

Technische Hochschule Deggendorf

Fakultät Angewandte Informatik

(Bachelor-Studiengang Angewandte Informatik)

**Erstellung einer psychoakustischen Rauschformungsfunktion für
Quantisierungsdither in der digitalen Audiosignalverarbeitung**

Seminararbeit für den Kurs

Grundlagen des wissenschaftlichen Arbeitens 2

an der Technischen Hochschule Deggendorf

vorgelegt von:

Prillwitz, Robin

00805291

am: 25. Januar 2023

Prüfer:

Prof. Dr. Kristina Wanieck

Abstrakt

Im letzten Schritt einer digitalen Audioverarbeitungs-Pipeline wird ein Signal mit hoher Bit-Tiefe, meistens im Floating Point Format, in eine niedrigere Bit-Tiefe konvertiert um das Signal praktisch verwenden zu können. Bei dieses Quantisierungsprozesses entstehen Quantisierungsfehler. Diese Verzerrungen können durch Dither gemindert und anschließend durch Rauschformung spektral manipuliert werden. In dieser Arbeit wird eine numerische Methode anhand des Häufigkeitsstichprobenverfahrens zur Ermittlung einer arbiträren Rauschformungsfunktion präsentiert. Dieses wird dann zur Erstellung eines psychoakustischen Rauschformungsfilters basierend auf *ISO226* angewendet. Diese Methode liefert offline berechnete Koeffizienten für einen Finite Impulse Response (FIR) Filter.

Inhaltsverzeichnis

Abbildungsverzeichnis	iii
Gleichungsverzeichnis	iv
Akronyme	v
Glossar	vi
1 Quantisierung und Dither	1
2 Rauschformung	3
2.1 Ziele und Konzept	3
2.2 Arbiträre Rauschformungsfunktion	4
2.3 Psychoakustisch Ideale Funktion	6
3 Diskussion	6
Literaturverzeichnis	8

Abbildungsverzeichnis

1	Quantisiertes Signal	1
2	Blockdiagramm der Rauschformung	4
3	Rauschspektren mit und ohne Rauschformung	5
4	Normalisierter Bodeplot des Filters	7

Gleichungsverzeichnis

1	“Mid-tread” Quantisierungsfunktion	2
2	Quantisierungsfehler	2
3	FIR Filter Übertragungsfunktion	4
4	Auswertung der Zielfunktion für das Häufigkeitsstichprobenverfahren	4
5	Inversefouriertransformation im kontinuierlichen Zeitbereich	5
6	Diskrete Inversefouriertransformation für FIR Koeffizienten	5
7	Kosinusfensterfunktion	5
8	Gainoptimierung der Zielfunktion	6
9	Psychoakustische Rauschforungszielfunktion	6

Akronyme

ATH	Absolut Threshhold of Hearing.
DAC	Digital Analog Converter.
FIR	Finite Impulse Response.
IIR	Infinite Impulse Response.
LSB	Least Significant Bit.
SNR	Signal to Noise Ratio.
TPDF	Triangular Probability Density Funtion.

Glossar

dBFS	Dezibel Full-Scale. Der Maximale Wert, den ein digitales System verarbeiten kann liegt bei 0dBFS.
Fixed Point	Ganzzahl.
Floating Point	Gleitkommazahl, oft nach IEEE754 implementiert.
Hard-Clipping	Harte, nichtlineare Begrenzung.
Phon	Logarithmischer Lautstärkepegel, drückt die empfundene Lautstärke aus (Auch bekannt als Fon).
Soft-Clipping	Weiche Begrenzung an einen Maximalwert.

