

Klarenz Barlow

Von der Musiquantenlehre

Band I: Texte

Feedback Papers 34

FEEDBACK PAPERS 34
2008

Alle Rechte vorbehalten
Copyright 2008 by Johannes Fritsch, Rolf Gehlhaar & Co.
ISBN 978-3-9812713-2-4

Feedback Studio Verlag, Genter Straße 23, 60672 Köln
Tel. +49 221 527763 Fax 5104139

Von der Musiquantenlehre Band I – 1. Auflage

Layout, Design, Text und Grafik: Klarenz Barlow
Druck und Verarbeitung: Hundt Druck GmbH Köln
Mit freundlicher Unterstützung der Universität Kalifornien in Santa Barbara

Vorwort

Diese Arbeit ist aus Materialien gewachsen, die ich für meinen Unterricht im Fach Computermusik an der Hochschule für Musik Köln von seiner Einrichtung 1984 bis zu seiner Beendigung 2005 zusammenstellte. Auch am Königlichen Konservatorium Den Haag war von 1990-2006 diese Thematik fester Bestandteil meines Lehrstoffs, wie sie es seitdem an der Universität Kalifornien in Santa Barbara ist. Von der Faszination der Verbindungen von Musik zu Mathematik, Akustik, Phonetik und Informatik, und weiter von meiner eigenen Beschäftigung mit der Quantifizierung von Harmonik und Metrik angetrieben, erstellte ich nach und nach 32 Kapitel mit einer Vielzahl von Abbildungen.

Von der Musiquantenlehre ist bewusst sehr komprimiert geschrieben, jedes Kapitel auf zwei Seiten, mehr als eine konzentrierte und umfassende Begleitung zum Unterricht, weniger als ein autonomes Lehrbuch. Hinzu kommt eine Rolle als Nachschlagewerk, vor allem dann, wenn sein Inhalt bereits assimiliert wurde. Es ist meine Hoffnung, dass meine ehemaligen Schüler, die soviel Geduld mit dieser Materie bei manchmal sogar mehrfach wiederholten Kursbesuchen aufbrachten, darin alles finden was sie gelernt haben, und dass sich die *Musiquantenlehre* für diejenigen, die in den Lehrbetrieb gegangen sind, für ihren eigenen Unterricht als nützlich erweist. Ich hoffe auch, dass diese nach 24 Jahren nunmehr fertiggestellte Arbeit meine jetzigen und künftigen Studenten die Materie leichter erlernen lässt als es ihren Vorgängern möglich war.

Zum Schluss möchte ich all jenen Freunden und Kollegen danken, die im Laufe der Jahre in Kontakt mit dem langsam wachsenden Buch kamen und mir dazu wertvolle Anregungen geliefert und Hilfe geleistet haben. Meinem Verleger und Freund Prof. Johannes Fritsch gebührt besonderer Dank; er hat mit vorbildlicher Geduld viele Jahre auf die Herausgabe der *Musiquantenlehre* gewartet.

Klarenz Barlow
Köln 2008

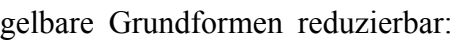
Von der Musiquantenlehre

Inhalt

Kapitel		Seite
1	Mathematik 1 : Zahlensysteme	2
2	Mathematik 2 : Zweidimensionale Kurven	4
3	Harmonik 1 : das Tonhöhenintervall	6
4	Mathematik 3 : Lineares & Logarithmisches	8
5	Mathematik 4 : Umrechnung linear-logarithmisch	10
6	Akustik 1 : Tonhöhe & Lautstärke	12
7	Harmonik 2 : Frequenzverhältnisse	14
8	Akustik 2 : Cent & Dezibel verglichen	16
9	Akustik 3 : Lautheit & Lautheitspegel	18
10	Harmonik 3 : Reine Stimmung	20
11	Mathematik 5 : Trigonometrie & Analytische Geometrie	22
12	Akustik 4 : Schallwelle und Spektrum	24
13	Informatik 1 : Hardware & Software	26
14	Informatik 2 : Digital & Analog	28
15	Informatik 3 : Programmierung	30
16	Informatik 4 : Die Programmiersprache C	32
17	Informatik 5 : Allgemeine Funktionen in C	34
18	Informatik 6 : MIDI	36
19	Harmonik 4 : Harmonizität	38
20	Harmonik 5 : Tonhöhenrationalisierung: Theorie	40
21	Harmonik 6 : Tonhöhenrationalisierung: Praxis	42
22	Metrik 1 : Quantifizierte Metrik	44
23	Metrik 2 : Metrische Kohärenz	46
24	Mathematik 6 : Stochastische Analyse & Synthese	48
25	Akustik 5 : Spektralanalyse & -synthese	50
26	Akustik 6 : Frequenzmodulation & Phasenverzerrung	52
27	Akustik 7 : Klang & Geräusch	54
28	Akustik 8 : Tonhöhenempfindung	56
29	Akustik 9 : Maskierung	58
30	Akustik 10 : Konsonanz und Dissonanz	60
31	Phonetik 1 : Physiologische Phonetik	62
32	Phonetik 2 : Akustische Phonetik: Formanten	64

01 Mathematik 1 – Zahlensysteme

Ein Blick auf das Anzeigefenster einer Digitaluhr zeigt, dass sämtliche an derselben Stelle erscheinende Ziffern aus verschiedenen Kombinationen derselben sieben Striche bestehen – durch das individuelle Ein- oder Ausschalten von diesen wird die jeweils erwünschte Gestalt erreicht. Sind andere Zeichen innerhalb dieser Gegebenheiten möglich? Wenn ja, wie viele und was für welche sind das?

Die Antwort auf diese Fragen sind $\Gamma 01$ zu entnehmen: zu sehen sind 128 verschiedene Zeichen, manche relativ vertraut (vor allem jene 73 schraffierten, ganz verbundenen, auf 28 spiegelbare Grundformen reduzierbar: ; bis auf letztere 7 sind alle zudem schlangenförmig).

Ob darüber hinaus wirklich keine weiteren Zeichen mit den sieben wie oben angeordneten Strichen möglich sind, kann wie folgt geprüft werden:

Die sieben Striche seien zunächst in eine feste Betrachtungsreihenfolge gebracht (z.B. | | - | - | | | |), so dass wir sie bei jeder zu prüfenden Kombination immer auf gleiche Weise der Reihe nach durchnehmen. Der ein- bzw. ausgeschaltete Zustand der Striche kann mit 0 bzw. 1 angegeben werden; sind alle ausgeschaltet, so schreiben wir 0000000 – das Einschalten des letzten Striches ergibt die Darstellung 0000001. Wenn wir ganz systematisch alle solche Variationen von 0 und 1 auflisten (so etwa: 0000000, 0000001, 0000010, 0000011, 0000100, 0000101, ... 1111001, 1111101, 1111111), so sehen wir, dass diese genau 128 an der Zahl sind. Wären statt der sieben Striche nur vier zu untersuchen, so sähe unsere Zählung so aus: 0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111, 1000, 1001, 1010, 1011, 1100, 1101, 1110, 1111.

Diese Anordnung der Ziffern 0 und 1 wird eine *binäre* Zahlendarstellung genannt, weil sie nur zwei Grundziffern benutzt: 0 und 1 (,binär‘ kommt vom lat. *bīnī* = ,paarweise‘). Die geläufigere, mit den zehn Ziffern 0 bis 9 gebildete Zählweise nennt man die *dezimale* (vom lat. *decem* = ,zehn‘). Die dezimale Zahl 5741 bedeutet genau $5000+700+40+1$ oder $5 \cdot 10^3 + 7 \cdot 10^2 + 4 \cdot 10^1 + 1 \cdot 10^0$ (für jede Zahl n gilt $n^0=1$); die binäre Zahl $ijklmn$ bedeutet $i \cdot 2^5 + j \cdot 2^4 + k \cdot 2^3 + l \cdot 2^2 + m \cdot 2^1 + n \cdot 2^0$ oder $32i+16j+8k+4l+2m+n$ – man bedenke, dass die Ziffern i bis n nur den Wert 0 oder 1 haben können. 101101 (binär) gleicht also $32+0+8+4+0+1=45$ (dezimal).

Nicht nur ist für das globale Dezimalsystem die biologische Beschaffenheit menschlicher Hände mit ihren zehn Fingern ausschlaggebend, man kann selbst an den Ziffern verschiedener Weltkulturen Finger und Hände als Zeichenursprung noch erkennen, wie z.B. bei den hierzulande geläufigen sogenannten ,arabischen‘:

1 zeigt einen nach oben, 2 und 3 zwei und drei nach links gerichtete Finger. Laut Victor Goldschmidt (1932) zeigt 5 eine offene, nach unten gehaltene Hand und 0 eine geballte Faust für die Zehn – erst spät wurde hier die Rolle der Null übernommen (vgl. auch die Wählscheibe älterer Telefonapparate, wo 0 zehn Impulse auslöst). Weiter bedeute 4 u.a. fünf minus eins ($\text{,}\mathfrak{4}\text{'}$), 6 zeige eine Hand mit aufrechtem Zeigefinger als $5+1$ ($\text{,}\mathfrak{6}\text{'}$) und 9 zehn minus eins ($\text{,}\mathfrak{9}\text{'}$). Zuletzt sehen wir in 7 und 8 – ähnlich 2 und 3 – wieder zwei und drei Finger wohl der anderen, hinzugezogenen Hand.

Das dezimale und das binäre Zählsystem unterscheiden sich in ihrer *Basis*, bei ersterem 10, bei zweiterem 2. Den 10 dezimalen Ziffern (0 bis 9) vergleichbar gibt es nur 2 binäre Ziffern (0 und 1). Bei zweistelligen Zahlen eines Zählsystems folgt auf jede mögliche Erstziffer jede mögliche Zweitziffer; daher gibt es $10^2=100$ zweistellige dezimale und $2^2=4$ zweistellige binäre Zahlen. So ist auch zu ersehen, dass die Anzahl aller n -stelligen Zahlen B^n beträgt, wobei B die Basis des Systems ist. Die größte n -stellige B -bezogene Zahl beträgt B^n-1 , z.B. die größte 7-stellige Binärzahl $1111111_b = (64+32+16+8+4+2+1)_d = 127_d = 2^7-1$ (die Zusätze $_b$ bzw. $_d$ bedeuten ‚binär‘ bzw. ‚dezimal‘).

Die kleinste nicht negative Zahl jeder Basis ist 0; die höchste 7-stellige Binärzahl beträgt 1111111_b oder 127_d . Es gibt $2^7=128$ unterschiedliche 7-stellige Binärzahlen; das zeigt, dass jede Zahl zwischen 0_b und 1111111_b (d.h. zwischen 0_d und 127_d) einer Binärzahl bzw. einem der 128 abgebildeten Kombinationen der sieben Striche entspricht.

Andere Basen als 2 und 10 sind nicht nur möglich, sondern (vor allem in der Datenverarbeitung) gebräuchlich – dem Oktalsystem liegt die Basis 8 zugrunde, dem verbreiteteren hexadezimalen die Basis 16: hierbei werden die Werte 10_d bis 15_d durch die Ziffern A B C D E und F dargestellt, noch un kreativer als der 1940er Vorschlag der amerikanischen Duodezimal-Gesellschaft, 10_d und 11_d als ‚X‘ und ‚E‘ (gesprochen *dek* und *el*) zu schreiben. Aufregender wären neue Ziffern wie z.B. 0 1 2 3 4 5 6 (0 1 2 3 4 5 6 in 7-Strichdarstellung; übrigens, 0 hieß 10 auch in Altägypten) sowie neue nichtdezimale Zahlennamen gewesen: generell tragen die Zahlwörter von 11–19 das Wort ‚zehn‘ sogar ausdrücklich, beispielsweise ab 13 in germanischen, ab etwas höher in romanischen Sprachen (‚elf‘ und ‚zwölf‘ bedeuten zu deutsch ‚eins‘ bzw. ‚zwei übriggelassen‘ – nach Abzug der ‚zehn‘). Zum Schluss seien auch die Basen 20 der Mayas und 60 der Altbabylonier erwähnt.

T02 zeigt Umrechnungen der Zahlen 0_d bis 255_d in binär, oktal (Zusatz $_o$) und hexadezimal (Zusatz $_h$): Ziffernsprünge ähnlich $9_d \rightarrow 10_d$, $99_d \rightarrow 100_d$ finden sich im Binären beispielsweise bei $11_b \rightarrow 100_b$, $1111_b \rightarrow 10000_b$, $11111_b \rightarrow 100000_b$ sowie in den anderen Systemen bei $7_o \rightarrow 10_o$, $77_o \rightarrow 100_o$, $F_h \rightarrow 10_h$ usw. wieder. Binär entspricht jede Oktalziffer stets der gleichen Dreier- (z.B. $3_o = 011_b$; $33_o = 011\ 011_b$; $333_o = 011\ 011\ 011_b$), jede Hexadezimalziffer stets der gleichen Viererkombination (z.B. $A0_h = 1010\ 1101_b$, $0A_h = 1101\ 1010_b$ oder $E3_h = 1110\ 0011_b$, $3E_h = 0011\ 1110_b$).

02 Mathematik 2 – Zweidimensionale Geometrische Kurven

Hängen variable Werte x und y voneinander derart ab, dass für x -Werte auch y -Werte existieren, so lässt sich in einem zweidimensionalen *kartesischen Koordinatensystem* (Erfinder: René Descartes [1596-1650]) eine Kurve zeichnen – s. $\Gamma 03$. Dieses System gibt Einzelpunkte als Zahlenpaar (x,y) an, wobei x den waagerechten und y den senkrechten Abstand von zwei Achsen angibt, der x -Achse $y=0$ und der y -Achse $x=0$, die sich am *Ursprung* $(0,0)$ kreuzen.

$\Gamma 03c$ zeigt, dass manch ein x -Wert keinen y -Wert hat, z.B. $x=\pm 1,57$ bei der Kurve mit *Funktion* $y=\hat{a}(x)$ (s. die senkrechten gestrichelten Linien) bzw. $x=0$ bei $y=1/x$; die Kurven ‚fliehen ins Unendliche‘, kehren dann wieder ‚von gegenüber‘ zurück. Auch bei $y=\pm\sqrt{x}$ ($y=\pm x^{1/2}$) in $\Gamma 03a$ und $y=\hat{a}(x)$ in $\Gamma 03b$ gibt es für $x<0$ kein y : Quadratwurzel und Logarithmen negativer Zahlen existieren nicht.

Kurvenkenntnisse erweisen sich als nützlich bei der Ermittlung algebraischer Annäherungen an empirisch gewonnene oder kompositorisch erdachte Verläufe. Betrachten wir $\Gamma 03a$: im von den beiden Achsen sowie $(1,1)$ eingegrenzten Rechteck sehen wir, dass $x^{1/2} > x > x^2$ gilt, bzw. rechts oben umgekehrt. $\Gamma 03b$ zeigt, dass $y=e^x$ nach links immer näher an die x -Achse, $y=\hat{a}(x)$ nach unten immer näher an die y -Achse kommt; die Kurven sind *asymptotisch* (vom gr. a [= ‚nicht‘] + syn [= ‚mit‘] + $ptōtós$ [= ‚fall-fähig‘, vom $píptein$ =fallen]), wobei hier die x - bzw. y -Achse als Asymptote dient (e und π [s.u.] sind die Konstanten 2,7182818284.. bzw. 3,1415926536..).

$\Gamma 03c$ zeigt Kurven mit achsenparallelen Asymptoten: z.B. bei $y=1/x$ ist die Asymptote zum Ursprung hin die y -Achse, nach außen die x -Achse; bei $y=\hat{a}(x)$ gibt es viele parallele Asymptoten $x=1,57n$ (richtiger: $=n\cdot\pi/2$) bei ungeradem n . Die Asymptote bei $y=\hat{Y}(x) [(e^x - e^{-x})/(e^x + e^{-x})]$ ist nach links $y=-1$, nach rechts $y=+1$. Bei den beiden nichtgeraden Kurven in $\Gamma 03a$ gibt es keine Gerade, zu der sie tendieren – obwohl rechts oben alle drei Linien sichtlich auseinandergehen, gilt für alle: $y \rightarrow \infty$ (unendlich) wenn $x \rightarrow \infty$. Sie sind also nicht asymptotisch. Alle Asymptoten in $\Gamma 03b$ und c sind senkrecht oder waagrecht; jedoch kann es auch nicht-achsenparallele Asymptoten geben.

