Repräsentation und Modifikation von spektralen Hüllkurven auf Basis von Formantmodellen

Robin Schmidt

June 29, 2006

Contents

1 Spektrale Hüllkurven **1**

1.1 Das Problem 2

1.2 Schätzung über Lineare Prädiktion 3

1.3 Schätzung über Discrete Allpole Modeling 4

1.4 Schätzung über Cepstral Smoothing 5

1.5 Schätzung mit dem True Spectral Envelope Algorithmus 6

2 Formantmodelle **7**

2.1 Serielle Allpolfilter 8

2.1.1 Das Modell 8

2.1.2 Transformation der Formanten 11

2.2 Parallele Allpolstruktur 12

2.2.1 Das Modell 12

2.2.2 Zusammenhang zu FOF's (Formantwellenformen) 13

2.2.3 Transformation der Formanten 14

2.3 Serielle Parametrische Equalizer 15

2.3.1 Das Modell 15

2.3.2 Herleitung des Betragsfrequenzgangs einer Biquad-Filterstufe 18

2.3.3 Der Gradient der Kostenfunktion 20

2.3.4 Transformation der Formanten 29

Im diesem Text wird die folgende Notation benutzt:

Variable	**Bedeutung**
bw	Bandbreite (in Oktaven) eines Peak- oder Bandpass-Filters
γ	Bandbreitenparameter eines Peak-Filters
$H(z)$	Transferfunktion eines Filters (allgemein)
$H_p(z)$	Transferfunktion des p-ten Teilfilters in einer Zusammenschaltung von P gleichartigen Filtern (seriell oder parallel)
$H_P(z)$	Transferfunktion einer aus P Teilfiltern bestehenden Filterstruktur, evtl. erweitert um einen globalen Gain-Faktor G
E	Fehler- bzw. Kostenfunktion
G	globaler Gain-Faktor für eine Filterstruktur
g	Gain-Parameter eines Peak-Filters als Faktor
g_{dB}	Gain-Parameter eines Peak-Filters in dB
θ	ein generischer Parameter
Ω	normierte Kreisfrequenz (allgemein)
Ω_c	normierte Mittenfrequenz eines Peak- oder Bandpass-Filters

Chapter 1

Spektrale Hüllkurven

1.1 Das Problem

Viele Spektren von natürlichen Klängen enthalten zweierlei Arten von spektraler Information die einander überlagern. Einerseits sieht man im Spektrum von z.B. harmonischen Klängen die Spektrallinien der einzelnen auftretenden Frequenzen. Zwischen diesen diskreten Frequenzen ist das Amplitudenspektrum im Prinzip gleich Null (das ist natürlich eine stark idealisierte Vorstellung). Andererseits hat jede solche Spektrallinie eine bestimmte Amplitude - verbindet man die spektralen Peaks auf geeignte Weise miteinander, so erhält man eine spektrale Hüllkurve, die die grundsätzliche Form des Spektrums beschreibt ohne jedoch einzelne Frequenzen aufzulösen. Am Beispel des Spektrums einer männlichen Stimme, die ein langgezogenes "a" spricht, wird dies deutlich:

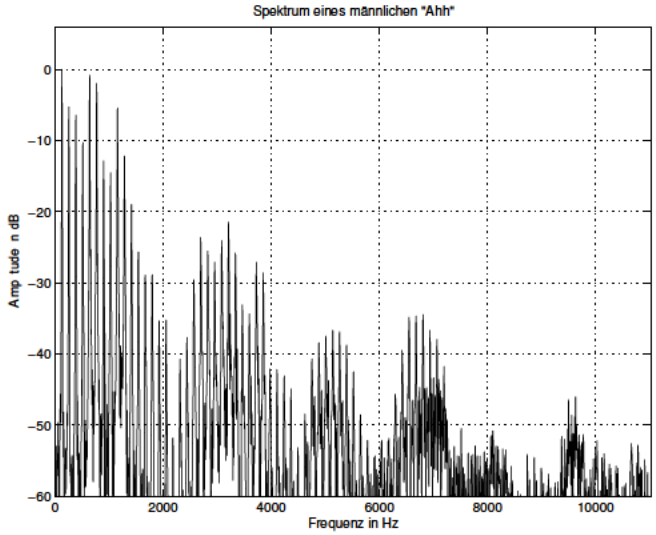

Figure 1.1: FFT-Spektrum eines männlichen "Ahh"

Wir sehen einerseits die (mehr oder weniger) diskreten Spektrallinien bei ganzzahligen Vielfachen der Grundfrequenz, andererseits sehen wir, dass z.B. die Frequenzbereiche um 700 Hz und 1200 Hz mehr Energie enthalten als umliegende Frequenzbereiche. Solche Frequenzbereiche mit besonders viel Energie nennt man Formanten. Diese entstehen physikalisch häufig durch Resonatoren, deren Resonanzfrequenzen oft unabhängig von der Grundfrequenz des Klangs sind. Bei akustischen Saiteninstrumenten besipielsweise werden diese Resonanzfrequenzen vor allem durch Form und Größe des Korpus bestimmt - und diese Dinge sind schließlich völlig unabhängig davon, welche Note gerade gespielt wird. Bei der menschlichen Stimme sind die Resonatoren durch die Geometrie der Hohlräume im Mund- und Rachenbereich bestimmt während das Anregungssignal bei stimmhaften Lauten durch das periodische Öffnen und Schließen der Glottis entsteht. Hier kann es zwar durchaus zu gewissen Wechselwirkungen zwischen Grundfrequenz und Formantbereichen kommen, näherungsweise kann man die Trennung jedoch auch hier vollziehen. Bei der menschlichen Stimme sind die Lagen der Formantbereiche ein wichtiges Kriterium zur Unterscheidung von Stimmen und auch zur Unterscheidung verschiedener Laute. Für viele interessante Signaltransformationen ist es wünschenswert, die spektrale Hüllkurve von den (durch die Periodizität des Signals gegebenen) spektralen Peaks zu trennen, denn wenn man diese beiden Informationen getrennt zu Verfügung hat, kann man sie auch getrennt voneinander modifizieren. Anwendungsbeispiele sind Pitch-Shifting unter Erhaltung der spektralen Hüllkurve und unter Verschiebung bzw. Skalierung der spektralen Hüllkurve unter Erhaltung der Tonhöhe.

1.2 Schätzung über Lineare Prädiktion

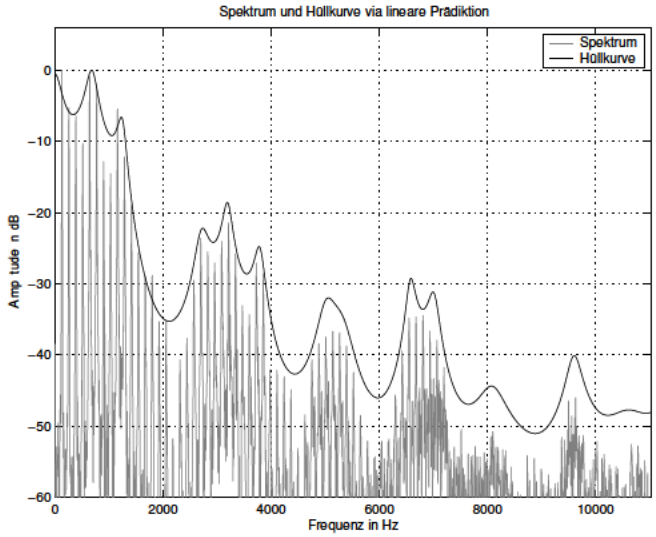

Figure 1.2: FFT-Spektrum und spektrale Hüllkurve (geschätzt über Lineare Prädiktion)

1.3 Schätzung über Discrete Allpole Modeling

1.4 Schätzung über Cepstral Smoothing

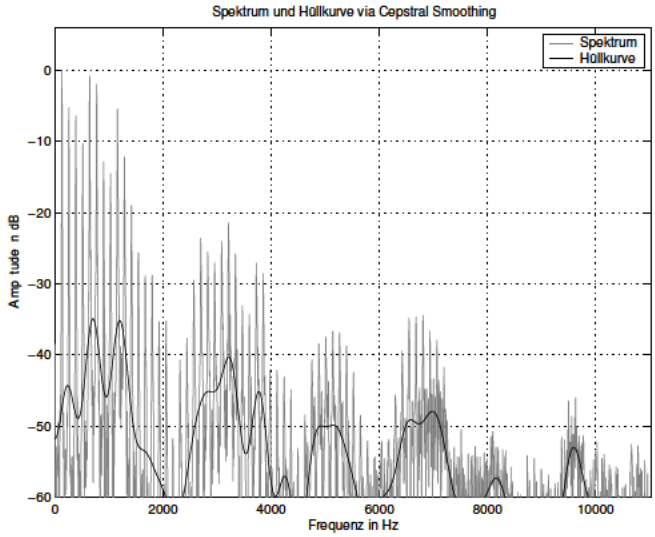

Figure 1.3: FFT-Spektrum und spektrale Hüllkurve (geschätzt über Cepstral Smoothing)

