Joachim Heintz

KI und Musik

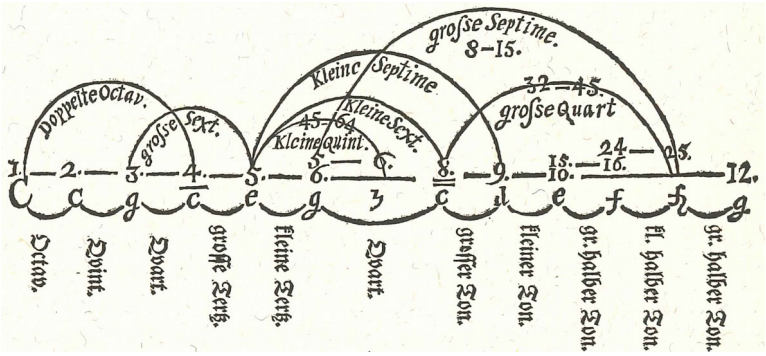
In einer Zeit, in der selbst seriöse Bücher seriöser Autoren in Wissenschaftsverlagen unter Titeln wie *Künstliche Intelligenz — Wann übernehmen die Maschinen?* [Mainzer 2019] erscheinen, in der Computerprogramme die stärksten Spieler in Schach und Go schlagen und KI in verschiedenen Ausprägungen Teil des Alltags wird, erscheint es vielen nur eine Frage der Zeit, bis Algorithmen den Menschen auch in der Komposition von Musik erreichen, wenn nicht übertreffen. "More Bach than Bach" heisst es bereits in einem aktuellen Titel [Miller], und es scheint, als habe man nur die Alternative, sich entweder der marktgetriebenen Euphorie über die jeweils jüngste *disruption* anzuschließen oder sich gegen alle Google-Logik, mit dem geheimen Wissen um einen auf Dauer verlorenen Posten, in der Trutzburg gestrigen Lebens zu verschanzen, das Bild des Menschen als Krone der Schöpfung gegen alle übermächtige Technik verteidigend, notfalls durch Unterbrechung der Stromversorgung [vgl. Heintz 2019].

Gerade in dieser Situation erscheint es für einen Handbuchbeitrag besonders angezeigt, auf einige Grundlagen zurückzugehen, die das Verhältnis von Musik und KI bestimmen. Das betrifft zum einen die Nähe der Musik zur Mathematik, zum anderen die geschichtliche und geografische, soziale und kulturelle Prägung des musikalischen Materials. Diese wird an einigen Stationen angedeutet und ihre Nachschaffung durch den Computer an Hand verschiedener Methoden diskutiert, bevor die Neue Musik der Gegenwart in den Blick kommt. Dabei werden vor allem drei frühe Werke aus den 1950er Jahren unter der Fragestellung algorithmischen Komponierens detaillierter besprochen, weil sich daran besonders anschaulich zeigen lässt, wie unterschiedlich der Einsatz von Algorithmen — und somit potentiell KI — hier verglichen mit der Situation der Stilkopie ist. Abschließend wird versucht, einige der sehr verschiedenen Aspekte des Themas *KI und Musik*, unter deren Vermischung die Diskussion oft leidet, voneinander zu trennen und in ihrem Zusammenspiel zu betrachten.

1. Musik und Mathematik

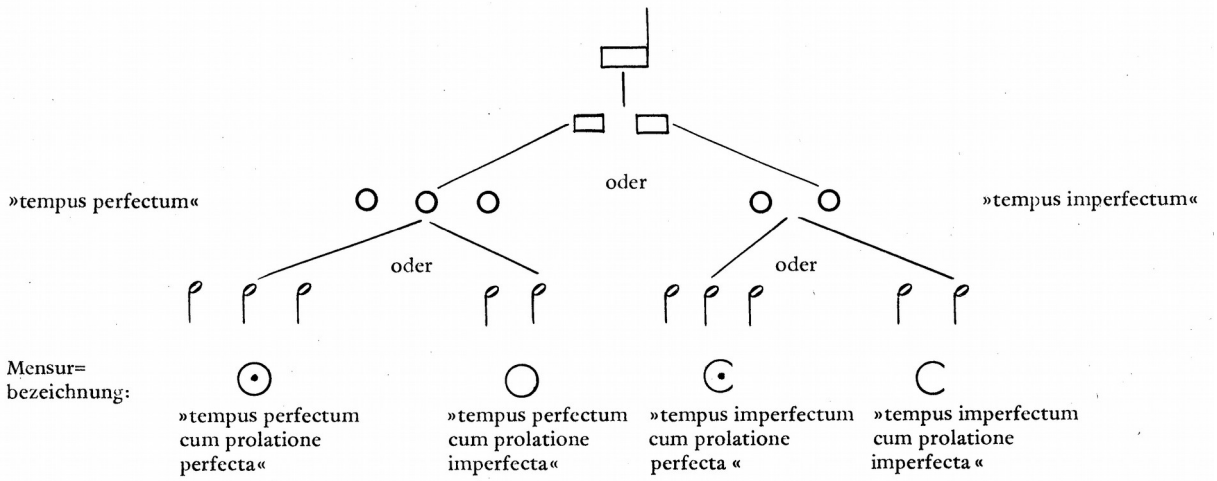
Alan Turing beschrieb die nach ihm benannte Maschine, die nach wie vor die allgemeine Form eines Computers charakterisiert, als Beitrag zu einer mathematischen Grundlagendiskussion [Turing 1936]. Die vielfältigen Beziehungen zwischen Informatik und Mathematik, die sich beispielsweise im Begriff des Algorithmus niederschlagen, sind offensichtlich. Das gilt aber auch, gleichsam von der anderen Seite her, für das Verhältnis von Musik und Mathematik, und zwar lange vor dem Computerzeitalter. Sobald es um den Vergleich von Tonhöhen und Tondauern geht, muss gemessen werden. Das gilt beispielsweise für die chinesische oder die indische Musik ebenso wie für die europäische Musik, die in ihrer Begrifflichkeit auf die griechische Musiktheorie zurückgeht und mit dem Mathematiker-Philosophen Pythagoras als Referenz und Symbol verbunden wird: "Cum igitur ante Pythagoram consonantiae musicae partim diapason partim diapente partim diatessaron [...] vocarentur, primus Pythagoras hoc modo repperit, qua proportione sibimet haec sonorum concordia iungeretur." [Boethius De Musica I, 10] "Da nun vor Pythagoras die musikalischen Consonanzen Octave, Quinte, Quarte [...] genannt wurden, so fand also Pythagoras zuerst die Art und Weise, in welchen Proportionen diese Consonanzen ausgedrückt werden konnten." [Boethius 1985, 16]

Tonhöhen werden gemessen, indem Proportionen bestimmt werden. In der westlichen Tradition geschieht das vor allem mit dem Monochord. Wenn ein bestimmter Ton A mit einem Ton B verglichen wird, der die Hälfte der Saitenlänge von A aufweist, ergibt das ein bestimmtes Intervall, das bei den Griechen *diapason*, im lateinischen Mittelalter dann *octava* genannt wird. Das Verhältnis A:B ist somit, was die Saitenlängen angeht, 2:1. So lassen sich die Töne, auf deren Auswahl und Fixierung die Musik beruht, beschreiben, es sei bei [Aristoxenos], bei Boethius und das ganze Mittelalter hindurch, im Barock noch ganz selbstverständlich bei [Mattheson] (siehe die folgende Abbildung) und in der Neuen Musik etwa bei [Grisey] oder [Zimmermann].



<Abbildung 1: [Mattheson]>

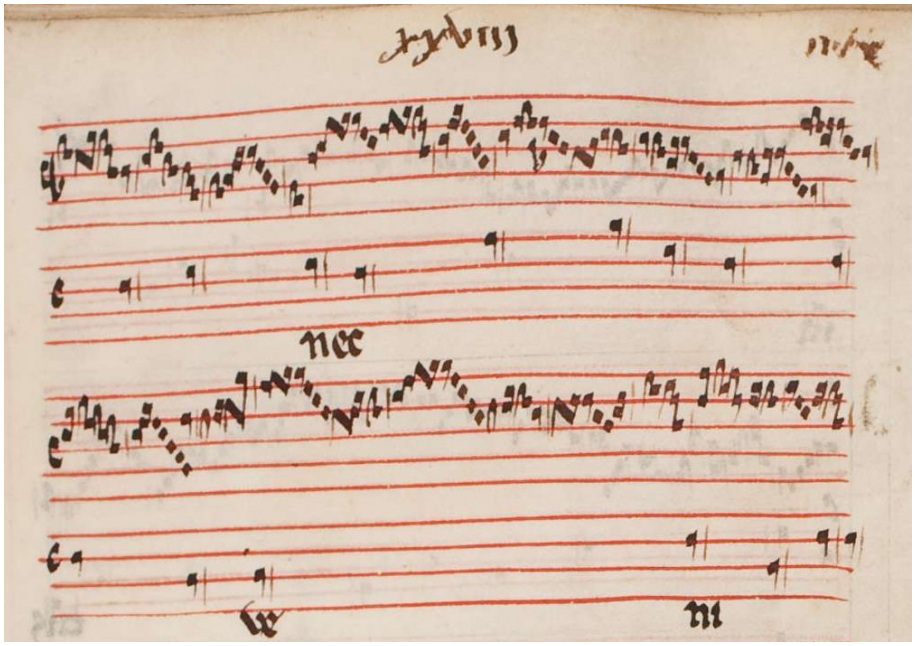
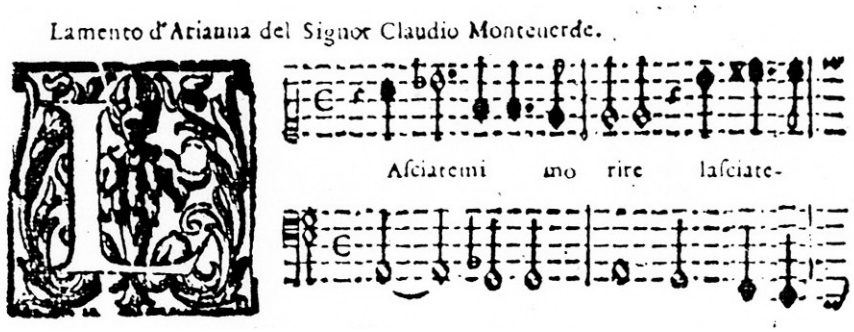
Tondauern lassen sich ebenfalls durch Proportionen bestimmen. Auch ohne mechanische Zeitmessgeräte lässt sich mit hinreichender Genauigkeit bestimmen, ob eine Tondauer halb so lang ist wie eine andere. Die beiden Bestimmungen 2:1 und 3:1 als Verhältnis einer langen zu einer kurzen Note genügte in der westlichen Tradition über viele Jahrhunderte. Als Beispiel hier ein Überblick über das Mensuralsystem des 15. Jahrhunderts [de la Motte]:



