Musik Frequenzen Formeln

Prof. i. R. Dr. Bruno Klingen, Fakultät für Informations-, Medien- und Elektrotechnik an der Technischen Hochschule Köln

Inhalt

Vorbemerkung

Die folgenden Ausführungen sollen eine Verbindung zwischen Musik und Mathematik herstellen, vermittelt durch die Frequenzen von Tönen.

Die Entwicklungen führen zu dem Ziel dieser Ausarbeitung, der Frequenzanalyse von Audiosignalen unter Verwendung der Fouriertransformation.

Für das Verständnis der Zusammenhänge ist es nicht erforderlich, die Formelumstellungen im Einzelnen nachzuvollziehen.

Zu den beigefügten Musikbeispielen - vorwiegend aus dem "Messiah" von Georg Friedrich Händel - lassen sich durch Links die entsprechenden Videos herunterladen.

1. Von Tönen zu Frequenzen

Die verbreitete rechnerische Grundlage eines musikalischen Stimmungssystems basiert auf sogenannten “gleichstufigen” oder “gleichtemperierten” Intervallen[1]. Üblicherweise geht man hierzu vom Kammerton a’ (“eingestrichenes” a) mit 440 Hz aus. Diese Schreibweise ist für die folgende Ausarbeitung weniger gut geeignet. Um eine systematische Nummerierung der Töne mit Bezug auf beliebige Oktaven zu benutzen, soll für den Kammerton die Notierung a_1 verwendet und andere Töne entsprechend geschrieben werden.

Es ist bekannt, dass die nächsthöhere Oktav, das zweigestrichene a_2, die doppelte Frequenz nämlich 880 Hz hat. Weitere Oktaven ergeben sich entsprechend durch Verdopplung oder Halbierung der Frequenzen. Anders als gebräuchliche Bezeichnungen (zum Beispiel “kleine Oktave”, “große Oktave”) werden hier alle Oktaven zum Kammerton a durchgehend durch den Buchstaben k gekennzeichnet mit der Schreibweise a_k. Zum Ton a_0 gehört folglich die Frequenz 220 Hz und zu a_{-1} die Frequenz 110 Hz.

Die 12 Halbtöne einer Oktav werden entsprechend der Tabelle 1 den Zahlen 1, 2, ... ,12 zugeordnet und zwar hier einheitlich mit den erhöhten Halbtönen :

c	cis	d	dis	e	f	fis	g	gis	a	ais	h
1	2	3	4	5	6	7	8	9	10	11	12

Tabelle 1: Töne und Tonnumern

Für die folgenden Entwicklungen ist es zweckmäßig, mit dem ersten Ton zur Oktav $k = 0$, nämlich c_0, zu beginnen. Die Frequenz hierzu wird durch das Symbol f_0 festgelegt. Zu der Frequenz[2]

$$f_0 = 130.81 \tag{1}$$

wird später die Übereinstimmung mit der Frequenz des Kammertons a_1 gezeigt.

Die Angabe der Frequenzeinheit Hz ist im weiteren entbehrlich, da sie in entsprechenden Zusammenhängen die einzig vorkommende Einheit zu Zahlenwerten ist.

Für die Frequenzen beliebiger Halbtöne der Oktav k und der Tonnumer n soll die Notierung $f_{k,n}$ benutzt werden. Damit können zum Beispiel die Oktaven zu c_0 (mit der Tonnumer $n = 1$) nach der genannten Verdopplungsvorschrift durch die Potenzformel

$$f_{k,1} = f_0 \cdot 2^k \tag{2}$$

berechnet werden. Für c_0 ergibt sich hieraus als Rechenbeispiel unter Anwendung der Potenzregel $s^0 = 1$ (zu einer beliebigen Basis s mit $s \neq 0$) in Übereinstimmung mit (1)

$$f_{0,1} = f_0 \cdot 2^0 = f_0$$

Die durch die Potenzen in (2) erfassten Oktaven k lassen sich analog auf die Halbtöne einer Oktav erweitern. Hierfür wird zu dem Exponenten k in (2) der Summand $(n - 1)/12$ für den n-ten

[1] vergl. Wikipedia “gleichstufige Stimmung”; in diesem Beitrag wird auch auf die Bedeutung von Andreas Werckmeister (1645 - 1706) und den von ihm eingeführten Begriff der “wohltemperierten Stimmung” hingewiesen, veröffentlicht 1707.

[2] Im Hinblick auf Dezimalbrüche in automatisch erstellten Tabellen dieser Ausarbeitung werden Dezimalzahlen hier in der angelsächsischen Dezimalpunktnotierung wiedergegeben.

Ton innerhalb der Oktav addiert. Daraus resultiert (unter Anwendung der Potenzregel $s^{r+t} = s^r \cdot s^t$ sowie der Wurzelschreibweise für gebrochene Exponenten)

$$f_{k,n} = f_0 \cdot 2^{k+(n-1)/12} = f_0 \cdot 2^k \cdot \sqrt[12]{2^{n-1}} \tag{3}$$

Auch hier ergibt natürlich der Fall des Tons c_0 mit $k = 0$ und $n = 1$ die Übereinstimmung

$$f_{0,1} = f_0$$

Ein weiteres Beispiel zur Formelauswertung lässt sich mit $n = 12$ in (3) durchführen, da sich nach 12 Schritten von der Tonnummer n aus derselben Ton in der nächsthöheren Oktav ergibt. Das führt (mit $s^1 = s$) auf:

$$f_{k,n+12} = f_0 \cdot 2^{k+(n+12-1)/12} = f_0 \cdot 2^{k+(n-1)/12+1} = 2 \cdot f_0 \cdot 2^{k+(n-1)/12} = f_{k+1,n}$$

Unter Anwendung von (3) wurden die in der Tabelle 2 wiedergegebenen Frequenzen zu den Tönen der Oktav $k = 0$ berechnet. Die dritte Zeile der Tabelle enthält die Zahlenwerte der Potenzen zur Basis 2 in (3) (gelegentlich auch als Frequenzfaktoren bezeichnet). Diese werden weiter unten benutzt.

1	2	3	4	5	6
130.81	138.59	146.83	155.56	164.81	174.61
1.000	1.059	1.122	1.189	1.260	1.335

7	8	9	10	11	12
185.00	196.00	207.65	220.00	233.08	246.94
1.414	1.498	1.587	1.682	1.782	1.888

Tabelle 2: Tonnummern, Frequenzen, Potenzfaktoren zur Oktav $k = 0$

Es ist noch die Kontrolle zu dem Kammerton a_1 (Tonnummer $n = 10$) in Verbindung mit der Festlegung (1) durchzuführen. Mit (3) wird hier (nach Runden)

$$f_{1,10} = f_0 \cdot 2^{1+9/12} = f_0 \cdot 2^{1.75} = 440$$

Ersetzt man den Exponenten in (3) durch die Variable x, dann erhält man eine von x abhängige Potenzfunktion:

$$\left.\begin{array}{r} y \\ f(x) \end{array}\right\} = f_0 \cdot 2^x \tag{4}$$

Die graphische Darstellung dieser Funktion ist in der Abbildung 1 wiedergegeben.

Durch die Gleichsetzung der Exponenten in (3) und (4), $x = k + (n-1)/12$, ergibt sich die Skalenteilung der horizontalen Achse oder "Abszisse" in der Graphik. Zwischen den Oktavnummern 0, 1, 2 befinden sich die jeweils 12 Halbtonunterteilungen entsprechend der Tabelle 1. Der linke Endpunkt des Funktionsgraphen gehört zu dem Ton c_0 mit der Frequenz $f_{0,1} = f_0$, die man auf der vertikalen Achse oder "Ordinate" ablesen kann. Die rechte Begrenzung führt mit dem Ton c_2 auf die Frequenz $f_{2,1} = f_0 \cdot 2^2 = 523.24$ (Anwendung von (2)). Zu den Tönen a_0 und a_1 (Tonnummern $n = 10$) lassen sich angenähert die Frequenz 220 und 440 ablesen.

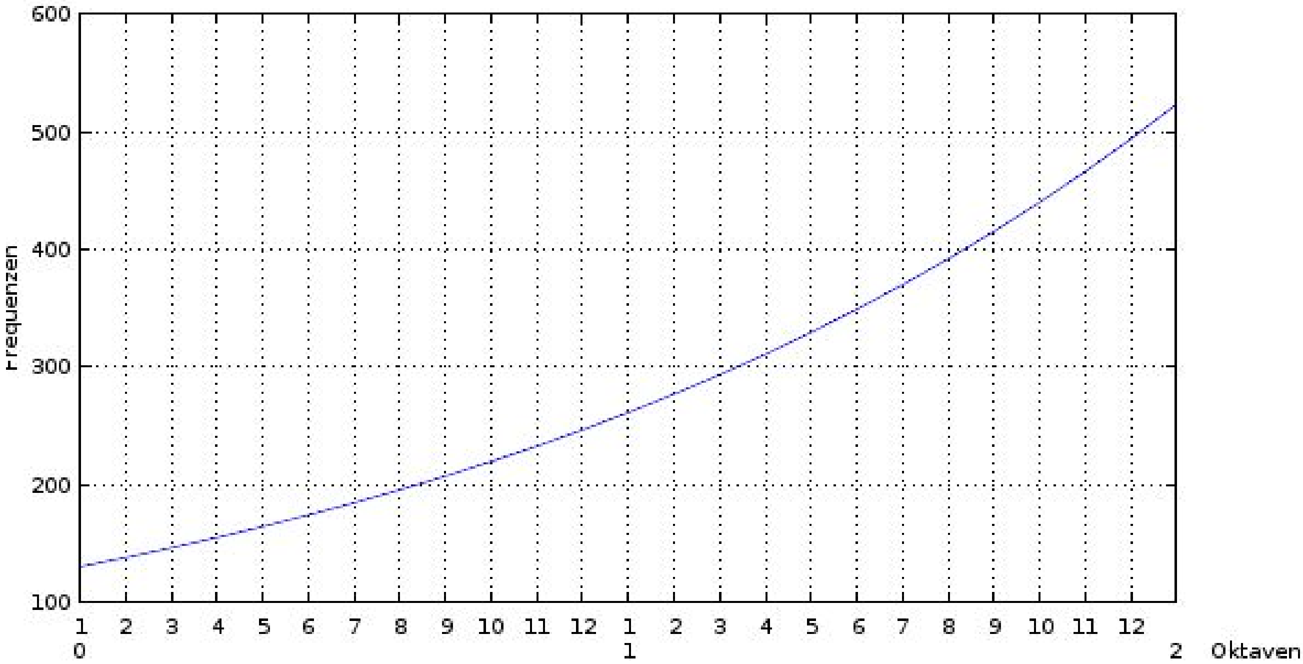


Abbildung 1: Frequenzen zu Oktaven und Tonnummern

Aus dieser Graphik wird deutlich, dass auch die x-Werte zwischen den Skalenteilungen der Abszisse einer Interpretation bedürfen. Das ist unter anderem Gegenstand des nächsten Abschnitts.

Weiterhin bleibt noch zu erwähnen, dass sich mit Potenzen wie in (3) beliebige Intervalle bilden lassen. So entstehen zum Beispiel mit den folgenden Gleichungen (5) die Quarte (fünf Halbtonschritte) beziehungsweise die Quinte (sieben Halbtonschritte) zum Ton mit der Frequenz f. Beide Potenzen zusammen ergeben die Oktav $2f$ wie in (6).

$$f_{\text{Quarte}} = f \cdot 2^{5/12}, \quad f_{\text{Quinte}} = f \cdot 2^{7/12} = f \cdot 1.4983071... \approx 1.5f \tag{5}$$

$$f_{\text{Oktav}} = f \cdot 2^{5/12} \cdot 2^{7/12} = f \cdot 2^{12/12} = 2f \tag{6}$$

Für die Quinte wurde in (5) als Näherung der Frequenzfaktor 1.5 anstelle des Wertes 1.498 aus Tabelle 2 Ton 8 eingesetzt. Im Vorgriff auf den im nächsten Abschnitt vorgestellten Centbegriff lässt sich zeigen, dass die Näherung von 1.498 durch 1.5 auf eine Frequenzdifferenz von 2 Cent führt und somit im Allgemeinen vernachlässigt werden kann. Zum Ton a_0 mit der Frequenz $f_{0,10} = 220$ ergibt sich zum Beispiel für die Quinte e_1 mit dem Frequenzfaktor 1.5 die Frequenz 330 anstelle von $329.62 = 2 \cdot 164.81$ nach Tabelle 2.

Abschließend wird noch der Begriff "Viertelton" vorgestellt. Der um einen Viertelton höhere Ton als der zu der Frequenz $f_{k,n}$ wird unter Verwendung der Gleichung (3) gebildet. Es ist in einem gleichtemperierten Stimmungssystems zwangsläufig, hierzu die Klammer im Exponenten von (3) durch (n + 1/2 - 1) zu ersetzen. Das führt mit (3) auf die Frequenz

$$f = f_0 \cdot 2^{k+(n+1/2-1)/12} = f_0 \cdot 2^{k+(n-1)/12} \cdot 2^{(1/2)\cdot(1/12)} = f_{k,n} \cdot 2^{1/24} \tag{7}$$

Eine weitere Multiplikation mit dem Vierteltonfaktor liefert dann

$$f_{k,n} \cdot 2^{1/24} \cdot 2^{1/24} = f_{k,n} \cdot 2^{2/24} = f_{k,n} \cdot 2^{1/12} = f_0 \cdot 2^{k+(n+1-1)/12} = f_{k,n+1}$$

Das ist also der auf dem Ton zur Frequenz $f_{k,n}$ folgende Halbton. Der Viertelton nach (7) liegt somit rechnerisch genau in der Mitte zwischen den Tönen zu $f_{k,n}$ und $f_{k,n+1}$.