1.5 Schätzung mit dem True Spectral Envelope Algorithmus

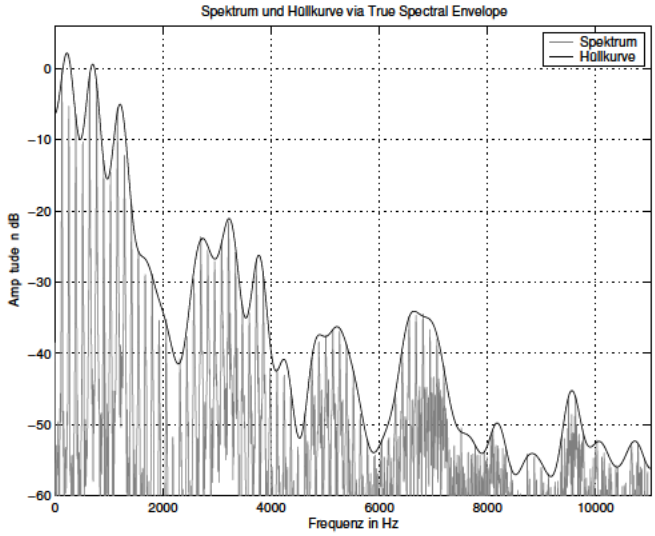

Figure 1.4: FFT-Spektrum und spektrale Hüllkurve (geschätzt mit dem True Spectral Envelope Algorithmus)

Chapter 2

Formantmodelle

2.1 Serielle Allpolfilter

2.1.1 Das Modell

In diesem Abschnitt soll eine serielle Verschaltung von Allpolfiltern zweiter Ordnung als Modell für unser Synthesefilter dienen. Wir modellieren unser Synthesefilter also mit folgender Filterstruktur:

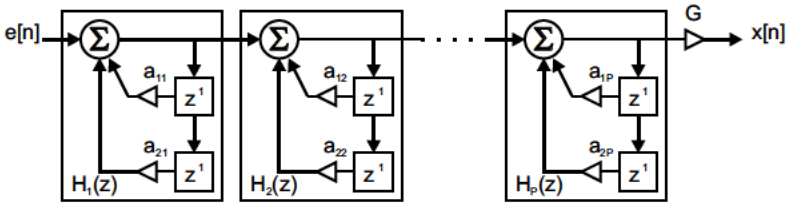

Figure 2.1: Serielle Verschaltung von P Allpolfiltern 2.Ordnung

Für die Berechnung der Filterkoeffizienten können wir von einem Allpolfilter in Direktform ausgehen - so wie es z.B. durch Lineare Prädiktion oder besser durch Discrete Allpole Modeling geliefert wird. Dieses Allpolfilter müssen wir dann in eine serielle Verschaltung von Filtern zweiter Ordnung zerlegen. Dazu müssen zunächst die Nullstellen des Nennerpolynoms der Transferfunktion des Allpolfilters gefunden werden (das sind die Polstellen des Filters). Da diese Polstellen entweder reell sind, oder als konjugiert komplexe Paare auftreten, können wir je zwei solcher konjugiert komplexer Pole zu einem quadratischen Polynom mit reellen Koeffizienten zusammenfassen - und dies sind genau die gesuchten Koeffizienten für eine Filterstufe. Einzelne reelle Polstellen werden in dieser Filterstruktur als Spezialfall mit erfasst, denn sie können durch ein Filter realisiert werden, bei dem der Koeffizient a_2 zu Null wird.

Die Transferfunktion einer einzelnen Filterstufe mit dem Index p gegeben ist durch:

$$H_p(z) = \frac{1}{1 + a_{1p}z^{-1} + a_{2p}z^{-2}} \tag{2.1}$$

Die Transferfunktion einer Reihenschaltung P solcher Filterstufen ergibt sich als Multiplikation der Transferfunktionen der einzelnen Stufen, außerdem wird ein globaler Gain-Faktor G eingeführt:

$$\begin{aligned} H_P(z) &= G\prod_{p=1}^{P} H_p(z) \\ &= G\prod_{p=1}^{P} \frac{1}{1 + a_{1p}z^{-1} + a_{2p}z^{-2}} \end{aligned} \tag{2.2}$$

Jede dieser Zweipol-Filterstufen hat demnach zwei Freiheitsgrade (oder auch nur einen, wenn $a_2 = 0$ ist). Es jedoch intuitiv besser verständlich, wenn man nicht die Koeffizienten selbst als einstellbare Parameter betrachtet, sondern den Betrag und Winkel des jeweiligen Polstellenpaares, das durch das Filter realisiert wird. Dieses Polstellenpaar wollen wir mit

den komplexen Zahlen $re^{j\varphi}$ und $re^{-j\varphi}$ bezeichnen. Der Winkel φ des Polpaares entspricht nämlich (näherungsweise) derjenigen normierten Kreisfrequenz Ω, bei der im Frequenzgang des Filters ein Peak auftritt (die genaue Frequenz des Peaks ist etwas verschoben, da die beiden Pole auch noch miteinander wechselwirken). Der Radius r hingegen bestimmt, wie stark das Peak ausgeprägt ist. Der Zusammenhang zwischen Radius r und Winkel φ einerseits und den Filterkoeffizienten a_1 und a_2 andererseits ist wie folgt:

$$\begin{aligned} a_1 &= -2r\cos\varphi \\ a_2 &= r^2 \end{aligned} \tag{2.3}$$

In Abbildung 2.2 sind beispielhaft einige Betragsfrequenzgänge solcher Zweipol-Filter mit verschiedenen Winkeln und verschiedenen Radien der Polstellen dargestellt. Im ersten Plot ist der Winkel $\varphi = \pi/4$, dies entspricht bei einer Samplerate von 44.1 kHz einer Frequenz von 5512.5 Hz. Wir werden die zu modellierenden Formanten in der Regel im Frequenzbereich unterhalb von 5 kHz suchen - wie wir sehen, zeigen die Zweipol-Frequenzgänge in diesem Bereich (von φ) eine Art Tiefpasscharakter. Wenn wir also beispielsweise einen Formanten durch ein solches Zweipolfilter modellieren wollen, wird dies zu beachten sein (wir werden jedoch noch sehen, dass eine einfache "Ein Formant entspricht einem Polpaar"-Zuordnung ohnehin nicht so ohne weiteres möglich ist).