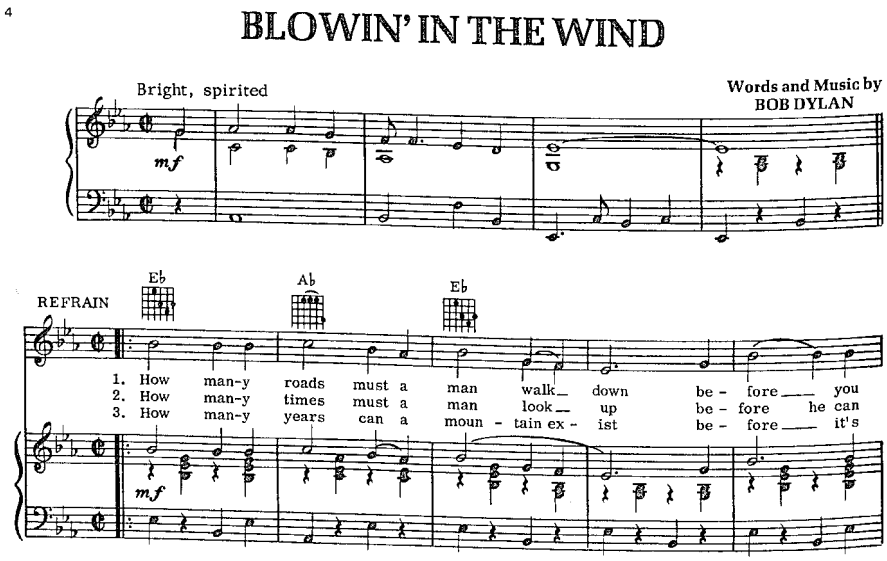
<Abbildung 2: [de la Motte]>

Nun bestand Musik gewiss zu allen Zeiten aus mehr als dem Treffen einer richtigen Tonhöhe und einer richtigen Dauer. Doch die Komposition von Musik, die als Notenschrift festgehalten wurde, beschäftigt sich in der europäischen Musik bis ins 18. Jahrhundert hinein nahezu ausschließlich mit diesen beiden Parametern:

1. Der Tonhöhe als "Stufe" einer "Tonleiter" oder "Skala", die wiederum auf Proportionen zurückgeht, letztlich also mathematisch konstruiert wird.
2. Der Tondauer, die als Proportion zu anderen Tondauern notiert wird, gegebenenfalls unter Angabe einer Referenzdauer.

Dieses Paradigma erfüllen die folgenden Ausschnitte aus dem 13., 17. und 20. Jahrhundert.

<Abbildung 3: [Notre Dame]>

<Abbildung 4: [Monteverdi]>



<Abbildung 5: [Dylan]>

Bei der Diskussion von KI und Musik sollte diese besondere Nähe von Musik und Mathematik immer im Bewusstsein bleiben. Im Unterschied zu Sprach- und Bildkunst ist das Grundmaterial der Musik, Tonhöhen und Rhythmen, von vornherein mathematisch beschreibbar, sei es traditionell durch Proportionen, oder modern durch Frequenzen und Zeitdauern. Es ist kein Zufall, dass die Musik im System der *septem artes liberales* [sieben freien Künste] zusammen mit Arithmetik, Geometrie und Astronomie zu einer Gruppe, dem *quadrivium*, gehörte.

2. Musik als Regelsystem

So erscheint Musik als geradezu prädestiniert dazu, von Maschinen komponiert, also berechnet zu werden. Das gilt auch dann, wenn die Ebene des Materials im Sinne eines bloßen Vorrats von Tonhöhen und Rhythmen überschritten wird und dessen künstlerische Formung in den Blick kommt. Komponieren konnte zu allen Zeiten insofern gelehrt und gelernt werden, als es auf einem — historisch je spezifischen — Regelsystem beruht, das klar formuliert war und entsprechend angewandt werden konnte. Das Regelsystem konnte von jedem gelernt werden, der über die entsprechenden Grundlagen in der musikalischen Elementarlehre verfügte. Insofern konnte auch jeder komponieren lernen; die Unterschiede in Individualität und Qualität entstehen auf der Basis der Beherrschung dieses Regelsystems.

Einige Situationen sollen im folgenden angedeutet werden, um ein Bild davon zu geben, wie in traditioneller Musik komponiert wurde und wie ein Computerprogramm eine solche Musik konstruieren kann.

2.1 Klassischer Kontrapunkt

Der 1725 erschienene *Gradus ad parnassum* [Stufen zum Parnass, d.i. dem Sitz der Musen] des Wiener Hofkomponisten Johann Joseph Fux diente Generationen von Komponisten als Lehrbuch des sogenannten strengen Kontrapunkts. Die historisch unzutreffenden Beschreibungen des zur Zeit der Abfassung schon 150 Jahre zurückliegenden *Palestrinastils* interessieren hier nicht. An der Fux'schen Didaktik lässt sich vielmehr nachvollziehen, wie Musik als Regelsystem gelehrt und prinzipiell von einem Programm erzeugt werden kann. Das wird im folgenden am ersten der Fux'schen Lernschritte demonstriert, dem sogenannten Satz *Note gegen Note* [Fux 45-55]. Dabei wird eine neue Tonfolge, der sogenannte Kontrapunkt, zu einer bereits bestehenden Tonfolge, dem *cantus firmus,* komponiert.

Die wesentlichen Regeln sind:

1. Die Tonart und der maximale Stimmumfang, die benutzbaren Töne definierend.
2. Eine Klassifikation der Intervalle, also der Tonabstände zwischen der vorgegebenen und der neu zu komponierenden Stimme, in
   a. Perfekte Konsonanzen (Prime, Quinte, Oktave), wobei die Prime nur beim ersten oder letzten Ton verwendet werden kann;
   b. Imperfekte Konsonanzen (Terz, Sexte);
   c. Dissonanzen (Sekunde, Quarte).
3. Die möglichen Bewegungsformen:
   a. Beide Stimmen in gleiche Richtung (Parallelbewegung).
   b. Beide Stimmen in unterschiedliche Richtung (Gegenbewegung).
   c. Eine Stimme bleibt liegen (Seitenbewegung).

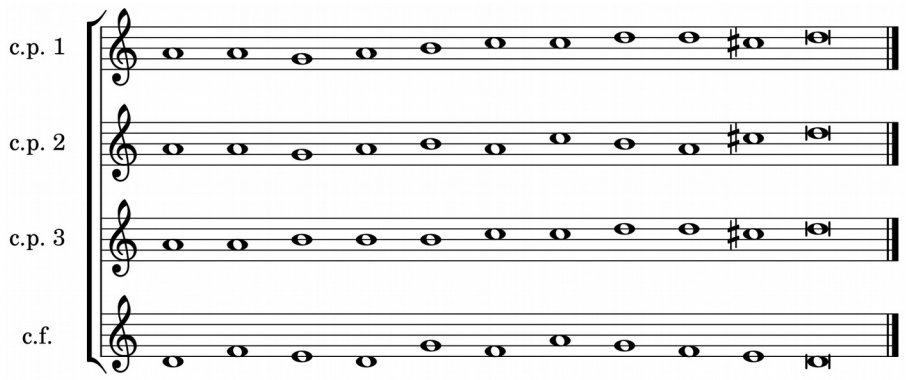
   Da Dissonanzen im Satz *Note gegen Note* ausgeschlossen sind, können die Regeln reduziert werden auf den Ausschluss von Parallelbewegung in eine perfekte Konsonanz hinein.

4. Regeln über den Abstand von einem Ton zum nächsten innerhalb des Kontrapunkts. Bevorzugt werden kleine Bewegungen.
5. Regeln über den Schluss:
    a. Das letzte Intervall (*ultima*) muss eine Oktave oder Prime sein.
    b. Das vorletzte Intervall (*paenultima*) muss imperfekt sein und erreicht die *ultima* schrittweise im Halbton, wenn der *cantus firmus* hier einen Ganzton hat, und umgekehrt.

