Wieso erkennt Alexa meine Stimme? Soundanalyse im Mathematikunterricht

Martin Vogt, Abrar Ahmed, Stefan Balke, Simone Bast und Saskia Laufer

„Musik ist für die Seele, was Gymnastik für den Körper.“ Dieses Zitat, welches Platon zugeschrieben wird (siehe etwa Musikschule Wertheim, o. D.), verdeutlicht die Bedeutung von Musik, aber auch Tönen, Sprache und Geräuschen in unserer Lebenswirklichkeit. Neben Texten (siehe Nonnenmann et al., 2023) und Bildern (Bast et al., 2022) ist *Sound* deshalb eine spannende Datenquelle für den Mathematikunterricht mit Bezug zur Lebenswirklichkeit der Lernenden und kann entsprechend gewinnbringend eingesetzt werden.

Dieser Beitrag stellt ein Lernarrangement bestehend aus 4 Lernsituationen vor. Grundlage für die Konzeption der Lernsituationen ist der Lehrplan für das berufliche Gymnasium in Rheinland-Pfalz für das Unterrichtsfach Mathematik (vgl. Pädagogisches Landesinstitut Rheinland-Pfalz, 2014) Konkret wird die Erlangung der Kompetenzen im Lernbereich 1 Denken in funktionalen Zusammenhängen ermöglicht. Inhaltlich bewegen sich die Lernsituationen im Bereich Nichtrationale Funktionen.

Die erste Lernsituation stellt die Grundlagen von Audiosignalen bzw. Tönen ins Zentrum der Betrachtung. Eine wichtige Erkenntnis aus diesem Einstieg in das Lernarrangement ist, dass unsere Wahrnehmung von Tönen „logarithmisch funktioniert“. Die zweite Lernsituation fokussiert die mathematische Modellierung von Audiosignalen durch Sinusfunktionen. Im Rahmen der dritten Lernsituation erarbeiten sich die Lernenden diverse Tools zur Visualisierung, Analyse und Darstellung von Sound. Die abschließende vierte Lernsituation zeigt auf, wo Soundanalyse in der Praxis eine Rolle spielt. Der Blick auf verschiedene Anwendungsfälle verdeutlicht hier die Relevanz der Behandlung von Sound im Mathematikunterricht.

Audiogrundlagen: Oder warum unser Ohr logarithmisch ist

> Lernsituation 1: *„Guten Morgen.“ Wie jeden Morgen, begrüßt Alexa Sie freundlich. Aber nicht nur Alexa, auch Autos oder Handys können Sprache und Musik erkennen, Töne erzeugen oder Stimmen imitieren. Wie geht das? Oder konkreter: Was ist eigentlich ein Ton?*

Im Rahmen dieser ersten Lernsituation kann der Umgang mit Potenzen und Logarithmen eingeübt bzw. vertieft werden. Zusätzlich ist das Erkennen exponentiellen Wachstums Bestandteil der Lernsituation. Nicht zuletzt können Fachtermini, die im Zusammenhang mit der Sinusfunktion auftauchen, von den Lernenden herausgearbeitet werden.

Audiosignale, bzw. Töne, entstehen durch die Vibration oder Schwingung eines Objektes. Dabei kann es sich beispielsweise um die Vibration einer Gitarrensaite handeln, oder um die der menschlichen Stimmbänder. Durch diese Vibrationen werden Moleküle in der Luft in eine wellenförmige Bewegung gesetzt und erzeugen dadurch einen Ton (vgl. Sueur, 2018). Dieser verkörpert gewisse Eigenschaften, die ihn zu einem definierbaren und darstellbaren Objekt machen.

Eine wichtige Eigenschaft ist die oben erwähnte *Wellenbewegung*, die *periodisch* verläuft und graphisch einer Sinusfunktion ähnelt (vgl. Abschnitt „Audiodarstellung“ unten). Zudem ist die *Lautstärke* eines Tons eine wichtige Charakteristik. Bezüglich der Sinuswellen gilt: Je höher die maximale Auslenkung der Welle ist, desto lauter ist ein Ton. Diese Auslenkung wird Amplitude genannt. Zuletzt kann anhand des periodischen Verlaufs des Tons seine *Frequenz* ω ermittelt werden. Diese wird in Hertz (Hz) gemessen und gibt an, wie viele Sinusperioden der Ton innerhalb einer Sekunde durchläuft. Sie ist der Kehrwert der Dauer einer Sinusperiode (T) (vgl. Wagner 2015):

$$\omega := \frac{1}{T}\,[\mathrm{Hz}]$$

Eine Frequenz von 2 Hz bedeutet also, dass in einer Sekunde genau zwei Perioden durchlaufen werden. Generell gilt, je höher die Frequenz ist, desto höher ist der Ton und desto kürzer ist eine Periode auf dem Graphen. Bei der Analyse von Musiksignalen wird ausgenutzt, dass Frequenzen und Tonhöhen verknüpft werden können. So entspricht das in Abbildung 1 zu sehende C4 (technische Notation; deckt sich mit dem mittleren c' auf einer Klaviertastatur) zum Beispiel einer Frequenz von 261,6 Hz und das eine Oktave höhere C5 der doppelten Frequenz von 523,2 Hz. Das eine weitere Oktave höhere C6 entspricht wieder der doppelten Frequenz, also 1046,4 Hz (vgl. Müller, 2015).