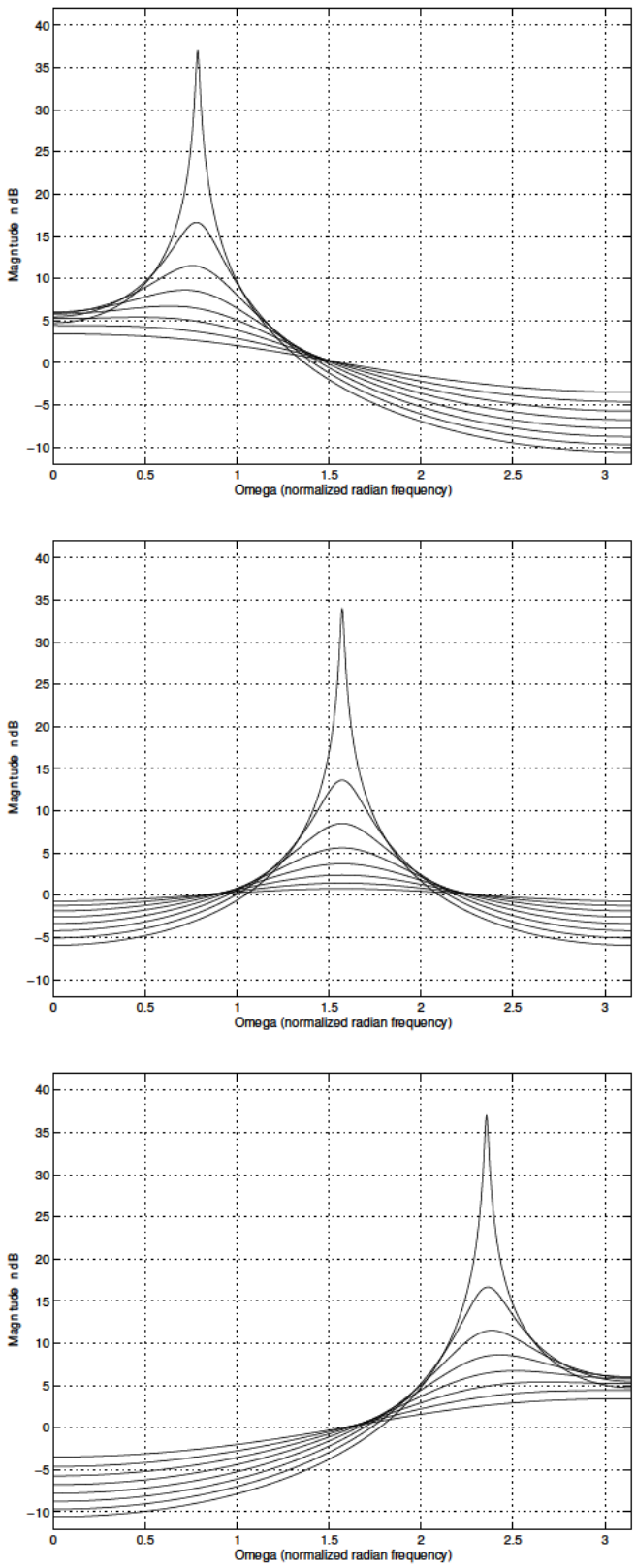

Figure 2.2: 3 Familien von Betragsfrequenzgängen von Zweipol-Filtern, die Winkel der Polstellen betragen $\pi/4$, $\pi/2$, $3\pi/4$ die Radien gehen jeweils von 0.29 bis 0.99 mit einem Inkrement von 0.1

2.1.2 Transformation der Formanten

2.2 Parallele Allpolstruktur

2.2.1 Das Modell

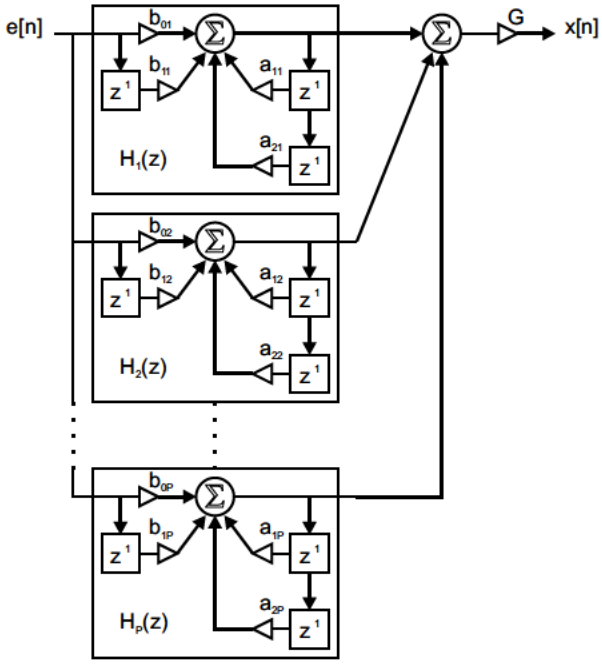

Figure 2.3: Parallele Verschaltung von P Filtern 2. Ordnung zu einer Allpolstruktur

2.2.2 Zusammenhang zu FOF's (Formantwellenformen)

2.2.3 Transformation der Formanten

2.3 Serielle Parametrische Equalizer

2.3.1 Das Modell

Für die Modellierung von Formantspektren wird in diesem Abschnitt eine Reihenschaltung von parametrischen Equalizern untersucht. Diese einzelnen parametrischen Equlizerstufen werden auch hier durch Filter zweiter Ordnung realisiert, allerdings enthalten diese Filter zweiter Ordnung jetzt jeweils ein Polstellenpaar und ein Nullstellenpaar (sind also "echte" Biquads). Als Blockdiagramm sieht die Struktur unseres Modellfilters folgendermaßen aus:

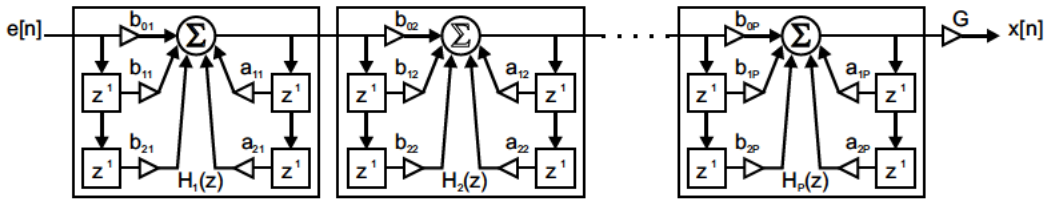

Figure 2.4: Serielle Verschaltung von P Filtern 2.Ordnung (Biquads)

wobei die Transferfunktion einer einzelnen Biquad-Stufe mit dem Index p gegeben ist durch:

$$H_p(z) = \frac{b_{0p} + b_{1p}z^{-1} + b_{2p}z^{-2}}{a_{0p} + a_{1p}z^{-1} + a_{2p}z^{-2}} \tag{2.4}$$

mit $a_{0p} = 1 \; \forall p$. Die Transferfunktion einer Reihenschaltung P solcher Equalizer ergibt sich als Multiplikation der Transferfunktionen der einzelnen Stufen, außerdem wird ein globaler Gain-Faktor G eingeführt:

$$\begin{aligned} H_P(z) &= G \prod_{p=1}^{P} H_p(z) \\ &= G \prod_{p=1}^{P} \frac{b_{0p} + b_{1p}z^{-1} + b_{2p}z^{-2}}{a_{0p} + a_{1p}z^{-1} + a_{2p}z^{-2}} \end{aligned} \tag{2.5}$$

wobei der Index p über die einzelnen Equalizerstufen läuft.

Als Grundlage für das Design der einzelnen Biquad-Stufen werden die Gleichungen für die Berechnung der Filterkoeffizienten von Robert Bristow Johnson (-¿ Literaturangabe) benutzt. Nach Robert Bristow Johnson ergeben sich die Koeffizienten für einen parametrischen Equalizer zu:

$$
\begin{aligned}
b_0 &= \frac{1+\gamma\sqrt{g}}{1+\gamma/\sqrt{g}} \\
b_1 &= \frac{-2\cos(\Omega_c)}{1+\gamma/\sqrt{g}} \\
b_2 &= \frac{1-\gamma\sqrt{g}}{1+\gamma/\sqrt{g}} \\
a_0 &= 1 \\
a_1 &= \frac{-2\cos(\Omega_c)}{1+\gamma/\sqrt{g}} \\
a_2 &= \frac{1-\gamma/\sqrt{g}}{1+\gamma/\sqrt{g}}
\end{aligned} \tag{2.6}
$$

wobei, wie man sieht $b_1 = a_1$ gilt. Dabei ist Ω_c die normierte Mittenfrequenz des Equalizers. Diese ergibt sich aus der Mittenfrequenz f_c (in Hz) und der Samplerate f_s (in Hz) wie folgt:

$$\Omega_c = 2\pi\frac{f_c}{f_s} \tag{2.7}$$

Der Parameter g ist der Vertärkungsfaktor - und zwar nicht in dB, sondern als Faktor mit der die Amplitude eines Sinus mit einer Frequenz von Ω_c vertärkt wird. Oft ist es jedoch intuitiver, den Verstärkungsfaktor in dB anzugeben - hierfür werden wir die Bezeichnung g_{dB} benutzen. Für die Umrechnung gilt:

$$g = 10^{g_{dB}/20} \tag{2.8}$$

Der Parameter γ ist der Bandbreitenparameter, der sich wie folgt aus der (in Oktaven gegebenen) Bandbreite bw berechnet:

$$\gamma = \sinh\left(\frac{\ln(2)}{2}bw\frac{\Omega_c}{\sin(\Omega_c)}\right)\sin(\Omega_c) \tag{2.9}$$