Der folgende Code in der Programmiersprache Python zeigt eine Möglichkeit, diese Regeln von einem Computer ausführen und so einen Kontrapunkt "komponieren" zu lassen. Dabei wird selbst auf der elementaren Ebene, um die es hier geht, vieles vereinfacht, um den Code möglichst knapp zu halten. Es geht also bei diesem kleinen Programm nur um den *proof of concept*, nicht um eine konkurrenzfähige KI.

```
cantus_firmus = [1, 3, 2, 1, 4, 3, 5, 4, 3, 2, 1]
perf = [5, 8] #perfekte konsonanzen
imp = [3, 6] #imperfekte konsonanzen
dissonanz = [2, 4, 7]
cp_ambitus = [4, 5, 6, 7, 8]
from random import random, choices
def intervall(ton1, ton2):
    return abs(ton1-ton2)+1
def richtung(ton1,ton2):
    if ton2 > ton1: return 1
    elif ton1 > ton2: return -1
    else: return 0
def auswahl(fall_1,fall_2):
    if random() < 0.5: return fall_1
    else: return fall_2
def naechster_ton(cf_jetzt, cf_vorher, cp_vorher, dist=0):
    neuer_ton = cp_vorher+dist
    intv = intervall(cf_jetzt,neuer_ton)
    if intv not in dissonanz and intv is not 1 and neuer_ton in cp_ambitus:
        return neuer_ton
    else:
        dist_neu = choices([0,1,-1,2,-2],[3,2,2,1,1])[0]
        return naechster_ton(cf_jetzt, cf_vorher, cp_vorher, dist_neu)
def ultima(cf_ambitus):
    if 1 in cf_ambitus and 8 in cf_ambitus: auswahl(1,8)
    elif 8 in cf_ambitus: return 8
    else: return 1
def paenultima(ultima,cf_dir):
    return ultima+cf_dir
def cp(cf,startton):
    l = len(cf)
    cp = [0] * l
    cp[l-1] = ultima(cp_ambitus)
    cp[l-2] = paenultima(cp[l-1],richtung(cf[l-2],cf[l-1]))
    cp[0] = startton
    for i in range(1,l-2):
        cp[i] = naechster_ton(cf[i],cf[i-1],cp[i-1])
    return cp
for x in range(5):
    print(cp(cantus_firmus,5))
```

Die Ausführung dieses Programms ergibt unter anderem diese drei Varianten:



2.2 Andere historische Paradigmen

Um wenigstens anzudeuten, wie verschieden einerseits musikalische Regelsysteme sind, wie sie sich aber andererseits prinzipiell nachprogrammieren lassen, soll ein Blick auf zwei andere Stationen der westlichen Musikgeschichte geworfen werden. Die erste gehört historisch dem frühen Mittelalter an und verbindet sich — bei allen Unterschieden im Detail — mit vielen "modalen" Musiksystemen in Asien. Die zweite gehört ins 19. Jahrhundert und ist beispielsweise über einige Weihnachtslieder weitläufig bekannt. Obgleich beide in der europäischen Musikgeschichte verortet sind, die riesige Vielfalt aller anderen Musikkulturen also ausgeblendet ist, kann schon an den drei besprochenen Beispielen verdeutlicht werden, welch gänzlich unterschiedliche Systeme sich hier ausgeprägt haben, die völlig verschiedene Beschreibungen fordern und je andere Regeln mit sich bringen.

2.2.1 Melodie eines gregorianischen Gesangs

Eine musikalische Beschreibung des gregorianischen Gesangs — dessen Textbezug ignorierend — hätte von den verschiedenen *Modi* auszugehen. Jeder Modus bringt nicht nur einen Vorrat an Stufen und einen verbindlichen Schlusston (*finalis*) mit sich, sondern vor allem auch verschiedene Qualitäten der Haupt- und Nebentöne. Insbesondere gehört zu jedem Modus ein *tenor*, auf dem besonders verweilt wird. In manchen Gruppen von Gesängen gibt es zusätzliche "sekundäre Rezitationsstufen", die je nach Modus "eine bedeutende strukturelle Rolle" ausüben. [Agustoni, 45]

Die melodischen Bewegungen folgen einem Wechsel von *arsis* [Hebung, Impuls] und *thesis* [Senkung, Entspannung]. "Kein Ton steht für sich allein. Er wird [...] von der auf einen Sammelpunkt hinzielenden Spannung oder der von einem Sammelpunkt ausgehenden Lösung [erfasst]." [Johner, 57] Die melodischen Formen beruhen auf einer unendlichen Vielfalt kleiner Bewegungen im großen Kräftefeld eines Modus. Dieses Feld kann sich unter Umständen verschieben, indem ein als *tenor* angesteuerter Ton durch modale Modulation für eine Zeit wie ein Grundton behandelt wird, oder es kann von vornherein eine modale Vielfältigkeit geben: "Modale Centonisation ist in diesem Sinn die Zusammensetzung zweier oder mehrerer modal unterschiedlich geprägter Melodieteile zu einer neuen, übergreifenden modalen Einheit." [Agustoni 55]

Diese wenigen Anmerkungen mögen genügen um anzudeuten, ein wie komplexes und verglichen mit der "normalen" Dur-Moll-Harmonik andersartiges System im gregorianischen Gesang vorliegt. Solche anderen Welten finden sich auch im arabischen Maqam, im persischen Dastgah oder im indischen Raga. Die Feinheiten dieser Systeme zu erlernen braucht bei Menschen Jahre intensivsten Lernens. Sie in Regeln zu fassen und von einem Programm reproduzieren zu lassen, erscheint prinzipiell möglich, aber sehr aufwändig. Und dabei wäre der Wort-Ton-Bezug, der unter anderem für die Gregorianik zentral ist, noch gar nicht berührt.

2.2.2 Melodie und Akkordbegleitung

Eine musikalische Beschreibung einer Melodie mit Akkordbegleitung des 19. Jahrhunderts — als Beispiel kann "Stille Nacht" dienen — als Grundlage einer regelbasierten Programmierung könnte bei den Zusammenklängen beginnen. Diese "Akkorde" umfassen drei Töne, teils noch mit einem Zusatzton. Das Grundmaterial einer Komposition besteht aus lediglich drei Akkorden, die in bestimmten funktionalen Verhältnissen zueinander stehen. Der Tonika genannte Akkord stellt das Gravitationszentrum dar; der als Dominante bezeichnete Akkord fungiert als Spannung zum Zentrum hin; der als Subdominante bezeichnete Akkord schließlich stellt eine spannungsarme Entferung vom Zentrum dar.

Durch diese Funktionalität ergeben sich bestimmte Möglichkeiten und Unmöglichkeiten, Normalitäten und Überraschungen bei der Aufeinanderfolge der Akkorde. Eingebunden sind sie in eine oft symmetrische Struktur von Taktgruppen, der auch die Melodie in ihren "Perioden" folgt: Vier Takte bestehen aus zwei Zweitaktgruppen; Zwei Takte bestehen aus zwei Eintaktgruppen. Melodie und Harmonie sind kongruent; die Melodie besteht also aus Harmonietönen, angereichert durch bestimmte Formeln wie Wechselnote oder Durchgang.

3. Regelbasierte Programmierung versus Machine Learning bei Stilkopien

Bei der oben am Beispiel des klassischen Kontrapunkts beschriebenen Komposition durch ein Computerprogamm wird versucht, die Tätigkeit eines komponierenden Menschen nachzustellen bzw. zu modellieren. Wird nach dem nächsten Ton eines Kontrapunkts gesucht, werden mehrere Möglichkeiten durchgespielt. Einige scheiden aus, weil sie gegen Regeln verstoßen. Andere Möglichkeiten bleiben; aus ihnen wird eine ausgewählt, sei es willkürlich oder nach übergeordneten Kriterien. Bei einem komponierenden Menschen kann es beispielsweise der "melodische Wille", eine "Sanglichkeit", ein Gefühl für die Ausgeglichenheit der Linienführung oder für den Wechsel von perfekten und imperfekten Konsonanzen sein, was zur — meist intuitiven — Entscheidung führt. Sind schon die oben (2.1.) aufgeführten, für einen Kompositionsschüler recht einfach zu lernenden Regeln nicht ganz trivial nachzuprogrammieren, weil sie einen mehrdimensionalen Entscheidungsraum bilden, gerät ein Programm, das die eben genannten "höheren" Kriterien nachbildet, leicht sehr komplex. So haben sich die Versuche, Musik eines Stils algorithmisch nachzubilden, bald anderen Verfahren zugewandt, die teils speziell für Musik entwickelt, teils aus anderen Bereichen übernommen wurden. Im folgenden werden zwei Beispiele diskutiert, die sich beide mit der Musik Johann Sebastian Bachs befassen.

