

## **Musik und Computer**

Werner Kaegi

*i.A. der Schweizerischen Musikzeitung verfasst (123.Jhg Jan/Feb 1983)*

### **Vorgeschichte**

Der digitale Computer ist eine programmierbare Rechenmaschine. Jeder endliche und diskrete Inhalt, der sich auf Zahlen abbilden lässt, kann durch den Computer dargestellt und manipuliert werden. Davon macht auch die Musik keine Ausnahme. Versuche, Musikalisches auf Zahlen abzubilden, sind sehr alt. Durch Plato und andere Autoren der Antike ist uns die Pythagoräische Tonleiter bekannt. Aus Quintschlägen aufgebaut, lässt sie sich durch Verhältnisse von kleinen ganzen Zahlen darstellen (rationale Zahlen) und bildet einen Teil der Proportionslehre, die der Wissenschaft der alten Griechen zugrunde lag und das musikalische Denken bis in die Neuzeit beherrschen sollte. Zwar kannte bereits die Antike auch Konzepte wie  $\sqrt{2}$  und  $\sin \alpha$ , die sich nicht oder nur unvollkommen durch ganze Zahlen darstellen lassen; allein ihre adäquate Repräsentation wurde erst mit dem Aufkommen der Naturwissenschaften der Neuzeit Allgemeingut, als man Ende des 16. Jahrhunderts in Sinustabellen und etwas später in Logarithmentafeln zum Gebrauch von Dezimalbrüchen überging und den physikalischen Raum als ein durch reelle Zahlen definiertes Kontinuum begriff. Jetzt war man imstande, die wohltemperierte Tonleiter explizit zu formulieren (Werckmeister 1697), obwohl sie schon seit der Ars nova auf Lauten praktiziert worden war. Derartige metrische Repräsentationen sagen nichts aus über die absolute Tonhöhe, sondern sie bringen lediglich Verhältnisse zwischen Tonhöhen zum Ausdruck (Intervalle); für absolute Aussagen bedarf es einer Standard-Referenz, welche die Musiker seit alters aus der gesprochenen Sprache bezogen und in Längenmassen ausgedrückt hatten. Im 19. Jahrhundert gelang es, die absolute Tonhöhe mit der physikalischen Zeitmessung zu verknüpfen (Frequenz (Hz)), und für Wissenschaftler und Musikinstrumentenbauer war damit die Tonhöhe auf eine physikalische Basiseinheit zurückgeführt; für die Musiker dagegen blieb sie bis in die Mitte unseres Jahrhunderts eine Erfahrungsqualität, und erst unter dem Einfluss der elektronischen Musik begannen die Komponisten in den 50er Jahren Tonhöhen in Form von Frequenzen zu manipulieren und schliesslich auch in ihren Partituren zu repräsentieren.

Was nun die Tondauer anbetrifft, hat sich ihre metrische Darstellung in der Musik zwar sehr früh vollzogen, aber doch nur sehr langsam weiterentwickelt. Von alters her kennt der Musiker aus der Verknüpfung seiner Kunst mit der Poesie die Begriffe der langen und der kurzen Silbe (*longa, brevis*) und den metrischen Formaufbau von Fuss, Zeile, Vers und Versfolge. Auch ist das Verhältnis von lang und kurz durch das Verhältnis 2:1 festgelegt (M.F. Quintilianus, Augustinus). Wo sich die Musik "Note gegen Silbe" an die Poesie hält, ist sie in ihren Zeitwerten denn auch völlig durch rationale Zahlen darstellbar. Als Standard-Zeitdauer wird die Emissionsdauer der kurzen Silbe der gesprochenen Sprache angesehen. Diese Repräsentation der temporellen Organisation von Musik blieb bis ins späte Mittelalter erhalten und fand in der Ars nova eine Formulierung, die eigentümlicherweise die musikalische Notation der westlichen Musik bis hinein ins 20. Jahrhundert beherrschte. Zwar wurde das Tempo durch Mälzels Metronom mit der physikalischen Zeitmessung verknüpft, aber Schönberg, Strawinsky oder Messiaen schrieben ihre Musik in Symbolen, die ebenso wie jene ihrer Kollegen der mensuralen Musik des Mittelalters Zeitdauerwerte durch rationale Zahlen ausdrücken (Halbe, Viertel, Achtel usw.). Von reellen Zahlen und rhythmischem Kontinuum keine Spur, obwohl doch die Agogik der Musik des 19. Jahrhunderts und vorallem Debussys hochdifferenzierte Rhythmik mit aller Macht