Aus der Transferfunktion $H_P(z)$ ergibt sich der Frequenzgang des Equalizers durch Auswertung der Transferfunktion entlang des Einheitskreises. Dies geschieht durch die Substitution $z \leftarrow e^{j\Omega}$. Also:

$$H_P(e^{j\Omega}) = G\prod_{p=1}^{P}\frac{b_{0p}+b_{1p}e^{-j\Omega}+b_{2p}e^{-2j\Omega}}{a_{0p}+a_{1p}e^{-j\Omega}+a_{2p}e^{-2j\Omega}} \tag{2.10}$$

Der Parametervektor

Wir haben also P Filterstufen, von denen jede einzelne über einen Verstärkungsparameter (g bzw. g_{dB}), einen Mittenfrequenzparameter (Ω_c bzw. f_c) und einen Bandbreitenparameter (γ bzw. bw) verfügt, wobei sich die verschiedenen Darstellungsformen der Parameter ineinander

umrechnen lassen. Zusätzlich haben wir noch den globalen Verstärkungsfaktor G (bzw. G_{dB}, wenn man diesen in dB ausdrücken will). Insgesamt haben wir also $3 * P + 1$ einzustellende Parameter. Oft ist es übersichtlicher, diese Parameter in einem Vektor zusammenzufassen. Wir defineren also den Parametervektor unserer Filterstruktur als:

$$\boxed{\theta := (G\ g_1 \dots g_P\ \Omega_{c1} \dots \Omega_{cP}\ bw_1 \dots bw_P)^T} \tag{2.11}$$

Die Kostenfunktion

Die spektrale Hüllkurve des Eingangssignals, dessen Formanten mit Hilfe einer solchen Equalizer-Kette nachgebildet werden soll, sei mit $X(e^{j\Omega})$ bezeichnet. Als Kostenfunktion bietet sich eine gewichtete mittlere quadratische Abweichung zwischen der logarithmierten spektralen Hüllkurve des Eingangssignals und dem logarithmierten Betragsfrequenzgang des Equalizers an:

$$E = \int_0^{\pi} w(\Omega) \left(\ln|X(e^{j\Omega})| - \ln|H_P(e^{j\Omega})|\right)^2 d\Omega$$

Hierbei ist $w(\Omega)$ eine Gewichtungsfunktion die es ermöglicht, in bestimmten Frequenzbereichen eine bessere Anpassung des Equalizers an das Spektrum zu bewirken wobei in anderen Frequenzbereichen unter Umständen eine Verschlechterung der Anpassung in Kauf genommen wird. Die oben genannte Kostenfunktion ist jedoch in erster Linie von theoretischen Wert, da hier angenommen wird, dass unsere zu approximierende Funktion $\ln|X(e^{j\Omega})|$ als kontinuierliche Funktion in Form eines mathematischen Ausdrucks vorliegt (und dass das Integral gelöst werden kann). In der Praxis haben wir es jedoch mit diskreten Spektren zu tun - d.h. unsere spektrale Hüllkurve ist nur an bestimmten Stützstellen Ω_k mit $k = 0, ..., K-1$ definiert, wobei K die Anzahl dieser Stützstellen (Bins) ist. Daher geht das Integral über in eine Summe:

$$\boxed{E = \frac{1}{K} \sum_{k=0}^{K-1} w(\Omega_k) \left(\ln|X(e^{j\Omega_k})| - \ln|H_P(e^{j\Omega_k})|\right)^2} \tag{2.12}$$

Außerdem wurde $\frac{1}{K}$ als Normierungsfaktor eingeführt - hauptsächlich deshalb, damit später sinnvolle Werte für Lernraten nicht von der Anzahl der Bins K abhängen. Diese Kostenfunktion sollte bezüglich der Parameter Ω_{cp}, g_p, bw_p mit $p = 1, ..., P$ minimiert werden. Dazu werden in den folgenden Abschnitten Ausdrücke für die partiellen Ableitungen der Kostenfunktion bezüglich dieser Parameter hergeleitet. Wenn man diese partiellen Ableitungen kennt, können gradientenbasierte Optimierungsverfahren eingesetzt werden. Da man die Reihenfolge der P Equalizerstufen beliebig vertauschen (permutieren) kann, ergeben sich schonmal $P!$ äquivalente Lösungen - wir lösen dies auf, indem wir festlegen, dass die erste Stufe für den untersten Formanten, die zweite für den nächsten usw. zuständig ist.

2.3.2 Herleitung des Betragsfrequenzgangs einer Biquad-Filterstufe

In diesem Abschnitt soll ein allgemeiner Ausdruck für den Betragsfrequenzgang einer einzelnen Biquad-Filterstufe werden, da dies für spätere Betrachtungen benötigt wird. Zunächst betrachten wir das Betragsquadrat $|H(e^{j\Omega})|^2$ des Biquads. Hierfür gilt:

$$\begin{aligned}
|H(e^{j\Omega})|^2 &= H(e^{j\Omega})H^*(e^{j\Omega}) \\
&= H(e^{j\Omega})H(e^{-j\Omega}) \\
&= \frac{b_0 + b_1 e^{-j\Omega} + b_2 e^{-2j\Omega}}{a_0 + a_1 e^{-j\Omega} + a_2 e^{-2j\Omega}} \cdot \frac{b_0 + b_1 e^{j\Omega} + b_2 e^{2j\Omega}}{a_0 + a_1 e^{j\Omega} + a_2 e^{2j\Omega}} \\
&=: \frac{Num}{Den}
\end{aligned}$$

wobei die Abkürzungen Num für den Zähler (englisch: Numerator) und Den für den Nenner (englisch: Denominator) eingeführt wurden mit dem Hintergedanken, das ganze Ausmultiplizieren nur für den Zähler Num durchzuführen - für den Nenner Den sieht die Rechnung dann ganz genauso aus. Es gilt also:

$$\begin{aligned}
Num &= (b_0 + b_1 e^{-j\Omega} + b_2 e^{-2j\Omega}) \cdot (b_0 + b_1 e^{j\Omega} + b_2 e^{2j\Omega}) \\
&= b_0^2 + b_0 b_1 e^{j\Omega} + b_0 b_2 e^{2j\Omega} \\
&\quad + b_1 e^{-j\Omega} b_0 + b_1 e^{-j\Omega} b_1 e^{j\Omega} + b_1 e^{-j\Omega} b_2 e^{2j\Omega} \\
&\quad + b_2 e^{-2j\Omega} b_0 + b_2 e^{-2j\Omega} b_1 e^{j\Omega} + b_2 e^{-2j\Omega} b_2 e^{2j\Omega} \\
&= b_0^2 + b_0 b_1 e^{j\Omega} + b_0 b_2 e^{2j\Omega} \\
&\quad + b_0 b_1 e^{-j\Omega} + b_1^2 \underbrace{e^{-j\Omega} e^{j\Omega}}_{e^{j\Omega - j\Omega} = 1} + b_1 b_2 \underbrace{e^{-j\Omega} e^{2j\Omega}}_{e^{j\Omega}} \\
&\quad + b_0 b_2 e^{-2j\Omega} + b_1 b_2 \underbrace{e^{-2j\Omega} e^{j\Omega}}_{e^{-j\Omega}} + b_2^2 \underbrace{e^{-2j\Omega} e^{2j\Omega}}_{1}
\end{aligned}$$

Vereinfachen und Zusammenfassen von Termen mit gleichen Koeffizienten:

$$\begin{aligned}
Num &= b_0^2 + b_0 b_1 e^{j\Omega} + b_0 b_2 e^{2j\Omega} + b_0 b_1 e^{-j\Omega} + b_1^2 + b_1 b_2 e^{j\Omega} + b_0 b_2 e^{-2j\Omega} + b_1 b_2 e^{-j\Omega} + b_2^2 \\
&= b_0^2 + b_1^2 + b_2^2 + b_0 b_1 e^{j\Omega} + b_0 b_1 e^{-j\Omega} + b_0 b_2 e^{2j\Omega} + b_0 b_2 e^{-2j\Omega} + b_1 b_2 e^{j\Omega} + b_1 b_2 e^{-j\Omega}
\end{aligned}$$