2. Von Frequenzen zu Tönen

Löst man die Funktionsgleichung in (4) nach x auf und ersetzt y durch die Frequenz f, dann erhält man die Umkehrfunktion zur Exponentialfunktion in (4) durch die Logarithmusfunktion und zwar den Logarithmus zur Basis 2:

$$x = \log_2\left(\frac{f}{f_0}\right) = \frac{\log(f/f_0)}{\log 2} \tag{8}$$

Die zweite Formel kann benutzt werden, um mit einem Taschenrechner unter Verwendung der log-Taste den Wert x zu berechnen.

Beispiel 1 Zunächst wird die Frequenz zu dem Ton mit der Oktav $k = 2$ und der Tonnummer $n = 4$, das ist also nach der vereinbarten Notierung der Ton dis_2, mit (3) berechnet:

$$f_{2,4} = f_0 \cdot 2^{2+(4-1)/12} = f_0 \cdot 2^{2.25} = 622.24 \tag{9}$$

Durch die Oktavierung dieser Frequenz mit der Division $622.24/4 = 155.56$ bestätigt sich die Frequenz zu dem Ton dis_0 in der Tabelle 2.

Umgekehrt führt diese Frequenz $f_{2,4}$ unter Verwendung des letzten Produktes in (9) mit (8) (unter Benutzung des Folgerungspfeils $\Rightarrow$) zurück auf

$$x = 2.25 = 2 + 0.25 \;\Rightarrow\; k = 2 \;\text{ und }\; 0.25 \cdot 12 + 1 = 4 \;\Rightarrow n = 4$$

Hier entspricht die Multiplikation der Nachkommazahl 0.25 mit 12 und die Addition einer 1 der Fortsetzung des Umkehrvorgangs zu (3) mit Bezug auf den Exponenten.

Es soll nun die Berechnung von Tönen zu beliebigen Frequenzen entwickelt werden. Hierbei kommt die Funktion "floor" zur Anwendung, die in den gängigen Programmiersprachen verfügbar ist. Hierzu gehört auch die Sprache "Octave", mit der die Berechnungen zu dieser Ausarbeitung durchgeführt wurden. In der Form $y = \text{floor}(x)$ wird die Zahl x auf die nächstniedrige ganze Zahl abgerundet. Beispiele: floor(1.7) = 1, floor(-1) = -1 und floor(-1.7) = -2.

Die Schritte der Berechnungsoperationen werden im Folgenden erklärt. Hierbei wird der Begriff des (musikalischen) "Cent" eingeführt. Dazu wird ein Intervall $I = [f_{k,n}\ ,\ f_{k,n+1})$ verwendet, begrenzt durch die Frequenzen zu zwei benachbarten Tönen. Das Intervall I wird logarithmisch in 100 Cent-Teile zerlegt. Dieser Vorgang ist in der zweiten und den weiteren Zeilen des folgenden Schemas enthalten.

Anschaulich kann man sich die hochauflösende Skalierung des Centmaßes an der Abszisse der Graphik in Abbildung 1 Seite 5 verdeutlichen. Hierdurch wird das Intervall zwischen zwei aufeinander folgenden Tönen in 100 Teilintervalle zerlegt.

Beispiel 2 Es wird die Frequenz $f = 761$ gewählt. Unter Verwendung der Hilfsvariablen f_h und h_1, h_2, h_3 ergibt sich:

$$\left\{\begin{array}{l}
f_h = 761 \cdot 2^{1/24} = 783.3 \\
x = \log_2(f_h/f_0) = 2.582 \Rightarrow k = \text{floor}(x) = 2 \\
h_1 = (x - k) \cdot 12 = 6.985 \Rightarrow h_2 = \text{floor}(h_1) = 6 \Rightarrow n = h_2 + 1 = 7 \\
h_3 = (h_1 - h_2) \cdot 100 = 98.5 \Rightarrow \text{Cent} = \text{round}(h_3 - 50) = \text{round}(48.5) = 49
\end{array}\right\}$$

Damit erhält man $k = 2$ und $n = 7$. Die Frequenz des zu f benachbarten Ton fis_2 ist $f_{2,7} = 2^2 \cdot 185.00 = 740.00$ (vergl. Tabelle 2). Die Abweichung zwischen $f = 761$ und $f_{2,7}$ wird nach dem Berechnungsschema mit 49 Cent ausgewiesen. Hiernach liegt diese Frequenz f näher an der Frequenz zu dem Ton fis_2 als an der zu g_2. Der Abstand zu g_2 beträgt zwangsläufig 51 Cent.

Die Verwendung der Funktion floor in dem Schema ist vorteilhafter als die zunächst naheliegende Benutzung der Rundungsfunktion. Bei round wird man jeweils in der zweiten und in der dritten Zeile auf etwas verwirrende Fallunterscheidungen geführt. Die Multiplikation in der ersten Zeile mit dem Vierteltonfaktor (vergl. Seite 5) ist Teil des Berechnungsverfahrens mit floor. Dieser Faktor wird am Ende durch die Subtraktion von 50 bei dem Centwert wieder ausgeglichen.

Beispiel 3 $f = 128$:

$$\left\{\begin{array}{l} f_h = 128 \cdot 2^{1/24} = 131.75 \\ x = \log_2(f/f_0) = 0.010307 \Rightarrow \ \mathrm{f}k = \mathrm{floor}(x) = 0 \\ h_1 {=} (x \text{ - } k) \cdot 12 = 0.12368 \Rightarrow h_2 = \mathrm{floor}(h_1) = 0 \Rightarrow n = h_2 + 1 = 1 \\ \ h_3 = (h_1 \text{ - } h_2) \cdot 100 = 12.37 \Rightarrow \mathrm{Cent} = \mathrm{round}(-37.63) = -38 \end{array}\right\}$$

Der nächstgelegene Ton ist also c_0, und der Unterschied zwischen $f = 128$ und f_0 in (1) beträgt -38 Cent.

Die[3] *Differenz der linken Klammer in der vierten Zeile des Schemas,* $h_3 = (h_1 - h_2) \cdot 100 = (h_1 - \mathrm{floor}(h_1)) \cdot 100$, *kann minimal den Wert* 0 *und maximal* 99.99 *annehmen. Zusammen mit der Subtraktion von 50 in der letzten Zeile ergibt sich nach diesem Berechnungsschema die Begrenzung* $-50 \leq \mathrm{Cent} <$ *50.*

Bei den Ergebnissen der Umrechnung von beliebigen Frequenzen in Oktav- und Tonzahlen nach dem vorgestellten Schema lässt sich also für die maximale Abweichung der Centwerte festhalten

$$-50 \leq \mathrm{Cent} < 50$$

Zu Kontrollrechnungen kann die Frequenzformel (3) erweitert werden auf die Einbeziehung von Centwerten:

$$f_{k,n,\mathrm{Cent}} = f_0 \cdot 2^{k+(n-1+\mathrm{Cent}/100)/12} \tag{10}$$

Diese Formel liefert für die Werte des Beispiels 3 ($k = 0$, $n = 1$, Cent = -37.63) wieder die Frequenz $f = 128$.

[3] Die in Kursivschrift wiedergegebenen Textteile sind für das Gesamtverständnis entbehrlich.

3. Frequenzanalyse

Musik wird heute fast ausschließlich in digitalisierter Form gespeichert und zur Hörwiedergabe verwendet. Ein charakteristisches Beispiel hierfür ist die CD. Eine solche Datensammlung eines Audiosignals, das man auch Sample nennt, enthält die abgetasteten Zahlenwerte des ursprünglichen analogen Signals.

Für die Bearbeitung solcher digitalen Audiosamples in Computerprogrammen ist das Format der wav-Datei die geeignete Form. CD-Samples lassen sich beispielsweise durch das Programm "Audiograbber" in wav-Dateien überführen.

3.1. Fouriertransformation

Eine vorrangige Möglichkeit der Untersuchung von Audiosamples ist die Frequenzanalyse. Damit lässt sich ein Audiosignal in die in ihm enthaltenen Frequenzkomponenten zerlegen. Die hierbei angewandte Methode ist die "diskrete Fouriertransformation", kurz DFT (zu den Transformationsformeln vergl. Internet unter "DFT"). Daraus wurde ein beschleunigtes Verfahren entwickelt, die sogenannte "fast fourier transformation" (FFT). Diese führt die Transformation auch bei größerem Datenumfang auf einem PC in kaum wahrnehmbarer Zeit aus.

Solche Berechnungen können beispielsweise mit dem kostenlosen und benutzerfreundlichen Programmiersystems Oktave durchgeführt werden. Mit diesem Programmpaket ist es möglich, wav-Dateien als gewöhnliche Arrays zu verarbeiten, und die FFT ist als Unterprogramm verfügbar.

Es wird davon ausgegangen, dass als Abtastresultat ein digitales Monosignal mit N Zahlenwerten vorliegt. Zur Veinfachung wird N als geradzahlig vorausgesetzt und hieraus $n = N/2$ gebildet. Speziell für einen Audioausschnitt mit einer Dauer von 1 Sekunde wird der Wert von N "Abtastrate" oder "Samplerate" genannt und mit s_r bezeichnet. Für die CD ist der Normwert $s_r = 44100$.

Das Ergebnis der FFT ist eine Folge von ebenfalls N Zahlen, die als Fourierkoeffizienten bezeichnet werden und hier durch die Variablen u_k dargestellt werden sollen[4]. Als Indizes wählt man üblicherweise die Zahlen $k = 0, 1, 2, \ldots, N - 1$.

Die Werte der u_k sind in der Regel komplexe Zahlen. Es lässt sich zeigen, dass die erste Hälfte der Fourierkoeffizienten mit der zweiten insofern übereinstimmt, als jeweils entsprechende Zahlen zueinander konjugiert komplex[5] sind, präziser:

$$u_{n-k} = \overline{u}_{n+k}, \quad k = 0, 1, 2, \ldots, n - 1 \tag{11}$$

Aus diesen Koeffizienten lassen sich die Amplituden zu den Frequenzen durch die Betragswerte $v_k = |u_k| = abs(u_k)$ berechnen.

Für die Frequenzanalyse interessieren vorrangig die positiven Amplituden v_k zu den Frequenzen des Signals. Für diese Amplitudenwerte gilt (wegen (11))

$$v_{n-k} = v_{n+k}, \quad k = 0, 1, 2, \ldots, n - 1 \tag{12}$$

Hiernach werden für die Auswertung der Fouriertransformation nur die $v_k,\ k = 0, 1, 2, \ldots, n$, verwendet. Speziell aus der Variablen v_n lassen sich keine Folgerungen ziehen, und v_0 enthält

---

[4] Mit der "inversen Fouriertransformation" lassen sich die ursprünglichen Abtastwerte exakt zurückgewinnen. Zumeist werden allerdings zuvor durch Manipulation der u_k-Werte gezielt Klangeffekte erreicht, beispielsweise Tiefpassfilter.

[5] *Zur komplexen Zahl $z = a + b\,\mathrm{i}$ mit der imaginären Einheit i lautet die konjugiert komplexe Zahl: $\overline{z} = a - b\,\mathrm{i}$. Zwischen diesen beiden besteht die Beziehung $|z| = |\overline{z}|$.*

die Summe aller Werte des Audiosamples, die für die Untersuchung des Samples in der Regel bedeutungslos ist. Häufig wird daher $v_0 = 0$ gesetzt, wie es auch in den folgenden Beispielen geschieht. Für die praktische Arbeit mit der Frequenzanalyse kommen daher im Weiteren nur die Werte v_k mit den Indizes $k = 1, 2, ...n - 1$ zur Anwendung.

Beispiel 4 Es werden zwei zeitabhängige Sinusschwingungen mit den Frequenzen $\nu_1 = 2$ Hz und $\nu_2 = 3$ Hz sowie unterschiedlichen Koeffizienten additiv überlagert (der griechische Buchstabe ν, gesprochen nü, wird häufig in physikalischen Formeln für die Frequenz verwendet). Die Funktion der additiven Überlagerung lautet[6]:

$$g(t) = v_1 \cdot \sin(2\pi\, \nu_1 t) + v_2 \cdot \sin(2\pi\, \nu_2 t) \tag{13}$$

In den Funktionsgraphen der Abbildung 2 ist ein Ausschnitt von 1 Sekunde zu dieser Funktion wiedergegeben. Der erste Summand der Funktion ist durch die Farbe rot (mit $\nu_1 = 2$ Perioden/Sekunde und $v_1 = 5$), der zweite Summand durch grün (mit $\nu_2 = 3$ Perioden/Sekunde und $v_2 = 2.5$). Die Summenfunktion $g(t)$ ist in blau dargestellt. Zusätzlich ist eine (gleichabständige) Abtastung der Funktion mit $s_r = 16$ eingezeichnet.