Abbildung 1 zeigt einen Ausschnitt einer Klaviertatstatur mit den Standardbezeichnungen der einzelnen Noten laut „MIDI-Notation“ (Musical Instrument Digital Interface). Die „MIDI-Nummer“ ordnet dabei jedem Ton von C0 bis G#9 (Gis 9) eine natürliche Zahl zwischen 0 und 127 zu (vgl. Dannenberg, 2006). C4 besitzt mit dieser Nummerierung beispielsweise die MIDI-Nummer 60 und der üblicherweise zum Stimmen von Instrumenten genutzte Kammerton A4 besitzt die MIDI-Nummer 69.

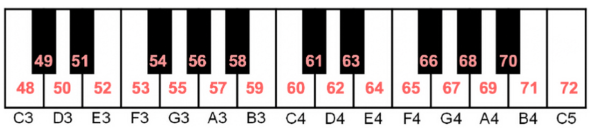

Abbildung 1. Klaviertastatur (www.audiolabs-erlangen.de/resources/MIR/FMP/C1/C1S2_MIDI.html, abgerufen am 4. 11. 2024)

Mithilfe dieser Informationen lässt sich die Mittenfrequenz jeder Taste auf der Klaviatur anhand der folgenden Formel berechnen (vgl. Müller, 2015):

$$\omega(p) = 2^{\frac{p-69}{12}} \cdot 440\,\text{Hz}$$

Hierbei ist $\omega(p)$ die Frequenz der Note p, wobei p durch den Wert der MIDI-Nummer angegeben wird. Wir beziehen uns in dieser Formel auf das bereits erwähnte A4 mit der MIDI-Nummer 69 und der Frequenz 440 Hz. Da eine Oktave in der westlichen Notenschrift in der Regel aus zwölf Halbtönen besteht, sorgt der Divisor 12 im Exponenten dafür, dass eine Note, die eine Oktave höher ist, auch eine genaue Verdopplung in der Frequenz verursacht. Betrachten wir beispielsweise die Note A5 mit der MIDI-Nummer 81, so erhalten wir

$$\begin{aligned}\omega(81) &= 2^{\frac{81-69}{12}} \cdot 440\,\text{Hz} = 2^{\frac{12}{12}} \cdot 440 = 2 \cdot 440\,\text{Hz}\\ &= 880\,\text{Hz}\end{aligned}$$

An der Formel lässt sich schnell erkennen, dass Frequenzen exponentiell steigen. Die Note mit der Notation A2 hat die Frequenz 110 Hz. A3 entspricht 220 Hz, A4 440 Hz, A5 880 Hz, und so weiter. Allerdings hört es sich für uns Menschen so an, als wäre der Abstand zwischen A2 und A3 genauso groß wie der zwischen A3 und A4. Die menschliche Wahrnehmung von Tonhöhen ist also logarithmischer Natur (vgl. Müller, 2015).

Mathematische Modellierung von Tönen

Lernsituation 2: *Sie haben noch immer ein Lied im Kopf, welches Sie gestern zum ersten Mal im Radio gehört haben. Sie kennen den Namen nicht, aber zum Glück die Melodie. Sie summen das Lied und eine App, wie bspw. Google Search, nennt Ihnen daraufhin den Namen des Liedes. Dabei kommt Ihnen folgende Frage: Wieso kann ein Computer mein Summen erkennen? Wie kann Sound mathematisch modelliert werden?*

In dieser zweiten Lernsituation liegt der Fokus auf dem im Lehrplan ausgewiesenen Kompetenzbereich *Funktionale Zusammenhänge modellieren und kommunizieren*. Die Lernenden erstellen die Sinusfunktion zu einem vorgegebenen Audiosignal, indem sie die Werte für Amplitude, Frequenz und Phase im Rahmen einer mathematischen Modellierung an den Graphen anpassen. Für besonders interessierte Lernende kann hier ein Ausblick auf die Fourieranalyse stattfinden.

Die wohl bekannteste Darstellungsform von Tönen zeigt eine „Wellenform“ (vgl. Sueur, 2018). In dieser sind insbesondere die zeitlichen Verläufe, z. B. die sinusartige Vibration von Tönen, sichtbar. Sie ist beispielsweise in Sprachnachrichten verschiedener Messenger-Dienste zu sehen. Ein auf dem Klavier angeschlagenes C5 ist in Abbildung 2 einmal dargestellt.