Am Ursprung in $\Gamma 03c$ berührt die Gerade $y=x$ die Kurven $y=\hat{Y}(x)$ und $y=\hat{a}(x)$ mit gleichgerichtetem Verlauf; sie ist die *Tangente* beider Kurven an diesem Punkt. Auf einer glatt verlaufenden Kurve kann überall der Kontaktpunkt einer Tangente sein. In $\Gamma 03d$ ist die Gerade (Allgemeinformel: $y=mx+c$) die Tangente zur Kurve, die sie am Punkt (x_p, y_p) berührt. Die Steilheit oder *Steigung* dieser Tangente wird durch m angegeben, der stets gleichbleibende Quotient a/b der Längen der zwei angezeigten Seiten a und b eines vorstellbaren rechtwinkligen Dreiecks beliebiger Größe; das c in der Formel entspricht dem y -Wert der Tangente an der y -Achse. Eine waagerechte Gerade hat die Steigung 0; je mehr sie sich im Gegenuhrzeigersinn aufrichtet, desto höher wird die Steigung, bis diese bei einer senkrechten Lage gegen ∞ tendiert. Die Steigung einer nach rechts unten verlaufenden Geraden ist negativ.

$\Gamma 03e$ zeigt den Bereich $0 \leq x \leq 1$, $0 \leq y \leq 1$ in 16 Kästchen mit jeweils drei Kurven, deren Tangentialverhalten durch die Buchstaben F, f, j und J angezeigt sind: die F-Reihe weist bei (0,0) senkrechte, die F-Spalte bei (1,1) waagerechte Tangenten auf. Umgekehrt sind alle (0,0)-Tangenten der J-Reihe waagrecht, die der J-Spalte bei (1,1) senkrecht. Kleinbuchstaben bezeichnen Nicht-Tangenten, wobei f eine Krümmung nach oben rechts und j eine nach unten links angibt, den Buchstaben ähnlich. In jedem Kästchen steht zudem die Quellfunktion der Kurven, wobei die nicht-tangentialen Steigungen bei (0,0) bzw. (1,1) willkürlich der Grundsteigung g bzw. $\delta (=1/g)$ gleichen: $g=2, 2\frac{1}{2}$ und 3. Für $g=2$ sind FF und JJ Kreisquadranten, Ff, ff, Jj und jJ Parabelsegmente (z.B. ‚Ff‘ bedeutet Reihe F, Spalte f; ff-artige Kurven sind *konvex*, jj-artige *konkav*).

Die neun in $\Gamma 03e$ unten links befindlichen Formeln dieser Kurvensammlung sind *kubische Funktionen* der Form $y=a+bx+cx^2+dx^3$, wie z.B. $y=2x-3x^2+2x^3$ für $g=2$ bei Kurve fj; ist $d=0$ (wie für $g=2$ bei fF und Jj), so sind sie *quadratische Funktionen* der Form $a+bx+cx^2$. Beide sind *Polynome* der Form $y=ax^0+bx^1+cx^2+dx^3+..$, wobei $a, b, c, d..$ *Koeffizienten* sind – die höchste mit einem Nicht-Null-Koeffizienten versehenen *Potenz* (d.h. $0^1 2^3 ..$) ist der *Grad des Polynoms*. Alle neun Formeln haben die Form $y=m_0x+(3-2m_0-m_1)x^2+(m_0+m_1-2)x^3$, wobei m_0 und m_1 die Steigungen bei (0,0) bzw. (1,1) sind. Von den restlichen sieben Formeln sind Fj, FJ und fJ Interpolierungen zwischen jenen des Neunerblocks, die anderen sind willkürlich gewählt.

Die *Differentialrechnung* betrifft den Veränderungsgrad von Variablen, wie z.B. mit der Steigung einer Kurve an einem gegebenen Punkt (d.h. der hier gedachten Tangente) als Maß ihrer y -Veränderung: je steiler die Tangente, desto schneller steigt y im Bezug auf x . Die Steigung wird die 1. *Ableitung* von y genannt und als y' (oder dy/dx) notiert; die 1. Ableitung der Steigung ist die Krümmung der Kurve sowie die 2. Ableitung von y und wird als y'' (oder d^2y/dx^2) notiert. Daher ist Geschwindigkeit (auf Zeit bezogen) die 1. Ableitung von Abstand, Beschleunigung dessen 2. Ableitung und die 1. Ableitung der Geschwindigkeit. Beim Polynom $y=a+bx+cx^2+dx^3+..$ gilt $y'=b+2cx+3dx^2+..$. Grundsätzlich heißt es: für $y=mx^n$ gilt $y'=mnx^{n-1}$ und daher $y''=mn(n-1)x^{n-2}$ usw. $\Gamma 03f$ zeigt Kurve fj ($g=2$) und deren 1. und 2. Ableitungen als Funktion der Steigung: je steiler fj ist, desto höher Kurve (i), je krummer fj, desto steiler (i) und desto höher oder tiefer (ii) – eine negative y -Steigung führt zu einem tiefen y' . Eine konvexe Kurve ist negativ gekrümmt, eine konkave Kurve positiv. Die Umkehrung der Ableitung ist das *Integral*, in der *Integralrechnung* so notiert: $y=\int y' dx$.

Eine Methode, mehrere Punkte mit einer sehr glatt wirkenden Kurve zu verbinden, ist der *Spline*, eine Aneinanderkettung kubischer Funktionen derart, dass an jedem der Eingabepunkte drei gleiche Wertpaare gegeben sind: der Wert beider angrenzenden Funktionen (y), sowie deren Steigungen (y') und Krümmungen (y'') – s. $\Gamma 03g$.

$F04$ zeigt eine für Komponisten nützliche Ansammlung von Formeln, die bei der Berechnung der Beschleunigung einer Folge von Schlägen anhand vom Anfangs- und Schlusstempo, der Gesamtzeit und der Anzahl der Schläge eingesetzt werden kann.

03 Harmonik 1 – Über das Tonhöhenintervall

Dass es zwei Arten gibt, Tonhöhen wahrzunehmen, steht seit einiger Zeit fest: die eine erlaubt uns, musikalische Intervalle zu erkennen (s. Beispiele in T05) und dadurch in den Genuss der in der Welt überwiegend vorhandenen tonalen Musik zu gelangen; die andere, weitaus häufiger eingesetzte hat mehr mit dem Gefühl von ‚hoch‘ und ‚tief‘ oder mit ‚hell‘ und ‚dunkel‘ zu tun, nicht aber mit intervallischer Auswertung. Es gibt Gründe, zu glauben, dass diese zweite Art im Ohr entsteht, die erste erst im Gehirn. Die zweite Art ist es, wenn wir eine ‚Sprachmelodie‘ – in manchen Sprachen wie z.B. Chinesisch oder Thai zur Verständigung unentbehrlich – nur als solche wahrnehmen; erst bei wiederholtem Zuhören (z.B. bei einer Tonbandschleife) kommen wir auf die darin enthaltene ‚Tonhöhenmelodie‘. Auch die Tonhöhenwahrnehmung der Sprachformanten, die zur Erkennung von Sprachlauten erfasst werden müssen – dieser Bereich schließt auch die Erkennung von Klangfarben ein – bringt uns noch lange nicht auf melodische Gedanken. Selbst in der Musik gibt es viele Situationen, in denen eine intervallische Deutung der vorhandenen Tonhöhen unerwünscht ist, z.B. bei langsamen, ausgedehnten Glissandi oder bei Punktuellen – hier wirkt vorwiegend diese zweite Art des Hörens.

Seit Jahrtausenden sind Grundintervalle wie Oktav, Quint, Quart usw. Musikologen der ganzen Welt bekannt; dass die Töne dieser Intervalle verschiedene Grade von Zusammengehörigkeit aufweisen, war auch bekannt – man sprach in Westeuropa von ‚Konsonanz‘ und ‚Dissonanz‘ (lat. *con+* = ‚zusammen-‘ bzw. *dis+sonāre* = ‚auseinander-klingend‘). Ein Merkmal der Entwicklung des Tonhöhenmaterials in Europa während seiner dokumentierten Musikgeschichte ist die allmähliche Einfügung von immer – so wie es heißt – ‚dissonanteren‘ Intervallen; dies kam durch die immer komplexer werdende Harmonik (z.B. durch neue, gewagte Intervallverbindungen) zustande. Der mögliche Rückschluss, das Urtonmaterial muss nur aus den ‚konsonantesten‘ Intervallen bestanden haben, ist offensichtlich falsch. Als melodische Bausteine sind kleinere Intervalle wie Sekunden und Terzen bestimmt älter als die weltweit für ihre Konsonanz bekannten, aber intervallisch sprunghafteren Quart, Quinten oder Oktaven (man stelle sich eine ausschließlich aus diesen zusammengesetzte Melodie vor!). Tonhöhenintervalle werden in der Musiktheorie in zwei Größenklassen unterteilt: *Schritt* (kleine und große Sekunden) und *Sprung* (Terzen und größer); es gibt Gründe, die Schwelle zwischen diesen Klassen in tieferen Tonlagen nach oben zu verschieben.

Es ist möglich, dass eine Interaktion zwischen den beiden oben beschriebenen Hörarten – ich möchte sie die ‚intervallische‘ bzw. die ‚tonlagenmäßige‘ nennen – zum Aufbau von Tonhöhenkalen führte und zwar wie in folgender Vorstellung dargelegt:

Phase 1 – Sprachmelodie als gefühlsmäßiges Element der Sprache wird ritualisiert und formalisiert zu einer Art Sing-sang. Die Tonhöhenbewegungen sind verhältnismäßig klein (vielleicht Schritte bis zu einer Terz) und intervallisch noch undefiniert.

Phase 2 – Durch das Erweitern des Tonhöhenumfangs des Sing-sangs kommt man in die Nähe von ‚konsonanteren‘ Stellen im Tonhöhenraum (z.B. Quart und Quint), die intervallische Querverbindungen suggerieren. Hier fängt intervallisches Hören an. Durch ein Rationalisieren der Gesangsstufen werden Skalen geformt. Wahrscheinlich sind Quart, Quint und Oktav intervallisch bekannt und teils fixiert, während die anderen Stufen zunächst nur tonlagenmäßig angedeutet bleiben, ein Mischsystem.

Phase 3 – Man findet manche der klareren, gestimmten Skalenstufen übertrieben ‚sauber‘ und beginnt sie durch leichte Verstimmungen zu ‚trüben‘, zu ‚färben‘, ‚anzureichern‘. Hier wirkt wieder das Tonlagenhören: neben den intervallisch definierten Skalenstufen werden tonlagenmäßig fixierte Stufen vielleicht doch auch intervallisch zurechtgehört, sie sind aber jetzt reicher, vibranter als wenn sie reingestimmt wären. Es mag interessieren zu wissen, dass die Einführung der ersten nicht-diatonischen Skalenstufen (z.B. B oder Fis in die Tonart C) *Chromatisierung* genannt wurde (vom griech. *chroma* = ‚Farbe‘): daher stammt die Bezeichnung *Chromatische Skala* für das Endergebnis dieses Prozesses.

Phase 4 – Die gefärbten Stufen rationalisieren sich erneut: z.B. die kleine Terz, die möglicherweise als Färbung der großen entstand, fixiert sich auf ihre heute bekannten reingestimmten Stellen.

Alles andere ist eine zyklische, eher spiralartige Wiederholung von Phasen 3 und 4. In chinesischer Musik werden oft an sich klar definierte Intervalle wie die Oktav gerade so weit verstimmt, dass sie zwar sehr vibrant klingen aber noch als Oktaven wirken; das wäre ein Beispiel von Phase 3. In arabischer Musik kommen durch solche Verfärbung Intervalle wie die etwa auf halbem Weg zwischen kleiner und großer Terz befindliche neutrale Terz zustande, zuerst als intervallisch undefinierte Abweichung der beiden anderen, dann in den Körper der reinen Stimmung hineinbezogen, ein Beispiel für Phase 4. Hier möchte ich an Theorien zur Entstehung des Mondes erinnern: hat sich der Mond neben, unabhängig von aber zeitgleich mit der Erde aus derselben Staubwolke geformt, oder wurde er aus der Erde herausgerissen? So könnte man sich die neutrale Terz als unabhängige Erscheinung vorstellen oder aber – wie hier dargestellt – als später fixierte Umstimmung der anderen Terzen.

Diese spekulative Darlegung deutet Vorgänge an, die wahrscheinlich – wenn überhaupt – zum größten Teil in der menschlichen Vorgeschichte, sogar in der vormenschlichen Geschichte stattfanden und zwar nicht der Reihe nach sondern ineinander verschachtelt. Überhaupt ist es vorstellbar, dass auch heute die als Leittöne fungierenden Stufen einer Skala mehr zur Tonlagen- als zur Intervallstimmung neigen und dass sie erst beim Rollenwechsel durch eine Modulation vorübergehend an intervallischer Relevanz gewinnen.

04 Mathematik 3 – Das Lineare und das Logarithmische

Eine *arithmetische Folge* ist eine Folge von Zahlen, die eine konstante Differenz zweier benachbarter Zahlen aufweist, z.B. 2, 5, 8, 11, 14,... Hier ist die Differenz immer 3. In einer *geometrischen Folge* ist der Quotient zweier benachbarter Zahlen konstant, z.B. 2, 4, 8, 16, 32,... Hier ist der Quotient immer 2.

In der *Obertonreihe* (s. $\Gamma 06a$) sind die Frequenzen sämtlicher Tonhöhen ein ganzzahliges Vielfaches von der Grundton-Frequenz, z.B. 100 Hz, 200 Hz, 300 Hz, 400 Hz, 500 Hz, usw., also eine arithmetische Folge; die diesen Frequenzen approximativ entsprechende Notennamen sind nach dem gängigen, doch umständlichen deutschsprachigen System angegeben: ab 130 Hz sind die Namen aufwärts mit kleinem, abwärts mit großem Anfangsbuchstaben angegeben; die sich jeweils von C bis H laufenden Oktaven werden ab 250 Hz aufwärts mit Strichen, ab 67 Hz abwärts mit Ziffern gezählt. Was an dieser Tonfolge auffällt, ist, dass nach oben hin die Tonhöhenintervalle immer kleiner werden, obwohl die Frequenzdifferenz immer gleich bleibt, nämlich 100 Hz. Der Grund dafür ist, dass nicht die Differenz, sondern der Quotient zweier Frequenzen für die Größe des dazwischenliegenden Intervalls ausschlaggebend ist.

Wollten wir eine Folge von Tönen gleichen Abstandes aufstellen, so müssten wir die Frequenz eines jeden Tones mit einem konstanten Faktor multiplizieren; bei dem Intervall der Oktav wäre dieser Faktor 2, so dass – von der Frequenz 27,5 Hz vom tiefsten A des Klaviers ausgehend – die folgenden gleichnamigen Töne die Frequenzen 55, 110, 220, 440, 880, 1760... Hz aufweisen – s. $\Gamma 06b$ (die Notennamen sind ebenfalls angegeben). Diese ‚exponentiell‘ zunehmende (geometrische) Frequenzkurve hat also die Empfindung einer konstant steigenden Tonhöhe zur Folge. Wenn wir mit der Taste ‚ln‘ eines Taschenrechners die *natürlichen Logarithmen* dieser Frequenzen errechnen (die Bezeichnung ‚Logarithmus‘ entstammt den griechischen Wörtern *lógos* [hier = ‚Verhältnis‘] sowie *arithmós* [= ‚Zahl‘]), so ergeben sich folgende Werte: 3,3, 4,0, 4,7, 5,4, 6,1 usw., Zahlen mit einer konstanten Differenz von 0,7. Diese arithmetisch-logarithmische Zahlenfolge bietet eine bessere Entsprechung zur Tonhöhenempfindung als die lineare Folge der Frequenzwerte, wie von den Wissenschaftlern Ernst H. Weber (1795-1878) und Gustav T. Fechner (1801-1887) im mittleren 19. Jahrhundert eindrucksvoll begründet: das Weber-Fechner-Gesetz besagt, dass eine gerade noch wahrnehmbare Reizzunahme einen festen Prozentsatz des Ausgangsreizes ausmacht, woraus folgt, dass wenn physikalische Reize geometrisch zunehmen, deren biologische Empfindungen sich arithmetisch verhalten. Dieses Gesetz hat im Seh-, Hör-, Tast- und praktisch allen anderen Sinnesgebieten Gültigkeit.

