

Christof Weiß, Rainer Kleinertz, Meinard Müller

# **Möglichkeiten der computergestützten Erkennung und Visualisierung harmonischer Strukturen – eine Fallstudie zu Richard Wagners *Die Walküre***

Beitrag zur Jahrestagung der Gesellschaft für Musikforschung Halle/Saale 2015 –  
»Musikwissenschaft: die Teildisziplinen im Dialog«

Veröffentlicht unter der Creative-Commons-Lizenz CC BY-NC-ND 4.0  
© 2016 | Schott Music GmbH & Co. KG

**gfm**  
GESELLSCHAFT FÜR  
MUSIKFORSCHUNG

Christof Weiß, Rainer Kleinertz, Meinard Müller

## **Möglichkeiten der computergestützten Erkennung und Visualisierung harmonischer Strukturen – eine Fallstudie zu Richard Wagners *Die Walküre***

### **Zusammenfassung**

In diesem Artikel versuchen wir anhand eines konkreten Beispiels zu skizzieren, inwiefern informatische Methoden gewinnbringend im Bereich der Musikwissenschaft eingesetzt werden können und inwiefern umgekehrt musikwissenschaftliche Fragestellungen zu neuen Herausforderungen in der Informatik führen. Konkret stellen wir computerbasierte Werkzeuge vor, die es erlauben, große Musikdatenbestände hinsichtlich harmonischer Strukturen auf interaktive Weise zu durchsuchen, zu visualisieren und zu analysieren. Die musikwissenschaftliche Relevanz dieser Konzepte soll anhand eines Beispielszenarios aus Richard Wagners *Die Walküre* erprobt werden. Dabei gibt es mehrere relevante Fragestellungen: Inwieweit können bekannte harmonische Strukturen maschinell nachvollzogen und visuell dargestellt werden? Gibt es bisher verborgene harmonische Strukturen und Bezüge? Exemplarisch sollen diese Fragen hier im Bezug auf Wagners Begriff der »dichterisch-musikalischen Periode« diskutiert und die Möglichkeiten der vorgestellten Visualisierungsmethoden angedeutet werden. Dabei soll überprüft werden, ob und in welcher Weise sich die charakteristischen Aspekte solcher Perioden in den automatisch berechneten harmonischen Analysen und deren Visualisierungen abbilden und ob ähnliche Strukturen an anderen Stellen entdeckt werden können. An diesem Beispiel sollen paradigmatisch Möglichkeiten des Dialogs zwischen Historischer Musikwissenschaft und Informatik auf der Basis ihrer jeweils unterschiedlichen Voraussetzungen und Methoden aufgezeigt werden.

### **1. Einleitung**

Innerhalb der Musikwissenschaft haben automatisierte Methoden in den letzten Jahren zunehmend Beachtung gefunden. Das Potenzial der Objektivierung, Quantifizierung und Ausweitung von Analysen auf größere Werkbestände wird zunehmend erkannt und umgesetzt. Idealerweise treten dabei Musikwissenschaft und Informatik in einen Dialog, der Fortschritte auf beiden Gebieten ermöglicht. Es stellt sich die Frage, inwieweit Methoden aus der Informatik gewinnbringend für die Musikwissenschaft genutzt werden können und wie in umgekehrter Weise musikwissenschaftliche Problemstellungen neue Ansätze in der Datenverarbeitung anregen können. In der interdisziplinären Kooperation, die auch diesem Artikel zugrunde liegt, haben wir bereits einige Fragestellungen und Verfahren anhand von Klaviersonaten Ludwig van Beethovens entwickelt und auf ihre musikwissenschaftlichen Relevanz überprüft. Im Fokus stand dabei die Frage, inwieweit bekannte harmonische Strukturen maschinell nachvollzogen und visuell dargestellt werden können.<sup>1</sup>

In diesem Artikel wollen wir eine erste Fallstudie vorstellen, die sich mit Richard Wagners Tetralogie *Der Ring des Nibelungen* beschäftigt. Dieses Werk stellt nicht nur aufgrund seines großen Umfangs eine Herausforderung dar. Bei der harmonischen Analyse des *Rings* wird auch musikwissenschaftlich weitgehend Neuland betreten. Eine der wenigen großformatigen Arbeiten zur tonalen Struktur des *Rings* sind die

---

<sup>1</sup> Verena Konz, Meinard Müller, Rainer Kleinertz, »A Cross-Version Chord Labelling Approach for Exploring Harmonic Structures – A Case Study on Beethoven's *Appassionata*«, in: *Journal of New Music Research* 42/1 (2013), S. 61–77.

Analysen von Alfred Lorenz, die in der Musikwissenschaft höchst umstritten sind.<sup>2</sup> Abbildung 1 zeigt beispielhaft eine Visualisierung von Tonartenverläufen in der *Walküre*. Im Kontext solcher Fragestellungen verwies bereits Lorenz auf den von Wagner selbst geprägten Begriff der »dichterisch-musikalischen Periode«,<sup>3</sup> der auf eine Verbindung von Dichtung und Tonartverläufen in der Musik hinweist.

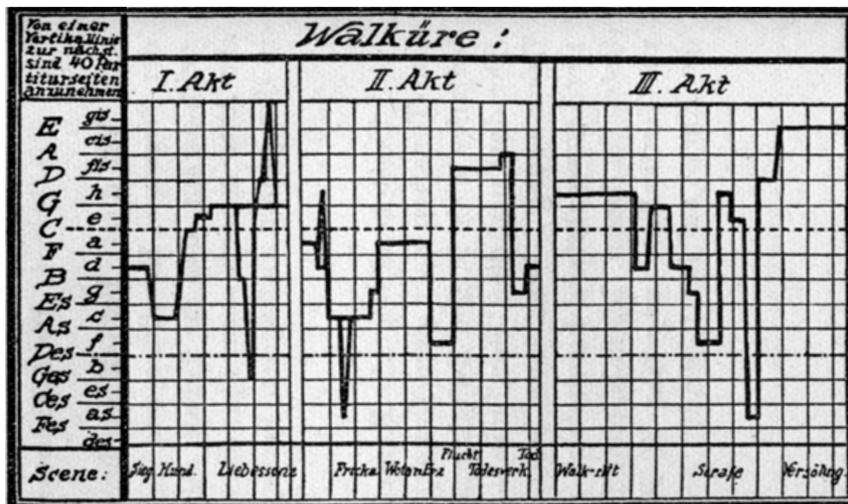


Abbildung 1 Tonartanalyse der *Walküre* von Alfred Lorenz.<sup>4</sup>

Um derartigen Theorien nachzugehen und einen Einblick in die harmonische Anlage der *Ring*-Musik zu erhalten, sind großflächige Analysen harmonischer Strukturen und Bezüge erforderlich. Dabei könnte eine in Teilen automatisierte Visualisierung auf Grund des »objektiv vorgegebenen Befundes« (der als solcher stets kritisch zu überprüfen ist) über die bisherigen Methoden hinaus zu neuen Fragestellungen sowohl in der Musikwissenschaft als auch in der Informatik führen. Zur Motivation unserer Fallstudie werden zunächst die musikwissenschaftlichen Hintergründe des Begriffs der »dichterisch-musikalischen Periode« vorgestellt (Abschnitt 2). Abschnitt 3 stellt die verwendeten Visualisierungsmethoden in den Grundzügen vor. Als konkretes Beispiel wird in Abschnitt 4 Sieglindes Erzählung »Der Männer Sippe saß hier im Saak« aus dem ersten Akt der *Walküre* (T. 955–1012) herangezogen. Anhand von Visualisierungen untersuchen wir exemplarisch die harmonischen Verläufe an dieser Stelle. Abschließend werden das Potenzial der vorgestellten Visualisierungsmethoden im Hinblick auf größere Zusammenhänge skizziert sowie die Grenzen der Methoden diskutiert (Abschnitt 5).