1 Quantisierung und Dither

Eindimensionale Signale, die dem Shannon Abtasttheorem unterliegen, sind temporal diskret. Aufgrund des Theorems kann die originale kontinuierliche Version des Signals ohne Informationsverlust rekonstruiert werden [1, S. 11 f.]. Die Wertemenge eines digitalen Signals wird, aufgrund nur endlichen Speichers, von einer theoretisch unendlich kontinuierlichen, auf eine definiert begrenzte Wertemenge abgebildet. Bei diesem Vorgang – Quantisierung – gehen Informationen irreversibel verloren. Ein solches diskret quantisiertes Signal und dessen kontinuierliches Original ist in Abbildung 1 gezeigt. Der Quantisierungsfehler, also die durch den Quantisierungsprozess entstehende Differenz, wird ebenfalls aufgeführt. In der digitalen Signalverarbeitung von Audiosignalen können sowohl Floating Point als auch Fixed Point Repräsentationen verwendet werden, um die zeit-diskreten Werte eines Signals zu speichern. Der größte Unterschied dabei ist der Dynamikbereich – die maximal mögliche Differenz zwischen dem Minimum und Maximum. Tabelle 3 führt den Dynamikbereich verschiedener Formate nach Smith [2] auf. Aufgrund der enorm großen Dynamik und der Einfachheit der Implementation ist Floating Point oft die bevorzugte Wahl für digitale Signalverarbeitung [2, S. 68 ff.].

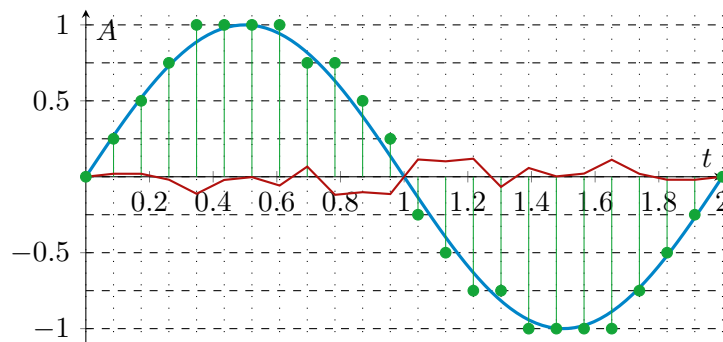


Abbildung 1: Quantisiertes Signal. Die grünen Datenpunkte weichen leicht von den originellen blauen ab. Die rote Linie zeigt den Quantisierungsfehler aus der Differenz zwischen dem kontinuierlichen und dem rekonstruierten diskreten Signal. Der Fehler kann maximal $\frac{1}{2}$ Least Significant Bit (LSB) betragen.

Zur Reproduktion von verarbeiteten Daten auf analogen Endgeräten, durch einen Digital Analog Converter (DAC), zur endgültigen Speicherung oder zur Distribution, wird jedoch eine Fixed Point Repräsentation wünschenswert bzw. erforderlich. Daher ergibt sich die Notwendigkeit für Quantisierung: die Konvertierung von einer hohen zu einer niedrigen Bit-Tiefe [3, S. 499]. Es verändert sich bei dieser Transformation nicht nur die Wertemenge, sondern

Tabelle 3: Dynamische Reichweite verschiedener Formate

Format	Min in dBFS	Max in dBFS	Δ in dB
8-Bit Fixed	−48	0	48
16-Bit Fixed	−96	0	96
24-Bit Fixed	−144	0	144
32-Bit Fixed	−192	0	192
32-Bit Float	−758	770	1528

auch die Auflösung. Die Wertemenge ist nur eine lineare Veränderung und kann hier ignoriert werden. Durch die Reduktion in Auflösung geht jedoch Präzision verloren. Daher wird eine Quantisierungsfunktion Q nach Wannamaker, Lipshitz, Vanderkooy et al. [3, S. 500] definiert:

$$Q[w] = \Delta \left\lfloor \frac{w}{\Delta} + \frac{1}{2} \right\rfloor \quad (1)$$

Diese gilt für einen Eingangswert w . Die Quantisierungsgröße Δ repräsentiert dabei die Größe eines LSB. Die Klammern deuten auf Runden nach $-\infty$. Der durch die Quantisierung entstehende absolute numerische Fehler e ist die Differenz zu dessen Eingangssignal $x[n]$ [3, S. 500]:

$$e[n] = Q[x[n]] - x[n]. \quad (2)$$

Dieser Fehler ist in direkter Abhängigkeit zum Eingangssignal $x[n]$ und ist daher periodisch und manifestiert sich als harmonische Verzerrung [3, S. 500][4, S. 147]. Diese ist in den meisten Anwendungsfällen eine unerwünschte Eigenschaft.