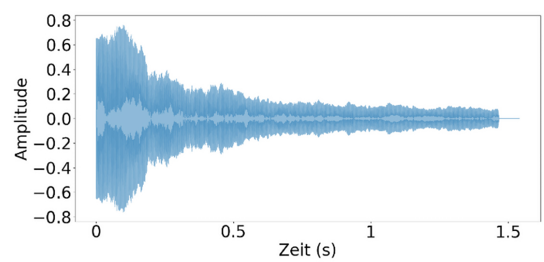

Abbildung 2. Ein auf dem Klavier angeschlagenes C5

Hier wird auf der x-Achse die Zeit in Sekunden angegeben und auf der y-Achse die Amplitude. Letztere wird im Verlauf des Tons kleiner, weil der Ton nach dem Anschlag abklingt und somit leiser wird. Da innerhalb weniger Sekunden mehrere hundert Perioden durchlaufen werden, erkennt man die eigentlichen Wellen dieser Form erst nach dem Heranzoomen in ein deutlich kleineres Zeitfenster (Abb. 3).

Wir erhalten eine wesentlich auseinandergezogene Darstellung eines kleinen Teils des Signals aus Abbildung 2. Weiterhin erkennen wir auf der x-Achse die Zeit in Sekunden und auf der y-Achse die Amplitude. Jedoch lässt sich aus dieser Darstellung noch mehr herauslesen: Wenn sich das periodische Signal in 0,01 Sekunden circa 5-mal wiederholt, dann wiederholt es sich in einer Sekunde ungefähr 500-mal. Wir haben es in diesem Fall also mit einer Frequenz von (etwas mehr) als 500 Hz zu tun. Tatsächlich ist in diesem Graphen der Ton C5 mit einer Frequenz von circa 523,2 Hz dargestellt. Mit unserer groben Abschätzung kamen wir demnach bereits ziemlich nah an das richtige Ergebnis heran.

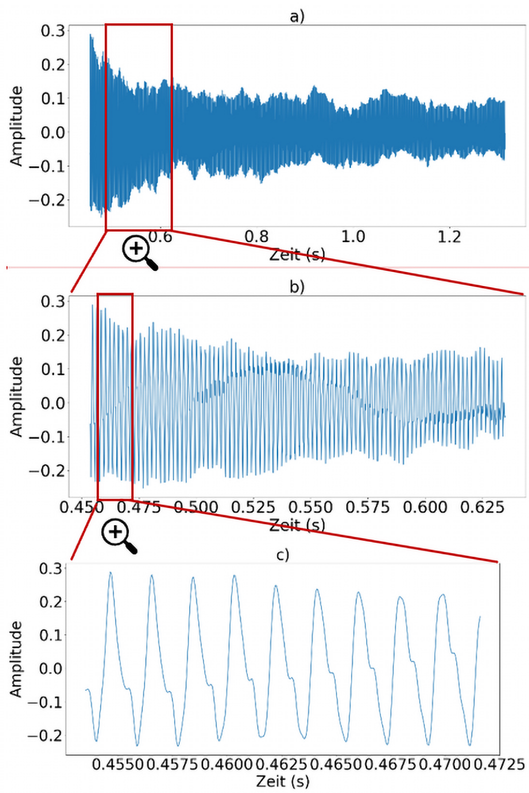

Abbildung 3. Klavier C5 – immer stärker werdender Zoom vom ersten (a) bis zum letzten (c) Graphen

Obwohl die Sinusform des Graphen in Abbildung 3c eindeutig erkennbar ist, lässt sich feststellen, dass es sich keinesfalls um eine „reine“ Sinusfunktion handeln kann. Diese würde schließlich wie in Abbildung 4 aussehen.

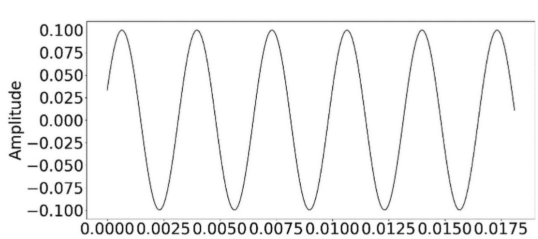

Abbildung 4. Reine Sinusfunktion mit einer Frequenz von 523,2 Hz

Das liegt daran, dass die meisten Töne keineswegs „reine“ Töne einer einzigen Frequenz sind. Eine auf einem Instrument gespielte Note ist deutlich komplexer und besteht aus mehreren Frequenzen, die auch nicht vom Anfang bis zum Ende des Tons gleichbleiben (vgl. Müller, 2015). Ein Ton kann daher als eine Art Überlagerung einzelner „reiner“ Töne, bzw. Frequenzen, gesehen werden. In Abbildung 3 spiegelt sich dies in den leicht unterschiedlichen Perioden wider. Mathematisch betrachtet handelt es sich also nicht um einen einfachen Sinusgraphen, sondern um den Graphen mehrerer aufaddierter Sinusfunktionen, wobei jede einzelne für eine bestimmte Frequenz steht.