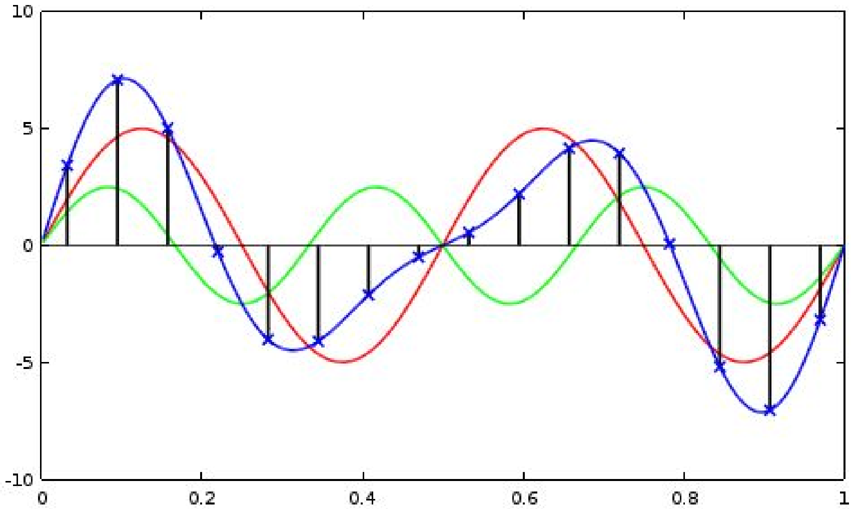


Abbildung 2: Summenfunktion zweier Sinusfunktionenen

Die Faktoren bei den Sinusfunktionen bestimmen die Lautstärke, mit der die Schwingung der entsprechenden Frequenz am gesamten Audiosignal beteiligt ist. Mit Frequenzen von 2 Hz bzw. 3 Hz gehört dieses Signal zu $g(t)$ natürlich zum tiefen Infraschallbereich.

Die Werte v_k der 16 Fourierkoeffizienten zu diesem Sample sind in der folgenden Abbildung 3 dargestellt. Hier ist anschaulich die charakteristische Eigenschaft der v_k-Werte erkennbar nämlich die Symmetrie der v_k, die als Formel in den Gleichungen (12) ausgedrückt ist. Diese Symmetrie ist bezogen auf die Stelle zu dem Wert $n = N/2$ (hier für $n = 8$). Somit ergibt sich die übliche Darstellung des Amplitudenspektrums einer Frequenzanalyse, wie sie mit Bezug auf das Beispiel 4 in der Abbildung 4 durch Halbierung des Graphen zur Abbildung 3 wiedergegeben ist.

[6] Die Formulierung des Argumentes $2\pi\nu t$ in der Summandenfunktion $\sin(2\pi\nu t)$ bewirkt, dass gerade ν Perioden in der Zeitspanne von 1 Sekunde auftreten.

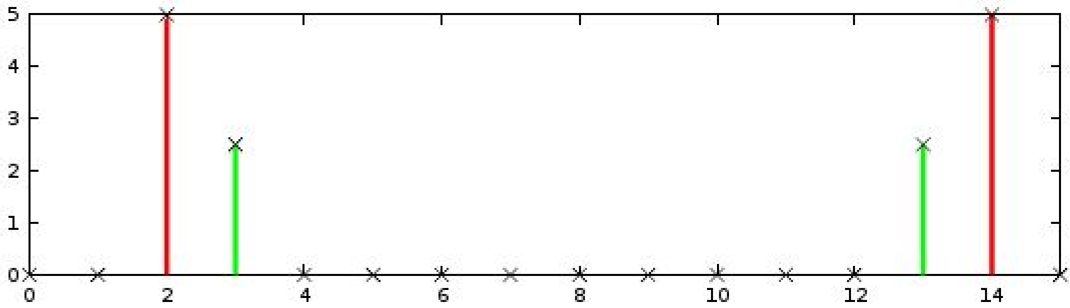


Abbildung 3: Frequenzzerlegung der Summenfunktion

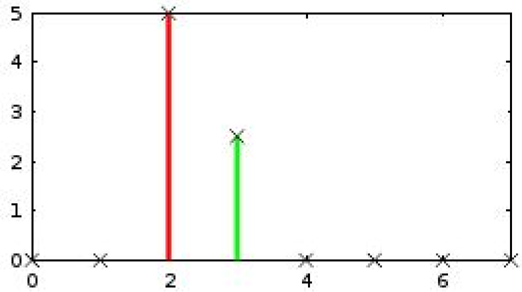


Abbildung 4: relevante Frequenzen

Hinweise 1:

1. Für ein Audiosample mit der Dauer von 1 Sekunde (wie in Beispiel 4 mit den Abbildungen 2 bis 4) besteht allgemein der Zusammenhang, dass die in dem Sample enthaltene Komponente zu der Frequenz ν identisch ist mit dem Index k der Amplitude v_k: $\nu = k$.

2. Zu den in der fogenden Abbildung 4 dargestellten Fourierkoeffizienten der Funktion $g(t)$ mit der zugehörigen Abbildung 2 ist noch anzumerken: Die Amplitudenwerte größer als Null aus der Frequenzanalyse ergeben sich in dieser Form prinzipiell nur dann, wenn die zugehörigen Frequenzkomponenten in dem Audiosample exakt mit einer ganzzahligen Anzahl von Perioden enthalten sind. Das wird in dem Beispiel 4 grundlegend deutlich. Als Gegenbeispiel zeigen die graphischen Darstellungen der folgenden Abbildung 5 die Ergebnisse der Frequenzanalyse zur Frequenz $\nu = 2.5$. Die Summe der Amplitudenwerte (rechte Abbildung) ist dann gleich der Amplitude der Sinusschwingung (linke Abbildung, hier = 1). Solche Effekte haben allerdings in üblichen Audiodateien nur geringe Bedeutung.

Die in (13) durch Sinusfunktionen eingeführten Frequenzkomponenten sollen durch Kosinusterme erweitert werden. Außerdem werden die komplexen Fourierkoeffizienten u_k zerlegt in den Realteil mit der Variablen r_k und dem Imaginärteil mit s_k, das heißt

$$u_k = r_k + \mathrm{i} \cdot s_k$$

Innerhalb des gesamten Audiosignals lautet dann die reellwertige Frequenzkomponente zu der Frequenz $\nu = k$:

$$g_k(t) = r_k \cdot cos(2\pi kt) + s_k \cdot sin(2\pi kt), \quad k = 1, 2, \ldots, n-1 \tag{14}$$

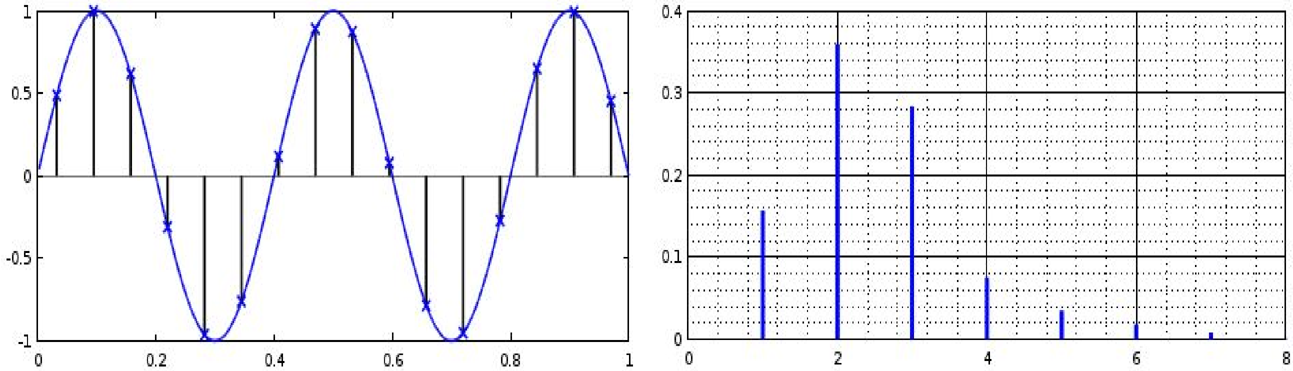


Abbildung 5: nicht ganzzahlige Periode in einem Sample

Umgekehrt liefert die Summe aller zeitabhängigen Frequenzkomponenten in (14) *als Ergebnis ein* <u>*analoges*</u> *Klangereignis zu dem abgetasteten Audiosignal.*

<u>Hinweise 2:</u>

1. Die graphische Darstellung des Wellenverlaufs eines Audiosamples liefert in der Regel wenig Informationen. Im ausgewählten Falle eines einzelnen Tones mit der Frequenz ν_1 wie in dem ersten Summand zur Funktion (13) *mit der Abbildung 2 Seite 10 kommt man zu einem einfachen und überschaubaren Wellenverlauf. Wenn bei demselben Zeitausschnitt die zugehörigen Oktaven dargestellt werden sollen mit den Frequenzen $\nu_1 \cdot 2^k$, dann ist wegen der exponentiellen Zunahme der Frequenzen nach wenigen Oktaven nur noch eine ausgefüllt Fläche zu sehen.*

Hier kommen die besonderen Möglichkeiten der Fouriertransformation mit der graphischen Frequenzdarstellung und den Amplituden darüber zum Tragen, wie sie in den Abbildungen 3 und 4 auf Seite 11 für den einfachen Fall der Funktion (13) *vorgestellt sind. Die Skalenteilungen der Abszisse zu den Frequenzen lassen sich allgemein flexibel an das jeweilige Audiosample anpassen. Beispiele hierzu finden sich in den Frequenzanalysen des nächsten Abschnitts.*

2. Der Algorithmus der Fouriertransformation zur Frequenzanalyse lässt sich als eine mathematische Entsprechung des Hörvorgangs im Ohr interpretieren. Im Innenohr befinden sich weit über 10000 Haarzellen unterschiedlicher Länge[7]. Jedes einzelne Haar besitzt eine von der Haarlänge abhängige Resonanzfrequenz. Ist diese Frequenz in einer eintreffenden Schallwelle enthalten, dann resultiert daraus eine Resonanzschwingung dieses Haares. Der Vorgang wird dadurch an entsprechende Neuronen des Gehirns weitergeleitet.

Bei der Fouriertransformation kommen diese Resonanzfrequenzen in den Amplitudengrößen der Fourierkoeffizienten zum Ausdruck. Die folgenden Abbildungen bieten hierzu Beispiele.

3. Im Weiteren wird für die Frequenz wieder der Buchstabe f verwendet. Der Buchstabe ν kommt vorwiegend in Verbindung mit sin- und cos-Ausdrücken - wie in diesem Abschnitt - vor.

3.2. Berechnungsergebnisse zu Frequenzanalysen

Es werden verschiedene Beispiele zu Klängen mit Obertönen vorgestellt. Durch Instrumente oder vokal erzeugte Töne enthalten durchweg zusätzlich zu dem hörbaren Ton einer bestimmten Frequenz auch Unter- und Oberschwingungen. Diese können als halbe sowie als doppelte, dreifache

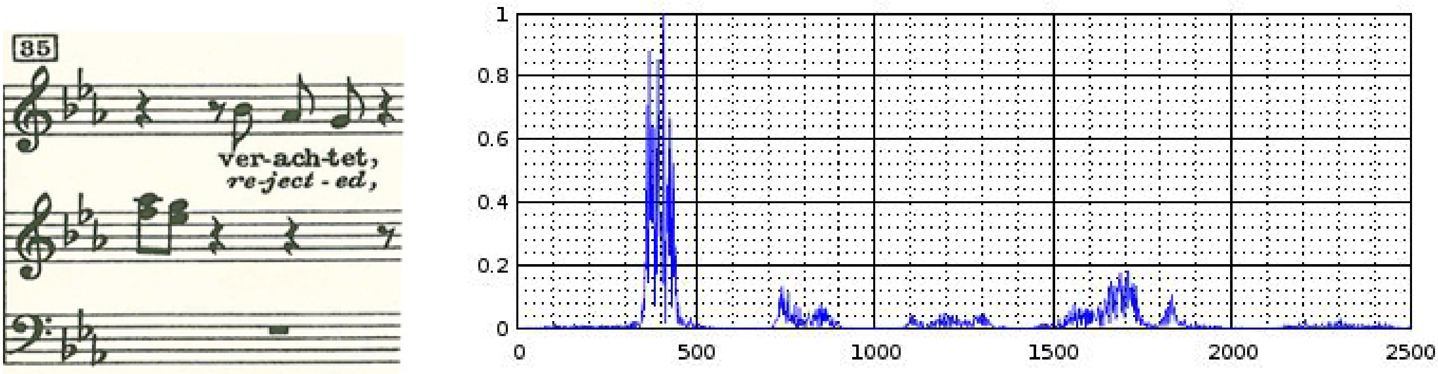
[7] vergl. im Internet “Anzahl Haarzellen im Innenohr”

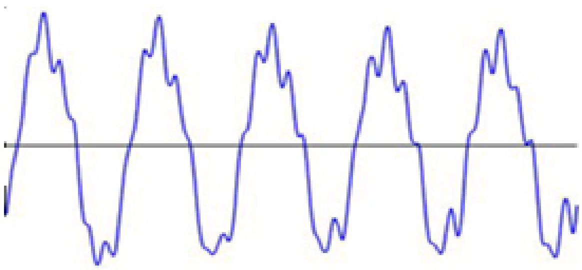
und mehrfache Frequenz der Hauptfrequenz vorkommen und zwar mit geringeren Amplituden. als die des hörbaren Tons (vergl. die folgenden graphischen Darstellungen als Ergebnisse der Frequenzanalysen).