Die Verzerrung kann nicht aus dem Signal eliminiert werden. Informationen werden durch das nichtlineare Runden irreversibel verloren. Durch das Hinzufügen von nichtdeterministischem Rauschen – dem Dither – vor der Quantisierung kann die Periodizität der Quantisierung reduziert werden. Das Signal bleibt verzerrt, jedoch ist die Verzerrung nicht mehr harmonisch, sondern bestenfalls ohne jegliche Korrelation zu dem Quantisierungsfehler. Das Dither ist spektral uniform und wird oft in einer Triangular Probability Density Function (TPDF) verwendet. TPDF bezieht sich auf die Form, die das Histogramm der Ditherwerte über einen längeren Zeitraum annimmt. Ein TPDF-Dither setzt sich aus zwei unabhängigen Rauschquellen mit je einer Amplitude von einem LSB zusammen [3, S. 507 f.][2, S. 23 ff. und S. 30 ff.]. Das Signal ist frei von unerwünschten Übertönen, jedoch nun mit einem stärkeren, spektral uniformen, Grundrauschen. Damit kann die harmonische Verzerrung des quantisierten Signals mit einem

durch das Dither induzierten, geringeren, Signal to Noise Ratio (SNR) „getauscht“ werden [4, S. 147]. Für viele Anwendungsfälle ist dieser Ausgangszustand zufriedenstellend [3, S. 514]. Jedoch ist dieser ein Kompromiss, welcher aus der Reduktion der Bit-Tiefe entsteht. Der Ideal-Zustand, ein Signal ohne Verzerrung, ist mit geringerer Bit-Tiefe dennoch unerreichbar.

2 Rauschformung

Ist ein uniformes Rauschen im Signal jedoch nicht zufriedenstellend kann dieses Rauschen auch spektral manipuliert werden. In den folgenden Abschnitten wird das Konzept der Rauschformung und eine Methode zur numerischen Kreation eines Rauschformungsfilters erklärt.

2.1 Ziele und Konzept

Ein Rauschformungsfilter ist ein digitaler Filter der auf dem absoluten Fehler $e[n]$ der Quantisierung operiert und ein gefiltertes Signal rückkoppelt, um eine gewünschte spektrale Form des Dithers zu erreichen. Die schematische Darstellung dieses Vorgangs ist in Abbildung 2 aus [4, S. 148] gezeigt. Ein spektral uniformes Rauschen wird so manipuliert, sodass es nur minimal in einem bestimmten Frequenzbereich auftritt. Abbildung 3 zeigt ein beispielhaftes Resultat von Rauschformung im Spektralbereich. In Audioanwendungen liegt die generell anerkannte maximal zu reproduzierende Frequenz bei 20kHz. Die Abtastrate f_s muss wenigstens das Doppelte davon betragen [1, S. 11 f.]. Bei standardmäßigen Abtastraten von 44.1kHz oder 48kHz überdeckt sich das Spektrum des Rauschens mit dem des Signals immer. Bei identischer Bandbreite des Signals, jedoch höheren Abtastraten, wie 96kHz oder 192kHz kann mehr Quantisierungsrauschen in höhere, für das Signal irrelevante, Oktaven verschoben werden. Dadurch wird der Effekt der Rauschformung verstärkt. Hohe Abtastraten sind jedoch speicherplatz- und rechenintensiver und somit oft nicht praktikabel.

Das Ziel der Rauschformung ist es, einen Filter $H[z]$ zu konzipieren, um das Rauschen in einem System kontrolliert zu manipulieren.

