreativ arbeiten
- KI ist im Arbeitsalltag etabliert
- Solange mit dem Material gearbeitet werden kann, macht es keinen Unterschied, ob dieses von Komponisten oder einer KI angeliefert wird

## Interview FM3

- Generierte Komposition: synthetischer Klangcharakter, repetitiv, fehlende Variation
- Funktioniert generell schon, wird jedoch nicht ganz abgekauft
- Im Gegensatz zum Original greift die generierte Version der Handlung nicht voraus
- Anstoß: würde die KI bei populären Tracks als Filmmusik besser funktionieren, als bei klassischem orchestralem Score?
- Zusammenspiel zwischen Musik und Soundeffekten fehlt in der KI-Version
- Erweiterungen der Funktionen von KI-Kompositionstools sind für den Einsatz zur Filmmusik nötig
- Nutzen: weniger in der Musik, dafür in anderen Bereichen wie Sprachbearbeitung
- Mehrwert durch Kostenersparnis, als Tempomusik für den Schnitt, aber nicht für die finale Tonpostproduktion
- KI als Chance für die Postproduktion, jedoch als Gefahr für den Musikbereich und das Hörverhalten der Menschen
- Selbst noch keine Kompositionstools ausprobiert
- Bei dem Export von Stems macht es keinen Unterschied, ob das Material von einer KI oder einem Menschen angeliefert wird

## Digitale Anlagen

Digitale Anlage 1: Originalfilmszene

Digitale Anlage 2: Liste mit Links zu den AIVA Kompositions-Dateien

Digitale Anlage 3: Soundalike Part 1

Digitale Anlage 4: Soundalike Part 3

Digitale Anlage 5: Inception-Szene mit generierter Filmmusik

Digitale Anlage 6: Anhand Referenztrack Time generierte Filmmusik

Digitale Anlage 7: Transkripte Experteninterviews

Hinweis: Alle digitalen Anlagen wurden in einem separaten Ordner in der angegebenen Reihenfolge dem beigelegten Speichermedium hinzugefügt.