Die relativen Amplitudenhöhen der Unter- und Obertöne (man nennt sie “Formanten”, vergl. Wikipedia) bestimmen maßgeblich die Klangfarbe einer Stimme oder eines Musikinstrumentes.

Beispiel 5 Der folgenden Abbildung 6 liegt aus dem Oratorium “Messiah” von Georg Friedrich Händel die Alt-Arie Nummer 21 “He was despised” zugrunde. Es handelt sich um eine Aufnahme aus Spanien mit dem Chor und Orchester “Ad Libitum” unter der Leitung von Francesc Gamon. Für die Frequenzanalyse wurde ein Sample zu Takt 35, Ton g_1, mit einer Sampledauer = 1/2 Sekunde gewählt (vergl. die Noten in Abbildung 6).

Das Notenblatt zu dem Ausschnitt in Abbildung 6 findet sich auf Seite 28. Hier lässt sich über einen Link das Video aus YouTube aufrufen. Der bei der vorliegenden Frequenzanalyse gewählte Ausschnitt ist nach 3 Minuten, 17 Sekunden zu hören. Entsprechendes trifft auch für die weiteren Notenbeispiele zu den Abbildungen 7, 11 und 15 zu.





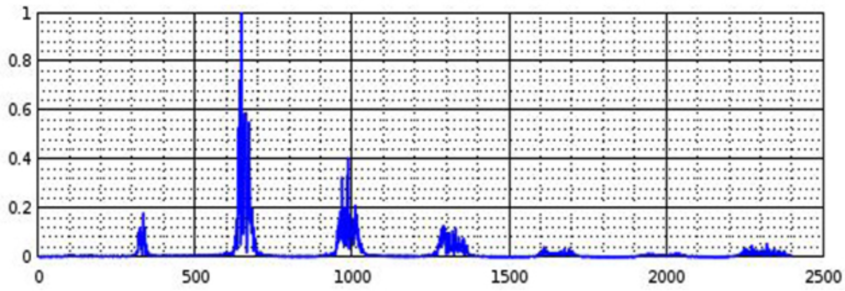
Frequenz	Amplitude	Ton
392	1.00	g1
776	0.09	g2
1196	0.06	d3
1558	0.09	g3

Abbildung 6: Messiah: “He was despised”

Durch das Vibrato der Altstimme werden zusätzlich benachbarte Frequenzen zu dem Grundton und seinen Obertönen ausgewiesen.

Die Wellenform der Graphik unten links in der Abbildung 6 stellt einen Ausschnitt aus dem Audiosample mit 565 Abtastwerten dar. Der hörbare Ton zu der Frequenz $g_1 = 392$ aus der Tabelle dieser Abbildung führt mit der Abtastrate auf (gerundet) $s_r/g_1 = 113$ Abtastungen pro Periode der Wellenform. Damit ergibt sich für die fünf dargestellten Perioden in der Abbildung die Übereinstimmung durch $5 \cdot 113 = 565$ Abtastungen.

Beispiel 6 Für die Abbildung 7 wird aus Messiah die Tenor-Arie Nummer 1, “Comfort ye”, verwendet, Takt 8 Fermatenton e_2, Sampledauer = 1 Sekunde Der Audioausschnitt entstammt dem entsprechenden Video mit “Ad Libitum” (vergl. das Notenblatt Seite 29, Audiosamle 43 Sekunden nach Start). Die deutlich ausgeprägten Frequenzstreuungen bei den Amplitudenmaxima in der Abbildung 7 gehen auf das Vibrato des Solisten zurück. Der Vergleich dieser Abbildung mit dem vorigen Beispiel zeigt, dass hier ein Frequenzspektrum mit Unterton und einer vielfältgen Obertonreihe vorliegt.



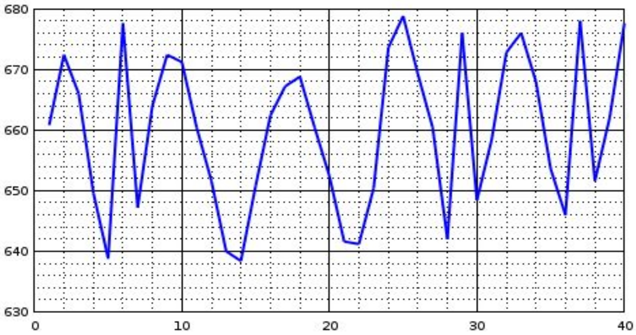
Frequenz	Amplitude	Ton
336	0.18	e1
648	1.00	e2
988	0.41	h2
1292	0.13	e3
1611	0.04	g3
2320	0.06	d4

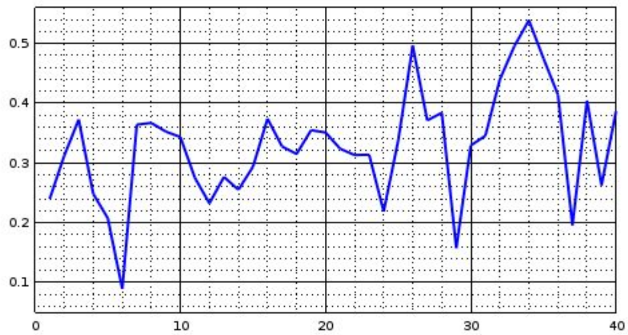
Abbildung 7: Messiah: “Comfort ye”

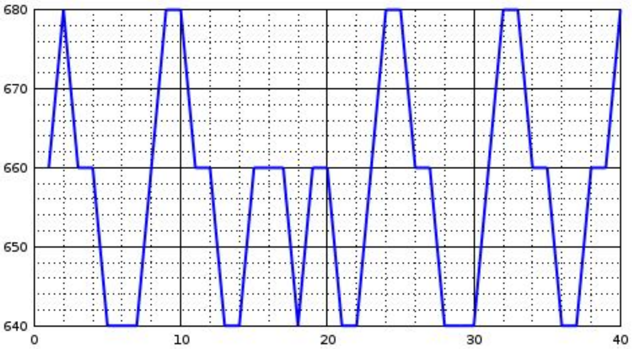
In der folgenden Abbildung 8 geht es um eine höhere Auflösung des Vibratoverlaufs. Aus dem Sample der Abbildung 7 (Samplelänge = 44100) wurden 40 aufeinander folgende äquidistante Teilausschnitte der Länge 2205 (Dauer 1/20 Sekunde) herausgenommen. Die zu diesen Untersamples gehörenden Frequenzanalysen sind durch das in Anhang 1 beschriebene Interpolationverfahren auf eine Frequenzgenauigkeit von 0.1 Hz präzisiert worden. Die Ergebnisse sind in den beiden oberen Graphiken der Abbildung 8 wiedergegeben.

An der Ordinate der beiden linken Graphiken in dieser Abbildung ist beispielsweise zu erkennen, dass die Frequenzbreite des Vibratos etwa 40 Hz beträgt; das ist in diesem Oktavbereich ungefähr plus/minus ein Viertelton. Die Vibratofrequenz selbst lässt sich mit etwa 8 Hz [8] ablesen (8 Perioden der Frequenzschwankungen in 1 Sekunde, vergl. Graphik oben links zur Abbildung 8).

[8] vergl. hierzu den Wikipedia-Eintrag “Vibrato”





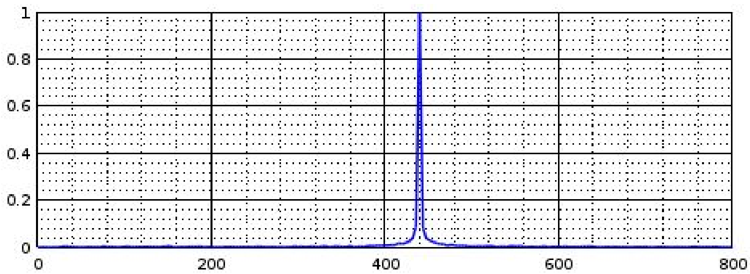


Nebenstehende Graphik: Frequenzergebnisse der Frequenzanalysen zu der darüber liegenden Abbildung bei einem Wert des Erweiterungsfaktors = 1. In Anhang 1 werden die hier auftretenden Frequenzsprünge von 20 Hz erklärt.

Abbildung 8: "Comfort ye" oben: Vibrato-Verlauf, links: Frequenzen, rechts: Amplituden

Beispiel 7 Ein klassisches Gegenbeispiel zu einem Klang mit Obertönen liefert die Stimmgabel, die eine reine Sinusschwingung der Frequenz $f = 440$ Hz erzeugt, wie die Abbildung 9 Stimmgabel mit den Tabellenwerten zeigt[9].

Die Resultate der Berechnungen zur Stimmgabel sind übrigens als Kalibrierung der diffizilen Programmierung einer Frequenzanalyse geeignet.



Frequenz	Ton	Centabweichg.
440.00	a1	0

Abbildung 9: Frequenzanalyse Stimmgabel

Bei diesem Beispiel und in vergleichbaren Fällen ist die Länge des Samples zu beachten. Hier wurde $N = 11025 = s_r/4$ *gewählt. Davon abweichende Zahlenwerte führen in der Regel zu Frequenzresultaten, die von* 440 *abweichen können. So ergibt ein Audiosample zu einem Ton mit* 440

[9] Audio Datei zur Frequenzanalyse: Auschnitt aus https://www.youtube.com/watch?v=BtYsR9Pe0Yk

Hz und einer Samplelänge von $N = 11075$ *das Berechnungsergebnis der Frequenzanalyse* 438. *- Mit der in Anhang 1 vorgestellten Formel lässt sich dieser Zusammenhang rechnerisch belegen.*

Hinweise 3:

1.) Verändert sich das Audiosignal innerhalb des untersuchten Samples (Beispiel: Koloratur-Ausschnitt), dann liefert das Ergebnis der Fourieranalyse sämtliche Frequenzen aller auftretenden Klänge wie in Beispiel 6, Abbildung 7. In Abhängigkeit von dem jeweiligen Audiosignal kann es zweckmäßig sein, Ausschnitte mit einer Dauer von etwa 1/4 Sekunde oder kürzer zu wählen. In den weiteren Beispielen sind nur solche mit konstantem Klang aufgenommen.

2.) Die graphischen Darstellungen von Frequenzanalysen sind in der Regel "normiert" wie in den bisherigen Abbildungen 6, 7 und 9 dargestellt. Das bedeutet, dass sämtliche Amplituden der jeweiligen Abbildungen mit demselben Faktor so multipliziert werden, dass die maximale Amplitude den Wert 1 erhält[10]. Dadurch lassen sich bei Bedarf Berechnungsergebnisse miteinander vergleichen.

Die Audiobeispiele der folgenden Seite mit den Frequenzspektren diverser Musikinstrumente stammen aus YouTube. Hier befindet sich über dem jeweiligen Instrumentennamen ein Link zu dem entsprechenden Video. Die Zeitangabe in der Klammer nach dem Namen gibt den Start des Samples innerhalb des Videos an, zu dem der solistische Instrumententon zu hören ist. Die Tonhöhe ist in der Tabelle durch die maximale Amplitude 1 erkennbar; die Sampledauer beträgt jeweils 0.3 beziwhungsweise 0.4 Sekunden.

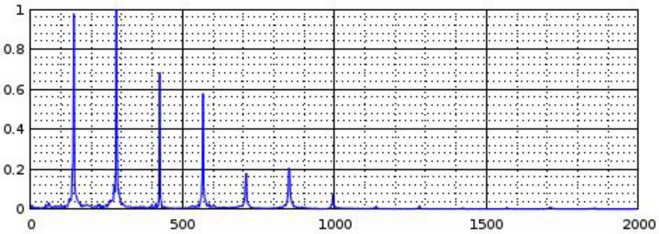
Bei dem ersten Beispiel handelt es sich um eine Monoaufnahme mit Sviatoslav Richter: Präludium und Fuge cis-moll aus dem "Wohltemperierten Klavier", BWV 849. Die Frequenzanalyse bezieht sich auf den erste Ton der Fuge.

[10] *Berechnung in zwei Schritten*

$$v_{\max} = \max(v_k), \quad k = 1, 2, \dots, n-1$$

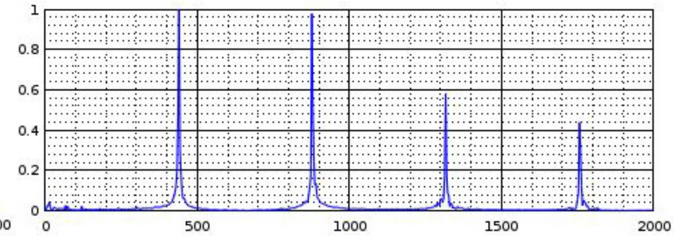
$$v_{k,neu} = v_k \cdot \frac{1}{v_{\max}}, \quad k = 1, 2, \dots, n-1$$

Abbildung 10: Frequenzanalysen



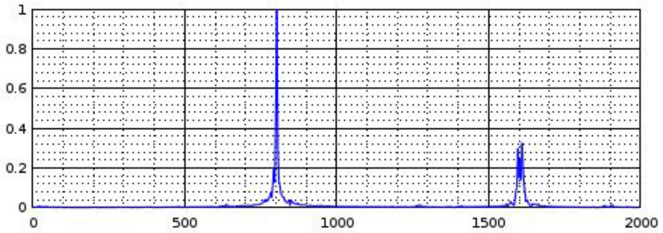
Frequenz	Amplitude	Ton
142	0.98	cis0
282	1.00	cis1
425	0.68	gis1
568	0.58	cis2

www.youtube.com/watch?v=ugSXVymv6b8
Klavier (nach 3.14 Min.)