Die tiefste in einem Ton vertretene Frequenz wird als Fundamentalfrequenz bezeichnet (vgl. Müller, 2015). Häufig ist sie auch die dominanteste Frequenz im Ton; etwa bei Saiteninstrumenten wie der Gitarre und dem Klavier. Sie bestimmt, als welche Note wir den Ton wahrnehmen. Neben der Fundamentalfrequenz können sowohl harmonische als auch unharmonische weitere Frequenzen im Ton vorkommen (die sog. Obertöne). Als harmonische Frequenz bezeichnen wir jede Frequenz, die aus der Multiplikation einer ganzen Zahl mit der Fundamentalfrequenz besteht (vgl. Müller, 2015). Beispielsweise stehen 261,6 Hz, 523,2 Hz und 784,8 Hz in einem harmonischen Verhältnis. Eine Frequenz, die nicht aus einer ganzzahligen Multiplikation mit der Fundamentalfrequenz gewonnen werden kann, wird als unharmonisch bezeichnet. Instrumente sind in der Regel so aufgebaut, dass sie nahezu ausschließlich harmonische Obertöne, also „Nebenfrequenzen“ produzieren (vgl. Müller, 2015).

Das Zählen von Periodendurchläufen ist nur für einfache Klänge ein schnelles Hilfsmittel. Sobald die Klänge komplexer werden oder durch weitere Instrumente überlagert werden, stößt diese Methode schnell an ihre Grenzen. In der Praxis wird daher die sogenannte Fourieranalyse verwendet, um Töne (oder im Allgemeinen Funktionen) in ihre Frequenzbestandteile zu zerlegen (vgl. Sueur, 2018). Hierbei werden Sinusfunktionen verschiedener Frequenzen mit dem zu testenden Signal verglichen. Passt eine Frequenz gut, bekommt sie einen hohen Wert zugewiesen. Normalerweise benötigt man für die vollständige Fourieranalyse Kenntnisse über den Umgang mit komplexen Zahlen (vgl. Müller, 2015). Da in diesem Fall aber mit Lernenden gearbeitet wird, die über dieses Wissen in der Regel nicht verfügen, wird hier eine vereinfachte Form des Vergleichens, bzw. Modellierens der Originalfrequenz vorgestellt. Zum Erstellen unserer Vergleichsfunktion wird das mathematische Modell einer Schwingung verwendet:

$$f(x) = A \cdot \sin\big(2\pi \cdot (\omega x - \varphi)\big)$$

Es steht dabei A für die Amplitude, ω für die Frequenz und φ für die sogenannte Phase, also die Verschiebung der Funktion entlang der x-Achse (vgl. Müller, 2015). Man kann nun die Parameter A, ω und φ variieren, um eine möglichst passende Sinusfunktion für den vorgegebenen Ton zu finden. Kriterien für die Ähnlichkeit einer Funktion sind hierbei beispielsweise eine gleiche Periodenlänge, gleiche Nullstellen, gleiche Amplitude und abschnittsweise Vorzeichengleichheit.

In Abbildung 5 wird das Signal aus Abbildung 3c mit den Sinusfunktionen verschiedener Frequenzen, Phasen und Amplituden verglichen. Dabei wird sich hier zunächst auf das Finden der Fundamentalfrequenz konzentriert, bevor Obertöne durch Addition hinzugefügt werden. Dazu verwenden wir die folgenden vier Funktionen:

$$f_a(x) = 0,1 \cdot \sin\big(2\pi \cdot (300x - 0)\big)$$
$$f_b(x) = 0,1 \cdot \sin\big(2\pi \cdot (523{,}2x - 0)\big)$$
$$f_c(x) = 0,1 \cdot \sin\big(2\pi \cdot (523{,}2x - 0,55)\big)$$
$$f_d(x) = 0,28 \cdot \sin\big(2\pi \cdot (523{,}2x - 0,55)\big)$$

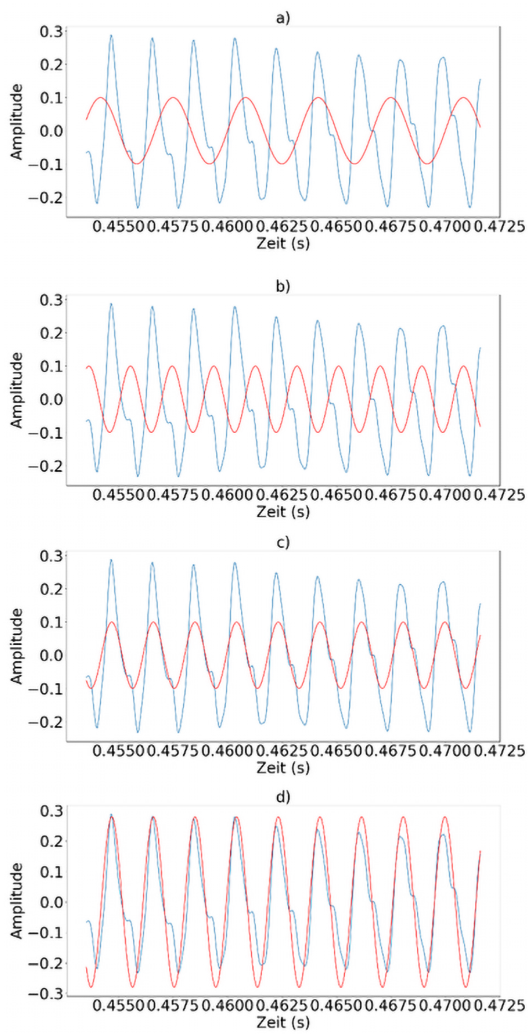

Abbildung 5. Vergleich des Signals aus Abbildung 3 mit den Sinusfunktionen $f_{\{a,b,c,d\}}$