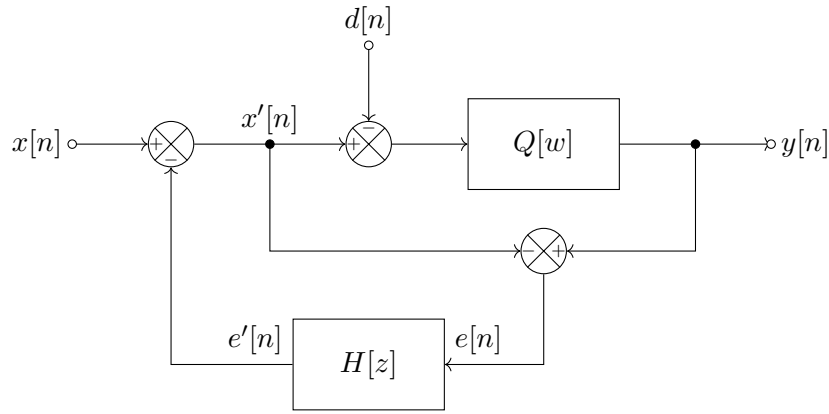


Abbildung 2: Blockdiagramm der Rauschformung aus [4, S. 148]. $x[n]$ bildet den Eingang, $y[n]$ den Ausgang und $d[n]$ das Dither. $Q[w]$ stellt den Quantizer (aus Gleichung 1) dar und $H[z]$ den Rauschformungsfilter, welcher hier als dessen Z-Transformation gezeigt ist.

2.2 Arbiträre Rauschformungsfunktion

Der Filter $H[z]$ ist ein herkömmlicher digitaler Filter. Dieser könnte sowohl als Finite Impulse Response (FIR) oder als Infinite Impulse Response (IIR) Filter implementiert werden. Eine IIR Implementation ist aufgrund der nicht-konstanten Verzögerung und Phasenverschiebungen für diese Anwendung nicht zu nutzen [5]. Da temporale Verzerrungen durch den Filter unerwünschte Interferenzen erzeugen könnten. Der Filter wird somit als FIR in folgender Form aus Verhelst und De Koning [4, S. 149] definiert als:

$$H[z] = \sum_{n=0}^{N-1} h[n]z^{-n} \quad (3)$$

Dabei ist N die Länge des Filters und $h[n]$ die Koeffizienten. Die Länge N kann frei gewählt werden, sollte jedoch eine Zweierpotenz sein.

Wie in [4, S. 149] wird hier auch eine kontinuierliche Zielfunktion $W(\omega)$ eingeführt. N sollte so gewählt sein, dass $W(\omega)$ für eine bestimmte Abtastrate nicht unterabgetastet wird [6]. Dieser Vorgang ist abhängig von der jeweiligen Funktion. Diese werden nach dem Häufigkeitsstichprobenverfahren numerisch zu Filter Koeffizienten umgewandelt [5]. Dabei wird die Funktion zuerst diskretisiert:

$$W[k] = W\left(\frac{2\pi k}{N}\right) \quad k \in 0, 1, 2, \dots, N-1 \quad (4)$$

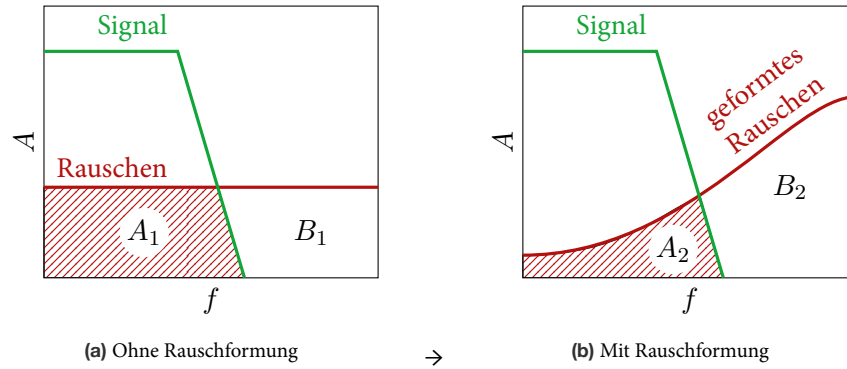


Abbildung 3: Rauschspektren mit und ohne Rauschformung. In Grün wird der Spektralbereich des Signals gezeigt. Der rote Graph zeigt das Quantisierungsrauschen. Der schattierte Bereich zeigt die Schnittmenge des Rauschens im Signal. A und B sind die Flächen unter dem roten Graphen: Es gilt $A_1 + B_1 \equiv A_2 + B_2$. Jedoch ist $A_2 < A_1$. Es wurde weder Energie hinzugefügt, noch reduziert. Die Summe der Energie bleibt konstant. Die Schnittmenge zwischen Rauschen und Signal verringert sich jedoch deutlich.