Ausklammern der Koeffizienten wenn möglich und Anwendung von $\cos(\varphi) = \frac{e^{j\varphi} + e^{-j\varphi}}{2}$:

$$\begin{aligned}
Num &= b_0^2 + b_1^2 + b_2^2 + b_0 b_1 \underbrace{(e^{j\Omega} + e^{-j\Omega})}_{2\cos(\Omega)} + b_0 b_2 \underbrace{(e^{2j\Omega} + e^{-2j\Omega})}_{2\cos(2\Omega)} + b_1 b_2 \underbrace{(e^{j\Omega} + e^{-j\Omega})}_{2\cos(\Omega)} \\
&= b_0^2 + b_1^2 + b_2^2 + 2\cos(\Omega) b_0 b_1 + 2\cos(2\Omega) b_0 b_2 + 2\cos(\Omega) b_1 b_2 \\
&= b_0^2 + b_1^2 + b_2^2 + 2\cos(\Omega)(b_0 b_1 + b_1 b_2) + 2\cos(2\Omega) b_0 b_2
\end{aligned}$$

Für den Nenner Den sieht die Rechnung, wie gesagt, genauso aus. Es ergibt sich also:

$$Den = a_0^2 + a_1^2 + a_2^2 + 2\cos(\Omega)(a_0a_1 + a_1a_2) + 2\cos(2\Omega)a_0a_2$$

Für unseren quadrierten Betragsfrequenzgang ergibt sich damit also insgesamt:

$$|H(e^{j\Omega})|^2 = \frac{b_0^2 + b_1^2 + b_2^2 + 2\cos(\Omega)(b_0b_1 + b_1b_2) + 2\cos(2\Omega)b_0b_2}{a_0^2 + a_1^2 + a_2^2 + 2\cos(\Omega)(a_0a_1 + a_1a_2) + 2\cos(2\Omega)a_0a_2}$$

und somit für den gesuchten Betragsfrequenzgang:

$$\boxed{|H(e^{j\Omega})| = \sqrt{\frac{b_0^2 + b_1^2 + b_2^2 + 2\cos(\Omega)(b_0b_1 + b_1b_2) + 2\cos(2\Omega)b_0b_2}{a_0^2 + a_1^2 + a_2^2 + 2\cos(\Omega)(a_0a_1 + a_1a_2) + 2\cos(2\Omega)a_0a_2}}} \tag{2.13}$$

Für P in Reihe geschaltete solche Stufen ergibt sich der Betragsfrequenzgang als Produkt über Betragsfrequenzgänge der einzelnen Stufen:

$$|H_P(e^{j\Omega})| = \prod_{p=1}^{P} |H_p(e^{j\Omega})|$$

bzw. mit zusätzlichem globalen Gain-Faktor G:

$$\boxed{|H_P(e^{j\Omega})| = G\prod_{p=1}^{P} |H_p(e^{j\Omega})|} \tag{2.14}$$

2.3.3 Der Gradient der Kostenfunktion

Für die Optimierung einer Funktion ist es im allgemeinen nötig, deren partielle Ableitungen bezüglich der zu optimierenden Parameter zu kennen. Da wir die Kostenfuntion

$$\boxed{E = \frac{1}{K}\sum_{k=0}^{K-1} w(\Omega_k)\left(\ln|X(e^{j\Omega_k})| - \ln|H_P(e^{j\Omega_k})|\right)^2}$$

bezüglich der Mittenfrequenzen, Bandbreiten und Verstärkungsfaktoren (die alle in $|H_P(e^{j\Omega_k})|$ stecken) der einzelnen Equalizerstufen optimieren wollen, benötigen wir also Ausdrücke für die Ableitung der Kostenfunktion nach eben diesen Parametern. Diese sollen im Folgenden hergeleitet werden. Danach werden die Ergebnisse benutzt, um einen Algorithmus zur praktischen Berechnung der Ableitungen zu erarbeiten. Werden die einzelnen partiellen Ableitungen zu einem Vektor zusammengefaßt, so wird dieser Vektor als Gradient der Funktion E bezeichnet. Unser Gradient hat demzufolge die Form:

$$\boxed{\nabla E = (\frac{\partial E}{G}\ \frac{\partial E}{\partial g_1}\cdots\frac{\partial E}{\partial g_P}\ \frac{\partial E}{\partial \Omega_{c1}}\cdots\frac{\partial E}{\partial \Omega_{cP}}\ \frac{\partial E}{\partial bw_1}\cdots\frac{\partial E}{\partial bw_P})^T}$$

Herleitung der partiellen Ableitungen der Kostenfunktion bezüglich der Equalizer-Parameter

Wir betrachten in diesem Abschnitt die Ableitung der Kostenfunktion nach einem beliebigen Equalizerparameter θ einer einzelnen Equalizerstufe, die mit dem Index i bezeichnet werden soll. Dabei ist i beliebig aber fest gewählt und der generische Parameter θ ist ein Platzhalter für entweder den Vertärkungsparameter g, den Bandbreitenparameter bw oder den Mittenfrequenzparameter Ω_c. Wir berechnen also:

$$\frac{\partial E}{\partial \theta_i} = \frac{\partial}{\partial \theta_i}\ \frac{1}{K}\sum_{k=0}^{K-1} w(\Omega_k)\left(\ln|X(e^{j\Omega_k})| - \ln|H_P(e^{j\Omega_k})|\right)^2$$

Aufgrund der Linearität der Ableitung, kann die Ableitung in die Summe hineingezogen werden, außerdem ist die Gewichtungsfunktion $w(\Omega_k)$ unabhängig von θ_i und kann deswegen als konstanter Faktor vor die Ableitung gezogen werden:

$$\frac{\partial E}{\partial \theta_i} = \frac{1}{K}\sum_{k=0}^{K-1} w(\Omega_k)\frac{\partial}{\partial \theta_i}\left(\ln|X(e^{j\Omega_k})| - \ln|H_P(e^{j\Omega_k})|\right)^2$$

Zur Vereinfachung der Schreibweise werden im weiteren Verlauf an einigen Stellen die Amplitudenfrequenzgänge von Eingangssignal und Equalizerkette anstatt mit $|X(e^{j\Omega_k})|$ bzw. $|H_P(e^{j\Omega_k})|$ abkürzend mit $|X|$ bzw. $|H_P|$ bezeichnet - wir wissen schließlich dass es sich hier um frequenzabhängige Größen handelt. Zur Ableitung des quadratischen Terms wird zuerst die Kettenregel angewandt:

$$\frac{\partial E}{\partial \theta_i} = \frac{1}{K} \sum_{k=0}^{K-1} w(\Omega_k) \cdot 2(\ln|X| - \ln|H_P|) \cdot \frac{\partial}{\partial \theta_i}(\ln|X| - \ln|H_P|)$$

Da unsere spektrale Hüllkurve $|X|$ nicht von θ_i (oder irgendwelchen anderen Parametern, die wir kontrollieren können) abhängt, wird die entsprechende Ableitung Null:

$$\frac{\partial E}{\partial \theta_i} = \frac{1}{K} \sum_{k=0}^{K-1} w(\Omega_k) \cdot 2(\ln|X| - \ln|H_P|) \cdot \frac{\partial}{\partial \theta_i}(-\ln|H_P|)$$

Umsortieren und Hineinziehen des Minuszeichens in die Klammer:

$$\boxed{\frac{\partial E}{\partial \theta_i} = \frac{1}{K} \sum_{k=0}^{K-1} w(\Omega_k) \cdot 2\frac{\partial}{\partial \theta_i} \ln|H_P| \cdot (\ln|H_P| - \ln|X|)} \tag{2.15}$$

Nun müssen wir also den logarithmierten Betragsfrequenzgang unserer Equalizerkette $\ln|H_P|$ nach θ_i ableiten. Dieser Betragsfrequenzgang ist gegeben mit Gleichung 2.14, somit gilt:

$$\frac{\partial}{\partial \theta_i} \ln|H_P(e^{j\Omega})| = \frac{\partial}{\partial \theta_i} \ln\left(G \prod_{p=1}^{P} |H_p(e^{j\Omega})|\right)$$

Eine angenehme Eigenschaft der Logarithmusfunktion ist es, dass sie Produkte in Summen verwandelt, denn es gilt: $\ln(a \cdot b) = \ln(a) + \ln(b)$, oder allgemeiner: $\ln \prod(\cdot) = \sum \ln(\cdot)$, wobei der Punkt in den Klammern als Platzhalter für irgendwelche Terme verstanden werden soll. Damit haben wir:

$$\frac{\partial}{\partial \theta_i} \ln|H_P(e^{j\Omega})| = \frac{\partial}{\partial \theta_i}(\ln(G) + \sum_{p=1}^{P} \ln|H_p(e^{j\Omega})|) \tag{2.16}$$