3.1. Copes *Experiments in Musical Intelligence*

David Cope beschreibt schon in den 1980er und 1990er Jahren ein Verfahren, das Stilkopien aus der Analyse einer Anzahl vorhandener Werke herstellt. Die in einer geeigneten Datenstruktur gespeicherten Partituren werden vom Programm zunächst in kleine Teile zerlegt. Auf der Basis von Signaturen, die beim Vergleich der Teile und ihrer Verbindungen entstehen, werden dann durch Rekombination neue Musikstücke erzeugt. In Copes Worten [Cope 2020]:

(1) deconstuction (analyze and separate into parts)
(2) signatures (commonality - retain that which signifies style)
(3) compatibility (recombinancy - recombine into new works)

Als Beispiel für eine frühe Form seines Programms *EMI* [Experiments in Musical Intelligence] sei hier die in der Programmiersprache LISP geschriebene Funktion zitiert, mit deren Hilfe aus zwei vorhandenen Bach-Inventionen eine neue erzeugt werden kann. Auch ohne tieferes Eindringen in den Code sind die Schritte der Datenzubereitung (`prepare-data`), der Analyse (`find-patterns`) und der Neuformulierung (`make-melody, add-second-voice, cadence-it`) gut erkennbar [Cope 1991, 99]:

```
(defun compose-invention (composition-1 composition-2)
  (setq meter (select-meter))
  (setq prepared-data
        (prepare-data composition-1 composition-2))
  (setq signature-dictionary (find-patterns))
  (setq melody (make-melody))
  (setq initial-new-composition (add-second-voice))
  (setq invention (cadence-it)))
```

3.2 Maschinelles Lernen mit *DeepBach*

Machine learning ist einer der Schlüsselbegriffe der gegenwärtigen KI-Entwicklung. Bei denen, die die Übernahme der Macht durch die Computer vor Augen haben, erscheint es als großer Schritt auf diese Zukunft hin. Denn es sind scheinbar nicht mehr Menschen, die in mühevoller Kleinarbeit die Programme mehr oder minder intelligent machen, sondern diese Programme lernen nun selbst. Und wenn sie auf diese Weise, womöglich unter Einbeziehung eines GAN-Sparringspartners, Go lernen und besser als die besten Menschen sind, warum sollen sie dann nicht Bach Choräle lernen und Bach schlagen, möchte man meinen.

Nun ist es gewiss nicht der Anspruch der Autoren von [*DeepBach*], besser als Bach zu sein. Sie wollen nur besser als zwei andere Verfahren bei der Rekomposition von Bachchorälen sein (Cope wird in diesem Zusammenhang nicht genannt) und evaluieren das durch einen Test mit 1272 Versuchspersonen [DeepBach 3.2]. Getestet wird, ob ein durch *DeepBach* generierter Choral für Bach gehalten wird. Auch wenn die Fragestellung im Rahmen der Untersuchung plausibel ist, ist es doch einen Gedanken wert, dass erfolgreiches maschinelles Komponieren hier mit erfolgreichem Betrug in eins gesetzt wird. Das Programm ist perfekt, wenn *fake* und *real* von niemandem mehr unterschieden werden können.

Da die Methoden des maschinellen Lernens vor allem in der visuellen Musterkennung entwickelt wurden, entbehrt *DeepBach* — anders als Cope — musiktheoretischer Begrifflichkeit. Benutzt werden stattdessen das Pseudo-Gibbs Sampling, das die Wahrscheinlichkeit eines Zustands, also hier beispielsweise eines Soprantons [DeepBach, fig. 4], beurteilt. Die Erläuterung des zentralen Algorithmus durch die Autoren sei hier zitiert, um einen Eindruck von der benutzten Begrifflichkeit zu geben [DeepBach, Algorithm 1]:

Algorithm 1 Pseudo-Gibbs sampling

Input: Chorale length L, metadata $\mathcal{M}$ containing lists of length L, probability distributions (p_1, p_2, p_3, p_4), maximum number of iterations M
Create four lists $\mathcal{V} = (\mathcal{V}_1, \mathcal{V}_2, \mathcal{V}_3, \mathcal{V}_4)$ of length L
{The lists are initialized with random notes drawn from the ranges of the corresponding voices (sampled uniformly or from the marginal distributions of the notes)}
for m from 1 to M **do**
    Choose voice i uniformly between 1 and 4
    Choose time t uniformly between 1 and L
    Re-sample $\mathcal{V}_i^t$ from $p_i(\mathcal{V}_i^t | \mathcal{V}_{\setminus i,t}, \mathcal{M}, \theta_i)$
end for
Output: $\mathcal{V} = (\mathcal{V}_1, \mathcal{V}_2, \mathcal{V}_3, \mathcal{V}_4)$

Es ist natürlich erlaubt, dass Informatiker ein aktuelles informatisches Modell an abstrahierten, also vor allem vom Wort-Ton-Bezug entkleideten, Bach-Chorälen testen. Doch ist die Frage, was unter "highly convincing chorales in the style of Bach" [DeepBach, Abstract] zu verstehen ist. Schaut man auf die drei kurzen Beispiele, die die Autoren in ihrem Artikel anbieten, findet man neben Fehlern wie Quintparallelen und einer ziellosen, schweifenden Harmonik eine große Anzahl von mindestens ungewöhnlich zu nennenden Wendungen. Das darin zum Ausdruck kommende Problem scheint einen zentralen Punkt zu berühren. Für Bach und seine komponierenden Zeitgenossen war der Raum des Komponierens nicht leer und neutral, sondern ein aus verschiedenen Qualitäten und Widerständen bestehendes Kräftefeld. In der Mitte des Raumes befindet sich die Normalität; zu den Grenzen hin wird der Widerstand gleichsam immer größer, was auch bedeutet: es wird Ungewöhnlichkeit signalisiert. Es gibt Wendungen, die man "eigentlich nicht macht", die aber an einer Stelle "gerade richtig" sind, weil hier eben etwas ungewöhnliches passiert — bei den Chorälen vor allem durch den Text motiviert.

Diese verschiedenen Qualitäten des musikalisch-kompositorischen Raumes zwischen Normalität und Unmöglichkeit, mit unendlich vielen Möglichkeiten der Spannung durch Abweichung, werden von *DeepBach* zu einer unterschiedslosen Homogenität reduziert. Der bei Bach tatsächlich tiefe Raum wird flach. Legt man diesen Maßstab an, dürfte es wohl zweifelhaft sein, ob "agnostic approaches (requiring no knowledge about harmony, Bach's style or music) using neural networks" wirklich "promising results" hervorbringen können [DeepBach, I], und ob die gegenwärtigen Beispiele, in Vorgehensweise wie im Resultat, irgendetwas mit Komposition zu tun haben. Die Zeitschrift *MIT Technology Review* ist sich da in der Besprechung von *DeepBach* allerdings sicher [TR 2016]:

That's interesting work that has fascinating implications. If it is possible for a deep-learning machine to produce chorales in the style of Bach, then why not also in the style of other composers and perhaps even other styles of music? [...] deep-learning machines from Sony's labs and elsewhere have begun to produce well-regarded pieces of music. It will come as no surprise if these machines soon begin to take on more ambitious works such as symphonies, operas, and more. Bach would surely have been amazed!

4. Algorithmisches Denken in Neuer Musik

Dass Johann Sebastian Bach über *DeepBach* Verblüffung und Begeisterung geäußert hätte, mögen die einen ebenso imaginieren, wie es für die anderen die späte Erfüllung von Bachs Verfluchung einer Musik ohne Seele als "teuflisches Geplerr und Geleyer" sein mag [Bach 1738]. Wissen kann es niemand, denn ein Computerprogramm wird hier auf eine Musik angewandt, die ganz anders entstand — in einer Zeit ohne Fotos, ohne Film, ohne Tonaufzeichnungen, von Internet und Computern zu schweigen.

Das unterscheidet Ansätze wie Copes *EMI* oder *DeepBach* grundlegend von der Präsenz, die algorithmisches Denken in der Komposition Neuer Musik seit den 1950er Jahren hat. Bei dieser handelt es sich nicht um eine Ersetzung des Menschen durch einen Automatismus, sondern um unterschiedliche Arten, wie algorithmisches Denken als Teil der kompositorischen Phanatasie, Praxis und Technik auftritt. Das soll im folgenden an drei Beispielen gezeigt werden, die von ganz unterschiedlichen Ansätzen ausgehen und zu ganz unterschiedlichen Resultaten führen, dennoch aber alle in einem ganz anderen Sinne algorithmisch sind, als dies in früherer Musik der Fall war. Die Maschine komponiert gleichsam mit; oder andersherum: die Phantasie des Komponisten benutzt maschinenähnliche, algorithmische Prozesse für bestimmte Stadien der Ausarbeitung. Dies ist möglich, weil das musikalische Material nicht mehr, wie in traditioneller Musik, als eine gewisse Ordnung der Töne, möglicher Rhythmen und benutzter Formen vorgegeben ist. Vielmehr muss es, als Teil der Komposition selbst, gefunden und entwickelt werden. Das gilt ganz besonders von elektronischer Musik, da es hier nicht einmal mehr die Grenzen des Instruments gibt.

4.1 Karlheinz Stockhausen: Studie II (1953)

An Stockhausens *Studie II* kann eine Methode nachvollzogen werden, die das musikalische Material durch deterministische Algorithmen erzeugt. Einige dieser Algorithmen sollen hier beispielhaft besprochen und nachprogrammiert werden. Für das Verständnis des Kompositionsprozesses in dieser Art Neuer Musik ist es entscheidend, die Motivation und das Resultat eines jeden Algorithmus zu verstehen. Die Situationen, aus denen heraus entwickelt wird, sind ebenso divers wie die erfundenen Methoden.

4.1.1 Generierung von Tonhöhen

Im Unterschied zu Cage's *Williams Mix* arbeitet die *Studie II* mit festen Tonhöhen. Diese lassen sich als Frequenz eines Signalgenerators prinzipiell beliebig einstellen. In vielen Aspekten lässt sich die *Studie II* als Modifikation eines Klavierstücks verstehen. Dieses Modell dient als Ausgangspunkt, wird aber durch das elektronische Medium spezifisch verändert. Bezogen auf die Tonhöhen heisst das: ähnlich wie beim Klavier besteht ein und dieselbe Proportion von einer Stufe zur nächstfolgenden, doch weist sie eine andere Größe auf als beim Klavier. Bei der "gleichschwebend temperierten" Stimmung des modernen Klaviers wird die Proportion 2:1 in 12 gleiche Abstände geteilt. Das ergibt den Faktor $\sqrt[12]{2}$ (= 1.059643...). Stockhausen entscheidet sich demgegenüber für die Proportion $\sqrt[25]{5}$, was mit dem Faktor 1.066494... einen geringfügig größeren Wert ergibt. Es ist also, verglichen mit einem Klavierstück, beides erreicht: Man kann in einer sehr ähnlichen Weise wie in der Tradition (der sogenannten Chromatik von zwölf Tönen pro Oktave) Zusammenklänge bilden, und dennoch klingen diese Zusammenklänge anders, neu, ungewohnt.