Γ03b zeigt drei Kurven: $y=e^x$, $y=x$, $y=\ln(x)$. Die Kurve von $y=\ln(x)$ sieht aus wie eine Krümmung nach unten rechts der Ausläufer der Kurve von $y=x$, deren Logarithmus sie darstellt. Nach den Logarithmus-Gesetzen ist $y=x$ der Logarithmus von $y=e^x$; obwohl erstere eine Gerade ist, ist sie auch eine ähnlich verlaufende Krümmung der zweiten Kurve. Man könnte sagen, die Logarithmisierung bewirke eine Art ‚Dämpfung‘ (Krümmung nach ‚unten‘ in den negativen y -Bereich, ‚Konvexisierung‘) einer gegebenen Kurve; der umgekehrte Weg – die Linearisierung – lässt die Kurve bei zunehmenden x schneller steigen. $y=e^x$ ist die lineare Form von $y=x$, welche in diesem Zusammenhang die logarithmische ist; $y=x$ ist die lineare Form von $\ln(x)$. Wie zu ersehen ist, ist die Bezeichnung ‚linear‘ nicht gleichbedeutend mit ‚geradlinig‘; der Begriff *Antilogarithmus* wird manchmal dafür verwendet (der ‚Antilog‘ von x ist e^x).

Der Ursprung des Logarithmus (allgemein abgekürzt \hat{a}) kann so veranschaulicht werden: gilt das Verhältnis $a=b^c$, so ist c der ‚Logarithmus vom Argument a zur Basis b ‘, als $\hat{a}_b a$ oder $\hat{a}_b(a)$ notiert. Der dekadische oder Zehner-Logarithmus von 2 (zur Basis 10, häufig abgekürzt als \bar{a} , welches hier Verwendung findet) ist 0,30103, weil $10^{0,30103}=2$ gilt. Z.B. könnte der Ausdruck $\bar{a}(u)$ bzw. $\hat{a}_{10}(u)$ als ‚10-hoch-wieviel-ist(u)?‘ verstanden werden. Die natürlichen Logarithmen (abgekürzt \bar{a}) sind solche zur Basis e , der nach dem Mathematiker Leonhard Euler [1707-83] sogenannten Euler'schen Konstanten $2,7182818284590452353602874\dots = (1+1/n)^n$ bei sehr großem n (bei den ersten Berechnungen von Logarithmen erwies sich die Basis e als die am einfachsten Handhabbare). Mit dieser Notationsweise ausgerüstet, lässt sich das Weber-Fechner-Gesetz wie folgt schreiben: $E=k+c.\hat{a}(R)$, wobei R den Reiz und E die Empfindung bezeichnen (k und c sind gebietspezifische Konstanten).

Einige Regeln:

$$\hat{a}(m.n) = \hat{a}(m) + \hat{a}(n)$$

$$\hat{a}(m/n) = \hat{a}(m) - \hat{a}(n)$$

$$\hat{a}(m^n) = n.\hat{a}(m)$$

$$\text{ist } n=\bar{a}(m), \text{ so ist } m=e^n$$

$$\hat{a}(1) = 0$$

$$\hat{a}_a(a) = 1$$

$$\hat{a}_a(b) = 1/\hat{a}_b(a)$$

$$\hat{a}_a(m) = \hat{a}_a(b).\hat{a}_b(m)$$

$$\bar{a}(m) = \bar{a}(e).\bar{a}(m) = 0,434294\ln(m)$$

$$a^{\hat{a}_a(m)} = m$$

Der natürliche Logarithmus lässt sich auch als unendliche Potenzreihe darstellen:

$$\bar{a}(1+x) = x - x^2/2 + x^3/3 - x^4/4 + \dots$$

05 Mathematik 4 – Die Umrechnung Linear↔Logarithmisch

Bei vielen Phänomenen bewirkt eine geometrische Veränderung eines Parameters eine arithmetische Empfindung beim Empfänger – messbare und fühlbare Werte bilden meist unterschiedliche Skalen. Ein Ton, dessen *Frequenz* als Maß der Tonhöhe fortlaufend mit einem konstanten Faktor multipliziert wird (eine geometrische Veränderung), steigt stets um ein konstantes *Tonhöhenintervall* (arithmetisch). Ein Ton, dessen *Schallintensität* als Maß der Lautstärke fortlaufend mit einem konstanten Faktor multipliziert wird (geometrisch), steigt stets (arithmetisch) um ein konstantes *Lautstärkenintervall*. Auch außerhalb der Musik erscheint diese Relation: Ein Stern, dessen Helligkeit fortlaufend mit einem konstanten Faktor multipliziert wird (geometrisch), wirkt stets (arithmetisch) um ein konstantes Intervall heller. $\Gamma 07$ zeigt das arithmetisch-geometrische Verhältnis anhand dieser drei Instanzen: jeweils sind auf der x -Achse unten Intervalle, auf der y -Achse links die entsprechenden Messwertfaktoren – nicht maßstabgetreu – angezeigt. Zur zusätzlichen Veranschaulichung sind oben Unterteilungen eines einzigen Intervalls und rechts die entsprechenden Faktoren – hier maßstabgetreu – gezeigt, einer optischen Vergrößerung der kleinen, jeweils unten links befindlichen, fett umrandeten grauen Rechtecke gleich. In jedem Bild ist also die Entsprechung zwischen der arithmetisch steigenden Intervallgröße und des resultierenden, geometrisch steigenden Faktors zu sehen, in $\Gamma 07a$ und b mit Frequenzfaktor 2 und Schallintensitätsfaktor 10 pro angezeigtes Intervall (rechts im Bild auch a) Frequenzfaktor $2^{1/5} = \sqrt[5]{2}$ pro Fünftelintervall bzw. b) Schallintensitätsfaktor $10^{1/10} = \sqrt[10]{10}$ pro Zehntelintervall); die Maßeinheiten dieser Intervalle (Oktav und Bel) werden später erläutert.

Zur Sternhelligkeit: Laut Definition ist ein Stern, der 100mal so hell ist wie ein anderer, um fünf *Helligkeitsgrade* heller; das ist die Einheit der in $\Gamma 07c$ gezeigten Intervalle. Ein Helligkeitsintervall von 5 Grad entspricht also einem Helligkeitsverhältnis von 1:100, ein Helligkeitsintervall von 10 Grad demnach einem Helligkeitsverhältnis von 1:10.000, weil jede einzelne Steigerung um ein gewisses Intervall dem gleichen geometrischen Faktor entspricht, hier 100fach pro 5 Grad. Gleichzeitig entspricht ein Helligkeitsintervall von genau 1 Grad einem Helligkeitsverhältnis von $1:100^{1/5} = \sqrt[5]{100}$, etwa 2,512 (s.u.).

Ist der einem Grad oder Intervall entsprechende Faktor F , so bedeutet eine Steigerung eines Messwertes um x solche Grade oder Intervalle (wobei x sowohl ganzzahlig als ein Bruch sein kann) eine Multiplikation des Wertes mit F^x . Wie oben gesehen, würde umgekehrt die Teilung des Intervalls in x gleiche Teile zu x aufeinander folgenden geometrischen Steigerungen führen, und zwar jeweils um die x -te Wurzel von F : deswegen ist der einem einzigen Helligkeitsgrad entsprechende Faktor $F=100^{1/5}$, weil in diesem Fall $F^5=100$ ist. Die Zahl x kann sowohl ganzzahlig als auch ein Bruch sein.

Allgemein gesagt: wenn die Steigerung eines Messwertes um G intervallische Einheiten (z.B. Tonhöhen-, Lautstärkenintervalle oder Helligkeitsgrade) eine Erhöhung des Wertes auf ein F -faches desselbigen verursacht, dann entspricht ein Intervall von g Einheiten einem Wertquotienten q_w von

$$q_w = w_2/w_1 = F^{g/G}$$

wobei die Variable g das Intervall in der gegebenen Einheit, w_1 und w_2 die Messwerte der das Intervall bildenden Elemente sind. Man merke: g und G werden in gleichen Einheiten gemessen, die Konstanten F und G bedingen sich gegenseitig! $F^{g/G}$ schreibt sich auch $(\sqrt[G]{F})^g$, die g -te Potenz der G -ten Wurzel von F . Als Erinnerungsstütze merke man auch, dass F für ‚Faktor‘ oder ‚-fach‘ und g sowie G für ‚Grad‘ stehen können.

Ein Beispiel aus dem Lichtbereich (die für Helligkeitsgrade vorgesehene Zahl sinkt bei zunehmender Helligkeit; Grad 1 ist also 100mal heller als Grad 6, negative Grade noch heller):

F und G sind durch 100 bzw. 5 zu ersetzen, woraus die Formel $q_h = 100^{g/5}$ erwächst. Die Helligkeit des Sterns Sirius beträgt $-1,58$ Grad, die des Sterns Alpha Centauri $0,06$ Grad. Das Helligkeitsintervall g beträgt somit $1,64$ Grad. Also ist Sirius $100^{1,64/5} = 100^{0,328} = 4,529$ mal heller als Alpha Centauri.

Umgekehrt kann man die Intervallgröße in der gegebenen Einheit durch die Angabe der Messwerte oder zumindest deren Quotienten gewinnen:

$$g = G (\hat{a}(q_w) / \hat{a}(F))$$

Auch hier sind Variable g und Konstante G einheitsgleich; ebenso bedingen sich die Konstanten F und G gegenseitig.

Noch ein Beispiel aus dem Lichtbereich anhand der entsprechenden Formel $g = 5 (\hat{a}(q_h) / \hat{a}(100))$:

Die Sonne ist 15,85 Billionen* mal so hell wie ein gerade noch sichtbarer Stern mit Helligkeitsgrad 6. Also ist sie um $5(\hat{a}(15.850.000.000.000) / \hat{a}(100)) = 33$ Grade heller; ihr Helligkeitsgrad beträgt -27 .

* Im beinahe weltweiten Sprachgebrauch bedeutet *Billion* eine Million Millionen (nach der hier verwendeten, sogenannten *langen Leiter* genannt); dagegen meint es in den USA und Brasilien, auch tendenziell in manchen anderen, hauptsächlich englischsprachigen Staaten eintausend Millionen (die *kurze Leiter*), wofür die Peletier'sche Variante der langen Leiter *Milliarde* vorsieht.

06 Akustik 1 – Grundeinheiten der Tonhöhe und der Lautstärke

Ein Sinuston ist ein ‚reiner‘ Ton, bei dem eine einzige Frequenz erklingt; diese wird allgemein in *Hertz* (nach dem Physiker Heinrich Hertz [1857-94], kurz ‚Hz‘) gemessen, der Schwingungsanzahl pro Sekunde. Eine arithmetisch steigende Hertzahlenfolge lässt nicht gleiche Tonhöhenabstände empfinden; vielmehr tut dies eine geometrische Folge, wie z.B. die Oktaven ...110 220 440 880... Hz. Ein Oktavsprung nach oben, intervallisch in 12 gleiche Halbtöne bzw. 1200 *Cent* (lat. *centum* = ‚hundert‘, kurz ‚Ct‘; 1 Halbton = 100 Ct) traditionell teilbar, entspricht genau der Verdopplung der Frequenz; wenn also die Steigerung der Frequenz um 1200 Cent eine Erhöhung derselben auf ein zweifaches verursacht, so entspricht ein Intervall von c Cent einem Frequenzquotienten q_f von

$$q_f = f_2/f_1 = 2^{c/1200}$$

Hier ist c das Intervall in Cent, f_1 bzw. f_2 die Frequenz in Hertz der das Intervall bildenden Töne. Man merke: die Einheit für c gleicht der des Nenners 1200 ($q_f = 2^{h/12}$ gilt für das Intervall h in Halbtönen, $q_f = 2^{\Omega}$ für das Intervall Ω in Oktaven).

Ein Beispiel: Finde die Frequenz von c'' , 300 Ct über dem Kammerton a' (440 Hz).

Die Lösung: $f_2 = f_1 \cdot 2^{c/1200} = 440 \text{ Hz} \cdot 2^{300/1200} = 440 \text{ Hz} \cdot 2^h = 440 \text{ Hz} \cdot 1,1892 = 523,25 \text{ Hz}$

Noch eins: Finde die Frequenz von c' , 900 Ct unter dem Kammerton a' (440 Hz).

Die Lösung: $f_2 = f_1 \cdot 2^{c/1200} = 440 \text{ Hz} \cdot 2^{-900/1200} = 440 \text{ Hz} \cdot 2^{-3/4} = 440 \text{ Hz} \cdot 0,5946 = 261,63 \text{ Hz}$.

Dass diese Frequenz die Hälfte von jener von c'' beträgt soll nicht verwundern, denn c' liegt genau eine Oktav darunter. Siehe auch T08a für eine 12-oktavige Liste von Frequenzen im Abstand von 100 Ct.

Die umgekehrte Formel: $c = 1200(\hat{a}(q_f)/\hat{a}(2))$

Ein Beispiel: Wie groß ist das Intervall zwischen 500 Hz und 600 Hz?

$q_f = 600/500 = 1,2$; $c = 1200(\hat{a}(1,2)/\hat{a}(2)) = 1200(0,1823/0,6932) = 315,6 \text{ Ct}$.

Man sieht hierdurch, dass das Frequenzverhältnis 5:6 grundsätzlich dem Intervall 315,6 Ct (etwas größer als die kleine Terz zwischen a' und c'') entspricht.

Noch ein Beispiel: Wie groß ist das Intervall zwischen a' (440 Hz) und der Frequenz der Erdumdrehung (einmal in 23,9345 Stunden, d.h. 0,000011605 Hz)?

$c = 1200(\hat{a}(440/0,000011605)/\hat{a}(2)) = 1200(17,45/0,6932) = 1200 \times 25,17616 \text{ Ct}$ oder 25,17616 Oktaven, d.h. 25 Oktaven und 2,114 Halbtöne.

Dies bedeutet, dass die Erdumdrehung frequenzmäßig 11,4 Ct unter dem Ton G_{23} liegt.

Zur Lautstärke: Ein gerade noch hörbarer Ton von 1000 Hz übt einen *Schalldruck* (ein Maß der Lautstärke) von etwa 20 *Mikropascal* (abgekürzt ‚ μPa ‘) oder 0,2 *Nanobar* (‚nbar‘) aufs Ohr aus. Wird dieser Druck mit einem konstanten Faktor fortlaufend multipliziert, so entsteht der Eindruck einer ebenmäßig zunehmenden Lautstärke. Beim Faktor z.B. 10 wird das entsprechende Lautstärkenintervall mit 20 *Dezibel* angegeben.

Ein *Bel* (genannt nach dem Erfinder Alexander Graham Bell [1847-1922]) entspricht per Definition der Verzehnfachung der *Schallintensität* (ein anderes Lautstärkenmaß, in Leistung, z.B. Watt, pro Flächeneinheit, z.B. Quadratmeter) gemessen; also entsprechen 2 Bel einer Verhundertfachung der Schallintensität. Die Schallintensität variiert proportional zum Quadrat des Schalldrucks (s. T08b und T09; Letzteres veranschaulicht dieses Verhältnis) – die Verhundertfachung der Intensität bewirkt eine Verzehnfachung des Schalldrucks: 1 Bel = 10 Dezibel, 2 Bel = 20 Dezibel (hundertfache Intensität, zehnfacher Druck). Die Zusammenlegung mehrerer (nicht-sinusartigen) Schallquellen lässt deren Schallintensitäten summieren; doch ist für manche Betrachtungsweisen (z.B. bei Schallwellen-Amplituden oder Stromspannung) der Schalldruck ein geeigneteres Maß. Bei der Summierung von Sinustönen spielt die Phase eine entscheidende Rolle – z.B. zwei Sinuskurven gegenteiliger Phase löschen sich bei der Summierung gegenseitig aus (s. Kapitel 29).