Alle vier Abbildungen in Abbildung 5 zeigen in Blau das Originalsignal und in Rot die jeweils variierende Vergleichsfunktion. In Abbildung 5a ist die Frequenz (300 Hz) unpassend zum Signal und daher weist auch der Graph optisch keine Ähnlichkeit zum Signal auf. Im Gegensatz dazu wurde im Graphen von Abbildung 5b bereits die korrekte Frequenz gefunden, die allerdings gegensätzlich zum Originalsignal verläuft: Wenn das Originalsignal positiv ist, ist das Testsignal negativ und andersherum. Daher wird nun die richtige Phase gesucht. Diese ist in Abbildung 5c dargestellt. In Abbildung 5d wurde schließlich noch die Amplitude an das Originalsignal angepasst, um die optische Ähnlichkeit zu maximieren.

Die Information über die Frequenz hilft, weitere Obertöne zu finden. In Abbildung 6a wird zunächst eine weitere Funktion mit doppelter Frequenz (vgl. Abschnitt „Audiogrundlagen“) und halber Amplitude auf f_d addiert, da die zweite „Welle“ im Graphen deutlich kleiner aussieht. Es ist visuell erkennbar, dass die Phase dieses Obertons noch nicht gut eingestellt ist. Deswegen haftet die Erhöhung der zweiten Welle an der falschen Seite. Durch weiteres Ausprobieren kann die Phase angepasst werden. Abbildung 6b zeigt das Ergebnis mit einer Phasenverschiebung von $\varphi = 0{,}1$.

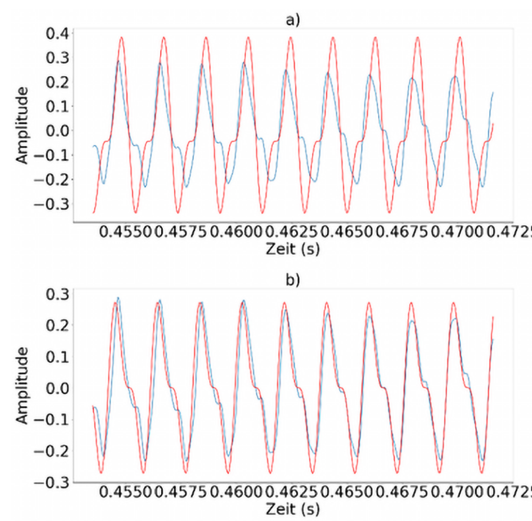

Abbildung 6. Addition weiterer Funktionen

Die Funktionsvorschriften zu den beiden Funktionen aus Abbildung 6 lauten im Ergebnis:

$$\begin{aligned} f(x) =& 0{,}28 \cdot \sin\big(2\pi \cdot (523{,}2x - 0{,}55)\big) \\ &+ 0{,}14 \cdot \sin\big(2\pi \cdot (2 \cdot 523{,}2x - 0{,}55)\big), \end{aligned} \tag{a}$$

$$\begin{aligned} f(x) =& \Big(0{,}28 \cdot \sin\big(2\pi \cdot (523{,}2x - 0{,}55)\big) \\ &+ 0{,}14 \cdot \sin\big(2\pi \cdot (2 \cdot 523{,}2x - 0{,}1)\big)\Big) \cdot 0{,}75. \end{aligned} \tag{b}$$

Der Faktor 0,75 am Ende von Funktionsvorschrift (b) trägt lediglich dazu bei, dass die Amplitude der gesamten Funktion weiterhin mit der des Originalsignals übereinstimmt. Auch er kann durch Ausprobieren ermittelt werden.

Abbildung 7 zeigt der Vollständigkeit halber das Ergebnis einer Fourieranalyse. Da es sich in der Regel um komplexwertige Amplituden handelt (dieser kodiert sowohl Frequenz, als auch Phase in einer komplexwertigen Zahl), wird hier das sog. Betragsspektrum gezeigt.