## 2. Musikwissenschaftlicher Hintergrund

Die Musik von Richard Wagners Hauptwerk *Der Ring des Nibelungen* ist von Anfang bis Ende von einer fast ununterbrochenen Reihe von Motiven geprägt, die – mal sich ähnelnd, mal kontrastierend – ein vielfältiges Geflecht bilden. Diese »Leitmotive« haben in der musikwissenschaftlichen Literatur vielfältige

<sup>2</sup> Alfred O. Lorenz, *Der musikalische Aufbau des Bühnenfestspiels Der Ring des Nibelungen. Das Geheimnis der Form bei Richard Wagner*, Bd. 1., Berlin 1924; Rainer Kleinertz, »Richard Wagners Begriff der »dichterisch-musikalischen Periode«, in: *Die Musikforschung* 67 (2014), S. 26–47.

<sup>3</sup> Richard Wagner, *Oper und Drama*, Leipzig 1852.

<sup>4</sup> Aus Lorenz, *Der musikalische Aufbau des Bühnenfestspiels*, S. 48.

Aufmerksamkeit in Form von Übersichten und Einzelstudien erhalten.<sup>5</sup> Jenseits dieses Motivgeflechts war die Musik Wagners vergleichsweise selten Gegenstand musikwissenschaftlicher Forschung.<sup>6</sup>

In den 1920er Jahren veröffentlichte Lorenz den ersten Band seiner vierbändigen Studie *Das Geheimnis der Form bei Richard Wagner*.<sup>7</sup> Ins Zentrum seiner auf die »Entschlüsselung« einer Großform zielenden Analysen stellte Lorenz einen von Wagner selbst in *Oper und Drama* geprägten Begriff: die »dichterisch-musikalische Periode«.<sup>8</sup>

Vereinfacht dargestellt geht es Wagner bei diesem Begriff der »dichterisch-musikalischen Periode« um die Motivation von Modulationen aus dem Drama heraus: Wenn der Text (die Handlung) in derselben Empfindung verbleibt, dann habe der Musiker (der Komponist) keine Veranlassung, die ursprüngliche Tonart zu verlassen. Wenn jedoch die Empfindungen wechseln – etwa von »Lust« zu »Leid« oder von »Wohl« zu »Weh« – dann habe er einen Grund, die Tonart zu verlassen. Kehrt dann später die ursprüngliche Empfindung zurück, so könne der Komponist wieder in die ursprüngliche Tonart zurückmodulieren und so die Einheit dieser Empfindung verdeutlichen.

Nachdem Werner Breig 2002 die Tauglichkeit dieses Begriffes für die Analyse des *Rings* grundsätzlich in Frage stellte,<sup>9</sup> hat Rainer Kleinertz 2014 einen Aspekt der Ausführungen Wagners in *Oper und Drama* in den Vordergrund gestellt, der bis dahin weitgehend unbeachtete geblieben war: den der Motivation von Modulationen.<sup>10</sup> Während Lorenz nur das für sich genommen periphere Merkmal einer wiederkehrenden Tonart berücksichtigte, dürfte – so die Hypothese von Kleinertz – für Wagner das zentrale Moment einer solchen »dichterisch-musikalischen Periode« in einer spezifischen Verbindung von Modulationen und tonaler Geschlossenheit bestanden haben.

### 3. Technische Grundlagen

Die automatisierte Erschließung von Musikdaten unter Berücksichtigung von musikalisch begründeten Variabilitäten ist eine zentrale Aufgabenstellung der Musikinformatik. Hierbei spielt die harmonische Analyse von Musik eine zentrale Rolle. Ausgangspunkt einer jeden Musikanalyse sind spezifische Darstellungen der betreffenden Musikstücke. Zentrale Darstellungsformen sind zum Beispiel Notentexte (grafische Darstellungen) oder Musikaufnahmen (akustische Darstellungen). In der Musikwissenschaft werden für die Analyse typischerweise gedruckte Partituren oder Klavierauszüge herangezogen. Um sie

---

<sup>5</sup> Vgl. Christian Thorau, *Semantisierte Sinnlichkeit: Studien zu Rezeption und Zeichenstruktur der Leitmotivtechnik Richard Wagners*, Bd. 50, Stuttgart 2003; Melanie Wald-Fuhrmann, Wolfgang Fuhrmann, *Abnung und Erinnerung. Die Dramaturgie der Leitmotive bei Richard Wagner*, Kassel 2013.

<sup>6</sup> Eine Liste signifikanter Ausnahmen findet sich u. a. bei Kleinertz, »Richard Wagners Begriff der »dichterisch-musikalischen Periode«.

<sup>7</sup> Lorenz, *Der musikalische Aufbau des Bühnenfestspiels*; vgl. Stephen McClatchie, *Analyzing Wagner's Operas: Alfred Lorenz and German Nationalist Ideology*, Rochester 1998.

<sup>8</sup> Wagner, *Oper und Drama*; Carl Dahlhaus, *Wagners Begriff der »dichterisch-musikalischen Periode«*, Regensburg 1965, S. 179–187; Kleinertz, »Richard Wagners Begriff der »dichterisch-musikalischen Periode«.

<sup>9</sup> Werner Breig, »Wagners Begriff der »dichterisch-musikalischen Periode«, in: *Schlagen Sie die Kraft der Reflexion nicht zu gering an. Beiträge zu Richard Wagners Denken, Werk und Wirken*, hrsg. von Klaus Döge, Christa Jost und Peter Jost, Mainz 2002, S. 158–172.

<sup>10</sup> Wie zentral für Wagner dieses Problem der Motivation von Modulationen durch das Drama war, lässt der 1879 in den *Bayreuther Blättern* erschienene Aufsatz »Über die Anwendung der Musik auf das Drama« erkennen: »Es scheint, daß schon jetzt einem sehr großen Theil des Publikums Manches, ja fast Alles in meinen dramatischen Musiken durchaus natürlich dünkt und demnach gefällt, worüber unsere »Professoren« noch Zeter schreien. Würden diese mich auf einen ihrer heiligen Lehrstühle setzen, so dürften sie dagegen vielleicht in noch größere Verwunderung gerathen, wenn sie wahrnahmen, welche Vorsicht und Mäßigung in der Anwendung, namentlich auch harmonischer Effektmittel, ich ihren Schülern anempfehlen würde, da ich diesen als erste Regel aufzustellen hätte, nie eine Tonart zu verlassen, solange als, was sie zu sagen haben, in dieser noch zu sagen ist«; Richard Wagner, *Sämtliche Schriften und Dichtungen, [Volks-Ausgabe]*, Leipzig 6[1911], Bd. 10, S. 193; vgl. Kleinertz, »Richard Wagners Begriff der »dichterisch-musikalischen Periode«.

einer automatischen Analyse zugänglich zu machen, müssen die Daten zunächst in maschinenlesbare Form überführt werden (Digitalisierung). Dazu werden die Notentexte im ersten Schritt gescannt. Die resultierenden digitalen Bilder müssen dann mittels geeigneter Software zur optischen Notenerkennung (*Optical Music Recognition*, kurz *OMR*) in symbolische Daten (wie z. B. das *MusicXML*-Format) konvertiert werden. Diese *OMR*-Techniken sind jedoch oft stark fehlerbehaftet und erfordern daher manuelle Nachbearbeitung.<sup>11</sup> Als alternative Herangehensweise hat sich in den letzten Jahren die direkte Analyse von Audiodaten aus einer oder mehreren Einspielungen entwickelt (ein Schwerpunkt des Forschungsgebiets *Music Information Retrieval*, kurz *MIR*). Hierzu werden Methoden der digitalen Signalverarbeitung eingesetzt. Dabei spielen akustische Besonderheiten und Artefakte eine Rolle wie z. B. der Einfluss von Obertönen, spielerische Unsauberkeiten in Tonhöhe und Rhythmik oder perkussive Komponenten.