Eine Steigerung der Lautstärke um 10 Dezibel (abgekürzt ‚dB‘) entspricht also genau der Verzehnfachung der Schallintensität; also entspricht ein Intervall von l dB einem Schallintensität-Quotienten q_i von

$$q_i = i_2/i_1 = 10^{l/10},$$

wobei l der Lautstärkenabstand in dB, i_1 bzw. i_2 die respektive Schallintensität (z.B. in W/qm) der das Intervall bildenden Töne sind.

Wie oben gesagt, entspricht eine 20 dB-Steigerung der Verzehnfachung des Schalldrucks; entsprechend gilt für einen Schalldruckquotienten q_p

$$q_p = p_2/p_1 = 10^{l/20},$$

l ist wie oben, p_1 bzw. p_2 der respektive Schalldruck (z.B. in μ Pa).

Ein Beispiel: Finde die Schallintensität eines Tones 120 dB über dem Pegel 1 pW/qm.
Die Lösung: $i_2 = i_1 \cdot 10^{l/10} = 1 \text{ pW/qm} \times 10^{120/10} = 1 \text{ pW/qm} \times 10^{12} = 10^{12} \text{ pW/qm} = 1 \text{ Watt}$.

Noch ein Beispiel: Finde den Schalldruck eines Tons 120 dB über dem Pegel 20 μ Pa.

Die Lösung: $p_2 = p_1 \cdot 10^{l/20} = 20 \mu\text{Pa} \times 10^{120/20} = 20 \mu\text{Pa} \times 10^6 = 20 \text{ Pascal}$.

Wie in T08b gezeigt, entspricht der Schalldruck 20 Pascal der Schallintensität 1 W/qm.

Umgekehrt gilt die Formel $l = 20(\hat{a}(q_p)/\hat{a}(10)) = 10(\hat{a}(q_i)/\hat{a}(10))$,
durch Zehnerlogarithmen auf $l = 20 \tilde{a}(q_p) = 10 \tilde{a}(q_i)$ zu vereinfachen.

Ein Beispiel: Wie viele dB entsprechen einer Schallintensitätsverdopplung?

Die Lösung: $l = 10 \tilde{a}(2/1) = 10 \times 0,30103 = 3,0103 \text{ dB}$.

Noch ein Beispiel: Wie viele dB entsprechen einer Schalldruckverdopplung?

Die Lösung: $l = 20 \tilde{a}(2/1) = 20 \times 0,30103 = 6,0206 \text{ dB}$, die Vervierfachung der Intensität bedeutend.

Aus obigem kann man folgern, dass Schalldruck und -Intensität (analog zu Frequenz) lineare und Dezibel (analog zu Cent) logarithmische Einheiten sind.

07 Harmonik 2 – Das Frequenzverhältnis eines rationalen Intervalls

Wird die Frequenz einer vorgegebenen Tonhöhe verdoppelt, so merken wir, dass die Tonhöhe sich erhöht und zwar um genau eine Oktav. Dies geschieht bei jeder beliebigen Ausgangsfrequenz, so dass wir sagen können: die Oktav ist ein Tonhöhenintervall, das einem Frequenzverhältnis von genau 1:2 entspricht. Wenn wir stattdessen das Verhältnis 2:3 (z.B. 100 Hz:150 Hz, der 2. und 3. Teilton einer auf 100 Hz basierenden Obertonreihe) betrachten, haben wir es hier mit der reinen Quint zu tun; weil 440 Hz dem Kammerton a' entspricht, ist die Frequenz vom e'', eine Quint höher, 660 Hz. Grundsätzlich können wir sagen, dass das Wesentlichste, das ein rationales Tonhöhenintervall charakterisiert, das Verhältnis zwischen seinen Frequenzen ist, eine Größe, die als Paar von teilerfremden Ganzzahlen der Form P:Q dargestellt wird.

Unsere musikalische Erfahrung sagt uns, dass eine Doppelquint einer großen None entspricht. Nehmen wir das Beispiel c'+Quint=g'; g'+Quint=d''. Ist f Hz die Frequenz von c', so beträgt die von g': $f \times 1,5 = 1,5f$ Hz. Da eine steigende Quint immer eine Frequenzzunahme von 50% verursacht, so ist die Frequenz von d'' $1,5f \times 1,5 = 2,25f$. Dass die Hälfte von diesem Wert ($2,25f/2 = 1,125f$) uns eine Oktav herunter auf d' bringt, hängt damit zusammen, dass eine fallende Oktav eine Halbierung der Frequenz zur Folge hat. Dieser Vorgang lässt sich auch so beschreiben:

$$\begin{array}{l} \text{oder} \quad \quad \quad c' \quad +\text{Quint} \quad +\text{Quint} \quad -\text{Oktav} \quad = d' \\ \quad \quad \quad \quad f \quad \times 3/2 \quad \times 3/2 \quad \times 1/2 \quad = (9/8)f. \quad \quad \quad * \end{array}$$

Daraus ist zu ersehen, dass die Steigerung einer Tonhöhe um ein Intervall durch die Multiplikation der Frequenz mit dem Bruch, der dem Frequenzverhältnis des Intervalls entspricht, erfolgt. Umgekehrt entspricht die Senkung um ein gewisses Intervall der Division durch den entsprechenden Bruch. Hier sind nun die Brüche von bisher vier Intervallen aufgezeigt worden, die bis zu einschließlich einer Oktav reichen; sie sind Prim (1/1), große Sekunde (9/8), Quint (3/2), Oktav (2/1).

Weitere Stufen können auf ähnliche Weise errechnet werden,
wie z.B. die große Sext = große Sekunde plus Quint, also Bruch = $9/8 \times 3/2 = 27/16$
oder die reine Quart = Oktav minus Quint, also Bruch = $2/1 \times 2/3 = 4/3$ *
oder die große Terz = große Sext minus Quart, Bruch = $27/16 \times 3/4 = 81/64$ *
oder die kleine Septim = Quart plus Quart, Bruch = $4/3 \times 4/3 = 16/9$

Obige Intervalle sind ausschließlich durch die Addition bzw. Subtraktion von Oktaven und Quinten gestimmt worden; diese Art von Stimmung heißt die *Pythagoreische* (nach ihrem Fürsprecher Pythagoras von Samos [ca.575-ca.500 v.Chr.]). Die darin befindliche große Terz 81/64 wird auch *Diton* genannt, weil sie sich aus der Summe zweier 9/8-Ganztöne ergibt ($9/8 \times 9/8 = 81/64$). Auch ersichtlich ist, dass die 16/9-Septim durch den Abzug eines 9/8-Ganztons von einer Oktav entsteht ($2/1 \times 8/9 = 16/9$).

(*die Teilung durch einen Bruch gleicht der Multiplikation mit der Bruch-Umkehrung)

Dass andere Stimmungen möglich sind, sieht man am Beispiel der selbigen großen Terz, die auch als 64:81 geschrieben werden kann (das Verhältnis P:Q – wobei hier die beliebige Konvention P<Q eingehalten wird – entspricht dem Bruch Q/P wenn das Intervall steigt, P/Q wenn es fällt). In der Obertonreihe, in der jeder Teilton ein Vielfaches der Grundfrequenz aufweist, ist das Intervall zwischen dem 1. und dem 2. Teilton (1:2) erwartungsgemäß eine Oktav, jenes zwischen dem 2. und dem 3. (2:3) eine Quint, jenes zwischen dem 3. und dem 4. (3:4) eine Quart (s. T06a). Zwischen dem 4. und dem 5. Teilton befindet sich ein Intervall, das mit 386 Ct etwas kleiner als die temperierte große Terz (400 Ct) ist; es klingt im Zusammenklang wie eine besonders reine, schwebungsfreie große Terz, weswegen es die *reine Terz* oder (wegen der Obertonreihe) *Naturterz* genannt wird. Vergleichen wir ihre Größe mit jener des pythagoreischen Namensvetters:

$$\begin{aligned} \text{bei } 5/4 \text{ beträgt das Intervall } 1200 \times (\hat{a} (5/4) / \hat{a} (2)) &= 386,31 \text{ Ct,} \\ \text{bei } 81/64 \text{ beträgt es } 1200 \times (\hat{a} (81/64) / \hat{a} (2)) &= 407,82 \text{ Ct.} \end{aligned}$$

Der 21,51 Ct-Unterschied, etwa ein Fünftel eines Halbtons, das *Syntonische Komma* genannt, entspricht dem Intervall $81/64 \times 4/5 = 81/80$.

Die Obertonreihe kann auf einer gespannten, schwingenden Saite (z.B. eines Klaviers oder Streichinstruments) durch Berührung von Knotenpunkten hörbar werden; für den n. Teilton beträgt der Abstand zwischen dem entsprechenden Knotenpunkt und dem Saitenende $1/n$ der Saitenlänge. Wird der 5. Teilton beispielsweise eines tiefen Cs auf dem Klavier zum Erklingen gebracht, so lässt sich feststellen, dass der mit der e'-Taste spielbare Ton tatsächlich nur 14 Ct höher liegt (vorausgesetzt, das Klavier ist gut gestimmt). Dass lässt sich auch so errechnen: das tiefe C hat die Frequenz 65.4064 Hz (s. T08a); folglich hat sein 5. Teilton die fünffache Frequenz 327,032 Hz. Jedoch ist laut T08a das genau 28 Halbtöne über C liegende e' höher: bei 329.628 Hz. Der Zusammenklang beider Töne ergibt aus akustisch erklärlichen Gründen eine deutlich hörbare, der Frequenzdifferenz entsprechende Lautstärkenschwebung von etwa 2,6 Hertz.

Natürlich lassen sich Intervalle auch durch die Addition bzw. Subtraktion von Oktaven, Quinten und (reinen) Terzen erstimmen, wie z.B.

die große Sext = reine Terz plus Quart; Bruch = $5/4 \times 4/3 = 5/3$ oder

die kleine Septim = zwei Quinten minus reine Terz; $3/2 \times 3/2 \times 4/5 = 9/5$.

Hier noch einmal die pythagoreische Ableitung (von C aus gemessen – hier sind Oktav (1:2), Quint (2:3) und Terz (4:5) mit , Q, T bezeichnet) der Töne

	D	E	F	G	A	B
	2Q-	4Q-2	-Q	Q	3Q-	2 -2Q
Alternativ dazu:		T			+T-Q	2Q-T

08 Akustik 2 – Cent und Dezibel verglichen

Ein Vergleich zwischen der Formel zur Umsetzung von Intervallgröße (Cent) in Frequenzverhältnis und jener zur Umsetzung von Lautstärkenabstand (Dezibel) in Intensitätsverhältnis zeigt interessante Parallelen, zum Beispiel wenn man den Frequenzquotienten einer temperierten großen Terz (400 Ct) errechnet:

$$q_f = 2^{400/1200} = 2^{1/3} = 1,25992105$$

Man kann also sagen: eine Frequenzsteigerung um 25,992% entspricht einer Erhöhung der Tonhöhe um eine temperierte große Terz.

Jetzt rechne man den Schallintensitätsquotienten eines Dezibels:

$$q_i = 10^{1/10} = 1,258925412$$

Man kann hier sagen: eine Schallintensitätssteigerung um 25,893% entspricht einer Erhöhung der Lautstärke um einen Dezibel.

Die Ähnlichkeit dieser beiden Werte kann man zunutze machen z.B. so: will man wissen, wie viele Dezibel einer Verdopplung der Schallintensität entspricht, so braucht man nur zu bedenken, wie viele große Terzen für eine Verdopplung der Frequenz nötig sind, oder in anderen Worten, wie viele große Terzen in eine Oktav passen – nämlich 3. Oder will man beispielsweise wissen, welchem Schallintensitätsverhältnis 10 dB entspricht, so braucht man nur zu überlegen, welchem Frequenzverhältnis 10 große Terzen (= 3 Oktaven plus Terz) entsprechen: $2 \times 2 \times 2 \times 5/4 = 10$ – wie per Definition (die Naturterz $5/4$ ist fast so groß wie die temperierte) – s. T10a für einen bildlichen Vergleich anhand einer Klaviatur der Tonhöhen- und Lautstärken-Einheiten (hier Schallintensitäts-Verhältnisse; für Schalldruck verdoppelt man die dB-Werte).

Die Tatsache, dass Cent und Dezibel logarithmische Ausdruckseinheiten für Messwertquotienten sind, die selber losgelöst von Maßeinheiten sind, ließe ihre Anwendung auch in anderen Bereichen zu, wie beispielsweise dem Bankwesen: würden wir annehmen, eine Geldsumme würde mit einem Zinseszins von jährlich 7,2% verzinst, so könnten wir wahrheitsgemäß behaupten, diese Summe wachse kontinuierlich um $1200(\hat{a}(1,072)/\hat{a}(2)) = 120$ Ct oder um $10 \lg(1,072) = 0,3$ dB pro Jahr. Um sich zu verdoppeln, braucht die Summe 1200 Ct/120 Ct bzw. 3 dB/0,3 dB = 10 Jahre. Diese Übung ist sicherlich eine Spielerei, aber die gewöhnlichen Anwendungsbereiche von Cent und Dezibel sollen nicht über die Tatsache hinwegtäuschen, dass sie logarithmisch errechnete Einheiten zur Messung von Verhältnissen sind.

So wie eine in Metern angegebene Zahl den räumlichen Abstand zwischen zwei Punkten beschreibt, werden Cent- oder Dezibelangaben dazu verwendet, den Abstand zwischen zwei Tonhöhen bzw. Lautstärken zu verdeutlichen. Berghöhen werden gewöhnlich in Metern über dem Meeresspiegel angegeben; hier wurde die Höhe des Meeresspiegels als ‚null Meter‘ definiert. Im Bereich der Temperatur wurde der Nullpunkt der Celsiusskala auf den Schmelzpunkt von Eis festgesetzt. Es gibt mehrere auf solche Weise fixierte dB-Skalen. Eine für akustische Vermessungen geläufige – dB(SPL) genannt (von engl. ‚Sound Pressure Level‘ = ‚Schalldruckpegel‘) – hat ihren Nullpunkt bei $20 \mu\text{Pa}$ ($= 1 \text{ pW/qm}$), laut allgemeiner Annahme dem schwächsten gerade noch mit den Ohren wahrnehmbaren Schalldruck bei 1000 Hz; T11 bietet eine Reihe von dB(SPL)-Beispielen aus dem Alltag. Eine andere Skala – dBu genannt – findet in Tonstudios Verwendung: der Nullpunkt ist hier 0,775 Volt (die bekannte, nach dem Physiker Alessandro Volta [1745-1827] genannte Einheit für Spannung, kurz ‚V‘) genormt; das ‚u‘ soll die Einheit vom dBV (Nullpunkt 1 V) unterscheiden. Von diesem maximalen Nullwert abwärts sind dBu-Werte negativ (vgl. die Kalibrierung des von Audiogeräten bekannten ‚Volume Unit [VU-] Meters‘ in T10b). Die Umrechnung von Spannungsverhältnissen in dB erfolgt nach der Formel für Schalldruck – die Spannung in Volt (V) verhält sich genau wie Schalldruck (Pa) und Amplitude (mm).

Wie aus T08b ersichtlich ist ein Pascal per Definition der Druck von 1 Newton Kraft (einem Gewicht von etwa 102 Gramm entsprechend) auf 1 Quadratmeter; folglich entsprechen $20 \mu\text{Pa}$ einer Gewichtverteilung von 2,04 Milligramm pro Quadratmeter oder etwa 20 gm pro Hektar (also: 1 Pascal = ca. 1 Tonne/Hektar); die Einheit (kurz ‚Pa‘) wurde nach dem Mathematiker und Philosophen Blaise Pascal (1623-1662) genannt. Eine andere Einheit für Druck ist ‚Bar‘ (über ‚Barometer‘ vom griech. *baros* = ‚Schwere‘ [der Luft]): 1 Bar = 100.000 Pascal oder 1.000 Hektopascal (kurz ‚hPa‘). Der Schalldruck einer Schallwelle ist das (als Effektivwert oder RMS – engl. *root mean square* = ‚Wurzel des Durchschnitts aller Quadrate‘) umgerechnete Ausmaß der durch sie verursachten Schwankung des Luftdrucks; da dieser normalerweise etwa 1 Bar beträgt, kann man sich vorstellen, wie eine schnelle Schwankung des Luftdrucks um nur $\pm 0,1\%$ (etwa 71 Pa RMS-Schalldruck bei einem Sinuston) zu ohrenbetäubenden 131 dB(SPL) führen kann.