Wir nehmen hier an, dass unser generischer Parameter θ_i nicht der globale Gain-Faktor G war - die Ableitung nach diesem Faktor wird später gesondert betrachtet. Damit hängt der G nicht von θ_i ab, und die entsprechende Ableitung wird zu Null, des weiteren können können wir die Ableitung in die Summe hineinziehen:

$$\frac{\partial}{\partial \theta_i} \ln|H_P(e^{j\Omega})| = \sum_{p=1}^{P} \frac{\partial}{\partial \theta_i} \ln|H_p(e^{j\Omega})|$$

An dieser Stelle ist anzumerken, dass die Ableitung von $\ln|H_p|$ nach θ_i nur für die Equalizerstufe p mit dem Index i einen Wert ungleich Null ergibt - alle anderen Übertragungsfunktionen hängen ja nicht von θ_i ab und sind somit als Konstanten bezüglich θ_i anzusehen - die entsprechenden Ableitungen werden also zu Null. Daher bleibt von der Summe nur der Term für $p = i$ übrig:

$$\frac{\partial}{\partial \theta_i} \ln |H_P(e^{j\Omega})| = \frac{\partial}{\partial \theta_i} \ln |H_i(e^{j\Omega})| \tag{2.17}$$

Hier ist etwas bemerkenswertes passiert: die Ableitung des logarithmierten Betragsfrequenzgangs der Equalizerkette nach θ_i ist somit einfach gleich der Ableitung des logarithmierten Betragsfrequenzgangs der Equalizerstufe mit dem Index i. Bezüglich der weiteren Herleitung entkoppeln die einzelnen Equalizerstufen demnach. Daher können wir die Schreibweise etwas vereinfachen und den Index i ab jetzt weglassen. Wenn wir später für eine bestimmte Equalizerstufe die Ableitung suchen, kann der Index i nach Bedarf wieder eingeführt werden. Die Kettenregel liefert:

$$\frac{\partial}{\partial \theta} \ln |H(e^{j\Omega})| = \frac{1}{|H(e^{j\Omega})|} \frac{\partial}{\partial \theta} |H(e^{j\Omega})| \tag{2.18}$$

wobei mit $|H|$ in diesem Falle der Betragsfrequenzgang einer einzelnen Biquad-Stufe ist (und zwar derjenigen mit dem Index i). Ein Ausdruck für den Betragsfrequenzgang einer Biquad-Stufe wurde zuvor mit Gleichung 2.13 bestimmt. Wir setzen dieses Ergebnis nun also ein:

$$\frac{\partial}{\partial \theta} |H| = \frac{\partial}{\partial \theta} \sqrt{\frac{b_0^2 + b_1^2 + b_2^2 + 2\cos(\Omega)(b_0 b_1 + b_1 b_2) + 2\cos(2\Omega) b_0 b_2}{a_0^2 + a_1^2 + a_2^2 + 2\cos(\Omega)(a_0 a_1 + a_1 a_2) + 2\cos(2\Omega) a_0 a_2}}$$

was sich auch schreiben läßt als:

$$\frac{\partial}{\partial \theta} |H| = \frac{\partial}{\partial \theta} \left(\frac{b_0^2 + b_1^2 + b_2^2 + 2\cos(\Omega)(b_0 b_1 + b_1 b_2) + 2\cos(2\Omega) b_0 b_2}{a_0^2 + a_1^2 + a_2^2 + 2\cos(\Omega)(a_0 a_1 + a_1 a_2) + 2\cos(2\Omega) a_0 a_2} \right)^{\frac{1}{2}}$$

Dabei entspricht das Innere der Klammer (ohne den Exponenten $\frac{1}{2}$) genau $|H|^2$. Wieder wird die Kettenregel angewandt:

$$\begin{aligned} \frac{\partial}{\partial \theta} |H| &= \frac{1}{2} (|H|^2)^{\frac{-1}{2}} \left(\frac{\partial}{\partial \theta} \frac{b_0^2 + b_1^2 + b_2^2 + 2\cos(\Omega)(b_0 b_1 + b_1 b_2) + 2\cos(2\Omega) b_0 b_2}{a_0^2 + a_1^2 + a_2^2 + 2\cos(\Omega)(a_0 a_1 + a_1 a_2) + 2\cos(2\Omega) a_0 a_2} \right) \\ &= \frac{1}{2|H|} \left(\frac{\partial}{\partial \theta} \frac{Num}{Den} \right) \end{aligned} \tag{2.19}$$

wobei wieder der Übersichtlichkeit halber die beiden Hilfsgrößen Num und Den benutzt wurden, die jeweils Zähler und Nenner des Bruchs repräsentieren sollen. Jetzt benötigen wir die Quotientenregel:

$$\frac{\partial}{\partial \theta} \frac{Num}{Den} = \frac{(\frac{\partial}{\partial \theta} Num) \cdot Den - (\frac{\partial}{\partial \theta} Den) \cdot Num}{Den^2} \tag{2.20}$$

Die partielle Ableitung des Zählers nach θ lautet:

$$\frac{\partial}{\partial\theta}Num = \frac{\partial}{\partial\theta}b_0^2 + \frac{\partial}{\partial\theta}b_1^2 + \frac{\partial}{\partial\theta}b_2^2 + 2\cos(\Omega)\frac{\partial}{\partial\theta}(b_0b_1 + b_1b_2) + 2\cos(2\Omega)\frac{\partial}{\partial\theta}(b_0b_2)$$

Ganz analog gilt für die Ableitung des Nenners Den:

$$\frac{\partial}{\partial\theta}Den = \frac{\partial}{\partial\theta}a_0^2 + \frac{\partial}{\partial\theta}a_1^2 + \frac{\partial}{\partial\theta}a_2^2 + 2\cos(\Omega)\frac{\partial}{\partial\theta}(a_0a_1 + a_1a_2) + 2\cos(2\Omega)\frac{\partial}{\partial\theta}(a_0a_2)$$

Hier treten Ableitungen von Produkten von Filterkoeffizienten auf. Diese werden nach der Produktregel wie folgt gebildet (hier nur für die b-Koeffizienten dargestellt):

$$\frac{\partial}{\partial\theta}b_0^2 = 2b_0\frac{\partial b_0}{\partial\theta}; \qquad \frac{\partial}{\partial\theta}b_1^2 = 2b_1\frac{\partial b_1}{\partial\theta}; \qquad \frac{\partial}{\partial\theta}b_2^2 = 2b_2\frac{\partial b_2}{\partial\theta};$$

$$\frac{\partial}{\partial\theta}b_0b_1 = \frac{\partial b_0}{\partial\theta}b_1 + \frac{\partial b_1}{\partial\theta}b_0$$

$$\frac{\partial}{\partial\theta}b_1b_2 = \frac{\partial b_1}{\partial\theta}b_2 + \frac{\partial b_2}{\partial\theta}b_1$$

$$\frac{\partial}{\partial\theta}b_0b_2 = \frac{\partial b_0}{\partial\theta}b_2 + \frac{\partial b_2}{\partial\theta}b_0$$

Wir sehen also, dass wir zunächst einmal die Ableitungen der Koeffizienten b_0, b_1, b_2 bezüglich des Parameters θ berechnen müssen - diese können dann später (zusammen mit den Koeffizienten selbst) mit der Produktregel zu den entsprechenden Ableitungen die in obiger Formel zur Ableitung des Zählers Num zusammengebaut werden. Daher ist es an dieser Stelle an der Zeit, für den generischen Parameter θ wieder die konkreten Filter-Parameter einzusetzen um die entsprechenden Ableitungen zu berechnen.