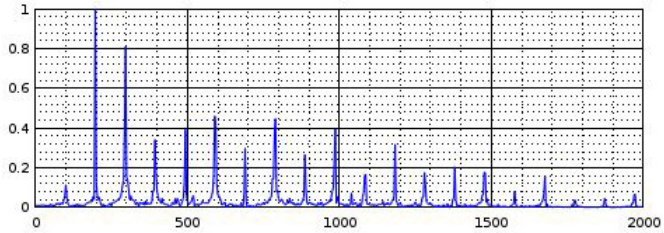
Frequenz	Amplitude	Ton
440	1.00	a1
877	0.98	a2
1317	0.58	e3
1757	0.44	a3

www.youtube.com/watch?v=E-pIudhSCSg
Oboe (nach 0.23 Min.)



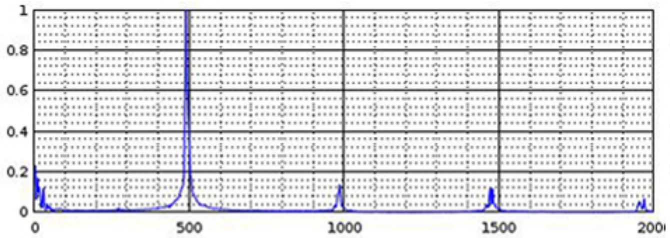
Frequenz	Amplitude	Ton
802	1.00	g2
1611	0.43	g3

www.youtube.com/watch?v=ZUZYoVw7moc
Trompete (nach 6.33 Min.)



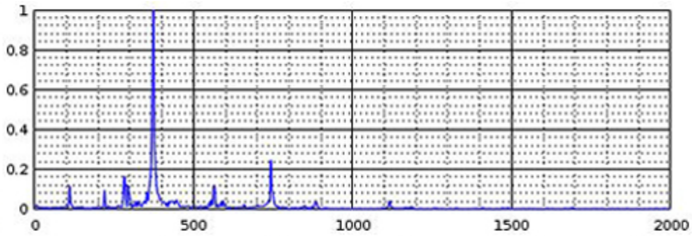
Frequenz	Amplitude	Ton	Frequenz	Amplitude	Ton
100	0.11	g-1	987	0.40	h2
197	1.00	g0	1087	0.17	cis3
297	0.81	d1	1183	0.32	d3
393	0.34	g1	1280	0.18	dis3
493	0.39	h1	1380	0.20	f3
590	0.46	d2	1477	0.18	fis3
690	0.30	f2	1577	0.08	g3
790	0.45	g2	1677	0.16	gis3
887	0.27	a2			

www.youtube.com/watch?v=1zgDS6SgK9s
Posaune (nach 0.35 Min.)



Frequenz	Amplitude	Ton
490	1.00	h1
987	0.14	h2
1473	0.12	fis3
1970	0.07	h3

www.youtube.com/watch?v=3VOkrddp6M8
Violine (nach 3.35 Min.)



Frequenz	Amplitude	Ton
283	0.17	cis1
373	1.00	fis1
563	0.12	cis2
743	0.25	fis2

www.youtube.com/watch?v=lNuJVfe-t3o
Horn (nach 6.14 Min.)

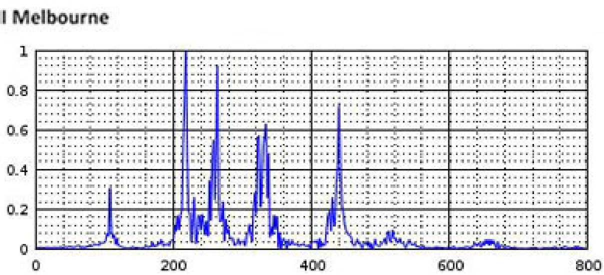
Bemerkenswert ist das vielfältige Frequenzspektrum der Posaune.

3.3. Frequenzstreuung

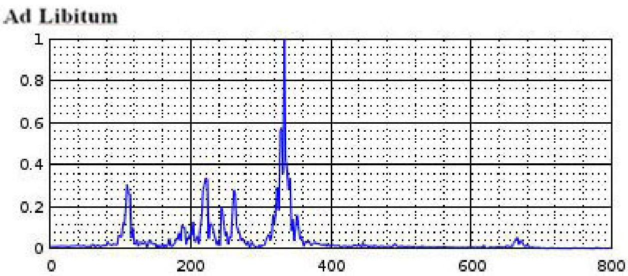
Die graphischen Darstellungen der Frequenzanalysen in den Abbildungen 6 (Seite 13), 9 (Seite 15) und 10 (Seite 17) weisen durchgehend eine Frequenzstreuung um die Frequenzen mit den großen Amplituden auf. Diese Streuungen sind dem Erscheinungsbild nach von anderer Art als die der Abbildung 7 (Seite 14) zu einem Vibrato. Die Form der Amplitudenverläufe zu den eingangs genannten Abbildungen in der Umgebung der Maxima lässt mit ihrer Annäherung an eine Gaußsche Glockenkurve auf stochastische Ursachen schließen. Somit bietet sich als Maßzahl zur Charakterisierung dieser Streuungen die Standardabweichung[11] an. In die Berechnungen dieser Maßzahl gehen jeweils die Amplitudenwerte der Umgebungen zu den Amplitudenmaxima ein[12].

Für die Abbildung 11 kam aus Messiah die Nummer 44 "Since by man came death" zur Anwendung. Im ersten Beispiel sind die Ergebnisse eines Laienchors wiedergegeben[13]. Für das zweite und dritte Beispiel wurden nacheinander einsetzend Orchester und Chor von "Ad Libitum" gewählt entsprechend den Noten in Abbildung 11. Auf Seite 30 befindet sich das Notenbild zu den ersten sechs Takten dieser Nummer mit einem Link zum Video in Youtube.

Frequenz	Ton	Stand.-abw.	MiWe=
			0.153
334	e1	0.173	
264	c1	0.158	
218	a0	0.217	
108	a-1	0.065	



Frequenz	Ton	Stand.-abw.	MiWe=
			0.097
334	e1	0.178	
266	c1	0.025	
222	a0	0.092	
110	a-1	0.093	



Frequenz	Ton	Stand.-abw.	MiWe=
			0.075
442	a1	0.006	
332	e1	0.024	
222	a0	0.161	
110	a-1	0.111	

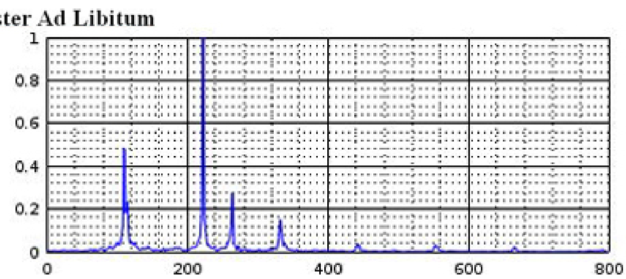


Abbildung 11: Instrumental- und Chor-Samples

11 *Die Standardabweichung wird durch die Wurzel aus der Varianz* s^2 *nach der üblichen Formel berechnet (vergl. zum Beispiel http://www.frustfrei-lernen.de/mathematik/varianz-berechnen.html).*

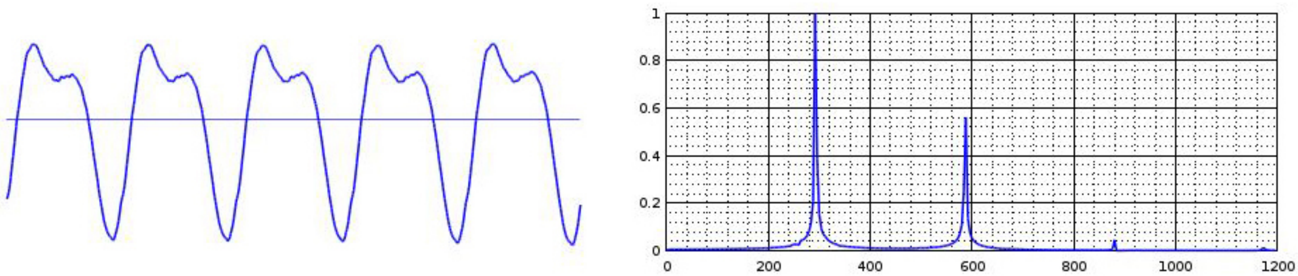
12 *Präziser ausgedrückt wird ein Frequenzintervall ausgewählt, das durch drei Halbtöne tiefer beziehungsweise höher als die Frequenz mit der maximalen Amplitude begrenzt ist. Diese Zahl hat sich aus zahlreichen Beispielrechnungen ergeben. Ziel war es, die sichtbaren Unterschiede in den graphischen Darstellungen der Audiobeispiele durch die Ergebnisse der entsprechenden Standardabweichung auszudrücken.*

13 https://www.youtube.com/watch?v=lrxDODZgT3w (Choir GRII Melbourne)

Die Mittelwerte der Standardabweichungen geben in der Abbildung von oben nach unten deutlich die abnehmende Tendenz wieder, die in den graphischen Darstellungen erkennbar und von den jeweiligen Klangkörpern zu erwartern ist.

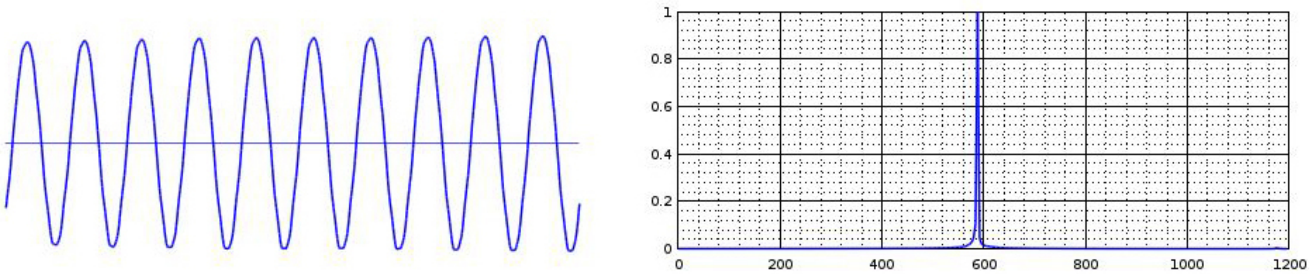
Die drei Audiobeispiele haben eine Sampledauer von 1/2 Sekunde.

Ein interessantes Instrumentenbeispiel bezüglich der Frequenzanalyse ist die Glasharmonika. In der Abbildung 12 finden sich zu dem Grundton d_1 die Obertöne mit der doppelten und der dreifachen Frequenz[14]. Der Wellenverlauf des Audiosignals hierzu weist fünf Perioden mit erkennbaren Oberschwingungen auf. Wie bei der Berechnung auf Seite 13 ergibt sich hier für die fünf Perioden ein Samplelänge von $s_r/f \cdot 5 = 750$, die dem Audioausschnitt zugrunde liegt. In der Abbildung 13 werden zu dem Ton d_2 keine Obertöne ausgewiesen.



Frequenz	Amplitude	Ton
294	1.00	d1
586	0.88	d2
880	0.06	a2

Abbildung 12: Glasharmonika, Tonlage tief



Frequenz	Amplitude	Ton
586	1.00	d2

Abbildung 13: Glasharmonika, Tonlage hoch

Bei der Glasharmonika (Beispiel einer Abbildung hierzu unter "Glasharmonika" im Internet) werden die Töne in der Regel aus Glasschalen durch das Reiben mit angefeuchteten Fingern erzeugt. Die Tonhöhe hängt von Durchmesser und Dicke der jeweiligen Schale ab. Lässt man die

[14] Audiodatei zur Frequenzanalyse: Klangarchiv von Ole Schmidt, Frankfurt am Main.

Abhängigkeit von der Dicke beiseite, dann entsteht eine zunehmende Tonhöhe durch Verkleinerung des Schalendurchmessers[15].

Für die Samples wurde eine Dauer von 1/4 Sekunde gewählt.

[15] *Es ist nachvollziehbar, dass durch den kleineren Durchmesser der Schale zu dem höheren Ton der Schwinguns-spielraum des Kristallgitters im Glas eingeschränkt ist. Daher kann beispielsweise keine Schwingung mit der halben Wellenlänge für die Frequenz zur Oktav* (etwa 1200 Hz) *entstehen. In der Graphik der Abbildung 12 bestätigt die sehr geringe Amplitude zu der Frequenz* 880 Hz *dieses Phänomen.*

Anhang 1: Verbesserung der Frequenzberechnungen zur Fouriertransformation

Die bestimmenden Größen eines Audiosamples sind: Samplerate s_r, Anzahl der in dem Sample enthaltenen Zahlen oder Samplelänge N, Nummer k der einzelnen Frequenzkomponente (mit ihrer Amplitude v_k, vergl. Seite 9) und die hierzu gehörige Frequenz f_k. Für den Zusammenhang zwischen diesen Größen lässt sich im Hinblick auf die Ergebnisse einer Frequenzanalyse folgende Gleichung herleiten:

$$f_k = \frac{s_r}{N} \cdot k \tag{15}$$

Aus dieser Formel kann man entnehmen, dass für den Fall einer Sampledauer von einer Sekunde, das heißt $N = s_r$, der Bruch sich zu dem Wert 1 kürzt. Hiernach sind dann die Frequenzen f_k exakt identisch mit den Indizes k der Fourierkoeffizienten. Beispielsweise gibt so v_{440} die Amplitude zu dem Kammerton a mit der Frequenz von 440 Hz in dem Audiosignal wieder. Durch diesen Zusammenhang ist somit die Genauigkeit der Frequenzergebnisse auf die Werte der ganzen Zahlen festgelegt.