T10c zeigt die im Raum Köln-Bonn während der Jahre 2000 und 2001 gemessenen Luftdruckschwankungen; diese Kurve stellt naturgemäß eine niederfrequente Schallwelle dar, unbeschadet der Kalibrierung der Zeitachse in Monaten und der Druckachse in Hektopascal. Hier beträgt die durchschnittliche RMS-Lautstärke 9,48 hPa oder 173,5 dB(SPL)! Dazu befinden sich die lautesten Frequenzkomponenten im einstelligen Mikrohertzbereich, wie in T10d sichtbar. Daraus wird klar, dass wenn mittels einer (wie vom Autor H. G. Wells geschilderten) Zeitmaschine dieser Zeitraum von 2 Jahren in z.B. 2 Sekunden durchflogen würde, der resultierende Schall – nun auf etwa 150 Hz erhöht – mit seinen 173,5 dB(SPL) den ungeschützten Ohren schwere Schäden zufügen würde.

Wir wissen, dass manche Tiere Frequenzen hören können, die für das menschliche Ohr zu hoch liegen – die untere bzw. obere Hörgrenze des Menschen wird gewöhnlich bei 20 Hz bzw. bei 20 kHz angesetzt – das entspricht einem Umfang von $1:1000 =$ etwa 10 Oktaven. Es ist aber nicht so, als würde das menschliche Hörvermögen bei 20 Hz plötzlich anfangen, um dann zehn Oktaven höher ebenso plötzlich zu erlöschen: die Hörbarkeit tritt an beiden Enden des Hörbereichs sehr allmählich ein und bleibt auch dazwischen unterschiedlich gut. Ein mit konstanter Schallintensität frequenzmäßig gleitender Sinuston erscheint viel lauter z.B. bei 1000 Hz als bei 100 Hz oder bei 10000 Hz. Wenn umgekehrt die Intensität eines Tones (z.B. durch einen Lautstärkenregler) verändert wird, um bei allen Frequenzen einen gleich lauten Eindruck zu vermitteln, so sind Pegelschwankungen die Folge, die 1933 von Harvey Fletcher (1884-1981) und Wilden A. Munson (1902-1982) veröffentlicht wurden; sie sind in (1956 von D.W.Robinson und R.S.Dadson) verbesserter, als DIN 45630 eingetragener Form als Kurven in $\Gamma 12$ zu sehen – gezeigt sind dB(SPL)-Verläufe von bei allen Frequenzen gleich laut wirkenden Sinustönen, deren Intensität bei 1000 Hz auf ein Vielfaches von 10 dB(SPL) gesetzt wurde. Zur optischen Hilfe sind hier auch mathematisch inter- und extrapolierte, punktiert gezeichnete Kurven hinzugefügt, die mit der empirisch ermittelten Realität nichts zu tun haben.

Vergleichen wir Messwerte bei den sechs ($1\frac{1}{2}$ Oktaven getrennt befindlichen) Tonhöhen 44, 125, 354, 1000, 2828 und 8000 Hz. Um gleich laut zu wirken gleiten dB(SPL)-Werte von z.B. 79 (44 Hz) über 63 (125 Hz), 56 (354 Hz), 60 (1 kHz), 54 (2,8 kHz) zu 67 (8 kHz). Diese Schwankungen sind bei leiseren Tönen noch extremer, wie z.B. bei dB(SPL)-Werten von 59 (44 Hz), 39 (125 Hz), 29 (354 Hz), 30 (1 kHz), 24 (2,8 kHz) und 39 (8 kHz). Auch die Hörschwelle, jene Lautstärke, bei der ein Sinuston gerade nicht mehr hörbar ist, zeigt den Verlauf 43, 20, 9, 3, $-2\frac{1}{2}$ (!) sowie 16 dB(SPL) bei den sechs genannten Peilfrequenzen. Der negative dB(SPL)-Wert bei 2,8 kHz entspricht einem Schalldruck von etwa 15 μPa , noch tiefer als der dB(SPL)-Nullpegel von 20 μPa ; die Verstärkung in diesem Frequenzbereich entspringt einer leichten Resonanz des zwischen Trommelfell und Ohrmuschel gelegenen äußeren Gehörgangs. Die je nach Tonhöhe und Intensität empfundene *Lautheit* [s.u.] wird mit einem Pegelwert (*Lautheitspegel*) in *Phon* (vom griech. *phōnē* = ‚ein Laut‘, kurz ‚Ph‘) beschrieben, d.h.: der Lautheitspegel von z.B. 60 Ph entspricht den im ersten Beispiel gezeigten, der – *Isophon* genannten (vom griech. *isos* = ‚gleich‘, *phōnē* s.o.) – Kurve entlang variierenden Intensitätspegeln. Per Definition sind dB(SPL)- und Phonwerte bei 1 kHz identisch. Aus den Kurven kann man also für jede Frequenz (x-Achse; in Hz) und jeden Intensitätspegel (y-Achse; in dB(SPL)) den entsprechenden Lautheitspegel (in Ph) ablesen; z.B. ein 500 Hz-Ton mit 50 dB(SPL) wirkt 54 Ph laut, also um 4 Ph lauter als bei 1000 Hz. Interessanterweise liegt die Hörschwelle bei 3 Ph (etwa 28 μPa bei 1 kHz). In gewisser Hinsicht können Isophone als Kurven ‚relativer Taubheit‘ betrachtet werden – je höher sie verlaufen, desto schlechter ist die Wahrnehmung.

Das Wort ‚Lautheitspegel‘ – analog zu ‚Schallintensitätspegel‘ – ist m.E. eindeutiger als die übliche Bezeichnung ‚Lautstärke‘ (im Englischen sagt man allgemein dafür *loudness level*); ‚Pegel‘ – wie ‚Intervall‘ – deutet immer eine logarithmische Skala an. ‚Lautstärke‘ bleibt hier eine Allgemeinbezeichnung, vergleichbar mit ‚Tonhöhe‘.

Eine Intensitätsverdopplung hebt dessen Pegel um etwa 3 dB – diese Feststellung leitet sich von der Definition des Dezibels ab. Bei 1 kHz entspräche dies einer Zunahme des Lautheitspegels um 3 Ph, die jedoch, wie es sich herausstellt, nicht den Eindruck der Lautheitsverdopplung erweckt; Messungen zeigen, dass für diesen Eindruck 10 Ph nötig sind. Das heißt, dass wenn die Schallintensität eines 1 kHz-Tons sich ver1024facht, der Lautheitseindruck nur auf das 8fache steigt (i=Intensität, l=Lautheit):

$$1024 \times i = 10 \times i\text{-Verdopplung} = i\text{-Pegel} + (10 \times 3 \text{ dB}) = i\text{-Pegel} + 30 \text{ dB}$$

$$\rightarrow l\text{-Pegel} + 30 \text{ Ph} = l\text{-Pegel} + (3 \times 10 \text{ Ph}) = 3 \times l\text{-Verdopplung} = 8 \times l\text{-Eindruck.}$$

Das lineare Verhalten des Lautheitseindrucks (schlicht *Lautheit*) verdient es, in einer linearen Einheit widergespiegelt zu werden: diese gibt es in der Form der Einheit *Sone* (vom lat. *sonus* = ‚ein Laut‘, kurz ‚Sn‘); 1 Sn entspricht per Definition 40 Ph. Daraus lässt sich folgende Umrechnungsformel erstellen: da die Steigerung der Lautheit um 10 Phon einer Erhöhung derselben auf ein zweifaches verursacht, so entspricht ein Lautheitsintervall von $\Phi_2 - \Phi_1$ Phon einem Lautheitsquotienten q_l von

$$q_l = s_2/s_1 = 2^{(\Phi_2 - \Phi_1)/10}$$

wobei Φ_1 und Φ_2 die Lautheitspegel in Phon, s_1 und s_2 die lineare Lautheit in Sone der das Intervall bildenden Töne sind. Per Definition gilt 1 Sn (für s_1 einzusetzen) = 40 Ph ($\rightarrow \Phi_1$), also lässt sich vereinfacht schreiben: $s = 2^{(\Phi - 40)/10}$ wobei Φ der Lautheitspegel in Phon und s die Lautheit in Sone sind.

Die umgekehrte Formel: $\Phi = 10(4 + (\hat{a}(s)/\hat{a}(2)))$

Beispiele:

Wieviel Sone sind 3 Phon? : $s = 2^{(3-40)/10} = 2^{-3,7} = 0,0769$ oder $1/13$ Sn (die Hörschwelle)

Wieviel Sone sind 100 Phon? : $s = 2^{(100-40)/10} = 2^6 = 64$ Sn

Wieviel Phon sind 100 Sone? : $\Phi = 10(4 + (\hat{a}(100)/\hat{a}(2))) = 10(4 + 6,64) = 106,4$ Ph

Laut allgemeiner Annahme entsprechen die Dynamikpegel *ppp*, *pp*, *p*, *mf*, *f*, *ff* und *fff* den Lautheitswerten 1, 2, 4, 8, 16, 32 und 64 Sn (40, 50, 60, 70, 80, 90, 100 Ph). Zur Summierungen von Lautheiten werden (mit Einschränkungen) die Sonewerte addiert, z.B. zwei 60 dB(SPL)-Rauschbänder um 44 und 1000 Hz, über deren ermittelten Lautheiten 30 und 60 Ph die entsprechenden linearen Werte 0,5 und 4 Sn sich zu 4,5 Sn addieren lassen.

Γ12b bietet eine Zusammenfassung aller bisher eingeführten linearen und logarithmischen Begriffe mit einem Versuch zur visuellen Verdeutlichung der gekrümmten Symmetrie der beiden Grundfunktionen $y=e^x$ und $y=\hat{a}(x)$.

10 Harmonik 3 – Die zweidimensionale klassische ‚reine Stimmung‘

Betrachten wir Γ_{13a} : jede waagerechte Reihe stellt einen offenen Quintenzirkel dar, jede senkrechte Spalte eine endlose Folge reiner Terzen. Im Herzen dieses Netzwerkes ist C (fett umrandet) als orientierender Kern gekennzeichnet, obwohl jeder andere Ton diese Rolle erfüllen könnte. In diesem Netz kann man sowohl pythagoreische als auch naturterzbezogene Intervalle wiederfinden, wie die rein pythagoreische C-Durskala: F-C-G-D-A-E-H (hellgraue Kästchen) oder die naturterzbezogene (manchmal stark vereinfacht *aristoxenisch*, nach Aristoxenos [350-300 v.Chr.] genannt, einem ihrer ersten Verfechter): A-E-H/F-C-G-D (dunkelgrau umfasste Kästchen). Aus obigem ist auch zu ersehen, dass das oktaven- und quintenbezogene pythagoreische System den höchsten beteiligten Primfaktor auf 3, das aristoxenische ihn auf die nächste Primzahl 5 setzt. Diese beiden Systeme werden daher als 3- bzw. 5-Limit-Tonsysteme bezeichnet.

An dieser Stelle erlaubt sich die berechtigte Frage: welche von den beiden ist ‚C-Dur‘? Es ist klar, dass die zwei Stimmungen und somit ihre Frequenzen verschieden sind: beim gemeinsamen C wären die Töne A, E und H bei der 5-Limit-Stimmung um ein syntonisches Komma (21,5 Ct) tiefer als bei der 3-Limit-Stimmung. Was hat das mit der verbreiteten *Gleichtemperierung* von Musikinstrumenten zu tun (wo alle Intervalle Vielfache eines einzigen Grundintervalls – des 100 Ct-Halbtons – sein sollen)?

Ein Beispiel: auf der Straße hören wir eine Marschmusik und wir versuchen, im Vorbeigehen nicht in den Takt der Musik zu geraten. Dies fällt schwer, weil obwohl das Schrittempo dem Musiktempo für nur einen Sekundenbruchteil wirklich gleich sein würde (präziser läuft wohl niemand), biegt der Zeitsinn die Schritt-Zeitpunkte zurecht, um alles für das Gehirn erfassbarer zu gestalten. Noch ein Beispiel: ein kleines Kind singt – tonhöhenmäßig verstimmt – ein uns unbekanntes aber offenbar in Dur komponiertes Lied: zu dieser letzten Erkenntnis gelangen wir durch die Fähigkeit, die schrägen Töne zurechtzuhören. Nachher würden wir die korrigierte Melodie auf einem Klavier wiedergeben können, wahrscheinlich mit Billigung des Sängers. Dennoch kommt einem die Temperierung eines sogar frischgestimmten Klaviers unmittelbar nach dem ausgiebigen Genuss mittelalterlicher Musik reichlich verstimmt vor, weil der Tonhörensinn durch diese Musik sensibilisiert und anspruchsvoller gestaltet wird.

Je nach momentaner musikalischer Sensibilisierung ist ein Zuhörer also unterschiedlich in der Lage, Tonhöhen in seiner bewussten oder unbewussten Vorstellung dahin zu biegen, wo sie den meisten musikalischen Sinn ergeben; dieser lässt auch temperiert gestimmte Musik vorstellungsmäßig in verschiedenen sogenannten ‚reinen‘ – d.h. durch ganzzahlige Frequenzverhältnisse darstellbaren – Stimmungsformen erscheinen. Das heißt: die Musik wird vom Gehirn rationalisiert (soweit sie diesem Vorgang vom Komponisten nicht bewusst entzogen wird: Schönberg gab sich ja viel Mühe, wirklich atonal‘ zu komponieren), jedes irrationale Frequenzverhältnis wird unbewusst in eines der quantitativ nahegelegenen rationalen umgewandelt. Die Harmonik ist das Studium von dem, was intervallisch gemeint ist oder zumindest verstanden wird.

Zu weiterer Veranschaulichung seien drei sprachwissenschaftliche Begriffe herangezogen: die Phonetik, die das tatsächlich Erklungene beschreibt, die Semantik, die versucht (ggf. durch Zurechthören) aus diesem einen Sinn zu gewinnen, sowie die Grammatik, die sowohl dem Verständnis des Gehörten als auch dessen (Verständnis bezweckender) Erschaffung dienen kann. Ein grammatisch effektiv eingesetztes Tonsystem kann semantisch begriffen werden, auch bei leichten ‚phonetischen‘ (sprachakzent-ähnlichen) Abweichungen von der Theorie. Ein Beispiel: Bach macht Sinn auch auf modern gestimmten Instrumenten. Knapp formuliert: Grammatik hilft Semantik bei Phonetik. Umgekehrt: Eine effektive semantische Deutung lässt aus ‚phonetisch‘ Rezipiertem heraus eine plausible Grammatik postulieren. Ein Diktat setzt Verständnis voraus; u.a. macht die weite Verbreitung tonaler Grammatiken ein tonales Diktat einfacher als ein atonales, ein 12-töniges einfacher als ein mikrotonales.

Es gibt zahlreiche Skalen, die anhand von 3-Limit-, 5-Limit- und noch weiter gehenden Erwägungen beschrieben und gestimmt werden können. $\Gamma 13b$ zeigt zwei weitere 5-Limit-Tonleiter, wovon wir zuerst die seit Jahrhunderten bekannte klassische europäische chromatische (hellgrau) betrachten möchten: die zwölf Töne bilden eine saubere Matrix aus drei Quintreihen mal vier Terzspalten. Auf dieser Basis sind die auf eine Oktav reduzierte, in steigende Reihenfolge gebrachte Frequenzverhältnisse wie folgt:

1:1 15:16 8:9 5:6 4:5 3:4 32:45 2:3 5:8 3:5 5:9 8:15

Selbst das klassische nordindische System der *Srutis*, 22 seit über zweitausend Jahren als umfassenden Tonhöhenvorrat musiktheoretisch beschriebene Intervalle, finden sich dunkelgrau umfasst in diesem Netz aus Quinten und Terzen wieder. Die leider verbreitete Annahme, es handle sich hierbei um 22 gleichtemperierte Intervalle (jedes also $1/22$ der Oktav = 54,5 Ct), gehört ins Reich der Mythen. Dass in der nordindischen Praxis auch andere Intervalle als diese 22 Verwendung finden, ist auch reichlich belegt.