Damit der Algorithmus starten kann, bleibt noch zu entscheiden, was der tiefste und der höchste mögliche Ton sei. Die Entscheidung für den tiefsten Ton ist relativ willkürlich; Stockhausen

entscheidet sich für 100 Hz, unter anderem wohl deshalb, weil tiefere Frequenzen bei Sinustönen an Hörbarkeit und vor allem an Durchsichtigkeit verlieren. Der höchste Ton befindet sich vernüftiger Weise an der oberen Grenze der Hörschwelle.

Der Komponist von 1953 saß an dieser Stelle mit einer Logarithmentafel versehen am Schreibtisch und berechnete die Reihe der möglichen Frequenzen: 100 Hz wurde mit 1.066... multipliziert und ergab gerundet 107 Hz; das wiederum wurde mit 1.066... multiplziert und ergab 114 Hz, und so fort. Mit einer Programmiersprache wie Python lässt sich das wie folgt erzeugen:

```
Alle_Frequenzen = [round(100*5**(n/25)) for n in range(81)]
-> [100, 107, 114, 121, 129, 138, ..., 16171, 17247]
```

4.1.2 Andere Parameter

Erscheint die Generierung möglicher Tonhöhen aus (wenn auch neuen) Verhältniszahlen noch recht konform zur Tradition, wird bei anderen Parametern der Komposition deutlich, dass es sich hier nur um einen von vielen Schritten eines neuen Denkens handelt. Das sei hier wenigstens angedeutet [ausführlich vgl. Silberhorn].

Als Grundmaterial für alle Fragen der Auswahl von Tonhöhen, Rhythmen und Lautstärken werden zunächst sogenannte Zahlenquadrate generiert. Dabei handelt es sich um fünf Reihen, deren jede eine Permutation der Ausgangsreihe 3-5-1-4-2 darstellt:

```
3  5  1  4  2
5  2  3  1  4
1  3  4  2  5
4  1  2  5  3
2  4  5  3  1
```

Von diesen Zahlenquadraten werden insgesamt zehn durch Anwendung verschiedener Permutationsmethoden erzeugt. Beispielsweise kann das obige Zahlenquadrat durch folgenden Code generiert werden:

```
def addMod(zahl1,zahl2,mod):
    return (zahl1+zahl2-1) % mod+1
reihe = [3, 5, 1, 4, 2]
quadrat = [reihe]
for zeile in range(1, len(reihe)):
    intervall = reihe[zeile] - reihe[0]
    reiheNeu = [addMod(wertNeu,intervall,len(reihe)) for wertNeu in reihe]
    quadrat.append(reiheNeu)
```

Aus diesen Reihenabläufen werden nun in mehr oder weniger komplizierten Verfahrensweisen die Klänge, Dauern, Einsatzabstände und Lautstärken bestimmt. So setzt sich, Schicht für Schicht, das ganze Stück mit seinen 380 Klängen zusammen. Unterschiedliche Texturen innerhalb von *Studie II* werden dabei dadurch gewonnen, dass die durch den Reihenablauf erzeugten Ereignisse je anders gelesen werden. So kommt es beispielsweise durch verschiedene Lesart der Einsatzabstände dazu, dass Teil 3 aus kurzen Klängen besteht, während sich die Klänge in Teil 4 zu großen Schichtungen auftürmen.

4.1.3 Phantasie und Algorithmus

Kompositorische Phantasie in *Studie II* findet sich vor, in und zwischen den Algorithmen. Bevor ein Algorithmus entwickelt wird, gibt es ein kompositorisches Bedürfnis. Die Algorithmen selbst sind Erfindungen, eigentümliche Formen, im Fall der *Studie II* geprägt von der Obsession, alles und jedes aus der Grundgestalt der Reihe 3-5-1-4-2 hervorgehen zu lassen. Und auch wenn die Algorithmen fertig sind und laufen, werden sie verschieden gelesen. Und wenn der Komponist merkt, dass im Verlauf der Komposition die bisherigen Reihenabläufe in den folgenden Teilen zu unbefriedigenden Resultaten führen, ändert er sie, beispielsweise in der Geschwindigkeit, in der sie ablaufen. So ändert Stockhausen beim Beginn des dritten Teils eine Methode, die zur Auswahl bestimmter Register führte, derart, dass nun alle Register benutzt werden. [Heintz 2014, 70]

Nicht Ersetzung der Phantasie durch Algorithmen also, sondern Spiel mit Algorithmen als Ausdruck der Phantasie. Und das Ohr des Komponisten, seine künstlerische Entscheidung, als begleitende, korrigierende, entscheidende Instanz.

4.2. John Cage: Williams Mix (1952)

Eine ganz andere Phantasie, ein ganz anderes Material ist der Ausgangspunkt von Cages *Williams Mix*, ebenfalls für Tonband komponiert. Ein gleichsam repräsentativer Ausschnitt aus der Totalität aller Klänge wird wie in einem Kaleidoskop neu zusammengesetzt, in acht selbständigen Tonbandspuren. Sie entsprechen den acht Lautsprechern, die beim Konzert um das Publikum herum postiert sind. "Diese 8 Lautsprecher zu hören ist eine außergewöhnliche Erfahrung. Nichts hat Platz, außer dem unmittelbaren Zuhören. Die Luft war so belebt, dass man ein Teil davon wurde. Jedenfalls ist das unsere Reaktion und die einiger weniger Leute im Publikum gewesen", schrieb Cage nach der ersten Aufführung an Pierre Boulez [Cage-Boulez]. Die Folge und Kombination der Einzelklänge wirkt zufällig, unvorhersehbar; sie erinnern an die Splitter, Sprünge und Überlagerungen der städtischen und medialen Wirklichkeit. Um diesen Eindruck zu erreichen, hat Cage sehr genau und sehr komplex konstruiert.

Auch hier laufen Algorithmen ab, aber im Unterschied zur Reihentechnik, sie sie Stockhausen in *Studie II* benutzt, haben diese Algorithmen einen prägnanten Zufallsanteil. Geht es beispielsweise um die Auswahl eines Klanges, wird — vereinfacht gesagt — bei der Herstellung der Partitur an einer bestimmten Stelle durch Zufallswürfe entschieden, aus welcher der folgenden Kategorien der entsprechende Klang stammen soll [Cage]:

A) Stadtklänge
B) Landklänge
C) Elektronische Klänge
D) Manuell produzierte Klänge einschließlich solcher der Musikliteratur
E) Mit Atem erzeugte Klänge einschließlich der Lieder
F) Leise Klänge, die elektrisch verstärkt werden müssen

Ebenfalls durch Zufallsauswahl wird entschieden, welche Dauer dieser Klang hat und ob bei ihm Tonhöhe, Klangfarbe und Lautstärke verändert werden oder nicht. Es gibt also, basierend auf den Grundentscheidungen, wie beispielsweise den sechs Klangkategorien, eine unermessliche Anzahl verschiedener Versionen für eine Partitur. Jede dieser Möglichkeiten wird bei allen Unterschieden im Einzelnen doch, bezogen auf die Bewegungsform, die Überlagerungen, die Fetzenhaftigkeit, eine ähnliche Charakteristik aufweisen wie die eine Version, die Cage realisiert hat.