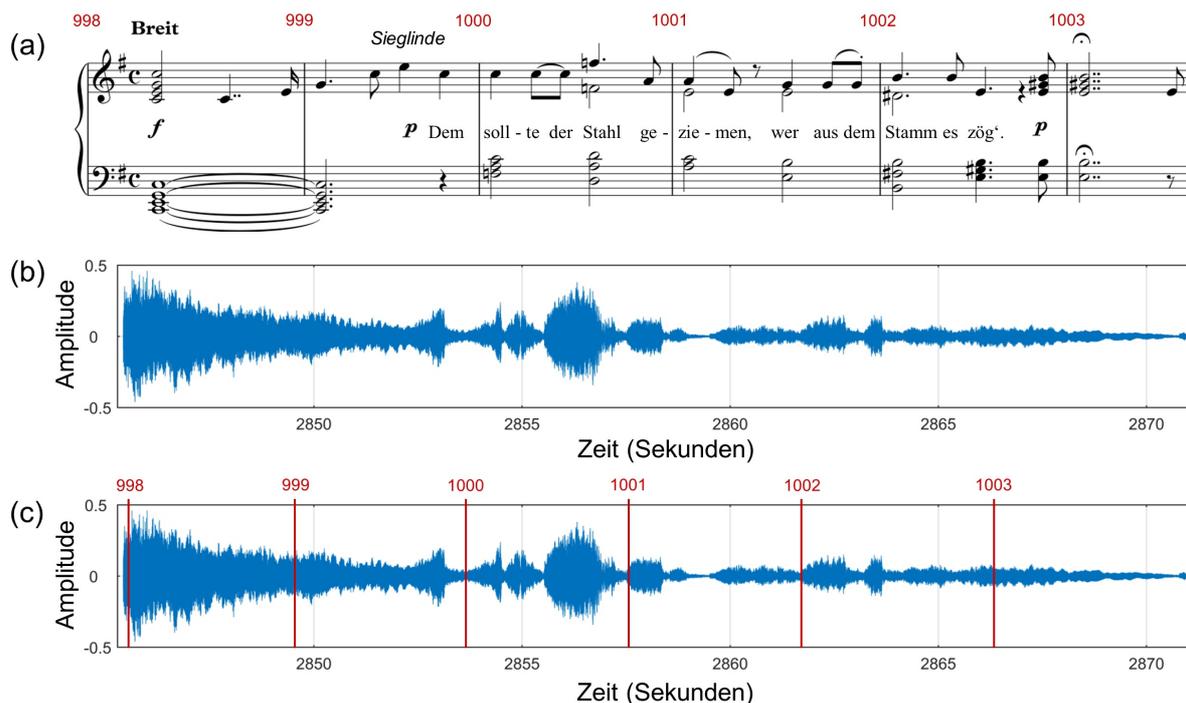
Abbildung 2 zeigt eine Stelle aus dem ersten Akt der *Walküre* in verschiedenen Darstellungsformen. Die obere Abbildung (a) zeigt eine Klavierreduktion von Karl Richard Kleinmichel (Mainz, ca. 1910).<sup>12</sup> In Abbildung (b) ist die Wellenform derselben Stelle aus einer Aufnahme mit Herbert von Karajan dargestellt, bei der die Amplitudenwerte über eine Zeitachse geplottet werden. Abbildung (c) enthält zusätzliche Informationen über die Positionen der Taktanfänge. Ein grundlegender Unterschied der Darstellungen ist die unterschiedliche Kodierung zeitlicher Informationen. Während Notentexte in musikalischen Einheiten (Takten) organisiert sind, weisen Musikaufnahmen eine physikalische Zeitachse (gemessen in Sekunden) auf. Die Verknüpfung dieser Zeitachsen stellt eine besondere Herausforderung dar, welche beispielsweise durch manuelles Markieren der Taktanfänge in den Musikaufnahmen gelöst werden kann. Die Automatisierung (*Musik-synchronisation*) dieser Prozedur stellt eine typische Aufgabenstellung der Musikverarbeitung dar.<sup>13</sup>

---

<sup>11</sup> Donald Byrd, Jakob G. Simonsen, »Towards a Standard Testbed for Optical Music Recognition: Definitions, Metrics, and Page Images«, in: *Journal of New Music Research* 44/3 (2015), S. 169–195.

<sup>12</sup> Diese Bearbeitung ist frei erhältlich unter <http://www.imslp.org> (aufgerufen am 20.6.2016).

<sup>13</sup> Siehe beispielsweise Roger B. Dannenberg, Ning Hu, »Polyphonic Audio Matching for Score Following and Intelligent Audio Editors«, in: *Proceedings of the International Computer Music Conference (ICMC)*, San Francisco 2003, S. 27–34; Simon Dixon, Gerhard Widmer, »MATCH: A Music Alignment Tool Chest«, in: *Proceedings of the International Society for Music Information Retrieval Conference (ISMIR)*, London 2005, S. 492–497; Sebastian Ewert, Meinard Müller, Peter Grosche, »High Resolution Audio Synchronization Using Chroma Onset Features«, in: *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Taipei, Taiwan 2009, S. 1869–1872; Meinard Müller, *Fundamentals of Music Processing*, Cham (Switzerland) 2015, Kapitel 3.