Im besagten Netz sind mehrere gleichnamige Noten zu sehen; beispielsweise kommen in $\Gamma 13b$ die Töne C, E, F, G, A und H dreimal, die Töne D und B gar viermal vor. Ein Vergleich zeigt, dass jeder dieser Töne über vier Quinten aufwärts und eine Terz abwärts (oder anders herum) mit seinem nächsten Namensvetter verbunden ist, d.h. $3/2 \times 3/2 \times 3/2 \times 3/2 \times 4/5 = 81/80$, das Syntonische Komma von 21,5 Ct. Auch sogenannte, der Theorie bekannte *enharmonische Äquivalenzen* wie die *kleine Diesis* 125:128 ($-3\bar{T}$: 41,1 Ct), das *Diaschisma* 2025:2048 ($3 -4Q-2\bar{T}$: 19,6 Ct) und das 524288:531441(!) große *Pythagoreische Komma* (12Q-7 $-$ s. die mittlere Reihe von Ges bis Fis: 23,5 Ct) sind im Netz vorhanden. Das Diaschisma und die kleine Diesis kämen in einer Musik vor, in der Cis als Leitton zu D unweit vom Des als Leitton zu C fungieren – s. $\Gamma 13c_1$ & c_2 ; in reingestimmter Musik sind die Töne Cis und Des zwar enharmonisch ‚äquivalent‘ aber klanglich 19,6 Ct bzw. sogar 41,6 Ct auseinander.

11 Mathematik 5 – Von der Trigonometrie zur Analytischen Geometrie

Betrachten wir das Dreieck ABC in $\Gamma 14a$: der Winkel an der Spitze C ist rechtwinklig. Die Seite AB gegenüber diesem Winkel wird die *Hypotenuse* (griech. *hypotēnūsa* = ‚unter [dem rechten Winkel] spannen‘) genannt. Von jedem der zwei anderen nichtrechten Winkeln (an den Spitzen A und B) aus gesehen, gibt es außer der Hypotenuse noch zwei Seiten, eine gegenüberliegende (*Gegenkathete* [griech. *káthetos* = Senkrechte] genannt) und eine benachbarte (*Ankathete* genannt) – gegenüber der Spitze A, zum Beispiel, liegt die Gegenkathete BC, neben ihr die Ankathete AC.

Den Quotienten Gegenkathete/Hypotenuse nennt man den *Sinus* ([lat. = Biegung, Busen, Falte] abgekürzt β) des Winkels am Blickpunkt, d.h. $\beta(A) = BC/AB$, $\beta(B) = AC/AB$ (der Winkel an einer Spitze ‚X‘ wird auch ‚X‘ genannt). Den Quotienten Ankathete/Hypotenuse nennt man den *Cosinus* (lat. = ‚Ergänzung zum Sinus‘, abgekürzt α) des Winkels, so dass $\alpha(A) = AC/AB$, $\alpha(B) = BC/AB$ (dabei bleiben Sinus und Cosinus – als Absolutwert gesehen – immer zwischen 0 und 1, denn Gegen- und Ankathete können niemals größer als Hypotenuse sein). Es bleibt noch eine dritte Kombination, nämlich, Gegenkathete/Ankathete: diesen Quotienten nennt man den *Tangens* (vom lat. *tangere* = ‚berühren‘, abgekürzt $\acute{\alpha}$): $\acute{\alpha}(A) = BC/AC$, $\acute{\alpha}(B) = AC/BC$. In der Tat gleicht die Steigung m der Geraden $y=mx$ an einem beliebigen Punkt (x,y) dem Quotienten y/x und damit dem Tangens des Winkels, der die Gerade mit der x -Achse bildet – vgl. $\Gamma 03d$.

Die Dreiecksrechnung, auch *Trigonometrie* genannt (vom griech. *treís* = ‚drei‘ und *gōnos* = ‚angewinkelt‘), zeigt, dass der Sinus, Cosinus oder Tangens eines gegebenen Winkels stets gleich bleibt. Zum Beispiel wieder $\Gamma 14a$: sind die drei Seiten des Dreiecks AB, BC, CA im Längenverhältnis z.B. 5:4:3, so ist die wirkliche Größe des Dreiecks egal – bleibt das Verhältnis 5:4:3 konstant, so bleiben alle Proportionen und somit alle drei Winkel konstant. Hier gilt $\beta(A) = BC/AB = 4/5 = 0,8$. Daraus kann man A errechnen, den Winkel *Arcus Sinus* ([lat. *arcus* = ‚Bogen‘] kurz $\hat{\cup}$) von 0,8 – der Taschenrechner zeigt für A: $\hat{\cup}(0,8) = 53,13^\circ$. Ähnlich gilt $B = \hat{\cup}(AC/AB) = \hat{\cup}(0,6) = 36,87^\circ$. Es verwundert nicht, dass $A+B = 90^\circ$ ist, denn die Summe aller drei Winkel eines ebenen Dreiecks beträgt immer 180° .

Ist ein nichtrechter Winkel sowie die Länge einer Seite eines rechtwinkligen Dreiecks bekannt, so lassen sich alle anderen Merkmale dieses Dreiecks bestimmen. Wenden wir uns $\Gamma 14b$ zu – wir sehen einen Kreis, der drei rechtwinklige Dreiecke OMN, OPQ und ORS umschließt; die Hypotenusen dieser drei Dreiecke stellen den *Radius* (vom lat. für ‚Radspeiche‘ oder ‚Strahl‘) des Kreises dar und sind folglich gleich. Beträgt nun der durch die drei Punkte R, O und S umrissene Winkel ROS 60° , so können wir rechnen: $RS = RO \cdot \beta(ROS) = r \cdot \beta(60^\circ) = 0,866r$, wobei r der Kreisradius ist. Ist Winkel POQ 45° , so gilt $PQ = PO \cdot \sin(POQ) = r \cdot \beta(45^\circ) = 0,707r$. Gilt $MON=30^\circ$, so gilt $MN = MO \cdot \sin(MON) = r \cdot \beta(30^\circ) = 0,5r$. Da $MN < PQ < RS$ gilt, gilt auch $\beta(30^\circ) < \beta(45^\circ) < \beta(60^\circ)$.

Zeichnen wir eine Kurve auf einem zweidimensionalen Koordinatensystem, auf dem die x -Achse von links nach rechts einen kontinuierlich steigenden Winkel und die y -Achse den entsprechenden Wert von $\beta(x)$ (wobei $r=1$) darstellt, so entsteht die in $\Gamma 14c$ ersichtliche Sinuskurve: rechts vom Ursprung erstreckt sich etwa ein Viertel mehr als eine (einer Kreisumrundung in $\Gamma 14b$ entsprechende) *Periode* davon. Die drei Werte $\beta(30^\circ)$, $\beta(45^\circ)$, $\beta(60^\circ)$ sind als senkrechte Linien rechts vom Ursprung eingezeichnet – sie verhalten sich zueinander genau wie die Linien MN, PQ, und RS in $\Gamma 14b$.

Der Kreis-Sinuskurve-Bezug wird an der Spirale (s. $\Gamma 14e$) verdeutlicht: oben rechts (0°) zeigt deren Mittenachse direkt auf den Betrachter, dann wird sie (nach unten und später nach links) um 10° , 30° , 60° , 75° und 90° nach rechts gedreht. Die erste Abbildung (0°) formt einen Kreis, die letzte (90°) eine Sinuskurve.

In $\Gamma 14c$ ist die x -Achse nicht vornehmlich in Grad geeicht – bei 30° (eingeklammert) sieht man $\pi/6$, bei 45° $\pi/4$, bei 60° $\pi/3$, bei 180° $\pi/1$ usw. Die x -Achse ist in *Rad* (vom lat. *radius*), einem sehr häufig in der Mathematik anzutreffenden Bogenmaß kalibriert, wobei $1 \text{ Rad} = 180/\pi \text{ Grad} = 57,29577951^\circ$, der Winkel, der zwischen zwei Radien eines Kreises liegt, die am Kreisumfang einen radiuslangen Kreisabschnitt einfassen. $\Gamma 14d$ vergleicht diese Definition mit einem gleichschenkligen Dreieck mit radiuslangen Seiten und 60° -Winkeln. Da π der Quotient Kreisumfang/Kreisdurchmesser ist, lässt eine komplette Kreisumrundung um 360° eine 2π -fache Radiuslänge zurücklegen; also entspricht die Zurücklegung einer einzigen Radiuslänge dem Winkel $360^\circ/2\pi = 1 \text{ Rad}$.

Es ist nicht notwendig, den Dreiecksursprung der trigonometrischen Funktionen immer vor Augen zu halten; normalerweise reicht die kartesische Darstellung aus.

Hier sind einige der wichtigsten trigonometrischen Regeln:

$$\beta(x) / \grave{\alpha}(x) = \acute{\alpha}(x)$$

$$\beta^2(x) + \grave{\alpha}^2(x) = 1$$

$$\beta(\pi/2 - x) = \grave{\alpha}(x)$$

(...diese drei Gleichungen sind an der

$$\beta(0) = 0; \beta(\pi/2) = 1$$

Verschiebung der in $\Gamma 14c$ gezeigten

$$\grave{\alpha}(0) = 1; \grave{\alpha}(\pi/2) = 0$$

Sinus- und Cosinusperiode ersichtlich)

Die Regeln können am Dreieck in $\Gamma 14a$ leicht bewiesen werden, z.B.:

$$\beta(A) / \grave{\alpha}(A) = (BC/AB) / (AC/AB) = BC/AC = \acute{\alpha}(A).$$

Da $AB^2 = BC^2 + AC^2$ gilt (der Satz von Pythagoras), so gilt auch Folgendes:

$$\beta^2(A) + \grave{\alpha}^2(A) = (BC/AB)^2 + (AC/AB)^2 = (BC^2 + AC^2) / AB^2 = AB^2 / AB^2 = 1.$$

$$\beta(90^\circ - A) = \beta(B) = AC/AB = \grave{\alpha}(A).$$

Trigonometrische Funktionen können als Reihen dargestellt werden, z.B. (x in Rad):

$$\beta(x) = x - x^3/3! + x^5/5! - x^7/7! + \dots, \text{ wobei } n! \text{ (n-Fakultät)} = n(n-1)(n-2)\dots \times 3 \times 2 \times 1.$$

12 Akustik 4 – Schallwelle und Spektrum

Ein Gegenstand, der ein Geräusch von sich gibt, befindet sich in einem Zustand rapider Vibration, die an die Luft (oder ein sonstiges Zwischenmedium) weitergegeben wird; diese Vibration überträgt sich dann aufs Trommelfell des Ohres und gelangt so in das auditive System. Betrachten wir $\Gamma 15a$: ganz links unten stellt der fette Strich einen zuerst ruhenden festen Gegenstand, die dünnen Striche darüber ruhende Luftmoleküle dar (diese kommen eigentlich nur gegen den absoluten Temperatur-Nullpunkt -273° Celsius zur Ruhe, aber ihre gewöhnlichen Bewegungen, weil weder gerichtet noch periodisch, haben keine Hörrelevanz). Fängt der feste Gegenstand an, sich hin und her zu bewegen, so ändert sich der z.B. durch den Wärmezustand bedingte Normalabstand zwischen ihm und dem ihm nächsten Luftmolekül; letzteres ‚versucht‘, diesen Ausnahmezustand auszugleichen, indem es sich auch entsprechend hin und her verschiebt. Solange der feste Gegenstand schwingt, muss das Luftmolekül mitschwingen; erst nach dem Stillstand des Gegenstands kehrt beim zweiten die Ruhe wieder ein. Aber: aus den gleichen Gründen wird auch das nächste Molekül in Bewegung versetzt, dann das nächste usw. – eine Schallwelle entsteht, die sich mit einer Geschwindigkeit von – je nach Temperatur und Feuchtigkeit – etwa 320 bis 360 Metern pro Sekunde durch die Luft fortbewegt; $\Gamma 15a$ stellt einen Zeitverlauf (links bis rechts) von 6 Sekunden sowie eine Entfernung (unten bis oben) von 330 Metern dar – natürlich fällt damit diese Darstellung der Molekulardichte der Luft (nur 1 Molekül alle 10 Meter) extrem mager aus: unter gewöhnlichen Umständen gibt es etwa 2.7 Trilliarden Luftmoleküle pro Liter, einer eindimensional-linearen Meter-Dichte von ca. 300 Millionen durchschnittlich $\frac{1}{3}$ Nanometer großen, $\frac{1}{21}$ Trilliardstelgramm schweren Molekülen entsprechend. Die Fortbewegungsart der Schallwelle nennt man *longitudinal*, weil sie in die Richtung stattfindet, in der sich die Moleküle bewegen. Bei Wellen auf einer Wasserfläche zum Beispiel, bewegen sich die Teilchen an der Oberfläche im rechten Winkel zur Wellenausbreitung – diese Art wird *transversal* genannt; Licht- und andere elektromagnetische Wellen gehören zu dieser Kategorie.

In $\Gamma 15a$ zeigt ein Blick auf den zwischen 90 und 100 Meter befindlichen mittelgrauen Streifen einen in der Zeit variierenden Abstand zwischen zwei benachbarten Molekülen – das Verhalten dieser Variation ist, wie die Schwingungen selbst, sinusförmig. Da der Luftdruck zum Molekularabstand umgekehrt proportional ist, verändert sich auch dieser sinusartig; diese Druckveränderung führt zum Phänomen des Schalldrucks, der anhand der Formel $p=2\pi f\alpha cd$ errechnet werden kann, wobei p der Schalldruck, f die Frequenz der Sinuswelle, α der maximale Molekularabstand von der Mittellage (meistens ‚Amplitude‘ genannt – s.u.), c die Schallgeschwindigkeit im Schall tragenden Medium sowie d die Dichte des Mediums sind. In dieser Formel ist die Schallschnelle (die Geschwindigkeit des Moleküls beim Durchgang durch die Mittellage) mit $2\pi f\alpha$ vertreten, die damit p/cd gleicht. Alle Maßeinheiten (Pa, Hz usw.) lassen sich auf Kilogramm, Meter und Sekunde reduzieren, wie in $\Gamma 08b$ ersichtlich.

Zur Terminologie: die Bezeichnung ‚Amplitude‘ wird für ‚Molekularabstand von der Ruhelage‘ leider sehr häufig verwendet, die nur auf die Hüllkurve der Maxima einer Welle bezogen Sinn macht; um die Verwirrung zu vergrößern wird ‚Amplitude‘ auch noch häufig im Sinne des allgemeinen Schalldrucks benützt.

Ein Beispiel der Berechnung der Schallschnelle: bei einer Schallgeschwindigkeit von 333 Metern pro Sekunde sowie einer Luftdichte von 1,3 Gramm pro Liter entspringt ein Schalldruck von 60 dB(SPL) (0,02 Newton/Quadratmeter bzw. $\text{kg}\cdot\text{M}^{-1}\cdot\text{s}^{-2}$) eines 1-kHz-Sinustons einer Schallschnelle von $p/cd = 0,02 \text{ kg}\cdot\text{M}^{-1}\cdot\text{s}^{-2}/(1,3 \text{ kg}\cdot\text{m}^{-3}\times 333 \text{ M}\cdot\text{s}^{-1}) = 0,000046 \text{ m/s}$ oder ca. $1/20 \text{ mm}$ (der 150.000-fache Molekulardurchmesser) pro Sekunde.

Eine Schallwelle wird gewöhnlich als zweidimensionale Kurve des veränderlichen Luftdrucks oder des – unter gleichen Bedingungen proportionalen – Molekularabstands zur Mittellage gegenüber der Zeit dargestellt. Was passiert, wenn ein beweglicher Gegenstand gleichzeitig zwei unterschiedlich schnellen Sinusschwingungen unterliegt? Diese werden addiert, wie $\Gamma 15b$ zeigt: zwei gleich stark ausschlagende Sinusschwingungen im Frequenzverhältnis von 2:3 wirken zusammen als eine wie die Sinuskomponenten ebenfalls periodische, wenn auch kompliziertere Kurve.