Da der Filterkernel (die Koeffizienten) eines FIR Filters $h[n]$ per Definition die Antwort des Filters auf einen Dirac-Impuls $\delta[n]$ ist, können wir die Koeffizienten durch die inverse Fourier Transformation der Übertragungsfunktion aus dem Bildbereich in den Zeitbereich konvertieren [5][6][7, S. 339 f.]. Es folgt, dass im kontinuierlichen Raum gilt:

$$h(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} W(\omega) e^{j\omega t} d\omega \quad (5)$$

Im diskreten Raum kann dies formuliert werden als:

$$h[n] = |IDFT(W[k])| \cdot c[n] \quad (6)$$

Das Fenster $c[n]$ aus Gleichung 6 kann frei gewählt werden. Damit Gleichung 5 und Gleichung 6 äquivalent sind, müsste $c[n] = 1$ gewählt werden. Zu beachten ist jedoch, dass manche bestimmte Fenster vor allem bei dieser Methode Trunkierungsartefakte verringern können [7, S. 340 f.]. Als Beispiel wird hier nur ein einfaches Kosinus-Fenster benutzt:

$$c[n] = \frac{1}{2} \cos\left(\frac{2\pi n}{N} - \pi\right) + \frac{1}{2} \quad (7)$$

Die Wahl der Funktion fällt hier nicht auf deren besonderen Leistung, sondern auf deren Einfachheit zur Demonstration.

2.3 Psychoakustisch Ideale Funktion

Zur Psychoakustik wird die Nichtlinearität des menschlichen Ohrs ausgenutzt. Eine ideale Rauschformungsfunktion kann aus einer modifizierten Abwandlung der 0 Phon Absolut Threshold of Hearing (ATH) Kurve aus ISO 226.2003 gewonnen werden. Diese kann numerisch generiert werden [8, S. 2 ff.][9]. Anschließend kann die Funktion auf die gesamte relevante Bandbreite extrapoliert und folglich normalisiert werden. Es entsteht die Funktion $W_\varphi(\omega)$. Diese kann durch eine Hard-Clipping oder Soft-Clipping Funktion auf ein Maximum m beschränkt werden. Die Clipping Funktion wird im folgenden als $S(m, x)$ bezeichnet. Durch das numerische Lösen eines “Least-Squares” Problems, wie in Verhelst und De Koning [4], kann eine optimale Verstärkung (Gain) gefunden werden, damit die Summe des Rauschens über den gesamten Spektralbereich betrachtet 0 ergibt. Dafür muss das globale Minimum des Integrals über die Zielfunktion in Abhängigkeit von dem Gain g gefunden werden [4, S. 149]:

$$g = \underbrace{\left(\frac{2}{f_s} \int_0^{f_s/2} S(m, W_\varphi(\omega)) + g \, d\omega \right)^2}_{\rightarrow 0} \quad (8)$$

Diese Optimierung kann auch numerisch erfolgen. Um den FIR Filter nach der Methode in Unterabschnitt 2.2 zu erstellen wird schließlich folgende Funktion verwendet:

$$W(\omega) = S(m, W_\varphi(\omega)) + g \quad (9)$$

3 Diskussion

Die konstruierte Filterfunktion wird in Abbildung 4 im Vergleich zur psychoakustischen Referenz $W[k]$ illustriert. Die Methode hat keine Garantie für alle Referenzen akkurat zu sein, jedoch kann sie relativ effizient berechnet werden. Auch durch die Modifikation und das Kriterium aus Gleichung 8 entsteht die abgebildete Diskrepanz zwischen den Systemen. Andere Referenzfunktionen, als die hier präsentierte, können bessere Ergebnisse liefern. Weiterhin ist die vorgestellte Methode auch nur Offline ausführbar; eine Online Version für temporale Rauschformung ist nicht vorgesehen, könnte jedoch möglicherweise auf leistungsstarken parallelen Systemen realisiert werden. Dazu müsste $W[k]$ in einer Abhängigkeit von $x[n]$ oder nur n bzw. t definiert werden. Dadurch könnte $W[k]$ kontinuierlich angepasst werden,