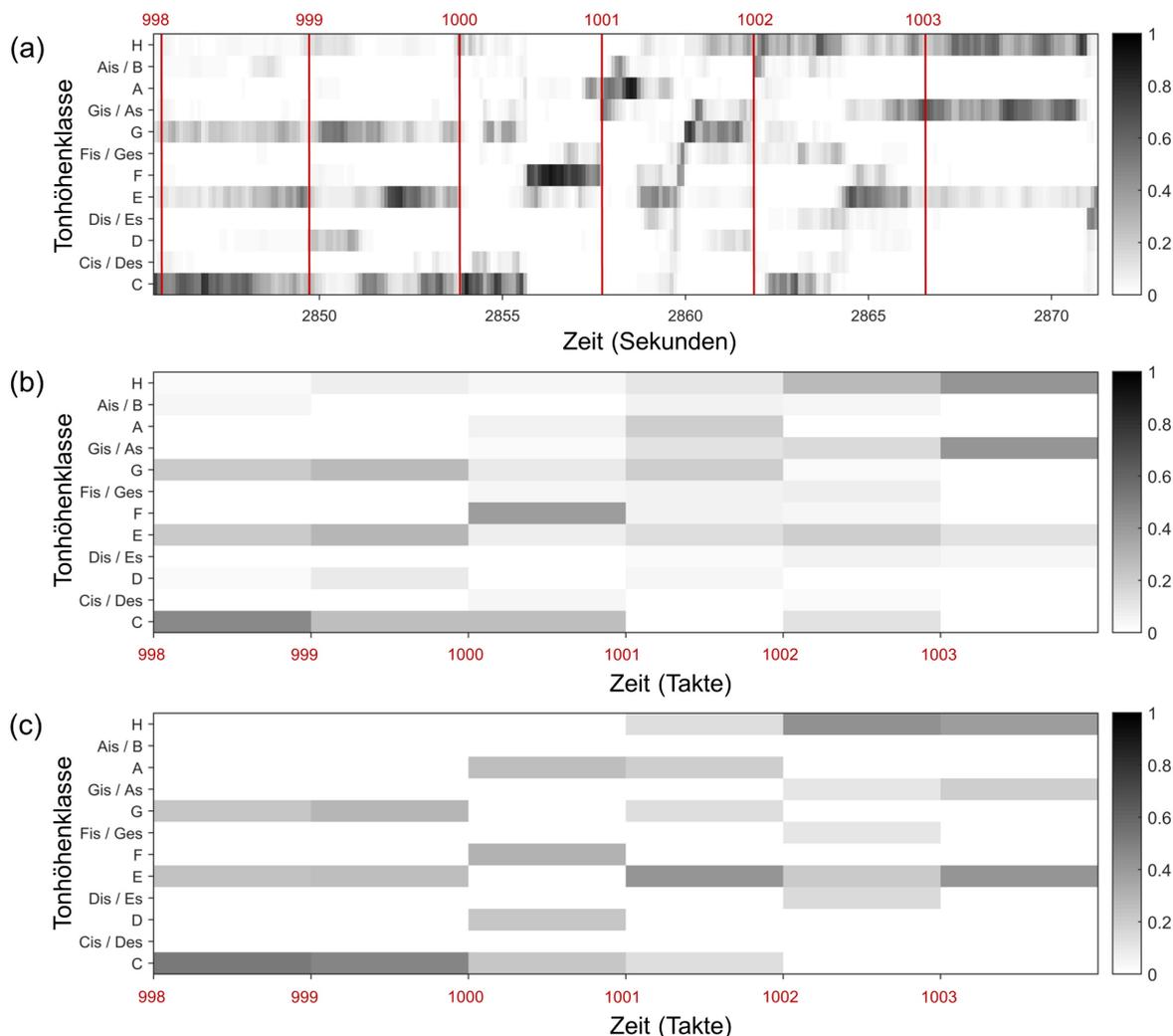


**Abbildung 2** Richard Wagner, *Die Walküre*, Akt 1, Takte 998–1003. (a) Notendarstellung. (b) Wellenform. (c) Wellenform mit markierten Taktanfängen.

Im vorliegenden Beitrag steht jedoch ein anderer Bereich des *Music Information Retrieval* im Fokus: die computergestützte Harmonieanalyse. Die technische Grundlage für die harmonische Analyse von Audiodaten stellt die Messung spektraler Energien dar, die in chromatischen Tonhöhenklassen zusammengefasst werden. Dieses Verfahren wird im Folgenden grob skizziert. Mittels Signalverarbeitungsmethoden (Fourieranalyse) zerlegen wir das Audiosignal zunächst in 88 Tonhöhenbänder. Die resultierenden Tonhöhenmerkmale beschreiben die spektrale Energie in diesen Bändern. Diese Tonhöhenmerkmale werden unter Berücksichtigung von Oktavverwandtschaften in zwölfdimensionale *Chromamerkmale* überführt, welche die Intensität der zwölf Tonhöhenklassen *c*, *cis* usw. bis *b* im Zeitverlauf angeben.<sup>14</sup> In unserem Verfahren erhalten wir zehn Chromavektoren pro Sekunde, wobei jeder Chromavektor grob gesagt 200 ms des ursprünglichen Audiosignals miteinbezieht (mit 100 ms Überlappung).

Abbildung 3a zeigt Chromamerkmale für die Beispielstelle aus dem ersten Akt der *Walküre*. Auf der horizontalen Achse ist die Zeitposition der Merkmalsvektoren in Sekunden aufgetragen. Die vertikale Achse gibt die chromatischen Tonhöhenklassen an. Die Auflösung enharmonischer Unterschiede (z. B. *cis* und *des*) ist hier nicht möglich. Die Intensität der jeweiligen Klassen wird durch Helligkeitswerte dargestellt. Dunkle Schattierungen implizieren hohe Energie in der betreffenden Tonhöhenklasse. In dieser Abbildung ist vor allem die Hauptmelodie gut zu verfolgen. Die übrigen Akkordtöne sind mit etwas schwächeren Energiewerten ebenfalls vertreten.

<sup>14</sup> Mark A. Bartsch, Gregory H. Wakefield, (2005): »Audio thumbnailing of popular music using chroma-based representations«, in: *IEEE Transactions on Multimedia* 7/1 (2005), S. 96–104; Meinard Müller, Sebastian Ewert, »Chroma Toolbox: MATLAB Implementations for Extracting Variants of Chroma-Based Audio Features«, in: *Proceedings of the International Society for Music Information Retrieval Conference (ISMIR)*, Miami, Florida 2011, S. 215–220; Müller, *Fundamentals of Music Processing*, Kapitel 2.



**Abbildung 3** Chromamerkmale für die Takte 998–1003 aus der *Walküre*, Akt 1. Die oberen Darstellungen wurden aus Audiodaten berechnet (Karajan, Deutsche Grammophon 1966) (a) in einer Auflösung von 10 Hz bzw. (b) takt-synchron. Abbildung (c) zeigt takt-synchrone Chromamerkmale, berechnet aus symbolischen Daten (Klavierreduktion).

Zur Vergleichbarkeit mit Notendarstellungen können Chromamerkmale auch taktweise berechnet werden (Abbildung 3b). Im Falle von harmonisch stabilen Takten konzentriert sich die Energieverteilung auf wenige Tonhöhenklassen. So zeigt sich zum Beispiel in Takt 998 die Struktur des C-Dur-Dreiklangs durch eine hohe Energie in den *c*-, *e*- und *g*-Bändern. Dagegen führen Akkordwechsel innerhalb eines Taktes zu eher verrauschten Merkmalen (siehe beispielsweise T. 1001). Abbildung 3c zeigt Chromamerkmale für dieselben Takte, gewonnen aus der symbolischen Notendarstellungen (Klavierreduktion), wo die Tonhöhen explizit gegeben sind. Im direkten Vergleich sind diese etwas weniger verrauscht als die aus dem Audiomaterial gewonnenen Merkmale. So ist z. B. in den Audiomerkmalen der Einfluss von Obertönen sichtbar, so in Takt 999 der Ausschlag in der Tonhöhenklasse *d*, vermutlich verursacht vom dritten Partialton des Melodietones *g'* (gespielt von der Trompete). Eine andere Abweichung tritt im Takt 1002 auf, wo in den Audiomerkmalen Energie in der Tonhöhenklasse *c* feststellbar ist. Dieser Wert rührt möglicherweise vom Vibrato der Singstimme (auf dem Ton *b'*) her. Insgesamt lässt sich aber ein sehr ähnliches Bild für beide Datentypen erkennen. Gerade für grobe zeitliche Auflösungen mag

daher die direkte Analyse von Audiodaten eine interessante Option sein, insbesondere wenn symbolische Daten nicht in ausreichender Qualität vorliegen.