Bei einer Samplelänge von beispielsweise $N = 2 \cdot s_r$ wird $f_k = k/2$. Für $k = 2, 4, 6, ...$ liefert die Frequenzanalyse die Amplituden zu den ganzzahligen Frequenz $f_1, f_2, f_3, ...$. Für $k = 1, 3, 5, ...$ werden die Amplituden zu den Frequenz $f_{1/2}, f_{3/2}, f_{5/2}, ...$ in dem Sample ausgewiesen. Analog lassen sich die Ergebnisse bei anderen ganzzahligen Vielfachen $N = m \cdot s_r,\ m = 3, 4, 5, ...$ interpretieren.

Die Folgerung aus einem kürzeren Sample soll am Beispiel der Sampledauer von 1/10 Sekunde dargelegt werden, d.h. $N = s_r/10$. Somit wird nach (15) $f_k = 10 \cdot k$. Das bedeutet, dass die Ergebnisse der Frequenzanalyse nur Aussagen über diejenigen Frequenzen in dem Audiosample machen, deren Zahlen Vielfache von 10 sind. Entsprechendes lässt sich bei einer Sampledauer von 1/2 Sekunde mit $f_k = 2 \cdot k$ folgern. Das trifft beispielsweise für die Tabellen zu den Abbildungen 6 und 7, Seite 13 und 14, zu mit den geradzahligen Frequenzergebnissen. Entsprechend sind in den Tabellen der Abbildung 10, Seite 17 die Vielfachen von 5 zu erkennen bei einer Sounddauer von 1/5 Sekunden, das bedeutet $s_r/N = 5$ und $f_k = 5 \cdot k$ (mit der Ausnahme des Musikbeispiels "Oboe").

Die Frequenzsprünge in der unteren Graphik der Abbildung 8, Seite 15 ergeben sich aus den Längen der Untersamples von 1/20 Sekunden, das bedeutet $s_r/N = 20$ und $f_k = 20 \cdot k$. Falls aus einem Sample präzisere Frequenzinformationen gewonnen werden sollen, lässt sich das mit einer einfachen Manipulation der Inputdaten erreichen. Durch Verlängerung des Samples mit Nullen beispielsweise auf eine Länge von $100 \cdot N$ ergeben sich entsprechend $100 \cdot N$ Amplitudenwerte zu dem verlängerten Sample mit demselben Audioverlauf. Dieser Vorgang trägt die Bezeichnung "Zero-Padding"[16]. Damit wird im Fall der ursprünglichen Sampledauer von einer Sekunde die Genauigkeit der Frequenzanalyse auf zwei Nachkommastellen verbessert[17]. Diese Ergebnisse der "erweiterten Fouriertransformation" (oder auch "erweiterte Frequenzanalyse") lassen sich entsprechend auf andere Zahlenbeispiele übertragen. Der Faktor bei N wird als Erweiterungsfaktor bezeichnet.

[16] vergl. im Internet unter "Zero-Padding DFT"

[17] *Es lässt sich zeigen, dass die Amplitudenwerte zu den zusätzlichen Frequenzen durch sogenannte "trigonometrische Interpolation" entstehen. Vergl. hierzu Bruno Klingen "Fouriertransformation für Ingenieur- und Naturwissenschaften", Springer Verlag 2001, Seite 271*

Beispiel 8 Es wurde mit einem Tongenerator der Firma NCH Software ein Audiosample zur Frequenz des eingestrichenen e erzeugt und in eine wav-Datei überführt. Mit (3) ergibt sich die Frequenz $f_{1,5} = 329.63$ (vergl. Tabelle 2 Ton 5 als Näherungswert zu $f_{1,5}$).

Die Ergebnisse der erweiterten Frequenzanalyse lassen sich aus der folgenden Tabelle 3 entnehmen. Die Zahlenwerte in der Kopfzeile geben die Erweiterungsfaktoren an.

Sampledauer	1	10	100	1000
1 s	330.00	329.60	329.63	
0.2 s	330.00	329.50	329.65	329.63
0.1 s	330.00	330.00	329.60	329.65

Tabelle 3: erweiterte Frequenzanalyse bei Tongeneratorauswertung

In dieser Tabelle erhält man beispielsweise den Wert 330.00 Hz zu der Sampledauer von 0.1 Sekunden bei dem Erweiterungsfaktor 1 mit (15) dadurch, dass hier der Bruch in der Formel den Wert 10 annimmt. Das bedeutet, dass die Werte der Fouriertransformation als Ergebnisse nur Frequenzen als Vielfache der Zahl 10 wiedergeben können. Damit wird zu der Frequenz $f_{1,5} = 329.63$ Hz der gerundete Wert 330.00 Hz geliefert.

Beispiel 9 Durch Anwendung von Zero-Padding bei den Fourierkoeffizienten im Frequenzbereich und anschließender inverser Fouriertransformation (vergl. Fußnote Seite 9) lässt sich ein Audiosample gezielt verändern. Die Folge hiervon ist eine Zeitdehnung des Samples.

In diesem Beispiel soll über den Chorsatz “Abendlied” von Josef Rheinberger eine Schwachstelle von Laienchören aufgezeigt werden: Das Problem des präzisen gleichzeitigen Einsatzes aller Sängerinnen und Sänger. In der folgenden Abbildung 14 sind links (Farbe blau) die Ergebnisse des Universitätschores München aus dem Jahr 2010 wiedergegeben[a]. Rechts (Farbe rot) befinden sich die des Chores “The Cambridge Singers”[b].

[a] *www.youtube.com/watch?v=dqzRWGhgGuI*
[b] *www.youtube.com/watch?v=6NpGibPW-7w*

Die Abbildung 14 gibt aus dem “Abendlied” ein Sample mit etwa 13000 Zahlen graphisch wieder; das entspricht einer Audiodauer von ungefähr 1/3 Sekunde. Damit kann der Verlauf des Choreinsatzes in den linken Graphiken der Abbildung 14 nicht durch ein Crescendo entstanden sein.

Unter Anwendung von Zero-Padding wurden diese Samples auf die achtfache Länge erweitert, wiedergegeben durch die beiden unteren Darstellungen der Abbildung.

Nach den Graphiken links mit dem Universitätschor München (Farbe blau) verläuft der Choreinsatz über die gesamte Zeit des Audiosamples von 1/3 Sekunde. Aus den graphischen Darstellungen rechts lässt sich ablesen, dass in diesem Sample der Chor angenähert zwischen den Abtaststellen 50000 und 60000 eingesetzt hat. Die Differenz entspricht in dem Audioausschnitt einer Zeit von

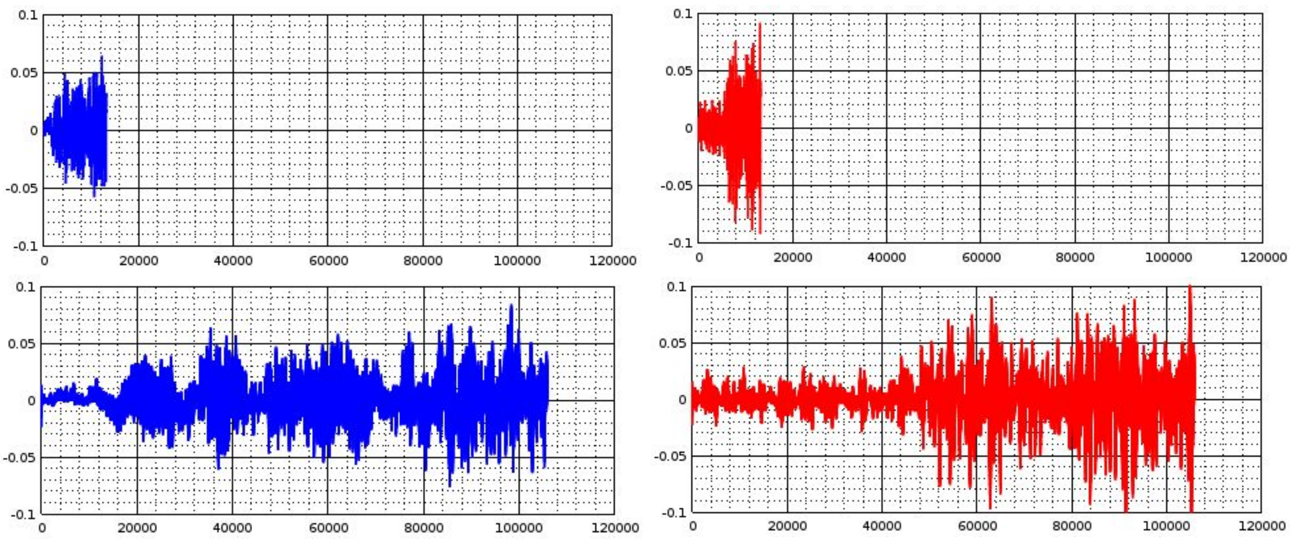


Abbildung 14: Beispiele zur Zeitdehnung von Audiosamples

10000 / 8 / Sampelrate = 0.03 = 30 Millisekunden, also etwa 1/10 der Dauer des Einsatzes mit dem Universitätschor.

Anhang 2: mp3 - Kompression von Audiodateien

Es liegt nahe, mit Hilfe der Frequenzanalyse die Datenkompression in mp3 -Dateien auf Qualitätsverluste zu überprüfen. Hierzu wurde aus Messiah die Nummer 42, “Hallelujah” (vergl Notenblatt Seite 30), untersucht. Die Abbildung 15 bestätigt die Erwartung, dass es sich um ein vielfältiges Frequenzspektrum handelt.

Ausgang war eine Messiah-CD mit der “Academy of St. Martin in the Fields” unter Neville Marriner. Durch die Anwendung des Programms “Audiograbber” wurde von der CD zu der Nummer 42 sowohl eine wav-Datei als auch eine mp3-Datei erzeugt. Das Ergebnis der Frequenzanalyse zur wav-Datei ist in der Abbildung 15 mit der Farbe blau dargestellt. Die mp3-Datei wurde mit Hilfe des Programms “Audacity” ebenfalls in eine wav-Datei überführt, dargestellt mit der Graphik der Frequenzanalyse in roter Farbe. Die Differenz dieser beiden Resultate von Frequenzanalysen zeigt die Abbildung in der Farbe magenta.

Aus diesem Beispiel lässt sich entnehmen, dass die mp3-Kompression die Qualität eines Audiosamples nur unerheblich verändert. Dieser Befund wird durch die Aussage eines Tontechnikers mit langjähriger Hörerfahrung bestätigt. Er kann einen Unterschied zwischen einem CD- und einem mp3-Klang höchstens dann erkennen, wenn er beide Sounds direkt nacheinander vergleicht.

Anhang 3: Quintenstimmung

Die Quintenstimmung war bis zum Aufkommen der elektronischen Hilfsmittel das vorherrschende Verfahren zur Stimmung von Tasteninstrumenten. Dieser Stimmungsmethode kommt zugute, dass man mit einigem Training ein Quintenintervall mit der hörbaren Schwebung auf die unten angegebene Schwebungszahl genau stimmen kann. Ein zentraler Begriff der Quintenstimmung ist das “pythagoräische Komma”[18].

[18] Die Informationen zur Urheberschaft des pythagoräischen Kommas sind uneinheitlich. In dem Artikel bei Wikipedia zu Pythagoras von Samos (etwa 570 bis 495 v. Chr.) wird sie zwar nicht ausgeschlossen, aber bezweifelt. In

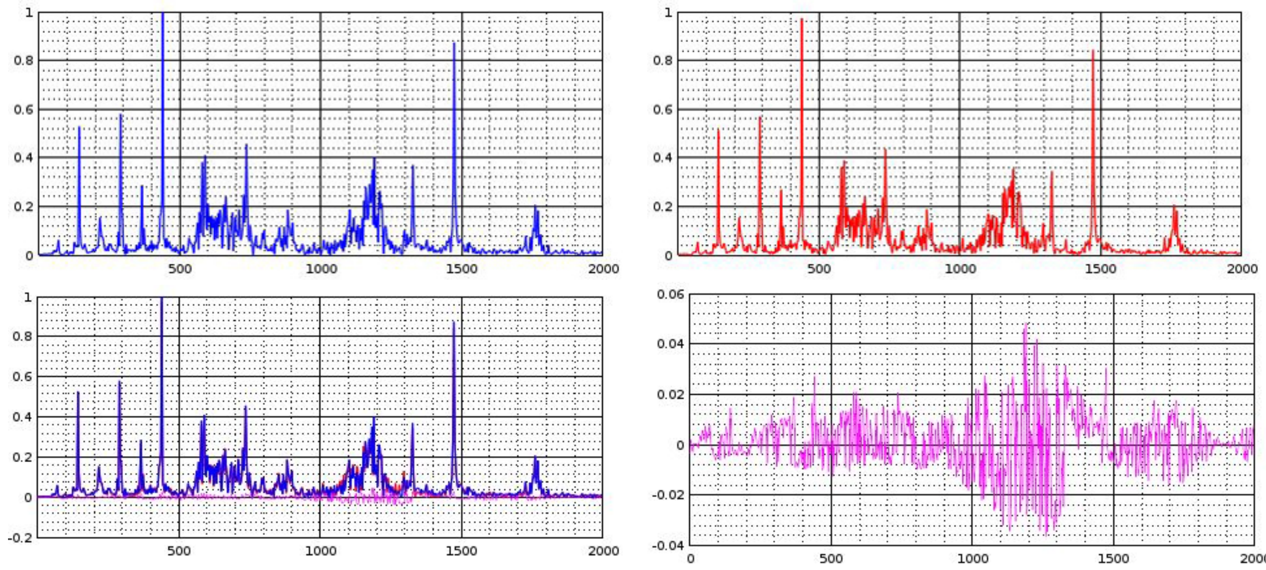


Abbildung 15: Messiah "Hallelujah" vor und nach Übergang mp3 zu wav

Die anschauliche Form der Quintenstimmung liefert - vom Ton der Stimmgabel a_1 ausgehend - der Quintenzirkel in Abbildung 16. Es wird zuerst mit sechs aufsteigenden Quinten bis zum Ton dis und dann mit sechs absteigenden Quinten bis zum Ton b gestimmt (der Ton es kommt in diesem Zusammenhang nicht vor).