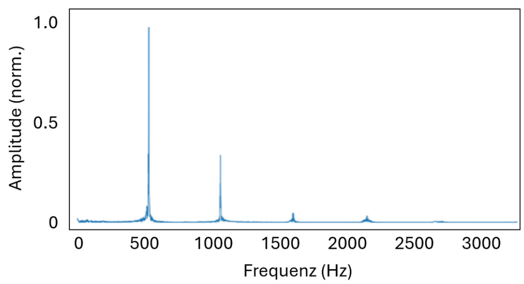

Abbildung 7. Betragsspektrum eines am Klavier gespielten C5

Auf der x-Achse in Abbildung 7 sind die Frequenzen von 0–3000 Hz abgebildet, während die y-Achse die Beträge der Amplituden angibt (normiert auf den höchsten Oberton). Es lässt sich ablesen, dass bei dem dargestellten Ton aus Abbildung 2 vor allem die Frequenzen 523,2 Hz, 1046,4 Hz, 1569,6 Hz und 2092,8 Hz vertreten sind. Erstere ist die am stärksten ausgeprägte Fundamentalfrequenz. Die anderen Frequenzen bilden die Obertöne und sind im Falle des Klaviersignals deutlich schwächer ausgeprägt als die Fundamentalfrequenz. In dieser Darstellungsform verlieren wir allerdings vollständig die Information über die zeitlichen Aspekte des gespielten Tons.

Die wohl bekannteste Darstellungsform, die sowohl zeitliche Aspekte als auch die Frequenz eines Tons angibt, ist das sogenannte Spektrogramm. Es handelt sich hierbei um eine Darstellung, die auf der x-Achse die Zeit betrachtet und auf der y-Achse die Frequenz angibt. Die Stärke der Frequenz lässt sich dann am Farbverlauf des Spektrogramms ablesen. In der Regel gilt: Je heller/deutlicher die Farbe ist, desto stärker ist die jeweilige Frequenz vertreten. Ein Beispiel für ein Spektrogramm des Tons C5 aus Abbildung 2 ist in Abbildung 8 aufgezeigt.

Ähnlich wie die Frequenzdarstellung zeigt das Spektrogramm besonders starke Frequenzausprägungen bei ganzzahligen Vielfachen der Hauptfrequenz 523,2 Hz. Außerdem lässt sich erkennen, dass beim perkussiven Anschlag der Klaviertaste (Hammer trifft Saite), also zu Beginn, sehr viele Frequenzen im Ton vertreten sind – wenn auch schwach (siehe Box 1 in Abbildung 8). Später dominieren die harmonischen Komponenten durch das Schwingen der Saite (siehe Box 2 in Abbildung 8), wobei diejenige(n) Frequenz(n)

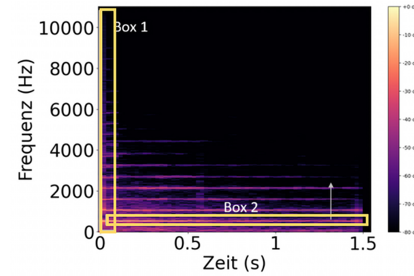

Abbildung 8. Betragsspektrogramm des Klaviertons C5

überwiegt/überwiegen, die zur Länge der Saite passt/passen.

Tools und Software

> Lernsituation 3: *Sie haben bereits viel im Internet zum Thema Sound recherchiert. Hierbei sind Ihnen immer wieder andere Darstellungsformen von Sound begegnet. Sie suchen nun nach Tools, die es Ihnen ermöglichen, Sound darzustellen und Ihre Erkenntnisse Ihren Freunden zu präsentieren.*

Diese dritte Lernsituation bewegt sich im Kompetenzbereich *Funktionale Zusammenhänge darstellen, analysieren und interpretieren*. Den Lernenden wird hier der Umgang mit diversen Tools zur Darstellung von Sound ermöglicht. IT-affine Lernende können eine Steigerung der Selbstwirksamkeit erreichen, indem sie mit diversen Programmiersprachen arbeiten, um Sound sichtbar zu machen.

Zum allgemeinen Verständnis ist es, ähnlich wie bei Bildern, wichtig, Töne nicht nur mathematisch darzustellen, sondern auch abspielen zu können. Im Idealfall gelingt es einer Lehrperson, ihren Lernenden den Zusammenhang zwischen Musik und Mathematik auch auf audiovisuelle Weise näherzubringen. Hierzu gibt es einige verschiedene Tools, die im Folgenden vorgestellt werden.

Zunächst können Audiosignale mithilfe von klassischen Programmiersprachen wie Python und R aufbereitet werden. So wurden beispielsweise alle Abbildungen, mit Ausnahme von Abbildung 1 mithilfe von Python erstellt. Der Code, der zur Erstellung der meisten Abbildungen genutzt wurde, basiert auf einer Vorlage auf Github (siehe Musicalkemist, o. D.). Mithilfe der Funktionen dieser Vorlage kann ein Audiosignal, beispielsweise eine von einem Klavier gespielte Note, importiert und das Signal auf verschiedene Weisen visuell dargestellt werden. Das Buch *Fundamentals of Music Processing using Python and Jupyter Notebooks* von Meinard Müller (2021) erklärt, wie Python hierzu verwendet werden kann. Wie auch R genutzt werden kann, erklärt unter anderem das Handbuch *Sound Analysis and Synthesis with R* von Jérôme Sueur (2018).