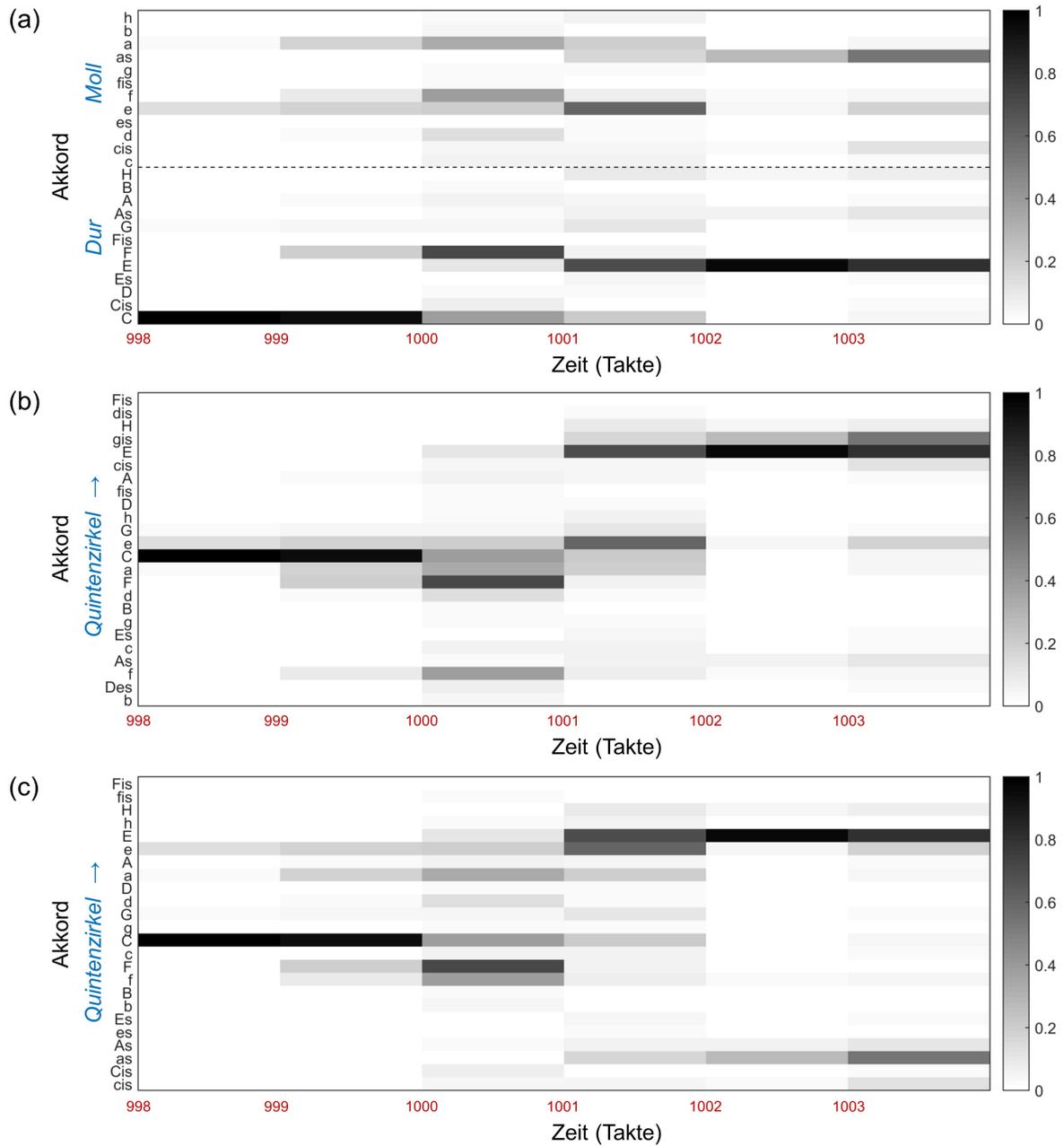
Chromamerkmale reflektieren also die lokale Energieverteilung der involvierten Tonhöhenklassen und stellen damit eine vergleichsweise grobe Messung des harmonischen Gehalts dar (sogenannte »Mid-level«-Darstellungen). Die Auflösung enharmonischer Unterschiede sowie von Oktavlagen ist so nicht möglich, was eine gewisse Einschränkung der Analysemöglichkeiten mit sich bringt.<sup>15</sup> Dennoch können sie weiterverarbeitet werden im Hinblick auf bestimmte Konzepte aus der Musiktheorie. Als ein einfaches Beispiel sei hier die Erkennung der 24 chromatischen Dur- und Moll dreiklänge genannt.<sup>16</sup> Hierzu werden die Chromavektoren für jeden Takt mit theoretischen Tonhöhenverteilungen für bestimmte Akkordtypen verglichen. Hieraus erhält man eine Art »Wahrscheinlichkeit« für die verschiedenen Akkorde.<sup>17</sup>

Abbildung 4a zeigt eine Akkordanalyse für unsere Beispielstelle aus der *Walküre*, basierend auf Audiodaten. Die Zeitachse (horizontal) ist in Takte gegliedert. Auf der vertikalen Achse sind die Akkorde in chromatisch aufsteigender Reihenfolge geordnet, getrennt nach Dur und Moll. Analog zu den Chromadarstellungen wird die Wahrscheinlichkeit für die verschiedenen Akkorde durch unterschiedliche Helligkeitswerte ausgedrückt. Im ersten Takt des Beispiels (T. 998) erzielt der C-Dur-Dreiklang klar den höchsten Wert. Es lassen sich aber auch Wahrscheinlichkeiten ungleich Null für verwandte Dreiklänge beobachten. Dies wird zum einen durch gemeinsame Tonhöhen verursacht, aber auch durch akustische Phänomene wie Obertöne, perkussive Signalanteile oder Vibrato, wodurch auch »nicht notierte« Tonhöhenklassen Energie erhalten. So ist im ersten Takt zum Beispiel ein leichter Grauwert für den e-Moll-Dreiklang erkennbar, welcher möglicherweise von der Präsenz der Tonhöhenklasse *b* als Oberton von *e* herrührt. Die taktweise Akkordbestimmung hat ihre natürliche Grenze bei jeweiligen Klassen wird durch Helligkeitswerte dargestellt. Dunkle Schattierungen implizieren hohe Energie in der betreffenden Tonhöhenklasse. In dieser Abbildung ist vor allem die Hauptmelodie gut zu verfolgen. Die übrigen Akkordtöne sind mit etwas schwächeren Energiewerten ebenfalls vertreten. Eine weitere Anordnungsmöglichkeit wird in Abbildung 4c realisiert, wo die Grundtöne im Quintenzirkel angeordnet sind, aber Akkorde in Variantbeziehung (z. B. C-Dur – c-Moll) nebeneinander stehen.

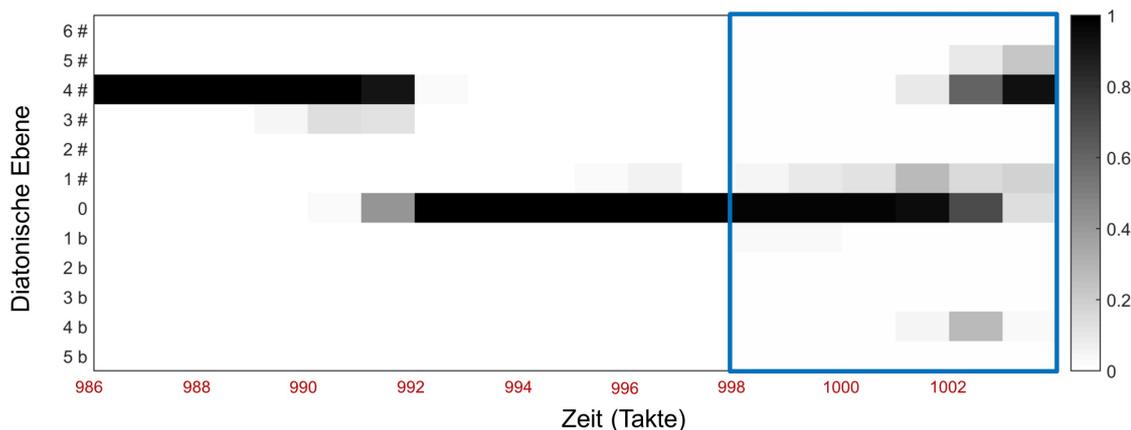
<sup>15</sup> Es kann beispielsweise nicht bestimmt werden, in welcher Lage oder Umkehrung ein Akkord auftritt.

<sup>16</sup> Alexander Sheh, Daniel P. W. Ellis, »Chord Segmentation and Recognition Using EM-trained Hidden Markov Models«, in: *Proceedings of the International Society for Music Information Retrieval Conference (ISMIR)*, Baltimore, Maryland 2003, S. 185–191; Nanzhu Jiang, Peter Grosche, Verena Konz, Meinard Müller, »Analyzing Chroma Feature Types for Automated Chord Recognition«, in: *Proceedings of the Audio Engineering Society Conference (AES)*, Ilmenau 2011, S. 285–294; Müller, *Fundamentals of Music Processing*, Kapitel 5.

<sup>17</sup> Technische Details zu einer solchen oder ähnlichen Berechnung finden sich bei Taemin Cho, Juan Pablo Bello, »On the Relative Importance of Individual Components of Chord Recognition Systems«, in: *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 22/2 (2014), S. 477–492.