in diese Richtung wiesen. Es ist wiederum die elektronische Musik, die den Komponisten das rhythmische Kontinuum eröffnet durch den Gebrauch von analoger elektronischer Mess- und Aufnahmeapparatur. Jetzt erst finden Tonbandlängen und die Einheit der Sekunde Aufnahme in die Partituren. Was schliesslich die Intensität anbetrifft, so hat sich ihre zahlenmässige Darstellung erst in unserem Jahrhundert unter dem Einfluss der Nachrichtentechnik und der elektro-akustischen Klangaufzeichnung entwickelt. Der Toningenieur kannte die Darstellung der Intensität durch Dezibel (dB) zwar seit den Anfängen des Rundfunks in den 20er Jahren, dem Komponisten aber wird sie erst 30 Jahre später mit dem Aufkommen einer selbständigen Tonbandmusik geläufig. Es sind die 50er Jahre, in denen eine junge Komponistengeneration Weberns Musik entdeckt und Stockhausen den Versuch einer totalen Serialisierung unternimmt, indem er "alle Parameter" der musikalischen Repräsentation in ein Kontinuum projizieren will. Weil sich die Klangfarbe diesem Wollen am hartnäckigsten widersetzt, verzichtet er im 'Gesang der Jünglinge' auf traditionelle und bedient sich analoger elektronischer Mittel der Klangerzeugung, für die Meyer-Eppler eine solide wissenschaftliche Grundlage bereitgestellt hatte.

### **Programme für die Organisation von musikalischen Symbolen**

In denselben 50er Jahren kommt der digitale Computer in das Stadium einer gebrauchsfähigen elektronischen Rechenmaschine mit interner Programmsteuerung. In ihm vereinigen sich die sehr alten Konzepte der zweiwertigen Aussagenlogik und der Automatisierung mit den Errungenschaften der modernen Wissenschaft und Technik. Gödels geniale Idee der Abbildung beliebiger finiter und diskreter Daten auf die natürlichen Zahlen wird fruchtbar im Konzept des Computers, zwischem gespeichertem Inhalt und Speicherplatz-Adressierung zu unterscheiden, wodurch der Computer den Charakter eines sehr allgemein anwendbaren Automaten erhält, der beliebige Inhalte nach Vorschriften in einer endlichen Zahl von Schritten mit grosser Schnelligkeit transformiert. Allerdings sind die Maschinen der 50er Jahre nicht nur gewaltige Kästen, sondern ihre Anschaffungs- und Unterhaltskosten gehen auch weit hinaus über die Mittel bescheidener Institutionen oder gar von Einzelpersonen. Computer finden sich in dieser Zeit nur in den Rechenzentren von Universitäten und Industrie, wo sie für reine Rechenaufgaben, noch nicht aber für die Steuerung von physikalischen Prozessen Verwendung finden. Auch die Musiker haben von alters her versucht, ihre Kunst durch Vorschriften zu umschreiben. Immer wieder, und vorallem im automatenfreundlichen 18. Jahrhundert, sind Versuche zur Algorithmisierung des musikalischen Komponierens gewagt worden. (Im Bereich der instrumentalen Musik ist P.Barbaud, der Gründer des '*Groupe de Musique Algorithmique de Paris*', zu künstlerisch ausserordentlich überzeugenden Resultaten gekommen ('*French Gagaku*', '*Mu-Joken*'). Jetzt, wo die programmierbare elektronische Rechenmaschine zur Verfügung stand, drängte der Gedanke sich dem Komponisten förmlich auf, und so entsteht die erste Generation von sogenannten "musikalischen Komponierprogrammen" für den digitalen Computer. Ihr Ziel ist es, Regeln zu formulieren, die musikalische Symbole für Tonhöhe, Zeitdauer und Intensität entlang der Zeitachse organisieren, rigoros und ohne Rücksicht auf ästhetische Kriterien. Seit 1956 experimentiert L.Hiller an der Urbana-Universität mit der Programmierung der Regeln des Palestrina-Kontrapunktes, indem er von einstimmigen bis zu polyphonen Sätzen fortschreitet. Das Verbot der Tonrepetition leitet ihn zu rekursiven Prozessen, wo jedes Ereignis  $E_k$  nach statistischen Gesetzen von seinem Vorgänger  $E_{k-1}$  abhängt. Musik wird als Folge von Ereignissen im Sinne einer 'Markow-Kette' verstanden (Illiac-Suite). I.Xenakis wirft der damals blühenden seriellen Musik ihre Beschränkung auf eine armselige und zudem nur auf die Tonhöhe angewandte Kombinatorik vor und wendet sich ausschliesslich statistischen Prozessen zu, denen er alle Dimensionen musikalischer Repräsentation unterwirft. Die nötigen Berechnungen