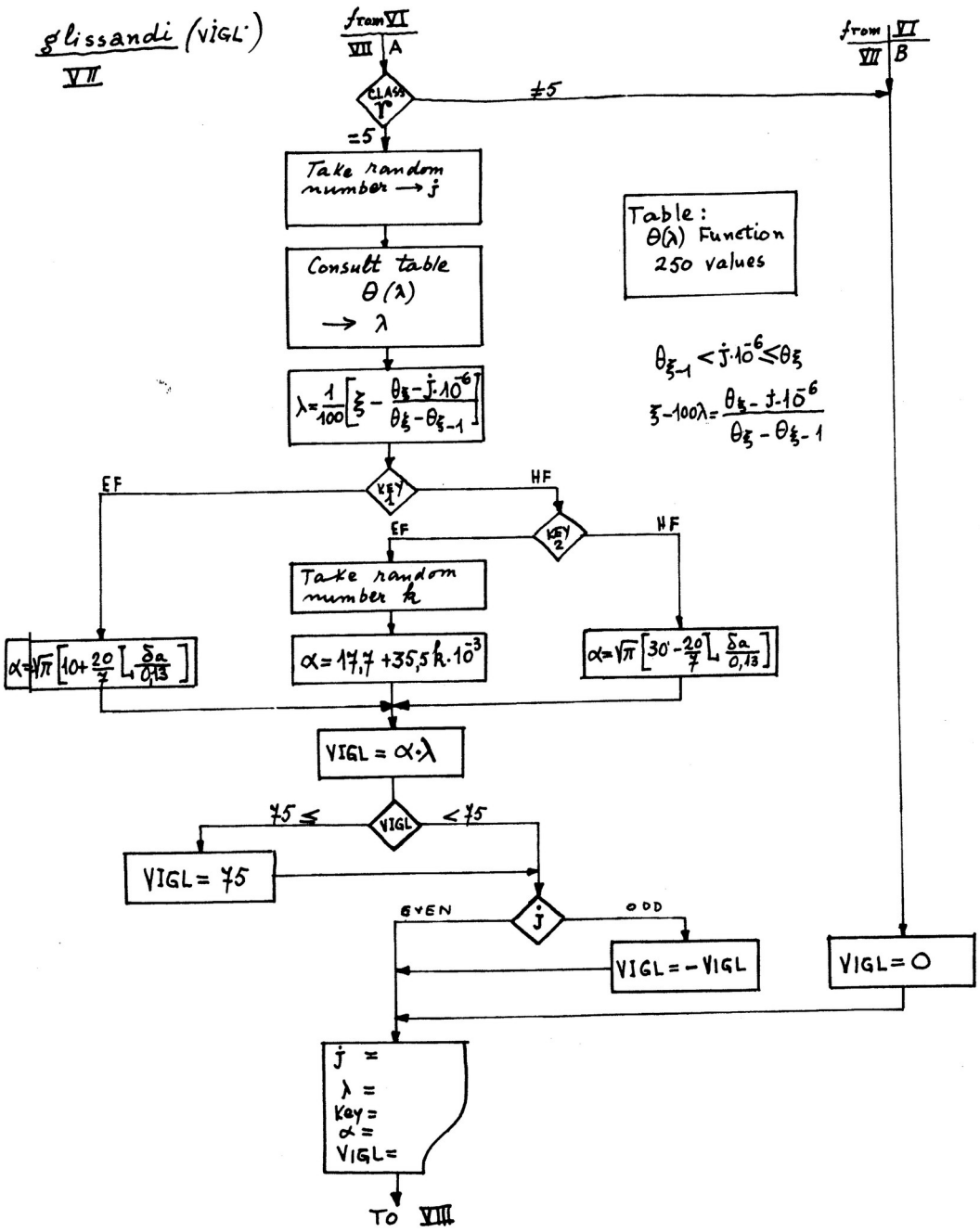
Auch hier ist also der Algorithmus gleichsam ein Partner des Komponisten. Wie Stockhausen entwickelt Cage Algorithmen, um dort hinzukommen, wo er hinkommen will. Da er ein anderes Ziel hat als Stockhausen, nicht die Ableitung eines Musikstücks aus einer einzigen kleinen Zahlenfolge, sondern eine offene, immer andere Struktur, die die Splitter der klingenden Wirklichkeiten einer starken Formung unterwirft, damit sie sich auf diese Weise als sie selbst zeigen können, deshalb sind auch Cages Algorithmen ganz andere als die von Stockhausen.

Cage und Stockhausen denken und handeln in diesen Kompositionen wie Programmierer — ohne etwas vom Computer zu wissen und zu einer Zeit, als die ersten höheren Programmiersprachen sich erst langsam herausbildeten. Ihre Phantasie bedient sich (auch) der Algorithmen, und die Algorithmen formen sich je nach der künstlerischen Phantasie.

4.3 Iannis Xenakis: Achorripsis und ST (1956-1962)

Iannis Xenakis war dank seiner mathematischen Bildung sehr früh bewusst, dass wesentliche Teile seiner kompositorischen Methode auch von einem Computer ausgeführt werden konnten. Dank persönlicher Kontakte konnte er mit dem zu seiner Zeit besten Computer, einem IBM-7090 arbeiten. Es entstanden sechs *ST* genannte Kompositionen, die 1962 in einem Konzert in der Zentrale von IBM-France aufgeführt wurden.

Ähnlich wie bei Stockhausen und Cage ist die große Nähe des kompositorischen Denkens zum Schreiben eines Programms nicht von der konkreten Kenntnis und dem Gebrauch eines Computers abhängig. So arbeitete Xenakis bereits mit stochastischen Algorithmen für seine Musik, bevor er hoffen konnte, zu einem Großrechner Zugang zu bekommen. Für sein 1956 komponiertes Stück *Achorripsis* entwarf er unter anderem diesen Ablaufplan, der aus einem Programmierlehrbuch stammen könnte:



<Abbildung 8: [Xenakis, 135]>

Bei *Achorripsis* wurden alle diese Prozesse noch mit der Hand ausgeführt, aber sie sind bereits ein fertig ausformuliertes Programm. Xenakis sagt 1980 im Rückblick [Varga 79]:

Wir haben eine Struktur vor uns, eine "black box". Wir füttern sie mit Daten, führen ihr Energie zu, und was herauskommt, ist Musik. Bei der Komposition von Achorripsis *hatte ich noch keinen Rechner zur Verfügung, um mir die Übergänge in der Musik als kontinuierliche Folge berechnen zu lassen. Die Struktur des* ST *Programms ist die gleiche wie in* Achorripsis*, aber da ich hier schon mit dem Computer arbeitete, verwendete ich auch die entsprechende Methode.*

Die *black box*, die mit Daten gefüttert wird und aus der dann Musik herauskommt — das könnte im Vorgriff auf eine komponierende KI gesagt sein. Und in der Tat ist Xenakis' Denk- und Vorgehensweise hier ein sehr frühes Beispiel für eine *creative machine* [Xenakis 1992, 131], sofern

der Ausdruck richtig verstanden wird. Denn gebaut ist diese Maschine natürlich vom Komponisten, nach seinen spezifischen Bedürfnissen, und in diesem Wissen [Varga 79]:

Wie bringt man es, bildlich gesprochen, fertig, eine "black box" zu bauen, aus der letztlich Musik herauskommen soll, und zwar interessante Musik?

Im Fall von Xenakis entsteht nicht nur interessante, sondern auch sehr individuelle, impulsive und leidenschaftliche Musik. So ist Xenakis' Entwurf einer *formalized music* sowohl ein Beispiel für die Selbstverständlichkeit, mit der in Neuer Musik mathematische Methoden und Computerprogramme benutzt werden können, als auch für den entscheidenden Anteil, den Individualität, Konzeption und Gebrauch dieser Programme für das Gesamtergebnis haben.

4.4 Folgerungen und Möglichkeiten

Algorithmisches Denken in der Arbeit mit Reihen (Stockhausen), in der Ermöglichung von variativen Strukturen (Cage), in der Anwendung stochastischer Methoden (Xenakis) — dies sind drei Beispiele aus den 1950er Jahren, die zeigen, wie in der Komposition Neuer Musik mit Maschinen gespielt werden kann. Diese Linie ist fortgesetzt worden und führt bis in die Gegenwart; der Gebrauch von KI im derzeit neuesten Gewand ist eine selbstverständliche Möglichkeit unter vielen, und dies kann eine seiner vielen möglichen Motivationen sein: "Der Einsatz nicht-trivialer Maschinen folgt [...] einem Bedürfnis, das an der Widerständigkeit jeglichen künstlerischen Materials ansetzt und mit einem veränderten Verständnis künstlerischer Subjektivität insgesamt zu tun hat." [Grüny]

Die Integration maschineller Verfahrensweisen als Teil oder Gegenstand schöpferischer Prozesse beruht auf einer musikalischen Sprache, die wie im Fall von Stockhausen, Cage und Xenakis ihre eigenen Regeln immer neu sucht und (er)findet. Das Zeitgemäße dieser Musik beruht darauf, dass Regeln nicht vorgefunden und angewandt, sondern gefunden und komponiert werden.

5. Dimensionen des Feldes *KI und Musik*

Kann eine KI Musik komponieren? "Natürlich" oder "nein" — beide Antworten sind richtig. Es hängt schlicht davon ab, welche Musik diese Frage meint, und was unter Komponieren verstanden wird. Es gibt nicht eine Musik, sondern viele Musiken; sowohl bezogen auf ihr Material, also ihre Klänge und Verbindungsmögichkeiten, als auch bezogen auf den Kontext, in dem sie komponiert und rezipiert wird. Wer kritisch darüber nachdenken will, welche Implikationen die technische Entwicklung für die Entstehung und die Rezeption von Musik hat, muss diese Unterschiede beachten und beschreiben. Das sei hier ansatzweise versucht.

5.1 Welche Musik wird komponiert

Welche Musik ist in dem Feld *Musik und KI* relevant? Es scheint sinnvoll, vorerst drei Gruppen zu unterscheiden:

1. Historische Musik mit einem abgeschlossenen Corpus an Werken. Beispiele wären die bereits besprochenen Kompositionen von Inventionen oder Chorälen im Stile Bachs.
2. Musik, die nicht als Stilkopie gemeint ist, sondern als aktuelle Musik, die aber auf einem eingegrenzten, gut beschreibbaren Stil, verbunden mit entsprechenden Regeln, basiert. Beispiele wären weite Teile von Popmusik oder Weltmusik.

3. Musik, deren Regeln prinzipiell von Komposition zu Komposition neu entwickelt werden und deren Ausarbeitung ein Teil des Kompositionsprozesses selbst sind. Beispiele wären die besprochenen Werke von Stockhausen, Cage und Xenakis und jüngere Kompositionen in dieser Tradition.

5.1.1 Stilkopien

Wird Musik der ersten Kategorie von einer KI komponiert, geht es um die — mehr oder minder gelungene, also dem Original mehr oder minder ähnliche — Kopie eines Stils. Solch ein Stil ist aber nichts abstraktes, sondern er trägt verschiedene "Handschriften": die Handschrift einer Epoche, die Handschrift eines Ortes, eines Landes, einer Kultur, und schließlich die Handschrift einer bestimmten Person. Wird solch eine Handschrift durch eine heutige Maschinenschrift nachgebildet, kommt es per se zur Kollision zweier Welten, sie mag bewusst sein oder unbewusst.