**Abbildung 4** Dreiklangsanalyse für die Takte 998–1003 aus der *Walküre*, Akt 1 (Karajan 1966). (a) Darstellung in chromatischer Anordnung. (b) Anordnung nach dem Quintenzirkel mit benachbarten Paralleklängen. (c) Anordnung nach dem Quintenzirkel mit benachbarten Variantklängen.



**Abbildung 5** Analyse diatonischer Skalen für die Takte 986–1003 aus der *Walküre*, Akt 1 (Karajan 1966).

Eine Herausforderung für die automatische Akkordanalyse liegt in der Erkennung und Bewertung akkordfremder Töne. Hierfür sind Stimmführungs- und Lagen-Aspekte von Bedeutung, die aus einer Chromadarstellung nicht ohne Weiteres rekonstruiert werden können. Um dieses Problem zu umgehen, kann der Fokus auf andere Aspekte gelegt werden. Als Beispiel seien hier diatonische Skalen angeführt, welche im Wesentlichen dem Tonvorrat einer Tonart entsprechen. Ein solches Analysesystem wird von Zsolt Gárdonyi und Hubert Nordhoff angeregt.<sup>18</sup> Hierbei sind die diatonischen Ebenen nach der zugehörigen Vorzeichnung benannt (»-2-Diatonik« steht z. B. für den Tonvorrat der B-Dur-Tonleiter). Zur Realisierung dieses Konzepts werden alle Tonhöhenklassen einer diatonischen Skala berücksichtigt, indem ihre Chromawerte multipliziert werden.<sup>19</sup> Der Übergang von der Akkordanalyse zu einer Skalen- oder Tonartanalyse erfordert zugleich eine zeitliche Vergrößerung, da sich die Analyse von Tonarten typischerweise auf größere Zeiträume bezieht. In unserem Beispiel mitteln wir über vier Takte. Abbildung 5 zeigt eine diatonische Analyse für einen etwas längeren Ausschnitt des *Walküre*-Beispiels. Auch hier ist die Zeitachse in Takte gegliedert, wobei sich die Ergebnisse auf einen viertaktigen Zeitraum um den betreffenden Takt beziehen. Die Helligkeitswerte geben die Wahrscheinlichkeit an, mit welcher ein bestimmter diatonischer Tonvorrat lokal realisiert ist. Der blau umrahmte Ausschnitt (entspricht Abbildung 4) zeigt klar die Rückmodulation von der 0-Diatonik (C-Dur) in die +4-Diatonik (E-Dur). In der Übergangsphase (T. 1002) mischen sich die Tonvorräte von E-Dur und C-Dur, was zu mehreren kleinen Ausschlägen führt. Im folgenden Abschnitt wollen wir den Nutzen solcher Analysen diskutieren.

#### 4. Anwendung und Ergebnisse

Als konkretes Beispiel wollen wir nun Sieglindes Erzählung »Der Männer Sippe saß hier im Saak« in der dritten Szene des ersten Akts der *Walküre* (T. 955–1012) im Sinne der Hypothese von Kleinertz (Abschnitt 1) diskutieren.<sup>20</sup> Abbildung 6 zeigt eine Klavierreduktion dieser Stelle. Dazu wollen wir den betreffenden Ausschnitt der Karajan-Aufnahme mit den vorgestellten Techniken analysieren und visualisieren.

<sup>18</sup> Zsolt Gárdonyi, Hubert Nordhoff, *Harmonik*, Wolfenbüttel 2002.

<sup>19</sup> Siehe Christof Weiß, Julian Habryka, »Chroma-Based Scale Matching for Audio Tonality Analysis«, in: *Proceedings of the 9th Conference on Interdisciplinary Musicology (CIM)*, Berlin 2014, S. 168-173. Hierbei sind die Töne der zugehörigen Tonikadreitöne stärker gewichtet, für die 0-Diatonik z. B. die Töne des C-Dur- und des a-Moll-Dreitons.

<sup>20</sup> Alfred Lorenz sieht in Sieglindes Erzählung einen Teilabschnitt der 12. Periode, (12 A a), den er von T. 926–1021 reichen lässt; siehe Lorenz, *Der musikalische Aufbau des Bühnenfestspiels*, S. 29.

Abbildung 7 zeigt eine Analyse der Stelle hinsichtlich diatonischer Ebenen. Die Chromamerkmale wurden hierzu über jeweils vier Takte gemittelt. Die grobe Modulationsstruktur ist bereits erkennbar. Die Tonart der Erzählung, in der sie auch beginnt, ist e-Moll (T. 955–962), angedeutet durch hohe Werte für die +1-Diatonik. Die ebenfalls deutliche Energie in der 0-Diatonik ist vermutlich auf den prominenten Melodieton *e*“ (T. 958 und 960) zurückzuführen. Im Verlauf der Erzählung sind drei Episoden in der +4-Ebene erkennbar (um T. 965, um T. 985 und kurz um T. 1003). Tatsächlich sind in der Partitur hier vier Kreuze vorgezeichnet, was auf E-Dur hinweist, die Varianttonart der Ausgangstonart e-Moll. Abbildung 8 zeigt eine Akkordanalyse für dieselbe Stelle. Die Anordnung nach dem Quintenzirkel zeigt hier ein ähnliches Muster. Die E-Dur-Episoden werden durch die Präsenz des zugehörigen Tonikadreitklangs bestätigt. Beim Blick auf den Text lassen sich nun jeweils Motivationen für die Wendungen nach E-Dur finden: In Takt 963 ist es der Eintritt des »Greises in grauem Gewand« (Wotan), in Takt 986 der Blick auf die Tochter (Sieglinde). Die Takte 1002–1003 bringen schließlich eine kurze Rückblende auf E-Dur, als angedeutet wird, für wen das Schwert bestimmt ist (Siegfried). Die Tonart E-Dur scheint also für die Wälsungen Sieglinde, Siegmund und ihren Vater Wotan zu stehen.

Wie in den Abbildungen angedeutet, wird diese Tonartenregion zweimal verlassen. Zunächst scheint die +3-Ebene erreicht zu werden (T. 970–977). Die Akkorddarstellung zeigt hier ein weniger klares Bild, was auf eine eher suchende Harmonik hinweist. In den Takten 973 und 976 ist aber tatsächlich fis-Moll erkennbar, was der Notentext bestätigt. Anlass für diese Modulation scheint die Angst der Männer vor Wotans Blick zu sein (»traf die Männer sein mächtiges Dräu'n«). Der Blick auf seine Tochter hingegen veranlasst die Rückmodulation über mehrere verminderte Septakkorde<sup>21</sup> und den Gis-Dur-Dreitklang, in den Abbildungen enharmonisch als As-Dur dargestellt.

---

<sup>21</sup> Eine Analyse solcher Klänge, die strukturell stark von Dur- oder Molldreiklängen abweichen, ist mit der vorgestellten Methode ebenfalls möglich (siehe Konz u. a., »A Cross-Version Chord Labelling Approach«), bleibt aber im vorliegenden Beitrag unberücksichtigt.