Aus den Filterdesign-Gleichungen (Gleichung 2.6) für die Koeffizienten ergibt sich für den Verstärkungsparameter g:

$$\begin{aligned}
\frac{\partial}{\partial g}b_0 &= \frac{\partial}{\partial g}\frac{1+\gamma\sqrt{g}}{1+\gamma/\sqrt{g}} \\
&= \frac{(\frac{\partial}{\partial g}(1+\gamma g^{\frac{1}{2}}))(1+\gamma g^{-\frac{1}{2}}) - (\frac{\partial}{\partial g}(1+\gamma g^{-\frac{1}{2}}))(1+\gamma g^{\frac{1}{2}})}{(1+\gamma/\sqrt{g})^2} \\
&= \frac{(\frac{1}{2}\gamma g^{-\frac{1}{2}})(1+\gamma g^{-\frac{1}{2}}) - (-\frac{1}{2}\gamma g^{-\frac{3}{2}})(1+\gamma g^{\frac{1}{2}})}{(1+\gamma/\sqrt{g})^2}
\end{aligned}$$

Der Zähler wird ausmultipliziert und noch etwas vereinfacht:

$$\begin{aligned}
\frac{\partial}{\partial g}b_0 &= \frac{\frac{1}{2}\gamma g^{-\frac{1}{2}} + \frac{1}{2}\gamma g^{-\frac{1}{2}}\gamma g^{-\frac{1}{2}} + \frac{1}{2}\gamma g^{-\frac{3}{2}} + \frac{1}{2}\gamma g^{-\frac{3}{2}}\gamma g^{\frac{1}{2}}}{(1+\gamma/\sqrt{g})^2} \\
&= \frac{\frac{1}{2}\gamma g^{-\frac{1}{2}} + \frac{1}{2}\gamma^2 g^{-1} + \frac{1}{2}\gamma g^{-\frac{3}{2}} + \frac{1}{2}\gamma^2 g^{-1}}{(1+\gamma/\sqrt{g})^2}
\end{aligned}$$

was uns schließlich zur endgültigen Form der inneren Ableitung des Koeffizienten b_0 nach dem Vertärkungsparameter g führt:

$$\boxed{\frac{\partial b_0}{\partial g} = \frac{\frac{1}{2}\frac{\gamma}{\sqrt{g}} + \frac{\gamma^2}{g} + \frac{1}{2}\frac{\gamma}{\sqrt{g^3}}}{(1+\gamma/\sqrt{g})^2}} \tag{2.21}$$

Auf die gleiche Art und Weise ergibt sich für die Ableitungen von b_1 und b_2:

$$\boxed{\frac{\partial b_1}{\partial g} = \frac{-\frac{\gamma}{\sqrt{g^3}}\cos(\Omega_c)}{(1+\gamma/\sqrt{g})^2}} \tag{2.22}$$

$$\boxed{\frac{\partial b_2}{\partial g} = \frac{-\frac{1}{2}\frac{\gamma}{\sqrt{g}} - \frac{\gamma^2}{g} + \frac{1}{2}\frac{\gamma}{\sqrt{g^3}}}{(1+\gamma/\sqrt{g})^2}} \tag{2.23}$$

Aus diesen 3 Ableitungen zusammen mit den Werten von b_0, b_1, b_2 werden wir später mit der Produktregel die Ableitung unserer Hilfsvariable Num nach g zusammenbauen. Für die Ableitungen der a-Koeffizienten (die wir für die Ableitung von Den brauchen) gilt:

$$\boxed{\frac{\partial a_0}{\partial g} = \frac{\partial}{\partial g}1 = 0} \tag{2.24}$$

$$\boxed{\frac{\partial a_1}{\partial g} = \frac{\partial}{\partial g}b_1} \tag{2.25}$$

da ja $a_1 = b_1$ gilt, und:

$$\boxed{\frac{\partial a_2}{\partial g} = \frac{\frac{\gamma}{\sqrt{g^3}}}{(1+\gamma/\sqrt{g})^2}} \tag{2.26}$$

An dieser Stelle haben wir uns also bis zu den inneren Ableitungen durchgearbeitet und können nun durch entsprechende Rücksubstitutionen und Anwendungen von Produkt- und Quotientenregel auch die Ableitungen der Kostenfunktion bezüglich eines Verstärkungsparameters g angeben. Wir wollen es uns jedoch hier ersparen, die gewonnenen Ergebnisse in einer einzigen großen und unübersichtlichen Formel zusammenzufassen. Die notwendigen Schritte werden stattdessen hoffentlich in der Zusammenfassung deutlich werden.

Für die partiellen Ableitungen der Filterkoeffizienten nach dem Bandbreitenparameter γ (die Ableitung nach der eigentlichen Bandbreite bw wird weiter unten betrachtet) ergibt sich :

$$\boxed{\frac{\partial b_0}{\partial \gamma} = \frac{\sqrt{g} - \frac{1}{\sqrt{g}}}{(1+\gamma/\sqrt{g})^2}} \tag{2.27}$$

$$\boxed{\frac{\partial b_1}{\partial \gamma} = \frac{\frac{2}{\sqrt{g}}\cos(\Omega_c)}{(1+\gamma/\sqrt{g})^2}} \tag{2.28}$$

$$\boxed{\frac{\partial b_2}{\partial \gamma} = \frac{-\sqrt{g}-\frac{1}{\sqrt{g}}}{(1+\gamma/\sqrt{g})^2}} \tag{2.29}$$

$$\boxed{\frac{\partial a_0}{\partial \gamma} = 0} \tag{2.30}$$

$$\boxed{\frac{\partial a_1}{\partial \gamma} = \frac{\partial b_1}{\partial \gamma}} \tag{2.31}$$

$$\boxed{\frac{\partial a_2}{\partial \gamma} = \frac{-\frac{2}{\sqrt{g}}}{(1+\gamma/\sqrt{g})^2}} \tag{2.32}$$

Für die partiellen Ableitungen der Filterkoeffizienten nach dem Mittenfrequenzparameter Ω_c ergibt sich:

$$\boxed{\frac{\partial b_0}{\partial \Omega_c} = \frac{\partial b_2}{\partial \Omega_c} = \frac{\partial a_0}{\partial \Omega_c} = \frac{\partial a_2}{\partial \Omega_c} = 0} \tag{2.33}$$

$$\boxed{\frac{\partial b_1}{\partial \Omega_c} = \frac{\partial a_1}{\partial \Omega_c} = \frac{2\sin(\Omega_c)}{1+\gamma/\sqrt{g}}} \tag{2.34}$$

Mit diesen Ableitungen könnten wir im Prinzip arbeiten. Für die Praxis bietet es sich jedoch an, nicht mit dem Parameter γ für die Bandbreite zu arbeiten, u.a. da dieser noch von Ω_c abhängt. Die Bandbreite bw (in Oktaven) hat dieses Problem nicht und ist außerdem intuitiv verständlich. Wenn wir die Kostenfunktion nach einer Bandbreite bw anstatt nach γ ableiten wollen, müssen wir im Prinzip nur noch die Ableitungen für die Filterkoeffizienten nach γ mit der inneren Ableitung von γ nach bw multiplizieren (nach der Kettenregel), das heißt zum Beispiel für b_0:

$$\boxed{\frac{\partial}{\partial bw} b_0 = \frac{\partial}{\partial \gamma} b_0 \cdot \frac{\partial}{\partial bw}\gamma} \tag{2.35}$$

wobei mit Gleichung 2.9:

$$\frac{\partial}{\partial bw}\gamma = \frac{\partial}{\partial bw}\sinh\left(\frac{\ln(2)}{2} bw \frac{\Omega_c}{\sin(\Omega_c)}\right)\sin(\Omega_c)$$

die Ableitung ergibt:

$$\boxed{\frac{\partial}{\partial bw}\gamma = \frac{\ln(2)\cdot\Omega_c}{2}\cosh\left(\frac{\ln(2)}{2}bw\frac{\Omega_c}{\sin(\Omega_c)}\right)} \tag{2.36}$$

Ebenso hätten wir anstatt nach dem Gain-Parameter g nach dessen Version in dB ableiten können. Dies entspricht einer Mulitplikation der Ableitungen der Filterkoeffizienten nach g mit der Ableitung von g nach g_{dB}. Diese ergibt sich als:

$$\begin{aligned}\frac{\partial}{\partial g_{dB}}g &= \frac{\partial}{\partial g_{dB}}10^{g_{dB}/20} \\ &= \frac{\ln(10)}{20}10^{g_{dB}/20}\end{aligned} \tag{2.37}$$

Und ganz analog ergibt sich für die Mittenfrequenz in Hz f_c:

$$\begin{aligned}\frac{\partial}{\partial f_c}\Omega_c &= \frac{\partial}{\partial f_c}2\pi\frac{f_c}{f_s} \\ &= \frac{2\pi}{f_s}\end{aligned} \tag{2.38}$$

wobei die letzten beiden Gleichungen nur der Vollständigkeit halber angegeben sind, in der Praxis werden wir nach den g_i, Ω_{0i} und bw_i ableiten. Die Entscheidung, ob z.B. die Ableitung nach γ oder bw betrachtet wird, ändert die Lage des Minimums nicht, da $\gamma = \gamma(bw)$ als Funktion von bw aufgefaßt, eine monoton steigende Funktion ist. Damit ist ein Minimum der Kostenfunktion E bezüglich eines bestimmten γ^* auch ein Minimum bezüglich der zu γ^* gehörenden Bandbreite bw^* (wobei Minimum bezüglich eines Parameters bedeutet, dass die entsprechende partielle Ableitung verschwindet). Gleiches gilt für $g = g(g_{dB})$ und $\Omega_c = \Omega_c(f_c)$. Es ändert jedoch die Art und Weise, wie wir die Kostenfunktion betrachten (also ob wir sie als z.B. Funktion von γ oder bw ansehen), und damit verändern wir die Form der Kostenfunktion um das gesuchte Minimum herum. Man kann sich diese Entscheidung im Prinzip wie ein Ver- bzw. Entzerrung der Kostenfunktion im Parameterraum vorstellen - und von der Form der Kostenfunktion in der Nähe des gesuchten Minimums hängen letztendlich die Konvergenzeigenschaften des eingesetzten Optimierungsverfahrens ab. Die Entscheidung für bw anstatt für γ liegt im Wesentlichen darin begründet, dass bei gegebener Bandbreite bw, γ von Ω_c abhängt (und das auch noch nichtlinear) und aufgrunddessen mehrere Größenordnungen überspannt - dies macht die Wahl einer passenden Lernrate (z.B. beim Gradientenabstieg) schwieriger.