werden erst mühsam von Hand, 1960 aber auf dem Computer von IBM-France durchgeführt. Während Xenakis' Programme sich auf den Einzelfall ausrichten ("ST10"), beginnt G.M.Koenig 1964 in Bonn mit dem Entwurf eines möglichst allgemeinen Komponierprogramms ('Projekt 1'). Gemeinsam ist all diesen Bestrebungen die algorithmische Manipulation von in Zahlen codierten musikalischen Notensymbolen.

### **Programme für die digitale Klangersynthese**

Aber das menschliche Ohr hört keine Symbole. Klänge werden dem Ohr durch physikalische Schwingungen vermittelt, oder genauer, durch Veränderungen des Luftdrucks in der Zeit. Um derartige Schwingungen zu erzeugen, muss der Computer mit geeigneter Randapparatur ausgerüstet sein, nämlich mit Digital-zu-Analog-Konvertern, welche die Zahlen in elektrische Gleichspannungen verwandeln, und mit Lautsprechern, die elektrische Spannungen in mechanische Schwingungen übertragen, welche via die Luft das Ohr erreichen. Im einfachsten Falle können die Lautsprecher nach Glättung des Signals mittels Filtern direkt an die Ausgänge der Konverter gehängt werden (Abb.1). Werden nun vom Computer in gleichmässigen

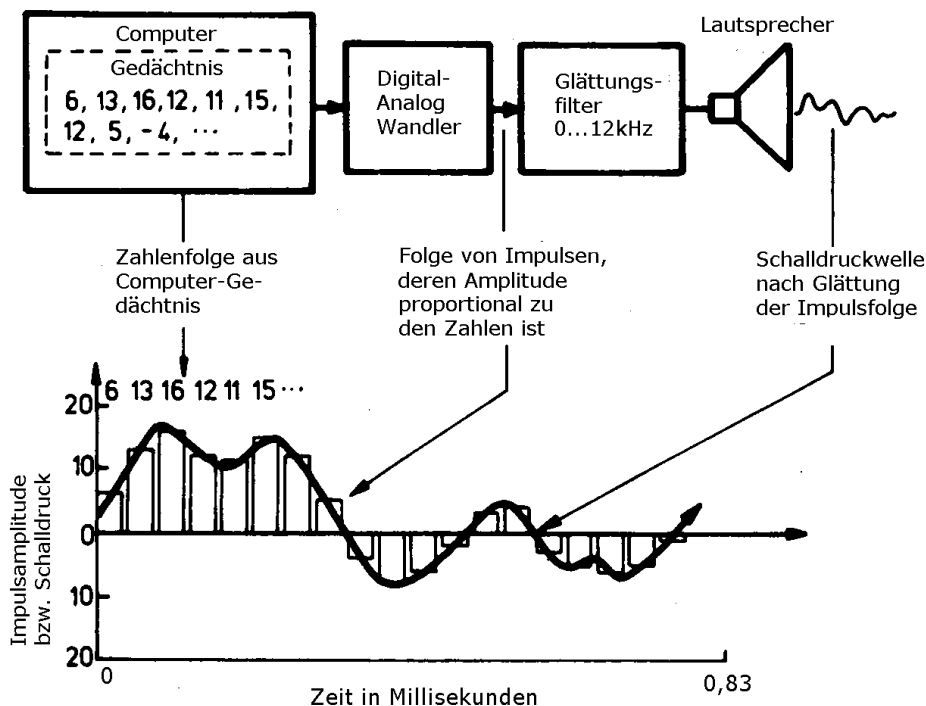


Abb.1 Klangerzeugung durch den digitalen Computer (nach M.Mathews)

sehr kurzen Zeitabständen Zahlen ausgelesen und konvertiert (*Samples*), so können diese nach Wandlung durch den Lautsprecher vom Ohr als Klänge gehört werden, falls die Frequenzen der erzeugten Schwingungen in den menschlichen Hörbereich fallen (ca.20-16000 Hz). Die Anzahl der vom Computer pro Sekunde ausgelesenen Zahlen (*Sampling Rate*) muss dabei mindestens zweimal grösser sein als die Frequenz des höchsten Tones, der erzeugt werden soll. Auf diese Weise ist es möglich, mit dem Computer Klänge zu generieren und -wenn gewünscht- durch Manipulieren von Zahlen zu verändern. Das Verfahren lässt sich auch umkehren. Willkürliche Klänge können durch das Mikrophon in elektrische Spannungen und