Welcher Art ist das Interesse an diesen Stilkopien? Seitens der Informatik können hier, beispielsweise bei *DeepBach*, Methoden validiert werden. Die Hausaufgabe im ungeliebten Pflichfach Musiktheorie könnte vielleicht auf diese Weise erledigt werden. Ein für die Musiktheorie, die Musikpsychologie oder die Kognitionsforschung lohnendes Feld könnte sich eröffnen, wenn die KI-Replikate mit den Originalen verglichen werden und so Spezifika menschlicher Kreativität besser erkannt werden können.

Was notwendiger Weise gänzlich fehlt, ist ein tieferes Eingehen auf die Musik an Hand einer teilnehmenden, mitvollziehenden Fragestellung. Warum hat Bach gerade bei diesem Wort die Harmonik bis zum Zerreissen gespannt? Warum benutzt Mozart hier ein einfaches Modell als "Thema ohne Thema" für den Beginn eines Satzes? Dass solche Fragen, die sinnvoll und notwendig bei historischer Musik zu stellen sind, an eine stilkopierende KI gestellt werden können, wird wohl auch der größte Phantast einer starken KI nicht behaupten wollen. Insofern beruhen aber alle diese Versuche auf einem reduzierten Musikbegriff. Um es polemisch auszudrücken: Belanglose Töne einer als Stil enthistorisierten Musik werden auf Knopfdruck hervorgebracht. Musik ohne Musik.

5.1.2 Gegenwartsmusik mit festem Stil

Wird ein Song komponiert, oder ein Stück für eine Gruppe traditioneller iranischer Instrumente geschrieben, gibt es eingermaßen klare Regularien, was zu einem bestimmten Genre gehört und was nicht. Wenn eine KI nun ein solches Stück Musik komponiert, ergeben sich teils ähnliche, teils unterschiedliche Fragen wie bei der klassischen Stilkopie. Ähnlich ist die Unmöglichkeit, mit Komponist oder Komponistin über das Werk zu sprechen und die Einbettung der Komposition in eine Situation zu erfahren. Selbstauskünfte wie "hier war ich verliebt" oder "das ist ein Protest gegen die Niederschlagung der Demonstrationen" können nicht mehr gegeben werden. Die damit verbundene Idolisierung des Künstlers als stellvertretendem Menschen (wie du und ich, aber hat es geschafft) ist obsolet.

Das mag verzichtbar sein, oder gerade besonders interessant, weil es das Verhältnis zur neuen, künstlichen Intelligenz thematisiert. Vielleicht wird auch verschwiegen, dass der Song nicht vom Komponisten geschrieben wurde, sondern von einem Programm. Es stellen sich dann eher Fragen auf der Ebene des Urheberrechts und des Marktes: wer lügt hier und wer verdient wieviel Geld mit was. Diese Fragen mögen wichtig sein und — beispielsweise für diejenigen die bisher in dieser Branche mit Komponieren ihr Geld verdienten — existenziell; für das Verhältnis von KI und Musik ergeben sich hier aber keine neuen Aspekte. Natürlich kann eine KI Musik dieser Art komponieren, und in Zukunft sicher immer besser, beispielsweise was den Personalstil betrifft.

5.1.3 Neue Musik

Wird eine Musik komponiert, die tatsächlich immer "neu" ist in dem Sinne, dass jedes Stück grundlegend neue Entscheidungen über das Material und die Verfahrensweisen mit sich führt, kann ein solches Stück Musik zumindest von einer gegenwärtigen KI nicht komponiert werden. Eine KI hat keinen Begriff davon, was es heisst, eine Verfahrensweise zu entwickeln. Es gibt für eine KI keine "Idee", keinen "Einfall" und kein "Konzept".

Natürlich könnte eine KI ebenso wie bei anderen Stilen eine Stilkopie herstellen. Diese würde vielleicht täuschend ähnlich klingen, wäre aber in dem eben bezeichneten Sinn etwas ganz anderes als neue Musik. Dennoch kann die Komposition neuer Musik, wie schon an den Beispielen aus den 1950er Jahren gezeigt, in vielen ihrer Ausprägungen einer KI sehr nahe sein. KI kann für Teile der Komposition eingesetzt werden, die — in der Tradition algorithmischen Komponierens — generiert werden. KI kann natürlich auch zum Thema der Komposition selber werden, beispielsweise über das Verhältnis einer Performance zu ihrem KI-Replikat.

5.2 Für welchen Kontext wird Musik komponiert

Die Frage, welche Folgerungen es hat, wenn eine Musik von einer KI statt von einem Menschen komponiert wird, hängt stark von dem Kontext ab, für den Musik geschrieben bzw. in dem sie aufgeführt wird.

5.2.1 Hintergrund und Illustration

Ein großer Teil von Musik, die heute jeden Tag zu hören ist, stellt eine Art Geräuschkulisse dar: Hintergrundmusik beim Straßencafé oder Friseur, Musik in Warteschleifen oder in Einkaufzentren. Niemand, der diese Musik komponiert, wird erwarten, dass ihr mit Aufmerksamkeit oder gar Hingabe zugehört wird. Hat es nicht eine große Plausibilität, dass solche Musik von einer KI geschrieben wird?

Ähnliche mag es sich mit funktionaler Musik verhalten, die illustrieren und klar definierte Wirkungen erreichen will. Wenn Klänge erwartet werden, die als Zeichen fungieren, um Angst oder Glück, Sieg oder Niederlage akustisch zu verbildlichen, warum soll das nicht eine KI machen? Welcher Konsument dieser Musik hätte ein Problem damit?

5.2.2 Konzerte

Irritation kann nur entstehen, wenn Musik, die für ein "wirkliches Hinhören" gedacht ist, von einer KI komponiert wurde. Gesetzt sei der Fall, in einem Sinfoniekonzert wird ein angeblich kürzlich aufgefunder Satz von Schumann gespielt. Später erfahren die Konzerbesucher jedoch, dass eine KI diesen Satz komponiert hat. Welche Reaktionen sind hier denkbar und worauf deuten sie hin?

Manche werden sich betrogen fühlen. Nicht nur durch die falsche Ankündigung, sondern vielleicht auch durch die Musik, die "Gefühle vorgetäuscht" hat. Vielleicht war man an einer Stelle ganz gerührt und schämt sich jetzt, dass eine Maschine das hervorrufen konnte. Vielleicht war man sich nach der vermeintlichen Uraufführung ganz sicher, dass dies ein "echter Schumann" war, und muss sich jetzt mit dem eigenen Fehlurteil abfinden. In beiden Fällen hätten Täuschung und Enttäuschung

verschiedene mögliche Effekte. Man könnte sich fragen, auf welcher Basis die eigene Kennerschaft beruht. Man könnte sich fragen, ob die eigenen Gefühle beim Hören von Musik wirklich die Gewissheit brauchen, dass der Komponist beim Schreiben der Musik dieselben Gefühle hatte.

5.3 Um welche Kultur geht es

Konzertbesucher, denen bei diesem Gedankenexperiment "ihr Schumann weggenommen" wurde, erleben womöglich zwei recht verschiedene Anteile in ihrer Enttäuschung. Zunächst die Zerstörung von Besitz und Heldentum. Der große Komponist, der doch mir in gewisser Weise gehört; mir, dem Konzertbesucher — er war meine Phantasie. Der Held, der mich über meinen Alltag erhebt, und den ich bewundern kann — er erwies sich als Maschine. Der andere Anteil, der sich möglicherweise schon in diesem ersten Anteil verbirgt und auf seine Weise äußert, ist der Wunsch nach einer besonderen Art von Kommunikation, die in einem Konzert stattfinden kann. Dort ist ein Mensch, der Musik hervorgebracht hat. An diesem Menschsein kann das Publikum eines Konzerts teilnehmen; es kann "Musik als existenzielle Erfahrung" teilen [Lachenmann].

Diese Begegnung eines produzierenden und mindestens eines rezipierenden Menschen wird nicht nur aufgehoben, wenn eine KI an Stelle des Produzenten tritt. Sie wird auch aufgehoben durch viele Eigenarten des klassischen Konzertbetriebs: die entleerte Feier beständiger Wiederholung, die Heroisierung und Einverleibung der großen klassischen Komponisten als bürgerliche Repräsentation, und schließlich die großenteils mediale Rezeption über Tonträger oder das Internet, die das Unpersönliche und Anonyme verstärkt.

So wendet sich die Frage nach KI und Musik zurück zu der Frage, welche Kultur, welche Kunst, welche Musik gemeint ist und erreicht werden soll. Zu wenige Menschen haben einen Begriff davon, dass und wie man eine Musik schaffen kann, die als Reaktion auf traumatische politische Ereignisse entsteht und diese im Gesang der Trauer begleitet, wie es Younghi Pagh-Paan in *Aa-ga* oder Luigi Nono im *Canto sospeso* getan haben. Dass man eine Musik schreiben kann, die "rohe" klangliche Materialitäten erkundet und komponiert, wie Helmut Lachenmann, oder eine Musik, die Klänge als Klänge thematisiert und "sein lässt", wie John Cage. Doch scheint der Wunsch nach einer in diesem Sinne menschlichen Kunst lebendig zu sein. Und so können vielleicht die Schockwellen beim Thema *KI und Musik*, die das Ende der Kultur nahen sehen und für sich genommen als alberne Übertreibung erscheinen müssen, auch als Zeichen des Erschreckens über eine fortschreitende Verdinglichung und den Verlust von Kultur als menschlichem Austausch und zwischenmenschlicher Verständigung verstanden werden. Insofern scheint das Thema *KI und Kunst* an die Wurzeln von Kunst und Kultur zurück zu gehen.