955 **Langsamer**  
*Sieglinde*  
*p*  
 Der Män-ner Sip-pe saß hier im Saal, von Hundung zur Hochzeit geladen: er frei-te ein Weib, das un-ge-fragt Schächer ihm schenkten zur Frau. Trau- rig

961  
 saß ich, wäh-rend sie tran-ken: ein Frem- der trat da her - **963** **Mässig**  
*pp*  
 ein: ein Greis in grau-em Ge- wand; tief

967  
 hing ihm der Hut, der deckt ihm der Au- gen ei- nes; doch des and- ren Strahl, Angst schuf es

973  
 al-len, traf die Män- ner sein mäch- ti- ges Dräu'n: **977**  
 mir al- lein weck- te das Au- ge süß

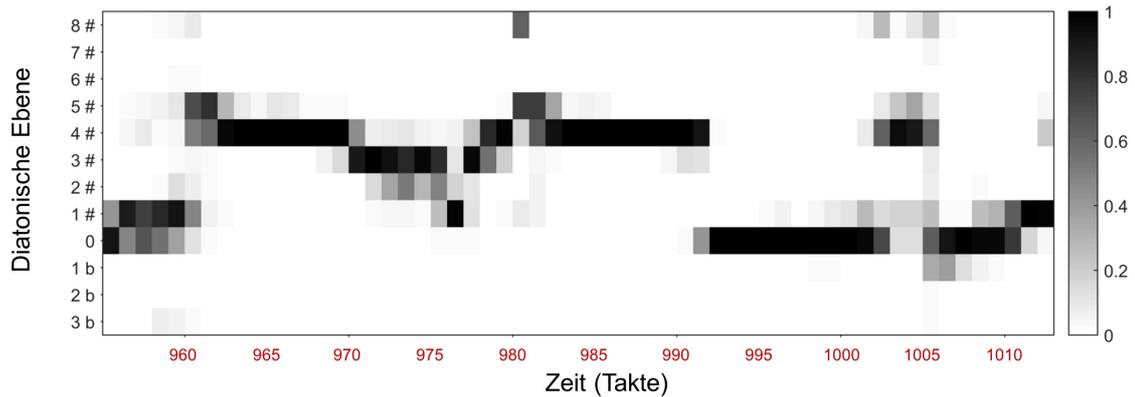
982  
 seh- nen- den Harm, Trä- nen und Trost zu- gleich. auf mich blickt' er, und blitzte auf jene, als ein Schwert in Händen er

991  
 schwang, das stieß er nun in der E- sche Stamm, bis zum Heft haf- tet' es **Breit**  
*f*  
 drin. dem

1000  
 soll- te der Stahl ge- zie- men, wer aus dem Stamm es zög'. **1004**  
*p*  
 Der Män-ner al- le, so kühn sie sich müh-ten, die Wehr sich keiner gewann.

1006  
 Gäs-te ka-men und Gäs-te gin-gen, die Stärksten zogen am Stahl; keinen Zoll ent-wich er dem Stamm: **1012** **Ruhig**  
*pp*  
 dort haf- tet schweigend das Schwert.

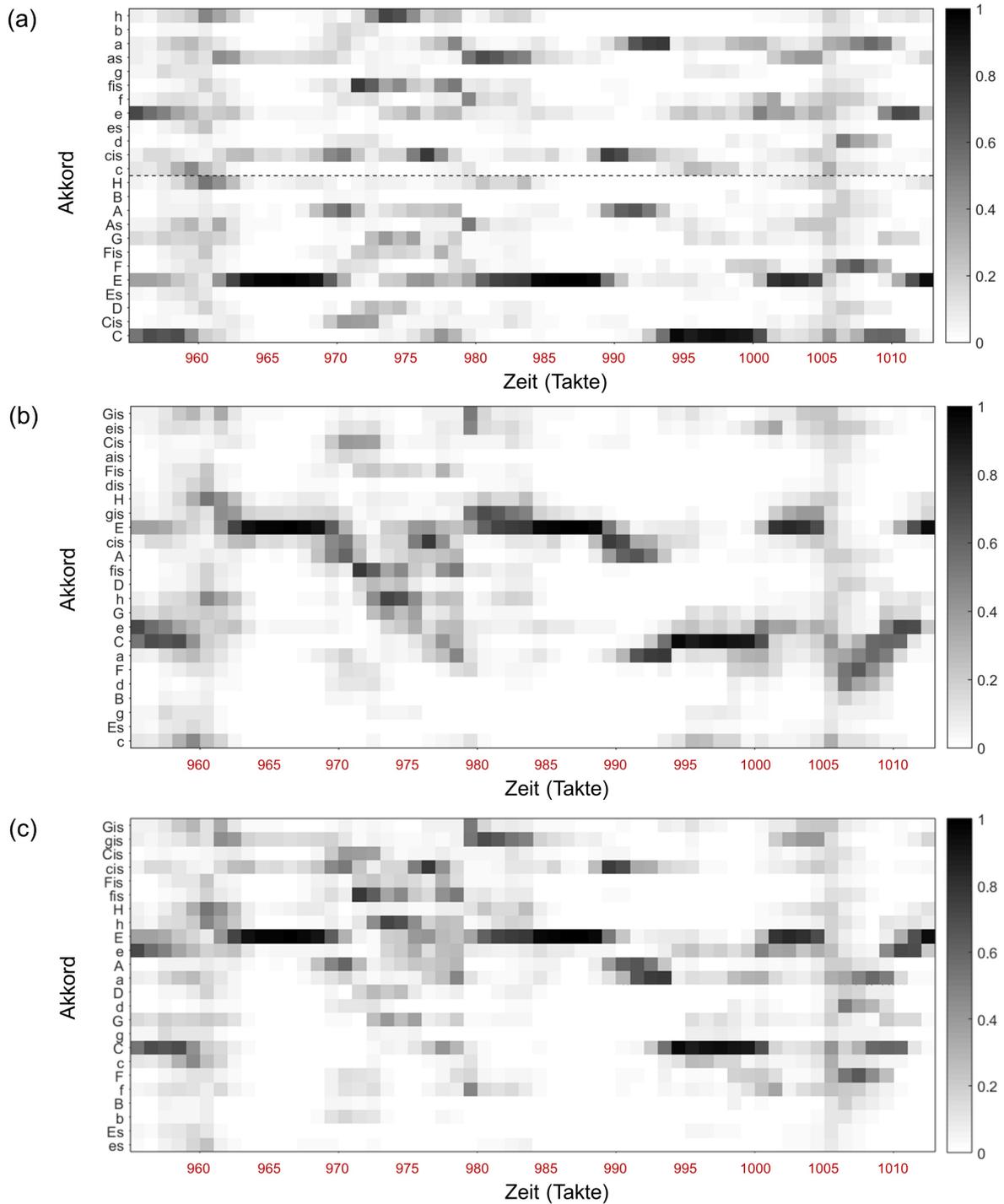
Abbildung 6 Klavierreduktion der Takte 955–1012 aus dem ersten Akt der *Walküre*.



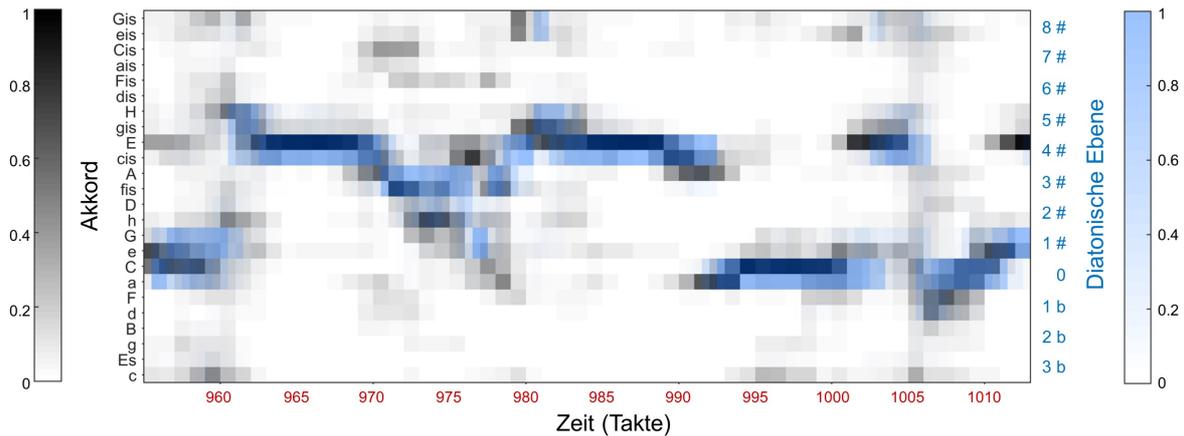
**Abbildung 7** Richard Wagner, *Die Walküre*, Akt 1, Takte 955–1005, Visualisierung diatonischer Ebenen (Karajan 1966). Die Fenstergröße beträgt 4 Takte, die Schrittweite 1 Takt. Die vertikale Achse wurde zur besseren Darstellung in Richtung der #-Vorzeichen verschoben.