diese durch Analog-zu-Digital-Konverter in Zahlen verwandelt werden, die der Computer in gleichmässigen Zeitabständen einliest und speichert. Erhalten wird eine digitale Registrierung des Klanges, der jederzeit vom Computer reproduziert werden kann. Durch Manipulieren der gespeicherten Zahlen und/oder Verändern der *Sampling Rate* kann der gespeicherte Klang beeinflusst werden.

Die Ausbreitung des Rechencomputers mit Konvertern kam in den 60er Jahren und ermöglichte die Steuerung von physikalischen Prozessen sowie die Registrierung und Verarbeitung von durch Sensoren aufgefangenen Daten. Die Maschinen wurden kleiner und billiger, während Speicherkapazität und Rechengeschwindigkeit rasch zunahm. (Für die Speicherung von  $10^6$  Bytes benötigte man 1953 einen Raum von  $220 \text{ cm}^3$ , 1960 waren es noch  $141 \text{ cm}^3$ , und heute (1982) strebt der Wert gegen  $2 \text{ cm}^3$ . Das menschliche Gehirn benötigt vergleichsweise  $1 \text{ mm}^3$ !) Damit wurde auch die digitale Klangsynthese möglich, der sich die zweite Generation von "musikalischen" Computerprogrammen nun zuwandte. Metrische Beschreibungen des Klanges im Sinne einer Zeitfunktion sind seit dem 19. Jahrhundert bekannt (Ohm, Helmholtz). Durch die Anwendung der Reihenentwicklung nach Fourier, wonach sich jeder willkürliche Klang bis hin zum Geräusch aus Sinustönen gewinnen lässt, war die Transformation der Zeitfunktion in den Frequenzraum möglich geworden, eine Repräsentation, die als "Spektrum" in die Klangforschung eingegangen ist. Die Verwendung der Transformation war jedoch für den elektronischen Rechner der 60er Jahre so zeitraubend, dass sich das Verfahren der Fourier-Klangsynthese erst in den 70er Jahren allmählich realisieren liess, nachdem schnellere Maschinen entwickelt und der einfachere FFT-Algorithmus (*Fast Fourier Transform*) gefunden worden war. So basieren denn die wesentlichen Programme für die digitale Klangsynthese auf dem Erzeugen und Manipulieren von Zeitfunktionen  $a(t)$  im oben vorgestellten Sinn, wobei auf Standard-Wellenformen zurückgegriffen wird, die entweder in Tabellen gespeichert (*Wave Tables*, Mathews, (*MUSIC V*, 1968) oder durch Algorithmen festgelegt sind (*F. Brown*, *AUDITU* 1976). (Die Flut der heute den Markt überschwemmenden Tisch- und Taschen-Synthesizer verwendet ausnahmslos Tabellen und variiert die Tonhöhe durch blosses Verändern der *Sampling Rate*.) Die Einsicht, dass der Klang entscheidend bestimmt wird durch seine spektrale Veränderung in der Zeit, ist durch die Arbeiten von J.C. Risset gefördert worden. Zwei Verfahren der Klangsynthese mögen besonders genannt werden: die "*Frequenzmodulation*" von J. Chowning (1973) und das *VOSIM*-Modell des Verfassers (1973), das inzwischen zur *VOSIM*-Klangsynthese in Soft- und Hardware-Versionen ausgebaut worden ist (Kaegi/Tempelaars 1978). Die Anwendungsmöglichkeiten ihrer Systeme sind durch musikalische Kompositionen der Autoren, ihrer Mitarbeiter und Studenten demonstriert worden (Chowning, Barbaud, Risset, Truax, Kaegi usw.). Programme zur Organisation von musikalischen Symbolen und klanggenerierende Programme können zu kompatiblen Paaren zusammengeschaltet werden. Damit wird den zur Stummheit verdamnten Symbolprogrammen sozusagen eine Stimme verliehen (P. Barbaud, *SCRIPTU* + F. Brown, *AUDITU*). Lösungen dieser Art bilden m.E. einen Kompromiss und ergeben in mancher Hinsicht unbefriedigende Resultate (P. Grossi, *TAU2*). Es ist auch versucht worden, die Ansteuerung des Klangprogrammes durch einen statistischen Algorithmus versorgen zu lassen (B. Truax, *POD5*).