In $\Gamma 15b$ oben rechts sind die zwei Sinuskomponenten auch als zwei parallele senkrechte Striche zu sehen – deren Frequenzverhältnis 2:3 (eine Quint) lässt sie als 2. und 3. Teilton eines *harmonischen Spektrums* darstellen – das Wort ‚harmonisch‘ bezieht sich auf das ganzzahlige Obertonverhältnis, das Wort ‚Spektrum‘ kommt vom lat. *spectrum* = ‚Erscheinung in der Vorstellung‘ (nach Newtons Licht-spektralen Experimenten). Die zwei Striche geben die hier gleich starke Amplitude, bzw. den Schalldruck der Töne an.

$\Gamma 15b$ veranschaulicht auch die Berechnung des Effektivwertes oder RMS, einer zu Bemessung des Schalldrucks verwendeten Größe, und zwar anhand der zwei gezeigten Sinuskomponenten sowie deren Summierung. Bei den zwei Komponenten wurden je 37 in regelmäßigen α -Abständen befindliche Sinuswerte ausgewählt (siehe die grauen Senkrechtlinien), deren Quadrate addiert und die Summe durch deren Anzahl 37 geteilt. Als Resultat wies jede der Kurven (beim willkürlich gesetzten Maximalwert 100) ein Durchschnittsquadrat von etwa 5000 auf, dessen Quadratwurzel – das RMS – 70,7 beträgt. Das den Schalldruck angegebende RMS der summierten Kurve (siehe die schwarzen Senkrechtlinien) beträgt 100 und impliziert damit eine durch die Schallwellensummierung bewirkte Schalldruckzunahme aufs $100/70,7 = 1,414 = \sqrt{2}$ -fache; dies entspräche einer damit verbundenen Zunahme der Schallintensität aufs 2-fache, welches zu erwarten ist: das gleichzeitige Erklängen zweier Schallquellen erwirkt generell die Summierung von deren Intensitäten. $\Gamma 15b$ zeigt zudem (links) einen hypothetischen ‚Molekularschnappschuss‘ der rechts abgebildeten Quintwelle: eine Teilchenwolke, deren Mitte die Schallquelle darstellt, zeigt Wellenmaxima als dichte, Wellenminima als dünne Ringe.

Dem Benutzer eines Computers fällt eine gewisse grundsätzliche Aufgefächertheit der Maschine auf (s. ¶16a): neben dem *Rechner* selbst gibt es zur Dateneingabe eine *Tastatur*, zur Datenkontrolle einen *Bildschirm* und – damit erarbeitete Informationen archiviert werden können – ein *Massenspeicher-Laufwerk* für die mechanische, magnetische oder optische und eventuell einen *Drucker* für die papierne Archivierung. Der Massenspeicher selbst hat zahlreiche Formen über die Jahre angenommen: u.a. Lochkarten und -streifen, Magnetbänder von der gewöhnlichen Tonbandkassette bis hin zum Spulenband, Disketten in Größen 3-, 5½- oder 8-Zoll (auch ‚Floppy-disk‘ genannt, vom engl. für ‚wackelig‘, nach der Beschaffenheit älterer Exemplare), bis zu mit Laserstrahlen les- und beschreibbaren Kunststoffscheiben oder sogenannten Wechsel- oder Festplatten.

Alle die obengenannten Geräte können unter dem Begriff *Hardware* (engl. f. ‚Eisenwaren‘) zusammengefasst werden. Damit sie aber funktionieren und untereinander kommunizieren, muss auch so genannte *Software* vorhanden sein, kodierte Informationen, die – einmal eingetippt – in der jeweils dafür vorgesehenen elektronischen Schaltung (dem *Chip*, engl. f. ‚Stückchen‘ oder ‚Splitter‘) residieren. Zusätzliche Software kann man kaufen: es handelt sich hierbei um fertige *Programme*, die vielleicht eine bequeme Textverarbeitung ermöglichen, Videospiele realisieren, oder – für den Musiker besonders interessant – spezielle Tongeneratoren klingen lassen oder Noten zeichnen. Käufliche Software ist leider meistens auf nur eine oder eine andere spezielle Computermarke gemünzt – die Anpassung eines Programms für einen ihm fremden Rechner ist äußerst schwierig.

Der Benutzer von Programmen befindet sich auf der ureigentlichen Ebene des Umgangs mit dem Computer. Wenn er aber selbst ein Programm schreiben will, muss er quasi eine Etage tiefer gehen (s. ¶16b): er muss seine Problemstellung in einer *Programmiersprache* formulieren (wie beispielsweise C, Pascal, Lisp, Basic, Assembler, Perl, Java usw.), das Programm mittels eines adäquaten *Editors* (vom lat. *ē+dare* = ‚aus-geben‘) in den Rechner eingeben, um es dann, mittels eines *Compilers* (vom lat. *com+pīla* = ‚Zusammen-legen‘, ursprünglich gar ‚-plündern‘) in *Maschinensprache* – die ‚eigene Sprache‘ des Rechnertyps – zu übersetzen. Oft gibt es einen Editor als integrierten Teil eines Compilers; da Fehler im Programm unvermeidlich sind, kann somit das häufige Hin und Her zwischen Editor und Compiler bequemer gestaltet werden. Im Falle von Sprachen wie Basic wird gewöhnlich jede Zeile des Programmtexts übersetzt und ausgeführt, direkt nachdem sie von einem *Interpreter* (ein ‚Simultanübersetzer‘, vom lat. *interpretarī* = ‚erklären‘ oder ‚übersetzen‘) gelesen worden ist. Der Vorteil ist, so kann man sehr viel direkter programmieren; der Nachteil: das interpretierte Programm läuft viel langsamer und unwirtschaftlicher als ein fertig kompiliertes.

Editor, Compiler und Interpreter sind auch Programme, die man installieren muss. Damit sie laufen, müssen sie dem im Rechner installierten *Betriebssystem* angepasst sein, ein Programm, das noch eine Ebene tiefer wirkt (Beispiele: CP/M, MS-DOS, RT-11, OSX, UNIX, Linux, Windows 98, 2000, XP, Vista usw). Das Betriebssystem ist es, das den normalen Umgang mit dem Computer ermöglicht – Daten können von einem Träger zum anderen kopiert oder sie können gedruckt werden; wichtiger noch, lauffähige Programme können gestartet und – wenn nötig – unterbrochen werden und wenn der Benutzer dem laufenden Programm Eingaben machen will, so nimmt das Betriebssystem diese entgegen und gibt sie ins Programm weiter. Für manchen Computer sind neuerdings mehr als ein Betriebssystem verfügbar; nach Bedarf kann vom einen zum anderen (nicht unumständlich) umgeschaltet werden.

Der Computer wird das Betriebssystem nur dann verstehen, wenn sein Hersteller entsprechende Grundsoftware (oft BIOS oder engl. *basic input/output system* genannt = ‚Grundsystem zur Ein-/Ausgabe‘) für den *Prozessor*, jenes ‚Hirn‘ im ‚Herzen‘ des Rechners, geschrieben hat (bekannte Prozessoren sind der Motorola 68000 Serie, die Intel 8086- und Pentium-Serien, usw). Hier befindet man sich auf der untersten Ebene der Software-Hierarchie.

Womit wir wieder bei der Hardware sind: neben dem Prozessor ist der Arbeitsspeicher RAM (engl. *random access memory* = ‚zufällsmäßig zugänglicher Speicher‘) von größter Wichtigkeit; der Prozessor kann an jede Ecke dieses Speichers blitzschnell gelangen, um dort eine Information zu erhalten oder abzulegen. Verglichen damit dauert das Lesen vom Massenspeicher (die ‚*Zugriffszeit*‘) – obwohl fast im Mikrosekundenbereich – eine kleine Ewigkeit; außerdem muss hier der Rechner oft einen ganzen Datensatz (‚*Datei*‘ genannt) abtasten bis er das findet, was er sucht.

Zusammenfassend lässt sich sagen:

Soll ein Computer betriebsbereit werden, so muss das Betriebssystem in den RAM geladen werden; nur dann können Programme laufen. Soll ein Programm geschrieben und getestet werden, so verfasst der Benutzer den Programmtext in einem Editor, wonach er den Text durch einen Compiler oder einen Interpreter in Maschinensprache übersetzen lässt – in dieser Form kann das Programm vom Betriebssystem gestartet werden. Die meisten Computermarken laden das Betriebssystem automatisch beim Einschalten.

Nennt man den Prozessor sinnbildlich das Hirn des Systems, so wäre der RAM das Gedächtnis, der Bildschirm der Gesichtsausdruck, das Laufwerk die Aktentasche. Dazu sollte noch ein Speicher genannt werden, ROM (engl. *read only memory* = ‚nur lesbarer Speicher‘) genannt, eine Art Hirnstamm, der für Programme wichtige, vorher vorbereitete und eingegebene Grunddaten enthält und daher nur gelesen, nicht beschrieben werden soll; es kann verschiedene Formen annehmen, die von eingebauten Chips bis zu CDs reichen.

Eine Musikaufnahme kann durch die (der Form der Schallwelle entsprechende) schwankende Magnetisierung von Eisenoxid (Tonband) oder Einritzung von Kunststoff (Schallplatte) bewerkstelligt werden – dies ist der so genannte analoge Aufnahmeweg; alternativ kann die Schallwelle als eine Folge von Zahlen dargestellt werden, die dann zu späterer Zeit in Stromschwankungen und dadurch in Schallwellen zurückversetzt werden können – das ist die digitale Lösung. Eine analoge Aufzeichnung hat gegenüber der digitalen gewisse Nachteile – die Magnetisierung eines Tonbandes kann mit der Zeit allmählich nachlassen oder sich auf benachbarte Windungen des Bandes als Vor- oder Nachecho durchdrücken; noch wesentlicher: das bei der Tonbandaufnahme entstandene leise Rauschen, auch das durch den Staub in den Rillen einer Schallplatte verursachte Knacken müssen mitangehört werden. Dagegen werden bei einer Digitalaufnahme lediglich die Zahlen, die die Abtastung der (vom Mikrophon weitergegebenen) Schallwellen ergibt, auf einer Vielzahl von möglichen Trägern gespeichert; eine etwaige Abnutzung von diesen Zahlen kann in der Regel nicht stattfinden – entweder werden sie gelöscht oder sie bleiben unverändert. Wenn durch eigentlich selten auftretende Materialfehler einige dieser Zahlen verfälscht werden, so fällt das Ergebnis eher als gänzlicher Verlust denn als Verschlechterung der Qualität auf. Die Hauptstörquelle bei digitaler Aufzeichnung liegt in der Auflösung der Kurvenabtastung – ist diese nicht hoch genug, ist die Verzerrung der Schallkurve hörbar.

Γ17a zeigt links eine Periode einer Sinuskurve in einer mit einer Auflösung von 8×8 erfolgten Abtastung, d.h. an acht gleichentfernten Stellen der x -Achse wird der auf eine der acht Zahlen 0 bis 7 auf- oder abgerundete y -Wert ermittelt. Diese Annäherung, als fette Linie zu sehen, ist offensichtlich um einiges gröber als eine mit der willkürlichen Auflösung 21×21 ermittelte (s. Γ17a Mitte). Bei 50×50 (Γ17a rechts) sieht die Kurve noch besser aus. Erst wenn die Auflösung etwa 500×500 erreicht, kann man relativ sicher sein, dass eine digitalisierte Welle von der ursprünglichen analogen akustisch kaum noch zu unterscheiden ist. Kommerziell käufliche Geräte zur Umwandlung von analogen Wellen in digitale und zurück, haben sogenannte *Analog-Digital-* bzw. *Digital-Analog-Umwandler* (kurz: ‚AD‘- bzw. ‚DA‘-Umwandler) inne – sie verfügen allgemein über eine Abtastrate von 44100, manchmal 48000 Hz und einen Zahlenumfang von -32768 bis +32767; bei einer Periode z.B. eines 20 Hz-Tons entspricht dies (bei 44100 Hz) einer Auflösung von 2205×65536 , bei 20000 Hz etwa nur 2×65536 – letztere Frequenz ist so hoch, dass die durch die Ungenauigkeit der niederen Auflösung entstehenden Obertöne unhörbar sind. Jede von einer Schallwelle derart abgetastete Zahl wird gewöhnlich ein *Sample*, das Verfahren *Sampling* genannt, und zwar aus dem Englischen für ‚Probe‘ oder ‚Muster‘, obwohl das Wort häufig auf missverständliche und bedauerliche Weise auch auf eine Klangprobe kurzer Dauer – also mehrere Samples enthaltend – angewandt wird.

Der Physiker Harry Nyquist (1889-1976) hat gezeigt, dass es für brauchbares Sampling einen Grenzwert der Relation zwischen Schallwellenfrequenz und Abtastrate gibt, und zwar 1:2; das heißt, dass für eine abzutastende Frequenz f die Abtastrate höher als das *Nyquist-Limit* $2f$ liegen sollte. $\Gamma 17c$ zeigt drei Sinuswellen, deren Frequenzen unter dem Limit N (bei $N/6$), am Limit bzw. über (bei $13N/6$) dem Limit liegen (bei der Abtastung 44100 Hz würden die drei Frequenzen 3675, 22050 [=N] und 47775 Hz betragen). Die Welle links mit Frequenz $N/6$ wird gut erfasst: die als Punkte gezeichneten Samples zeigen eine Linie, die der Vorlage sichtbar folgt. Jedoch zeigen die Samples der mittleren Welle mit Frequenz N ein falsches Bild: Nullwerte, Stille! Das rechts befindliche Beispiel mit Frequenz $13N/6$ zeigt ein Ergebnis, das jenem der $N/6$ -Welle gleicht: am Limit und höher liegende Frequenzen werden in den Samples fälschlicherweise als unter dem Limit befindliche wiedergegeben. Dieses Phänomen heißt *Aliasing* (vom engl. *alias* = ‚anders identifiziert‘) und wird durch das vorherige Wegfiltern von allen über dem Nyquist-Limit liegenden Frequenzen teilweise behoben.

Ob als analog-stufenlose Kurve oder als digitale Zahlenfolge gesehen, macht auf jeden Fall die Aufzeichnung einer Schallwelle eine Übertragung elektrischen Stroms von einem Gerät zu einem anderen notwendig. Dieser bildet bei der analogen Übertragung eines 1 kHz-Tons bei 0 dBu eine innerhalb $\pm 1,096$ Volt (V) schwankende Kurve mit RMS-Spannung $0,775 V = \sqrt{0,6} V$, die an einem Widerstand von 600 Ohm 1 Milliwatt Wechselstrom erzeugt. Bei digitaler Kodierung werden die abgetasteten Zahlen binär weitergegeben, wobei die Ziffer ‚0‘ 0 V und die Ziffer ‚1‘ etwa +5 V entspricht, ein Fall von Gleichstrom. In $\Gamma 17a$ sind die acht abgetasteten Zahlenwerte 4, 6, 7, 5, 2, 0, 1 und 4, binär ausgedrückt: 100, 110, 111, 101, 10, 0, 1 und 100. Für diese Zahlen braucht man nie mehr als drei binäre Stellen – es würde völlig reichen, die Folge als 100 110 111 101 010 000 001 100 auszugeben, vorausgesetzt, dass bei deren Rückverwandlung in Schall der DA-Umwandler weiß, dass es sich hier um dreistelligen Binärzahlen, um eine ‚3-Bit-Quantisierung‘ handelt: die Zahlen werden immer mit einer gleichbleibenden Stellenanzahl (hier 3) ein- und ausgegeben, egal ob sich Nullen zu Beginn einer jeweiligen befinden. Ein *Bit* (vom engl. *binary digit* = ‚binäre Ziffer‘), die kleinste Menge übertragbarer Information, enthält die einstellige Binärzahl 0 oder 1; die Daten (hier Bits) werden mit einer Geschwindigkeit übertragen, die man die *Baudrate* nach dem Ingenieur und Erfinder Jean-Maurice-Émile Baudot (1845-1903) nennt: 1 Baud = 1 Bit pro Sekunde bei der Bitübertragung. Die zur Ein- oder Ausgabe bestimmte Bitanzahl wird die *Bitbreite* genannt; die o.g. kommerziellen DA/AD-Wandler verwenden allgemein eine Bitbreite von 16. Werden die obigen acht 3-Bit-Zahlen mit einer willkürlich gewählten Geschwindigkeit von 2,4 Kilobaud befördert, so dauert die Übertragung 10 Millisekunden, genauso lang wie eine Periode der Frequenz 100 Hz. $\Gamma 17b$ vergleicht die analoge Stromübertragung dieser Periode (Sinuskurve) mit der digitalen (Rechteckskurve); in letzterer sind einige zusätzlich notwendige Kontrollbits nicht gezeigt worden. Die digitale Übertragung der Abtastung in $\Gamma 17a$ (Mitte) verlangt $21 \times 5 = 105$ Bits (statt der 24 für $\Gamma 17a$ links), denn die Darstellung der dazugehörenden Zahlen 16 bis 21 ist mit nicht weniger als 5 Bits möglich.