Das erneute Verlassen der Tonart E-Dur führt schließlich in eine längere Präsenz der 0-Ebene (T. 990–1001), über die Akkorde A-Dur und a-Moll nach C-Dur. In dieser Tonart wird Wotans Hineinstoßen des Schwertes in die Esche geschildert. Als schließlich angedeutet wird, für wen das Schwert bestimmt ist, erfolgt eine kurze Rückmodulation nach E-Dur (T. 1002–1003), bevor die Ausgangstonart der Erzählung (e-Moll) wieder erreicht wird. Diese wird zuletzt nochmals für wenige Takte in Richtung a-Moll verlassen (T. 1008–1009). Die Erzählung endet mit einer Kadenz in e-Moll, welche jedoch nach E-Dur aufgelöst wird (T. 1012).

Die Gegenüberstellung der Skalen- und Akkorddarstellungen (Abbildungen 7 und 8) zeigt, dass beide Methoden mehrdeutige Stellen aufweisen. Die diatonische Analyse scheint zunächst robuster zu sein. Durch das größere Analysefenster (4 Takte) lässt sich hier der grobe Modulationsverlauf gut erkennen, was insbesondere im Hinblick auf weiträumige Analysen nützlich sein kann. Diese Darstellung gibt allerdings keinen Hinweis auf das Tongeschlecht. Im Gegensatz dazu ist die Akkordanalyse detailreicher, weist aber mehr »Grauwerte« auf. Um die Vorteile beider Darstellungen auszunutzen, können diese kombiniert werden, was in Abbildung 9 mittels visueller Überlagerung geschieht. Als Basis ist in Grauschattierungen die Akkordanalyse aus Abbildung 8b (benachbarte parallele Dreiklänge) dargestellt. Sie wird nun mit der diatonischen Analyse überlagert (Blau-Schattierungen), wobei die diatonischen Skalen den jeweiligen Dur- bzw. natürlichen Molltonleitern entsprechen, deren Tonikadreiklänge sie überlagert sind. Hier zeigt sich, dass die Akkordanalyse (in entsprechender Anordnung) grob einen ähnlichen Verlauf wie die Analyse der diatonischen Skalen ergibt. Kleinere horizontale Abweichungen (wie z. B. in T. 980–982) sind hierbei auf die unterschiedliche Fenstergröße zurückzuführen. Alternativ zu diesem Vorgehen könnten die beiden Analysemethoden auch algorithmisch kombiniert werden, was eine Skalendarstellung mit Dur- und Moll-Unterscheidung ermöglichen würde. Diese und weitere Modifikationen der Analysemethoden sollen in zukünftigen Studien realisiert werden.



**Abbildung 8** Richard Wagner, *Die Walküre*, Akt 1, Takte 955–1005, Akkordanalyse. Die Fenstergröße beträgt 2 Takte, die Schrittweite 1 Takt. Die Teilabbildungen (a), (b) und (c) sind nach unterschiedlichen Akkordbeziehungen angeordnet.



**Abbildung 9** Richard Wagner, *Die Walküre*, Akt 1, Takte 955–1005, Visualisierung von Akkorden (schwarze Helligkeitsskala) überlagert mit diatonischen Ebenen (blaue Helligkeitsskala). Die Fenstergröße beträgt 2 Takte für die Akkorde und 4 Takte für die diatonischen Ebenen, die Schrittweite ist jeweils 1 Takt. Die diatonischen Skalen sind hier in Bezug gesetzt zu den zugehörigen Tonikadreiklängen der jeweiligen Dur- und natürlichen Molltonleiter. Eine diatonische Ebene bezieht sich daher auf zwei (benachbarte) Akkorde.

## 5. Grenzen und Möglichkeiten

Wie sich an diesem Beispiel zeigt, scheint der Begriff der »dichterisch-musikalischen Periode« durchaus brauchbar für die Analyse der *Ring*-Musik zu sein. Zweimal wird die Grundtonart vom Text motiviert verlassen; zweimal führt der Umschlag der Empfindungen zu einer abrupten Modulation, die in anderem Kontext (z. B. einer Symphonie) möglicherweise etwas willkürlich erscheinen würde.<sup>22</sup>

Um die Struktur solcher tonaler Verläufe darzustellen, kann eine computergestützte Analyse mit entsprechender grafischer Visualisierung hilfreich sein. Die technischen Parameter der Methoden bieten hierbei Raum für spezielle Anpassungen und Verbesserungen. So kann nicht nur die grafische, sondern auch die »musikalische Auflösung« der Darstellung verändert werden, indem mehr oder weniger Takte in die lokale Berechnung einbezogen werden. Diese Adaptionen bieten Potenzial für die Analyse ganzer Akte, Opern oder des gesamten *Rings* im Sinne eines ersten Überblicks.

Darüber hinaus lassen sich auch die musiktheoretischen Annahmen auf bestimmte Modelle hin anpassen. Beispielsweise kann die Anordnung der Akkorde oder Tonarten verändert werden, um bestimmte Beziehungen oder Symmetrieachsen zu verdeutlichen. Des Weiteren kann die Struktur der zu analysierenden Zusammenklänge adaptiv verändert werden, um z. B. verminderte Septakkorde oder auch nicht diatonische Skalen einzuschließen. Wie bereits erwähnt wäre zudem die algorithmische Kombination von Akkorden und Skalen zu einer Tonartdarstellung eine mögliche methodische Erweiterung.

Insbesondere sind die Analysemethoden auf ihre Genauigkeit und Robustheit zu untersuchen. Für audiobasierte Analysen spielt hierbei die Invarianz der Ergebnisse von der speziellen Interpretation (Einspielung) der Opern eine entscheidende Rolle. Konz u. a. haben gezeigt, dass die Untersuchung der Konsistenz von Analyseergebnissen über verschiedene Einspielungen einer Sonate zu robusten Ergebnissen führen kann.<sup>23</sup> Diese *Cross-Version*-Strategie soll in den folgenden Arbeiten an den Teilen des *Rings* überprüft werden. Dabei sind insbesondere Abweichungen zwischen notenbasierten und audiobasierten Ergebnissen von Interesse (wie in Abbildung 4 angedeutet).

<sup>22</sup> Vgl. hierzu Kleinertz, »Richard Wagners Begriff der »dichterisch-musikalischen Periode«, S. 45–47.

<sup>23</sup> Konz u. a., »A Cross-Version Chord Labelling Approach«.

Für die Musikwissenschaft bieten solche automatisierten Verfahren die Möglichkeit, auch umfangreiche musikalische »Texte« auf bestimmte Strukturen hin zu untersuchen. Dies erlaubt zum einen, bekannte oder hypothetische Strukturen zu erkennen und zu bestätigen, zum anderen aber auch, bisher unbekannte Strukturen aufzudecken und damit neue Hypothesen zu bilden. Eine automatisierte Auswertung größerer Musikcorpora könnte durchaus zu einem Paradigmenwechsel innerhalb der Musikwissenschaft führen. Bisherige Methoden würden dadurch nicht überflüssig, könnten aber aufgrund eines überprüfbareren und experimentell wiederholbaren Befundes, der weiterhin stets der historisch-hermeneutischen Diskussion bedarf, neue Fragestellungen eröffnen.

### **Danksagung**

Die International Audio Laboratories Erlangen sind eine Gemeinschaftseinrichtung der Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU) und des Fraunhofer-Institut für Integrierte Schaltungen IIS. Die vorgestellte Arbeit wurde durch die Deutsche Forschungsgemeinschaft gefördert (DFG MU 2686/7-1, DFG KL 864/4-1).