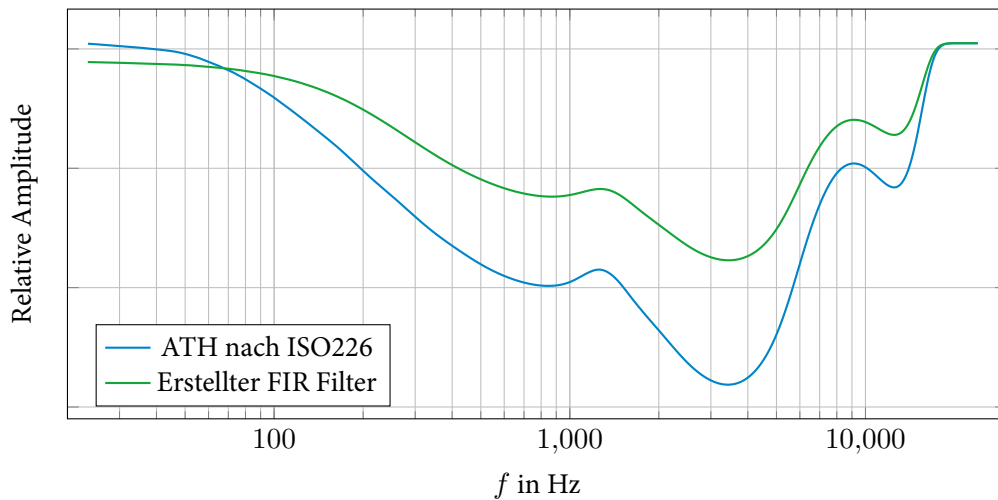


Abbildung 4: Normalisierter Bodeplot des FIR Filters und der modifizierten ATH. Der FIR Filter wurde mit 512 Taps für die Frequenzen von 20Hz bis 24kHz (äquivalent zu $\frac{f_s}{2}$) simuliert.

woraus dauernd neue Koeffizienten für $h[n]$ berechnet werden. Im Vergleich zu Verhelst und De Koning [4] liefert diese Methode vergleichbare Resultate. Der Unterschied liegt jedoch darin, dass hierbei arbiträre Funktionen verwendet werden können. Die ATH Kurve von ISO 226.2003 [8] und Tackett [9] wurde lediglich als Beispiel verwendet; arbiträre Kurven können auch für andere Anwendungsgebiete außerhalb von Audio, welche von Rauschformung profitieren, durch diese Methode konstruiert werden. Der Vorteil der vorgestellten Methode ist die effiziente Berechnung der Koeffizienten, solange eine hinreichend präparierte Zielfunktion $W[k]$ im Vorhinein existiert.

Für Applikationen, in denen starke Quantisierung notwendig ist, beispielsweise für die analoge Reproduktion eines digitalen Signals durch einen DAC mit nur geringer Bit-Tiefe (gewöhnlich 16-Bit), ist Quantisierung unumgänglich. Um jedoch die entstehende Verzerrung zu minimieren, ist Rauschformung eine äußerst ökonomische Maßnahme. Die Ausnutzung von psychoakustischen Effekten, durch die sorgfältige Wahl des Rauschformungsfilters, kann die wahrgenommene SNR die tatsächliche SNR übertreffen. Um ideale Resultate eines Systems zu erhalten, muss anwendungsspezifische Erfahrung in den Entscheidungsprozess mit einfließen. Eine pauschale Aussage über die spezifische Ausführung oder gar Notwendigkeit von Rauschformung für eine gegebene Applikation kann nicht gegeben werden und ist zur weiteren Forschung offen.