Neben diesen Programmiersprachen kann auch das Programm GeoGebra zur audiovisuellen Darstellung von Tönen verwendet werden. Unter folgendem Link (www.geogebra.org/m/tycxbu5U) kann zum Beispiel mit den Parametern verschiedener reiner Töne experimentiert werden, sodass die Schüler*innen schnell und spielerisch erkennen, wie die Veränderung von Tonhöhe und Amplitude den Graphen des Tons beeinflussen. Die Benutzeroberfläche des GeoGebra Tools ist in Abbildung 9 dargestellt.

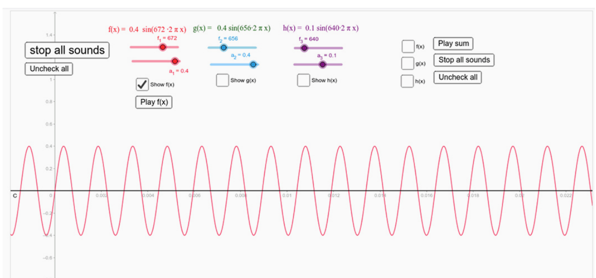

Abbildung 9. Benutzeroberfläche zum Experimentieren mit Sinusfunktionen in GeoGebra

Es können zudem mehrere Töne gleichzeitig abgespielt werden, um den Vergleich verschiedener Töne zu visualisieren. Dieses Tool eignet sich somit hervorragend dazu, den Lernenden einen Einstieg in die Tonanalyse zu bieten – insbesondere, wenn sie ein eigenes digitales Medium wie einen Computer oder ein Tablet zur Verfügung haben.

Zuletzt ist das Spektrogramm des Google Chrome Music Labs (musiclab.chromeexperiments.com/Spectrogram/) ein hervorragendes Tool, das sowohl als Einstieg als auch als Schluss genutzt werden kann. Anders als GeoGebra ermöglicht es dem Nutzer nicht nur den visuellen Vergleich verschiedener Töne, sondern auch den Vergleich verschiedener Instrumente, indem es ein Spektrogramm zeichnet. In Abbildung 10 werden beispielsweise die Frequenzen einer Trompete mit denen einer Harfe verglichen.

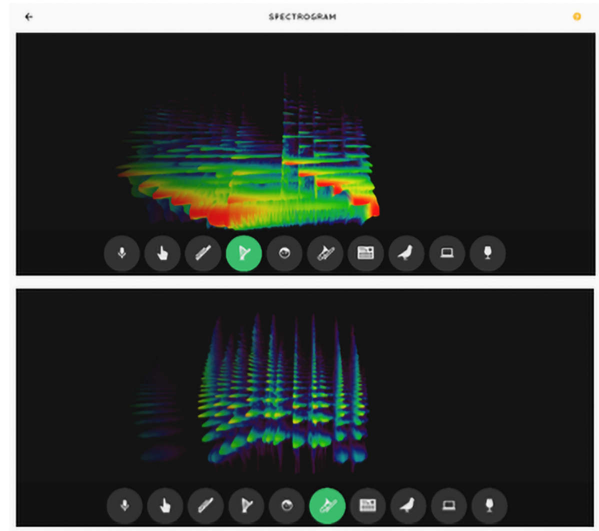

Abbildung 10. Spektrogramm einer Harfe (oben) und das einer Trompete (unten)

Das Spektrogramm des Google Chrome Music Labs ermöglicht es dem Nutzer zudem, selbst in das Gerätemikrofon zu sprechen und das Gesprochene in ein Spektrogramm zu übertragen. Das kann insbesondere dann interessant sein, wenn die Frage aufkommt, wie Alexa verschiedene Buchstaben identifizieren kann. Spricht man nämlich die fünf Vokale auf etwa gleicher Höhe und mit gleicher Lautstärke in das Mikrofon, so lassen sich im Spektrogramm bereits große Unterschiede erkennen. Die Mikrofonfunktion kann des Weiteren genutzt werden, um experimentell ein-, oder mehrstimmig zu singen. Auch hier kann die Veränderung vernommen und notiert werden.

Die Entwicklung künstlicher Intelligenz und digitaler Medien hat neue Werkzeuge zum Abspielen, Erstellen und Analysieren von Sound hervorgebracht. Zahlreiche dieser Tools sind kostenfrei verfügbar, wie Google Chrome Music Labs und können somit leicht in den Unterricht integriert werden.

Ausblick

> *Lernsituation 4: Die Lampe in Ihrem Zimmer „summt“ auf einmal merkwürdig. Kurz darauf geht sie kaputt. Dabei kommt Ihnen eine Idee: Können Daten, wie Sound, verwendet werden, um festzustellen, ob Maschinen oder Bauteile bald kaputt gehen?*

Diese abschließende Lernsituation stellt den Blick über den Tellerrand dar und liefert schlussendlich Antworten auf die Frage: *Wofür brauche ich das überhaupt?*