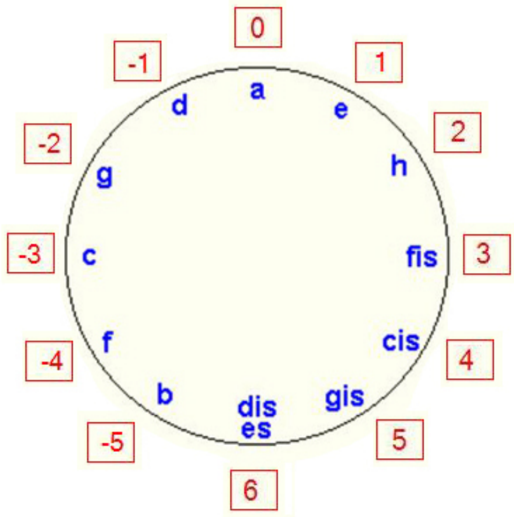


Abbildung 16: Quintenzirkel

Die Frequenzen zu den Quinten erhält man aus der folgenden Formel. Hierin wird für j die Quintennummer eingesetzt, die im Quintenzirkel der Abbildung 16 dem jeweiligen Tonnamen zugeordnet ist:

dem Wikipedia-Artikel zum pythagoreischen Komma heißt es "Als erster definierte der Pythagoreer Philolaos das pythagoreische Komma" (Philolaos von Korinth etwa 470 bis 399 v. Chr.).

$$f_{Quinte,j} = 440 \cdot \frac{1.5^j}{2^{n_j}}$$

Der Exponent n_j im Nenner der Formel ergibt sich dadurch, dass das Rechenergebnis durch Oktavieren im Bereich der eingestrichenen Oktav liegen soll.

Das enharmonische Tonpaar (es,dis) zu den Grenzfällen $j = 6$ und $j = -6$ führt auf

$$440 \cdot \frac{1.5^6}{2^4} = 313.242 = f_{1,4} + 11.73 \text{ Cent}$$

(Potenzregel $s^{-t} = 1/s^t$:)

$$440 \cdot \frac{2^3}{1.5^6} = 309.026 = f_{1,4} - 11.73 \text{ Cent}$$

Die Differenz von 23.46 Cent ist das "pythagoräisches Komma" - immerhin fast 1/8 Ton. Bei den Frequenzen beträgt die Abweichung im eingestrichenen Oktavbereich 4.22 Hz. Diese Differenz ergibt auf die zwölf Quintentöne verteilt etwa 0.35 Hz als Schwebung je Halbton. Diese musste früher beim Stimmen von Tasteninstrumenten zu jedem Ton beachtet werden, um eine gleichtemperierte Stimmung zu erreichen.

Von Interesse bei der Quintenstimmung ist die rechnerische Ermittlung der Frequenzabweichungen bei Intervallen und Dreiklängen (hier werden nur Dur-Dreiklänge behandelt). Grundlegend hierzu ist der Wert der Quintabweichung gegenüber der gleichtemperierten Stimmung nach der folgenden Tabelle 4. Die Centabweichung von 330.00 Hz zu der gleichtemperierten Frequenz lässt sich nach dem Schemaverfahren wie etwa in Beispiel 2, Seite 7, ermitteln.

	gleich-temperiert	Quinten stimmung	Cent-abweichung
e	329.63	330.00	1.96
a	440.00	440.00	0.00
gesamt			1.96

Tabelle 4: Centabweichung der Quinte a-e

Als Formel gilt für die Differenz

$$\text{Cent}_{\text{Quint}} = 12 \cdot 100 \cdot \log_2 \left(\frac{1.5}{2^{7/12}} \right) = 100 \cdot (12 \cdot \log_2 (1.5) - 7) = 1.955 \approx 2.0$$

Der Bruch im Argument zu $\log_2$ wird durch die Relation des Quintenfaktors 1.5 zu der Potenz der gleichtemperierten Quinte (vergl. (5) Seite 5) gebildet.

Auf dem Wert $\text{Cent}_{\text{Quint}} = 2$ Cent basieren alle Intervallabweichungen zweier Töne. Zur Ermittlung dieser Abweichungen geht man vom Grundton zum Intervallton über die Schritte des Quintenzirkels (und zwar nicht den Weg mit Überschreitung der (es,dis)-Grenze). Die Anzahl der Schritte führt dann durch die Multiplikation mit $\text{Cent}_{\text{Quint}}$ auf die Intervallabweichung. Die kleine Terz ergibt beispielsweise eine Abweichung von 6 Cent, zu der großen Terz gehören 8 Cent.

Aus der Richtung der Schritte - linksdrehend, rechtsdrehend - folgert das Vorzeichen zu dem Abweichungswert. Auf diese Art lassen sich auch die Abweichungen der Intervalle eines Dreiklangs bestimmen.

Mit Blick auf den Quintenzirkel wird ein Dreiklang bei dem Grundton (Beispiel Ton d) gestartet und um vier Quinten zur großen Terz (hier fis) geführt. Dann geht es zurück um drei Quinten zur kleinen Terz (hier a). Das ergibt dann zwangsläufig die Quinte zum Grundton.

Die berechneten Differenzen zwischen den Frequenzen durch Quinten- und gleichtemperierter Stimmung ergeben, dass hinsichtlich der Reinheit von Dreiklängen drei Gruppen gebildet werden können. Mit zunehmender Unreinheit enthalten diese Gruppen die Dreiklänge:

1.: c, d, e, f, g, a, b, h
2.: cis, fis, gis
3.: dis

Innerhalb dieser Gruppen stimmen die Dreiklangabweichungen wegen der identischen Quintenschritte präzise überein. Daher genügt es, jeweils ein Beispiel mit den Berechnungsergebnissen vorzustellen. Die dabei resultierenden Zahlenwerte lassen sich an den Dreiklängen der folgenden Abbildung 17 nachvollziehen. Die Summe der Terzabweichungen führt jeweils auf den Quintenwert.

Der Abweichungswert von 22 Cent bei der Quinte zum dis-Dreiklang resultiert aus der größtmöglichen Zahl von elf Quinten-Schritten.

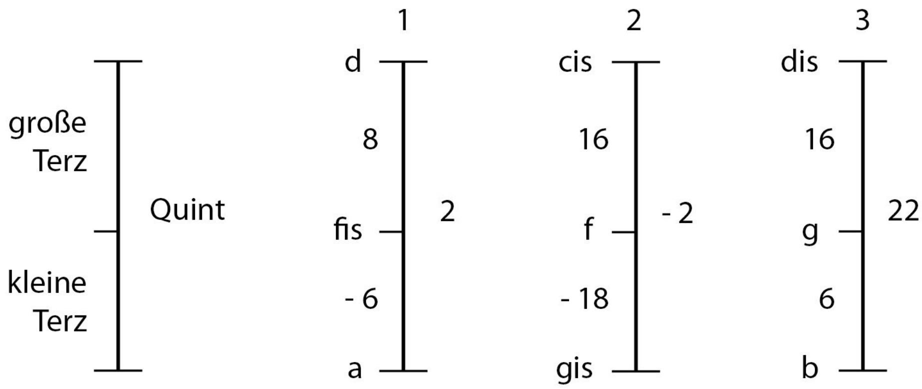


Abbildung 17: Centabweichungen bei typischen Dreiklängen

Texte zu den Sätzen der Musikbeispiele

1. Zu Beispiel 5, Abbildung 6 Seite 13:

21. Air (Alt)
He was despised and rejected of men, a man of sorrows and acquainted with grief. (Isaiah 53 : 3)
He gave His back to the smiters, and His cheeks to them that plucked off the hair: He hid not His face from shame and spitting. (Isaiah 50 : 6)

21. Arie
Er ward verachtet und von den Menschen zurückgestoßen, ein Mensch der Schmerzen und erfüllt mit Gram.
Er hält seinen Rücken hin denen, die ihn schlugen, und seine Wangen denen, die sein Haar ausrissen; er verbarg nicht sein Antlitz vor Schmach und Speichel.

2. Zu Beispiel 6, Abbildung 7 Seite 14:

1. Accompagnato (Tenor)
Comfort ye, comfort ye My people, saith your God. Speak ye comfortably to Jerusalem, and cry unto her, that her warfare is accomplished, that her iniquity is pardoned. The voice of him that crieth in the wilderness; prepare ye the way of the Lord; make straight in the desert a highway for our God.
(Isaiah 40 : 1-3)

1. Rezitativ
Tröste dich, mein Volk, spricht dein Gott. Redet trostreich mit Jerusalem und ruft ihr zu, dass ihr Kriegsdienst zu Ende, dass ihre Missetat vergeben ist. Es ist seine Stimme, die verkündigt in der Wildnis: Bereitet dem Herrn den Weg, ebnet in der Wüste einen Pfad für unsern Gott.

3. Zu Abbildung 11 Seite 18:

44. Choir
Since by man came death, by man came also the resurrection of the dead. For as in Adam all die, even so in Christ shall all be made alive.
(I Corinthians 15 : 21-22)

44. Chor
Da durch einen Menschen der Tod gekommen ist, so kam auch durch einen Menschen die Auferstehung der Toten. Denn wie in Adam alle sterben, werden sie in Christus ebenso alle lebendig gemacht werden.

4. Zu Beispiel 9, Abbildung Seite 22: Abendlied nach Lukas 24:29

Bleib bei uns, denn es will Abend werden, und der Tag hat sich geneiget.

5. Zu Abbildung 15 Seite 24: Schlussakkord

42. Choir
Hallelujah: for the Lord God Omnipotent reigneth.
(Revelation 19 : 6)
The kingdom of this world is become the kingdom of our Lord, and of His Christ; and He shall reign for ever and ever.
(Revelation 11 : 15)
King of Kings, and Lord of Lords.
(Revelation 19 : 16)
Hallelujah!

42. Chor
Halleluja, denn der Herr, der allmächtige Gott, herrschet.

Das Königreich dieser Welt ist zum Königreich unseres Herrn und seines Christus geworden; und er wird regieren auf immer und ewig.
König der Könige, Herr der Herren.

Halleluja!