Die Ableitung nach dem globalen Gain-Faktor G

Eine gewisse Sonderrolle spielt der globale Gain-Faktor G unserer Filterkette, da er nur einmal vorhanden ist (im Gegensatz zu den anderen Parametern, die jeweils pro Filterstufe existieren). Auch dieser Parameter soll optimiert werden - daher benötigen wir auch die partielle Ableitung der Kostenfunktion nach G. Diese ist besonders einfach. Wir steigen ein mit Gleichung 2.15 wobei wir θ_i durch G ersetzen:

$$\frac{\partial E}{\partial G} = \frac{1}{K} \sum_{k=0}^{K-1} w(\Omega_k) \cdot 2 \frac{\partial}{\partial G} \ln |H_P| \cdot (\ln |H_P| - \ln |X|)$$

wobei nach Gleichung 2.16:

$$\frac{\partial}{\partial G} \ln |H_P| = \frac{\partial}{\partial G} (\ln(G) + \sum_{p=1}^{P} \ln |H_p|)$$

Hierbei ist der einzige Summand, der von G abhängt $\ln(G)$, somit ist:

$$\boxed{\frac{\partial}{\partial G} \ln |H_P| = \frac{\partial}{\partial G} \ln(G) = \frac{1}{G}}$$

und damit haben wir alles, was wir brauchen.

Zusammenfassung der Ergebnisse der Herleitung

Wir wollen an dieser Stelle noch einmal die wichtigsten Zwischenergebnisse zusammenfassen, so dass es danach leicht fällt, einen Algorithmus zur Berechnung der partiellen Ableitungen anzugeben und zu implementieren. Ausgangspunkt ist Gleichung 2.15:

$$\frac{\partial E}{\partial \theta_i} = \frac{1}{K} \sum_{k=0}^{K-1} w(\Omega_k) \cdot 2 \frac{\partial}{\partial \theta_i} \ln |H_P| \cdot (\ln |H_P| - \ln |X|)$$

wobei nach Gleichung 2.17:

$$\frac{\partial}{\partial \theta_i} \ln |H_P| = \frac{\partial}{\partial \theta_i} \ln |H_i|$$

dabei ist (mit Gleichungen 2.18 und 2.19):

$$\frac{\partial}{\partial \theta_i} \ln |H_i| = \frac{1}{2|H_i|^2} \left(\frac{\partial}{\partial \theta_i} \frac{Num_i}{Den_i} \right)$$

wobei hier der Index i zur klaren Unterscheidung der Equalizerstufen wieder mitgeschrieben wurde. Die Ableitung des Bruchs ergibt sich mit der Qutientenregel:

$$\frac{\partial}{\partial \theta_i} \frac{Num_i}{Den_i} = \frac{(\frac{\partial}{\partial \theta_i} Num_i) \cdot Den_i - (\frac{\partial}{\partial \theta_i} Den_i) \cdot Num_i}{Den_i^2}$$

mit

$$
\begin{aligned}
Num_i &= b_{0i}^2 + b_{1i}^2 + b_{2i}^2 + 2\cos(\Omega)(b_{0i}b_{1i} + b_{1i}b_{2i}) + 2\cos(2\Omega)b_{0i}b_{2i} \\
Den_i &= a_{0i}^2 + a_{1i}^2 + a_{2i}^2 + 2\cos(\Omega)(a_{0i}a_{1i} + a_{1i}a_{2i}) + 2\cos(2\Omega)a_{0i}a_{2i} \\
\frac{\partial}{\partial\theta_i}Num_i &= \frac{\partial}{\partial\theta_i}b_{0i}^2 + \frac{\partial}{\partial\theta_i}b_{1i}^2 + \frac{\partial}{\partial\theta_i}b_{2i}^2 + 2\cos(\Omega)\frac{\partial}{\partial\theta_i}(b_{0i}b_{1i} + b_{1i}b_{2i}) + 2\cos(2\Omega)\frac{\partial}{\partial\theta_i}(b_{0i}b_{2i}) \\
\frac{\partial}{\partial\theta_i}Den_i &= \frac{\partial}{\partial\theta_i}a_{0i}^2 + \frac{\partial}{\partial\theta_i}a_{1i}^2 + \frac{\partial}{\partial\theta_i}a_{2i}^2 + 2\cos(\Omega)\frac{\partial}{\partial\theta_i}(a_{0i}a_{1i} + a_{1i}a_{2i}) + 2\cos(2\Omega)\frac{\partial}{\partial\theta_i}(a_{0i}a_{2i})
\end{aligned}
$$

wobei sich die Ableitungen der verschiedenen Produkte von Filterkoefizienten mit der Produktregel aus den Filterkoeffizienten selbst und den Ableitungen der Koeffizienten (Gleichungen 2.21-2.36) ergeben. Für die Ableitung nach G (d.h. θ_i wird durch G ersetzt) können wir direkt im zweiten Schritt $\frac{\partial}{\partial G}\ln|H_P|$ mit $\frac{1}{G}$ ersetzen.

Ein Algorithmus zur Berechnung des Gradienten

Nun sollen also die Ergebnisse der Herleitung in einen praktischen Algorithmus zu Berechnung des Gradienten übersetzt werden. Als erstes fällt auf, dass wir eine Summe über die Frequenzbins $k = 0, ..., K-1$ zu bilden haben - dies geschieht mit Hilfe einer Schleife über k. Innerhalb dieser Schleife benötigen wir die frequenzabhängigen Größen $w(\Omega_k)$, $\frac{\partial}{\partial\theta_i}\ln|H_P(e^{j\Omega_k})| = \frac{\partial}{\partial\theta_i}\ln|H_i(e^{j\Omega_k})|$, $|H_P(e^{j\Omega_k})|$ und $|X(e^{j\Omega_k})|$. Dabei sind die Größen $w(\Omega_k)$ und $|X(e^{j\Omega_k})|$ von außen vorgegeben, während die Gesamt-Transferfunktion $|H_P(e^{j\Omega_k})|$ von allen Parametern abhängt und $\frac{\partial}{\partial\theta_i}\ln|H_i(e^{j\Omega_k})|$ von den Parametern der jeweiligen Filterstufe i. Für die Berechnung der beiden letzen Größen benötigen wir u.a. die Filterkoeffizienten $b_{0i}, b_{1i}, b_{2i}, a_{0i}, a_{1i}, a_{2i}$ $i = 1, ..., P$, für die zweite Größe außerdem die inneren Ableitungen der Koeffizienten nach den entsprechenden Parametern - diese Größen sind jedoch nicht abhängig von der aktuellen Frequenz Ω und können daher vor der Schleife über Ω berechnet werden. Insgesamt bietet sich somit folgende Vorgehensweise an (hier in einer Art Pseudocode):

```
for i=1,...,P  % Schleife ueber die Filterstufen
 Berechne die Filterkoeffizienten fuer Stufe i aus den jeweiligen Parametern
end


for i=1,...,P  % Schleife ueber die Filterstufen
 Berechne die Ableitungen der Filterkoeffizienten fuer Stufe i nach den
 jeweiligen Parametern und die öbentigten Ableitungen von Produkten
 von Filterkoeffizienten
end

...
```

2.3.4 Transformation der Formanten