Neben den vorgestellten Anwendungsfällen in der Analyse von Instrumenten und generell in der Musik, gibt es zahlreiche weitere, wie etwa im industriellen Umfeld. So versuchen Firmen, wie Porsche oder Philips, zum Beispiel die Produktentwicklung oder die Produktion mithilfe von Sound-Messdaten zu verbessern (siehe MHP, o. D.). Hierzu werden die Geräusche einer Maschine aufgezeichnet und analysiert. Das kann auch bei der Fehleranalyse von mechanischen Bauteilen geschehen, indem etwa gegen diese „geklopft“ wird. Ein intaktes Getriebe klingt dann beispielsweise anders als ein defektes. Da die Geräusche eines Motors periodische Elemente besitzen (durch das Drehen des Motors), können diese vielfach ebenfalls wie oben gezeigt als (wellenförmige) Zeitreihe dargestellt werden. Diese industriellen Anwendungen können sogar so weit gehen, zu versuchen vorauszusagen, wann eine Maschine kaputt geht, bevor diese kaputt gegangen ist. Das wird als Predictive Maintenance, also vorausschauende Wartung, bezeichnet (vgl. etwa Bousdekis, 2019). Im Idealfall können dadurch größere Schäden verhindert werden, oder es können bereits Ersatzteile bestellt werden, kurz bevor diese benötigt werden.

Darüber hinaus bietet die aufkommende Verbreitung allgemein zugänglicher generativer Modelle zur Erzeugung von Musik oder dem Klonen von Stimmen, wie Suno AI, die Möglichkeit, mit den Lernenden die Potenziale, aber auch die Gefahren (z. B. in Form von „Fake News") zu diskutieren und ihre Wahrnehmung für diese zu schärfen.

Zusammenfassend lässt sich festhalten, dass der Datentyp Sound, neben Zahlen, Bildern und Texten, ebenfalls im Unterricht thematisiert werden sollte und die Möglichkeit für spannende Anwendungsfälle bietet.

Literaturverzeichnis

Bast, S., Vogt, M., & Wallerath, R. (2022). Autonomes Fahren – Ein Blick hinter die Kulissen der Mathematik der künstlichen Intelligenz. *GDM Mitteilungen, 112*, 7–10.

Bousdekis, A., Lepenioti, K., Apostolou, D., & Mentzas, G. (2019). Decision Making in Predictive Maintenance: Literature Review and Research Agenda for Industry 4.0. *IFAC-PapersOnLine, 52*(13), 707–612. DOI:10.1016/j.ifacol.2019.11.226

Dannenberg, R. B. (2006). The interpretation of MIDI velocity. *International Computer Music Conference*. 193–196.

MHP (o. D.). Automatische Erkennung von Geräuschanomalien mit KI. Abgerufen am 29. 11. 2024, von www.mhp.com/de/solutions/industrial-cloud-solutions/sounce

Müller, M. (2015). *Fundamentals of music processing: Audio, analysis, algorithms, applications* (Vol. 5). Springer Verlag.

Müller, M. (2021). *Fundamentals of music processing: Using Python and Jupyter Notebooks* (Vol. 2). Springer Verlag.

Musicalkemist (o. D.). AudioSignalProcessingForML. Abgerufen am 29.11.2024, von github.com/musikalkemist/AudioSignalProcessingForML/blob/master/10%20-%20Fourier%20Transform%3A%20The%20Intuition/Fourier%20Transform%20.ipynb

Musikschule Wertheim (o. D.). Musikzitate. Abgerufen am 29. 11. 2024, von www.musikschule-wertheim.de/weiterfuehrendes/musikzitate/index.html

Nonnenmann, M., Vogt, M., Bast, S., & Lübke, K. (2013). Mathematik und Sprache – Textanalyse im Mathematikunterricht. *GDM-Mitteilungen, 115*, 33-38.

Pädagogisches Landesinstitut Rheinland-Pfalz (2014). Lehrplan für das berufliche Gymnasium. Unter-richtsfach: Mathematik, Grund- und Leistungsfach, bildung.rlp.de/fileadmin/user_upload/bbs/Informationen_und_Materialien/Lehrplaene/Berufliches_Gymnasium/2015-01-08_LP_BG_Mathe.pdf

Sueur, J. (2018). *Sound analysis and synthesis with R*. Springer.

Wagner, J. (2015). *Physik für Wissenschaftler und Ingenieure* (7. deutsche Auflage). Springer Verlag.

Martin Vogt, Hochschule Trier
m.vogt@wir.hochschule-trier.de

Abrar Ahmed, Hochschule Trier und Universität Trier
a.ahmed@wir.hochschule-trier.de

Stefan Balke, International Audio Laboratories Erlangen*
stefan.balke@audiolabs-erlangen.de

Simone Bast, Staatliches Studienseminar für das Lehramt an berufsbildenden Schulen Trier
simone.bast@bbs-tr.semrlp.de

Saskia Laufer, Hochschule Trier und Universität Trier
laufer@uni-trier.de

*Die International Audio Laboratories Erlangen sind ein Zusammenschluss der Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU) und dem Fraunhofer-Institut für Integrierte Schaltungen IIS. Stefan Balke wird durch die Deutsche Forschungsgemeinschaft (DFG MU 2686/15-1; Grant No. 500643750) unterstützt.