Anhang 4: Notenblätter

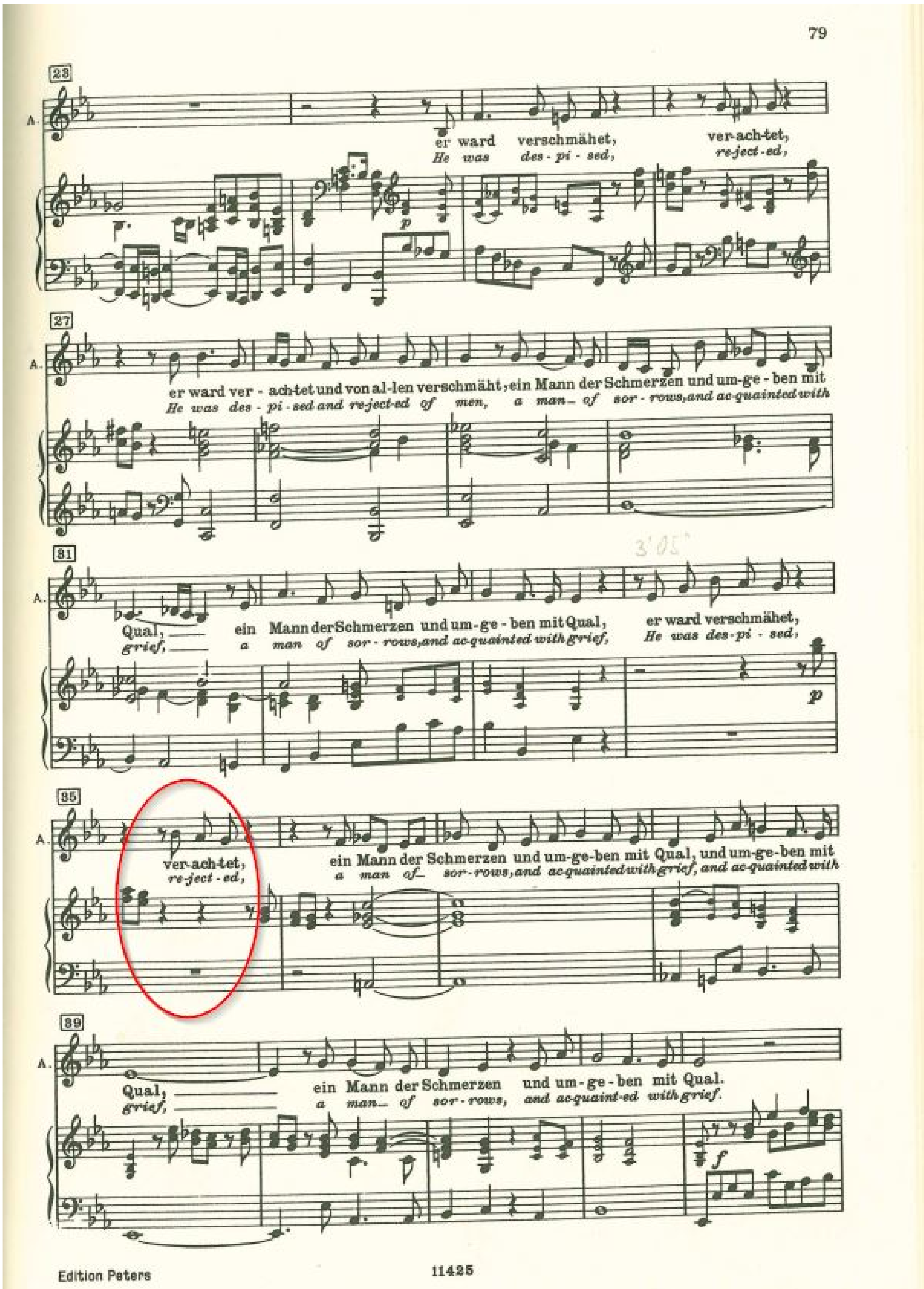


Abbildung 18: Messiah: "He was despised"

www.youtube.com/watch?v=1tC23A55eSU

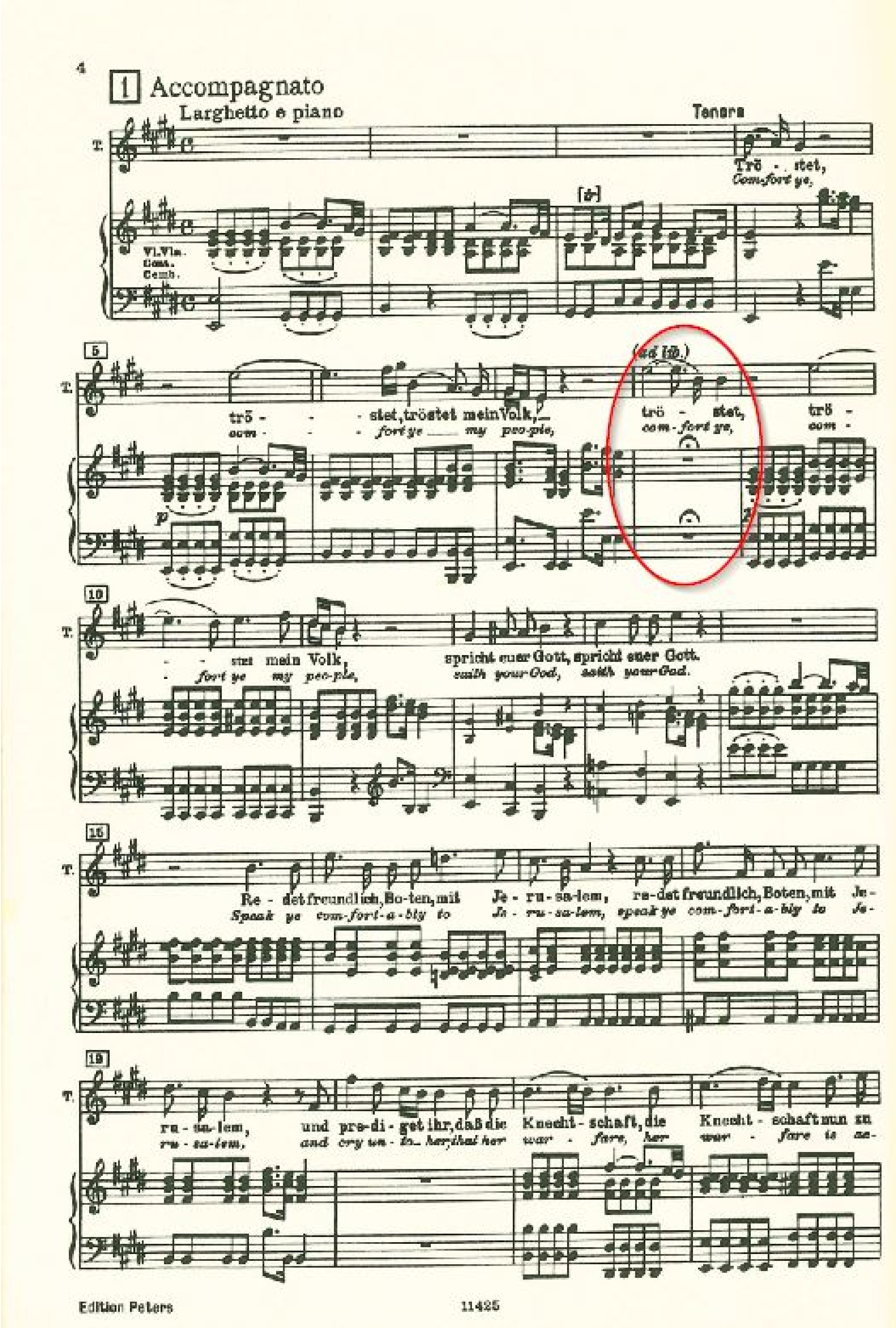


Abbildung 19: Messiah: “Comfort ye”

www.youtube.com/watch?v=iaIf44hNex0

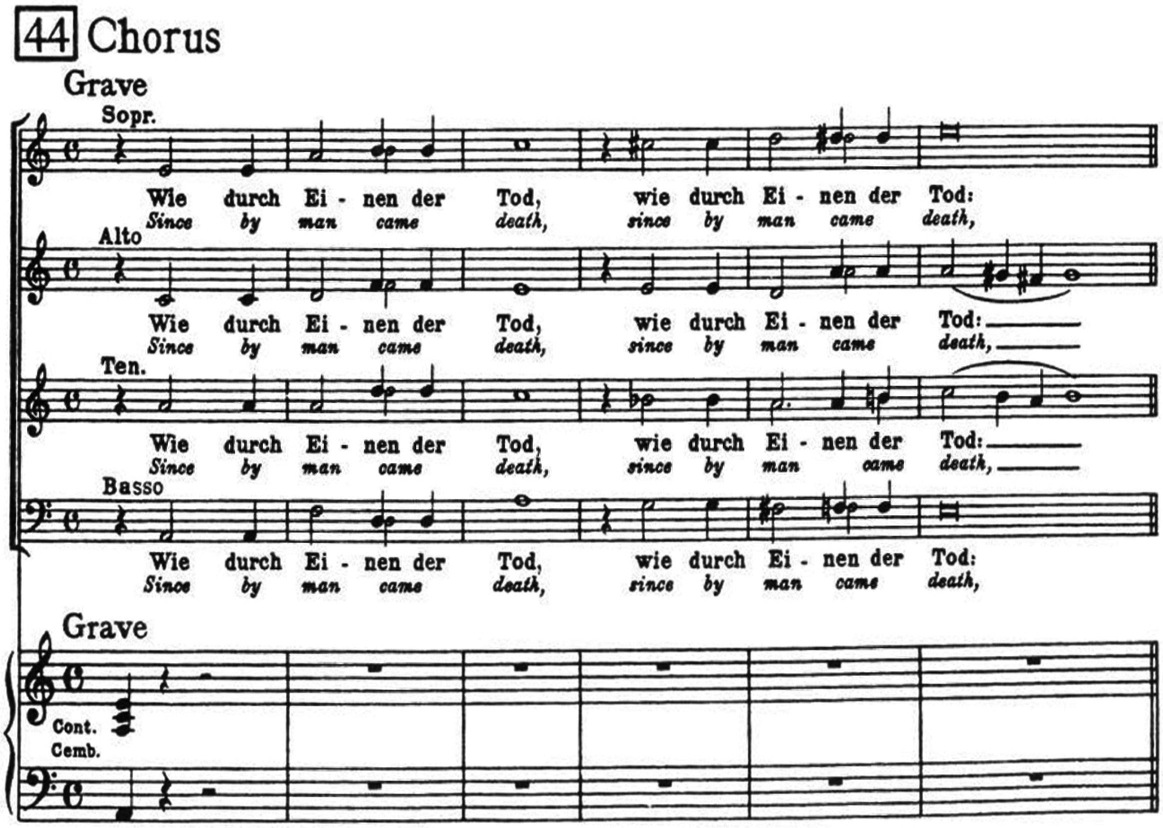


Abbildung 20: Messiah: "Since by man"

www.youtube.com/watch?v=jAYA4ZG0jnU

Abbildung 21: Messiah : "Hallelujah"

www.youtube.com/watch?v=DUC_U7IPhj8

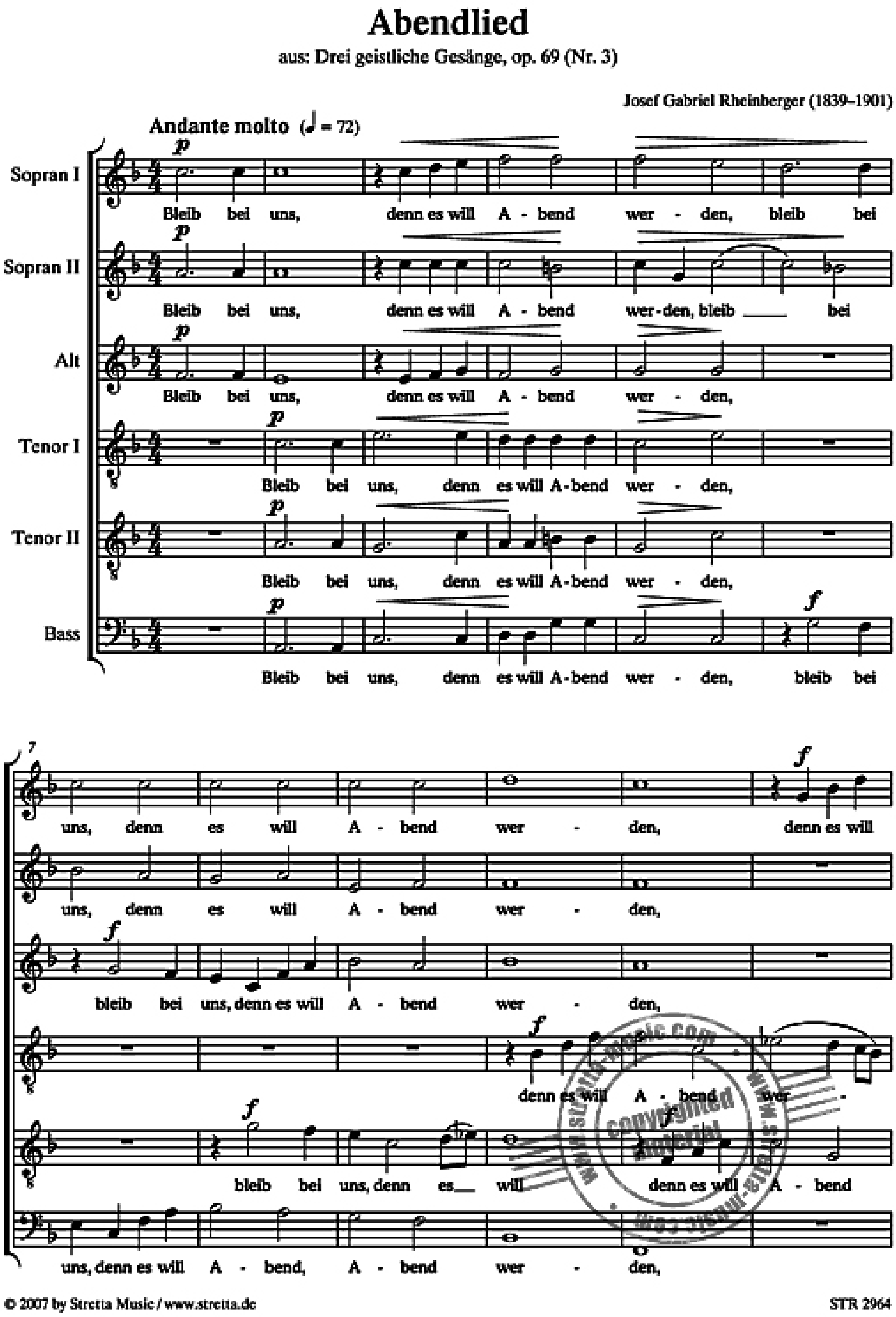


Abbildung 22: Rheinberger : "Abendlied"
www.youtube.com/watch?v=6NpGibPW-7w
Mit freundlicher Genehmigung der Stretta Music GmbH

Index