Literaturverzeichnis

- [1] C. Shannon, „Communication in the Presence of Noise,“ *Proceedings of the IRE*, Jg. 37, Nr. 1, S. 10–21, Jan. 1949. DOI: 10.1109/JRPROC.1949.232969. (besucht am 20. 11. 2022) (siehe Seiten 1, 3).
- [2] S. W. Smith, *Digital Signal Processing: A Practical Guide for Engineers and Scientists*. San Diego, CA, USA: California Technical Publishing, 2002, ISBN: 0-7506-7444-X. Adresse: <http://www.dspguide.com/pdfbook.htm> (besucht am 27. 12. 2022) (siehe Seiten 1, 2).
- [3] R. Wannamaker, S. Lipshitz, J. Vanderkooy und J. Wright, „A Theory of Nonsubtractive Dither,“ *IEEE Transactions on Signal Processing*, Jg. 48, Nr. 2, S. 499–516, Feb. 2000. DOI: 10.1109/78.823976. (besucht am 20. 11. 2022) (siehe Seiten 1–3).
- [4] W. Verhelst und D. De Koning, „Noise shaping filter design for minimally audible signal requantization,“ in *Proceedings of the 2001 IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics (Cat. No.01TH8575)*, New Platz, NY, USA: IEEE Press, Okt. 2001, S. 147–150. DOI: 10.1109/ASPAA.2001.969564. (besucht am 20. 11. 2022) (siehe Seiten 2–4, 6, 7).
- [5] S. Arar, „Design of FIR Filters Using the Frequency Sampling Method,“ All About Circuits. (Sep. 2017), Adresse: <https://www.allaboutcircuits.com/technical-articles/design-of-fir-filters-using-frequency-sampling-method/> (besucht am 20. 11. 2022) (siehe Seiten 4, 5).
- [6] J. O. Smith, *Spectral Audio Signal Processing, Frequency Sampling Method for FIR Filter Design*. Stanford, CA, USA: W3K Publishing, 2011, ISBN: 978-0-9745607-3-1. Adresse: https://ccrma.stanford.edu/~jos/sasp/Frequency_Sampling_Method_FIR.html (besucht am 21. 11. 2022) (siehe Seiten 4, 5).
- [7] G. Peceli und G. Simon, „Generalization of the frequency sampling method,“ in *Quality Measurement: The Indispensable Bridge between Theory and Reality (No Measurements? No Science!) Joint Conference - 1996: IEEE Instrumentation and Measurement Technology Conference and IMEKO Tec*, Bd. 1, Brüssel, Belgien: IEEE Press, Juni 1996, S. 339–343. DOI: 10.1109/IMTC.1996.507403. (besucht am 21. 11. 2022) (siehe Seite 5).
- [8] ISO, „Acoustics — Normal equal-loudness-level contours,“ International Organization for Standardization, Genf, Schweiz, Norm ISO 226:2003, Version 2, Aug. 2003. Adresse: <https://www.iso.org/standard/34222.html> (besucht am 21. 11. 2022) (siehe Seiten 6, 7).
- [9] J. Tackett, „ISO 226 Equal-Loudness-Level Contour Signal.“ Version 1.0.0.0. (Mai 2005), Adresse: <https://www.mathworks.com/matlabcentral/fileexchange/7028-iso-226-equal-loudness-level-contour-signal> (besucht am 21. 11. 2022) (siehe Seiten 6, 7).

Kolophon

Diese Arbeit ist ein \LaTeX 2 ϵ Dokument in Markus Kohm's KOMA-Script Dokumentenklasse. Der Hauptteil ist in Minion Pro, originell entworfen von Robert Slimbach, in einer Größe von 11pt gesetzt. Die Überschriften verwenden Neue Helvetica, eine Neuauflage des originellen Helvetica Schriftschnitts von Max Miedinger und Eduard Hoffmann. Alle Abbildungen und Darstellungen sind Eigenkreationen. Erstellt wurden diese mit TikZ, PGFPLOTS und Matplotlib. Kompiliert wurde das Dokument mit Lua \LaTeX und biber auf MacOS am 25. Januar 2023.

ROBIN PRILLWITZ
MMXXIII.