### **Formale Sprache und ihre Klanginterpretation**

Im einfachsten Falle kann man sich einen Zufallsgenerator denken, der die Argumente für die Steuereingänge des Klangprogrammes oder sogar nur die Amplitudenwerte der Zeitfunktion "automatisch" würfelt. Auch steht es jedermann frei zu behaupten, dass eben gerade dasjenige Musik sei, was sein System zu liefern vermag. Wir könnten mit Schweigen über derartige Kindereien hinwegsehen, würden

nicht gerade sie eine entscheidende Frage hell ins Licht rücken: Sind wir denn überhaupt frei, willkürlich zu setzen, was unter Musik zu verstehen sei? Ein Seitenblick auf die natürlichen gesprochenen Sprachen lehrt, dass davon keine Rede sein kann. Die lebendige Sprache ist der gefährlichste Feind für jenen, der sich der automatischen Sprachsynthese hingibt, denn allein schon das unschuldige Kriterium der Verstehbarkeit wird zur Richterin, deren Urteil unbarmherzig bejaht oder verneint. Und darum auch lassen sich Bedingungen formulieren, welche ein Sprachsynthesesystem mindestens zu erfüllen hat. Es liegt nun auf der Hand, sich auch im Bereich der klingenden Musik nach Kriterien umzusehen, um vielleicht dereinst die Gretchenfrage zu entscheiden, was denn Klänge zu *musikalischen* Klängen werden lässt? Diese Frage ist eng verknüpft mit dem Problem der Formerkennung, und das macht deutlich, wie glücklich der Computer gerade hier eingesetzt werden kann. Um im grossen Stil experimentieren zu können, bedarf es klanggenerierender Systeme von möglichst breiter Ausdruckskraft und den Mut, sie da in Gefahr zu bringen, wo sie am verletzbarsten sind: in der Nachahmung von allgemein anerkannten Klangkonzepten im weitesten Sinne. Man wird also von derartigen Systemen die Fähigkeit erwarten, nicht-triviale Klangklassen bilden zu können. (Die Klasse aller "lauten Klänge" des Systems ist trivial, jene aller "Oboenklänge" ist es nicht.) Für den Entwurf von Systemen dieser Art kann die Theorie der formalen Sprachen herangezogen werden. Die endliche Menge aller diskreten Ausgangssignale eines klangerzeugenden Systems lässt sich dann als "Alphabet" auffassen, über welches "Sprachen" gebildet werden können. Besteht nun eine eindeutige Abbildung zwischen Eingang und Ausgang des Klangsystems, dann lässt sich jede solche Sprache durch eine Menge von Steuerregeln oder "Grammatik" charakterisieren. Jetzt lassen sich auch Bedingungen formulieren, denen eine "musikalische" Sprache in jedem Falle zu genügen hat. Als Beispiele seien die Forderungen genannt, jedem Klang zwar eine variable Zeitdauer zuzuordnen zu können, ohne aber die Dauer seines Einsatzes mit zu verändern, und das gegenseitige Verhältnis von Tonhöhe und Klangfarbe durch Vorschriften zu definieren, da sonst die Erkennbarkeit des Klanges zerstört wird (feste und gleitende Formanten, Register usw.). In diesem Sinne hat der Verfasser 1976-78 in Utrecht das MIDIM-System für die musikalische Komposition entwickelt (*Minimum Description of Music*). Die heutige Technologie erlaubt es, umfangreiche Systeme auf kleinen Maschinen zu implementieren und der Arbeit in Realzeit zuzuführen; mit Hilfe von durch Mikroprozessoren gesteuerten VOSIM-Generatoren dreht MIDIM heute auf einem DEC LSI-11-Computer. Ähnliche Systeme werden den Komponisten ohne Zweifel sehr bald auf ihren Home-Computern zur Verfügung stehen.

Die Schnelligkeit, womit die Technologie sich entwickelt, erlaubt es dem Komponisten, seine Visionen heute besonders hochfliegend zu konzipieren, ohne im geringsten die Chance ihrer Realisierbarkeit zu verringern. Seine Sache ist es also, die Herausforderung anzunehmen und die Musik unserer Zeit sinnvoll zu formulieren.

Utrecht, im Dezember 1982