6. Rück- und Ausblick

Die Perspektive einer künstlichen Intelligenz, die Musik komponiert, reicht weit in die Geschichte der Musik zurück. Die Nähe der Musik zur Mathematik legt das nahe. Zusammen mit der Lehr- und Beschreibbarkeit der Musik als Regelsystem ermöglicht sie es, Komponiermaschinen mit relativ wenig Aufwand in der jeweiligen Technologie zu entwickeln. Dabei kann es sich um simple Entsprechungen handeln, wie bei Guido von Arezzo [Kirchmeyer], um mechanische Komponiermaschinen, wie bei Athanasius Kircher, um gesellige Komponier-Würfelspiele des 18. Jahrhunderts oder um die verschiedenen Programmierparadigmen des Computerzeitalters. Die tiefen neuronalen Netze als derzeit aktuellste Technik sind dabei nur eine Möglichkeit unter vielen. Neben einigen Vorteilen sind auch Nachteile mit ihnen verbunden, da ihre Modelle nicht aus

musikalischen Kontexten stammen und eine hohe Anzahl von Samples benötigen, die bei der Imitation eines historischen Stils oft nicht gegeben sind.

Bei der Diskussion um eine komponierende KI ist zu unterscheiden, welche Musik komponiert wird, mit welchem Ziel und Anspruch dies geschieht, und wie der Kontext von Produktion und Rezeption beschaffen ist.

Nachkompositionen klassischer Musik scheint nicht selten ein gänzlich ahistorisches Verständnis von Musik zugrunde zu liegen, als säße man heute mit Bach an einem Tisch und spielte Choralharmonisation statt Schach. Ziel mag es sein, eine Gebrauchsmusik auf Knopfdruck zu schaffen (Melodie rein, Harmonie raus); der Name Bach ist dann aber nichts anderes als ein Label, um einen Konsumartikel zu produzieren. Mit Komponieren hat dies nichts zu tun; es sein denn, man versteht unter Komposition nichts anderes als die Herstellung einer Anzahl von Tönen. Jenseits der Evaluation von informatischen Methoden durch eine Stilkopie klassischer Musik könnten sich allenfalls Forschungsaufgaben für die Musikpsychologie und die Kreativitätsforschung ergeben.

Für die Komposition von Musik, die einem festen Zweck dient und beispielsweise zum Erzeugen von Atmosphären und Gefühlen eingesetzt wird, ist KI sehr gut geeignet. Die bei Stilkopien klassischer Musik unbefriedigenden Aspekte spielen hier keine Rolle. Stattdessen geht es um die in anderen Kontexten diskutierten Fragen im Zusammenhang von KI-Produkten, wie beispielsweise dem Urheberrecht oder der Frage der Originalität.

Für die Neue Musik steht die KI in einem ganz anderen Kontext. Sie ersetzt nicht einen Prozess, der bisher von Menschen unternommen wurde, sondern sie kann Teil dieses Prozesses sein. Dies kann an frühe Verfahrensweisen der 1950er Jahre anschließen oder ganz neue Wege beschreiten. Da es in dieser Musik nicht um einen festen Stil geht, sondern um die immer neue Herausarbeitung von Verfahrensweisen im Zusammenhang mit einer künstlerischen Idee, kann die KI hier zwanglos integriert werden.

Eine sich nicht mit paranoiden oder enthusiastischen Projektionen begnügende Diskussion um Musik und KI darf die ganz verschiedenen Begriffe von Kultur, die bei den unterschiedlichen Musiken virulent sind, nicht in einen Topf werfen. Wird Kultur als fundamentaler Ausdruck von menschlicher Existenz in einer Gesellschaft und Teil ihrer Selbstverständigung verstanden, ist das Komponieren ein Teil davon. Ein Mensch komponiert, was andere Menschen hören und auf das sie in verschiedenster Weise reagieren. Keine KI kann das ersetzen, aber die Angstphantasien über die Ersetzung des Menschen durch die Maschinen scheinen darauf hinzudeuten, dass diese Kulturpraxis alles andere als selbstverständlich ist.

Literaturverzeichnis

Agustoni: Luigi Agustoni / Johannes Berchmans Göschl, Einführung in die Interpretation des Gregorianischen Chorals, Bd. 1, Kassel:Bosse 1995

Aristoxenos: Neues Handbuch der Musikwissenschaft, Hg. C. Dahlhaus, Bd. 1, Laaber 1989, S. 251

Bach 1738: Des Königlichen Hoff-Compositeurs und Capellmeisters ingleichen Directoris Musices wie auch Cantoris der Thomas-Schule Herrn Johann Sebastian Bach zu Leipzig Vorschriften und Grundsätze zum vierstimmigen spielen des General-Bass oder Accompagnement für seine Scholaren in der Music, Cap. 2, in: Philipp Spitta, Johann Sebastian Bach, Zweiter Band, Wiebaden: Breitkopf&Härtel 1962, 916

Boethius De Musica: Boethii De institutione musica libri quinque, ed. Gottfried F. Friedlein, Leipzig 1867

Boethius 1985: Anicius Manlius Severinus Boethius, Fünf Bücher über die Musik, Übersetzung Oscar Paul, Hildesheim:Olms 1985 (zuerst Leipzig 1872)

Cage: Richard Kostelanetz, John Cage, Köln:DuMont 1973, S. 159

Cage-Boulez: Pierre Boulez und John Cage, Der Briefwechsel, Hg. J.-J. Nattiez, Hamburg 1997, S. 159

Cope 1991: David Cope, Computers and Musical Style, Madison:A-R Edition 1991

Cope 2020: David Cope, Experiments in Musical Intelligence
http://artsites.ucsc.edu/faculty/cope/experiments.htm (5.2.20)

DeepBach: G. Hadjeres, F. Pachet, F. Nielsen, DeepBach: a Steerable Model for Bach Chorales Generation, 2016, https://arxiv.org/abs/1612.01010 (8.2.20)

Dylan: Bob Dylan, Blowin' in the Wind, Copyright © Warner Bros. Inc., © renewed 1990 Special Rider Music

de la Motte: Diether de la Motte, Kontrapunkt, München:dtv und Kassel:Bärenreiter 1981, S. 40

Fux: Johann Joseph Fux, Gradus ad Parnassum (lateinisch), Wien 1725

Grisey: Jérôme Baillet, Gérard Grisey, Paris:L'Itinéraire 2000, S. 81

Grüny: Christian Grüny, Mit Maschinen spielen, in: Kultur gestaltet Zukunft, Künstliche Intelligenz in Kunst und Kultur, Publikation der Tagung 15.-16. Mai 2019, S. 96
https://www.link-niedersachsen.de/damfiles/default/was-wir-foerdern/programme/link/link_tagungspublikation_stiftung_niedersachsen_einzelseiten.pdf-592ecf4c2fc6af14f41ee633aaf85a3a.pdf (28.02.2020)

Heintz 2014: Joachim Heintz, Läuft es sich in Kinderschuhen besser?, in: MusikTexte 141, 65-72

Heintz 2019: Joachim Heintz, Spielen statt Fürchtemachen, in: MusikTexte 162, 27-30

Johner: P. Dominicus Johner, Wort und Ton im Choral, Leipzig:Breitkopf&Härtel 1953

Kirchmeyer: Helmut Kirchmeyer, Vom historischen Wesen einer rationalistischen Musik, in: die reihe, Bd. 8 (1962), Wien:Universal Edition, S. 11-25

Lachenmann: Helmut Lachenmann, Musik als existenzielle Erfahrung, Wiesbaden:Breitkopf&Härtel 1996

Mainzer: Künstliche Intelligenz - Wann übernehmen die Maschinen?, Berlin:Springer 2019

Mattheson: Johann Mattheson, Der vollkommene Capellmeister, Hamburg 1739, S. 53

Miller: Arthur I. Miller, The Artist in the Machine, The World of AI-Powered Creativity, Cambridge/MA:MIT, S. 163

Monteverdi: Claudio Monteverdi, Lamento d'Arianna, Il sesto libro de Madrigali, Venedig 1614

Notre Dame: Herzog August Bibliothek Wolfenbüttel, Mss 628 Helmst, fol 24r <http://diglib.hab.de/mss/628-helmst/start.htm?image=00059> (10.02.2020)

Silberhorn: Heinz Silberhorn, Die Reihentechnik in Stockhausens Studie II, Rohrdorfer Musikverlag 1980

TR 2016: Deep-Learning Machine Listens to Bach, Then Writes Its Own Music in the Same Style (2016) https://www.technologyreview.com/s/603137/deep-learning-machine-listens-to-bach-then-writes-its-own-music-in-the-same-style/ (8.2.20)

Turing 1936: Alan Turing, On Computable Numbers, with an Application to the Entscheidungsproblem, in: Proceedings of the London Mathematical Society, Band 42 (1937), 230-265

Xenakis: Iannis Xenakis, Formalized Music, Thought and Mathematics in Composition, Hillsdale:Pendragon 1992

Zimmermann 1970: Bernd Alois Zimmermann, Tratto (II), in: Intervall und Zeit, Hg. C. Bitter, Mainz:Schott 1974, 113