Damit ein Rechner seinen Zweck erfüllt, muss er programmiert werden; eine Folge von Anweisungen muss ihm eingegeben werden, wonach er Daten von einem Massen- oder sonstigen Speicher in seinen Arbeitsspeicher lädt, um sie dort zu verarbeiten und sie anschließend von dort an eine vom Programmierer bestimmte Stelle zu leiten. Diese Daten, an und für sich Binärzahlen, werden wegen besserer Leserlichkeit auch als gewöhnliche Dezimalzahlen, Buchstaben (zusammen *alphanumerische Zeichen* genannt) und Interpunktionszeichen kodiert. Bedingt durch die Bauweise elektronischer Komponenten ist die Standardbitbreite für die Datenübertragung 8; acht Bits nennt man ein *Byte* (vom engl. *binary term* = ‚binärer Ausdruck‘), eine Informationskapazität, die die Darstellung von $2^8=256$ verschiedene Zahlen (0_d bis 255_d) oder Zeichen ermöglicht. Höhere Zahlen müssen durch die Zusammenlegung von mehreren Bytes verarbeitet werden, z.B. die Zahl 12345, binär 11000000111001: hierfür reichen zwei Bytes, in die die Bits in der Form 00110000 00111001 gefüllt werden (hexadezimal 3039_h; diese vier Ziffern stellen jeweils vier Bits oder je ein Halbbyte oder *Nibble* dar). Den Bezeichnungen Bit, Nibble und Byte – als ‚bisschen‘, ‚Knabbern‘ und ‚Beissen‘ (Letzteres über das englische *bite*) wohl übersetzbar – ist ein gewisser Sinn für Humor nicht abzuspüren.

Alphanumerische und Interpunktionszeichen werden für üblich nach dem sogenannten *ASCII-Code* („American Standard Code for Information Interchange“ = engl. für „[US-] amerikanischer Standardcode für Informationsaustausch“) kodiert: darin enthalten sind alle Einbytezahlen, bei denen das erste (d.h. linke, obere) Halbbyte einen der sechs Werte 2_d bis 7_d (= 2_h bis 7_h) belegt – das erste Bit (ganz links, auch engl. *most significant* = ‚bedeutsamste‘ genannt) ist also immer 0; dieser (daher 7-Bit-) Code ist in T18 zu sehen. Die Werte 0_h bis 1_h als linkes Halbbyte entsprechen Sonderzeichen, die man auf der Computertastatur gewöhnlich durch das gleichzeitige Drücken der je nach Computermarke sogenannten Ctrl bzw. Control (die ‚Kontrolltaste‘), Command (engl. f. ‚Befehl‘) oder Strg („Steuerungs“)-Taste mit einer anderen (einem der 26 Buchstaben A-Z plus sechs weiteren Zeichen zugeordneten) Taste erzeugt; diese Aktion ähnelt der Eingabe von Großbuchstaben durch das gleichzeitige Drücken der Shift („Umschalt“)-Taste und den (Klein-) Buchstabentasten.

Auch spezielle Tasten, wie z.B. Esc, Tab, Backspace, Return erzeugen selbständig diese Sonderzeichen – die ersten zwei Bezeichnungen sind Abkürzungen für Escape (= engl. f. ‚Flucht‘, häufig zur Programmbeendigung eingesetzt – ASCII-Nr.27) und Tablature (= ‚Tabulatur‘, häufig als Tabellentrennzeichen genutzt – ASCII-Nr.9), während Backspace = ‚Schritt rückwärts‘ wörtlich zu nehmen ist (ASCII-Nr.8) und Return = ‚Rückkehr‘ oder ‚Rückgabe‘ (ASCII-Nr.13) sich historisch auf den Hebel bezieht, die die Schreibmaschinenwalze zurück in die Ausgangslage zu bewegen vermochte. Diese Taste wird auch mit Enter (= ‚Eingabetaste‘) bezeichnet.

Die linken Halbbyte-Werte δ_n bis f_n sind für weitere Sonderzeichen vorgesehen, die allgemein nur durch Programme erzeugt und vom Programmierer mit verschiedenen Funktionen versehen werden können (dazu gehören z.B. die Umlaut- und Eszet-Tasten deutschsprachiger Tastaturen). Nur der Programmierer, der bewusst *maschinennah* arbeiten will, um besondere Einrichtungen des Geräts besser auszunützen, müsste seine Anweisungen binär eingeben; um Tippfehlern wegen der vielen Nullen und Einsen vorzubeugen, darf er – wenn die Rechnergrundsoftware dies erlaubt – die Zahlen auch hexadezimal eintippen. Weiter hilft ein dem Rechner beim Kauf oft beigelegtes, *Assembler* genanntes Programm, das die Eingabe der Hexadezimal-Anweisungen in der Form einfacher Worte ermöglicht. Im Gegensatz dazu setzt der Programmierer, der den Rechner einfacher handhaben will, sogenannte *Hochsprachen* wie Basic, Fortran, C oder Pascal ein (um nur einige wenige zu nennen), deren Compiler seine in relativ leserlicher Form geschriebenen Anweisungen in für den Rechner sinnvolle Binärzahlen – die für den jeweiligen Rechnertyp spezifische *Maschinensprache* – übersetzen.

Zum Beispiel: zwei Zahlen, sagen wir 25 und 40, werden multipliziert und das Ergebnis wird einer sogenannten *Variablen* (einem Zwischenspeicher) *zugewiesen*, die unter einem vom Programmierer erfundenen Namen immer erreichbar sind. Nennen wir sie produkt; in Basic und Fortran schreibt man die Multiplikation produkt=25*40, in Pascal produkt:=25*40;, in C produkt=25*40;. Die Anzeigeanweisung in Basic und in Fortran lautet print produkt, in Pascal write(produkt);, in C (hier vereinfacht) printf(produkt); (= ‚schreibe produkt auf den Bildschirm‘ – es wird fast immer Englisch verwendet). Das Ergebnis erscheint: 1000, bei Basic sofort, bei Fortran, Pascal und C nachdem das compilierte Programm zu laufen anfängt. Man kann produkt umgehen, indem man print 25*40 bzw. write(25*40); bzw. printf(25*40); schreibt. So aber kann produkt für andere Zwecke verwendet werden, z.B. für einen Vergleich: das Ergebnis der Multiplikation soll nur dann angezeigt werden, wenn es 500 übersteigt. Die entsprechende Anweisung lautet

```

in Basic:  if produkt>500 then print produkt
in Fortran: if (produkt.gt.500) print produkt
            (.gt. heißt engl. greater than für ‚größer als‘, vergleichbar ‚>‘)
in Pascal:  if produkt>500 then write(produkt);
in C:      if (produkt>500) printf(produkt);
            (sinngemäß ‚ist produkt größer als 500 dann zeige produkt an‘).
```

Ein Programm kann eine Schleife enthalten, in der eine Anweisung oder ein Block mit Anweisungen mehrmals ausgeführt werden kann, z.B. die Berechnung der bekannten *Fibonacci-Reihe*, in der die ersten zwei Zahlen 1 sind und jede weitere Zahl die Summe der zwei vorangegangenen ist. Diese Aufgabe wird in den vier o.g. Sprachen einfach gelöst (das Ergebnis: 1, 1, 2, 3, 5, 8, 13, 21, 34, 55...): s. L19 (die Zuweisungen, typisch wie a = 1 geschrieben, werden hier ohne Leerzeichen kompakt gezeigt). Zuerst wird 1 den Variablen a und b zugewiesen; dann erhält wiederholt c den Wert von a+b, wonach c angezeigt wird und der Wert von b bzw. c auf a bzw. b verschoben wird.

1972 von Dennis Ritchie (geboren 1941) – als Weiterentwicklung der Sprache *B* von Ken Thompson – eingeführt, entwickelte sich *C* zu einer der am meisten gebrauchten Sprachen. Ihre Maschinennähe ermöglicht eine sehr effiziente Programmierung. Modular strukturiert, lässt sie ganze in sich geschlossene Anweisungsblöcke an strategischen Stellen in den Programmtext einschieben. Ihrer Rolle als Ausgangspunkt für bekannte Sprachen wie C++, C# und Java zum Trotz ist C nach all den Jahren (nicht zuletzt durch das American National Standards Institute) überall einheitlich geblieben.

Ein C-Programm besteht grundsätzlich aus einem Hauptteil (genannt `main` – vom engl. f. ‚haupt‘) und ggf. weiteren *Funktionen* vom Entwickler selbst. L20 ein komplettes C-Programm. Es rechnet – auf Formeln aus Kapitel 6 basierend – Zahlenverhältnisse (z.B. 2:3) in Cent und Dezibel um. Zu Beginn stehen die *Präprozessor-Direktiven* `#include <stdio.h>` und `#include <math.h>` – C ist eine sehr kompakt definierte Sprache, deren ureigener Korpus nicht einmal die schon beschriebene Anweisung `printf` enthält. Diese steht in der einzubindenden *Bibliothek* `stdio.h`, einer vorgefertigten Datei, kodiert. Der Präprozessor ist ein Compiler-Teil, der den Programmtext für die Compilierung im Vorfeld bearbeitet. `stdio` heißt ‚Standard Input/Output‘ (= ‚Standard-Ein-/Ausgabe‘), `math` ‚Mathematik‘, wo eine Reihe fester Standardfunktionen wie Logarithmen vordefiniert sind; `.h` steht für ‚header‘, weil die `include`-Direktiven (= ‚einbinden‘) am Kopf (= engl. ‚head‘) des Programms stehen.

Der folgende zwischen `/*` und `*/` als Erinnerungstütze für den Entwickler eingefügte Kommentar ‚Umrechnung von Zahlenverhältnissen...‘, wird vom Compiler übergangen. Man merke, in auszuführenden Programmteilen sind ‚ä/ö/ü/Ä/Ö/Ü/ß‘ nicht zugelassen, im Kommentar ist das aber kein Problem.

Dann folgt `main()`. Die bisweilen *Parameter* enthaltende Klammern zeigen, dass dieses Modul (und ggf. andere vorhandene) als wertannehmende bzw. -zurückgebende Funktion (wie Sinus oder Logarithmus) verstanden wird. Der zurückgegebene Wert ist ein Maß des Durchführungserfolgs, dessen *Typ* auch entsprechend deklariert wird, hier mit `int` für *integer* (engl. f. ‚ganzzahlig‘), daher: `int main()`. Mit der Rückgabe (engl. `return`) dieses Werts durch `main` ans Betriebssystem – günstigstenfalls Null (= ‚alles war in Ordnung‘) – endet das Programm ordnungsgemäß.

Der Inhalt von `main` wird in geschweiften Klammern `{` und `}` eingefasst; zur besseren Leserlichkeit werden sie (wie in L20 links) übereinander geschrieben. Zuerst werden Variablen deklariert, `p` und `q` als `int`, `nLog2`, `nLog10`, `nLog_Quotient`, `ct` und `db` als Fließkommazahlen mit Dezimalkomponenten hinter dem Komma (`float`, vom engl. *floating point* = ‚Fließkomma‘). Diese Variablennamen sind frei erfunden und dienen zur Speicherung der im Verhältnis `p:q` stehenden Zahlen `p` und `q`, sowie der natürlichen Logarithmen von 2, von 10 bzw. vom Quotienten der o.g. Zahlen und der am Ende zu errechnenden Cent- bzw. Dezibelwerte.

Zuerst errechnet die Standardfunktion `log(...)` die natürlichen Logarithmen von 2 und 10 und weist sie den Variablen `nLog2`, `nLog10` zu, z.B. `nLog2 = log(2)`; (bei wiederholtem Einsatz spart es Rechenzeit, Variablen statt Funktionen einzusetzen). Die nächste Zeile erbittet die Eingabe der nicht-negativen Ganzzahlen `p` (zuerst = 1) und `q` durch `printf("Gib P:Q ein.. usw., gefolgt vom Hinweis, dass eine ungültige Eingabe das Programm beendet. Diese der Übersichtlichkeit wegen auf zwei printf-Funktionen verteilten, mit jeweils einer in Hochkommata ("..") eingefassten anzuzeigenden Zeichenketten könnten auch Formatanweisungen (z.B. %d) enthalten – s. unten. Das Steuerzeichen \n (für ‚neu‘ wie in ‚neue Zeile‘) am Ende einer Zeichenkette bewirkt bei der Anzeige einen Übergang zur nächsten Zeile.`

Die folgende `while`-Schleife soll einen – in abstandsmäßig gleich rechtsversetzten geschweiften Klammern eingefassten – Anweisungsblock mit der Bedingung (`p>0`) wiederholen – nur *während* (= engl. *while*) `p>0` gilt, wird der Block ausgeführt. Das Programm schreibt dann `P:`, setzt `p` provisorisch auf 0 (s. unten warum) und wartet auf die Tastatur-Eingabe von `p` mittels der Lesefunktion `scanf("%d", &p)`, wobei das Formatzeichen `%d` (für ‚dezimal‘) die Erwartung einer ganzen Dezimalzahl für `p` anzeigt; gewöhnlich braucht `scanf` ein `&` vor jeder einzulesenden Variablen.

Wird `p` ungültig eingegeben, z.B. als -1, 0 oder `x`, so bleibt der vorher gesetzte `p`-Wert Null bestehen; damit ist die Bedingung `if (p>0)` (engl. *if* = ‚wenn‘) unerfüllt. So wird der folgende, auf nunmehr dritter Ebene in `{` und `}` eingefasste Anweisungsblock übersprungen; das Programm kehrt zur `while`-Abfrage zurück. Aber auch hier ist (`p>0`) falsch; die letzte Anweisung `printf("Programm beendet.\n");` wird also jetzt ausgeführt, das Programm gibt mit `return` eine 0 aus und endet.

Ist aber `p` tatsächlich `>0`, dann wird der Anweisungsblock doch ausgeführt: das Programm schreibt `Q:`, wartet auf die `q`-Eingabe durch `scanf`, berechnet dann den Quotienten-Logarithmus `log(1.0 * p / q)` und weist diesen `nLogQuotient` zu. Der Grund für `1.0`: wird in C eine ganze Zahl durch eine andere geteilt, so ist das Ergebnis auch ganzzahlig: `1/2` würde nicht `0.5` sondern `0` ergeben, welches zu falschen Resultaten führe. Die (höherrangige) Multiplikation von `p` mit dem `float`-Wert `1.0` ergibt einen `float`, welcher durch `q` geteilt einen `float` erzeugt. `ct` und `db`, nach den Formeln berechnet (`1200` bzw. `6.021`), zeigen zwei `printfs` an: bei `p=1` und `q=2` lässt `printf("----Das Verhaeltnis %d:%d entspricht ",p,q)` die Zeichen `----Das Verhaeltnis 1:2 entspricht` schreiben, und `printf("%9.3f Ct oder %6.3f dB\nGib P:Q ein..\n",ct,db);` die Werte `ct` und `db` im unten beschriebenen Format schreiben, gefolgt – zwischen zwei Zeilenumbrüchen – von dem Hinweis `Gib P:Q ein...` Die *Formatanweisung* `%d:%d` weist die Ausgabe von `p` und `q` als Ganzzahlen (z.B. `1:2`) an. `%9.3f` besagt, dass `ct` 9-stellig geschrieben wird, mit 3 Stellen hinter dem (im Deutschen dem Komma entsprechenden) Dezimalpunkt: `1200.000` (ist hier ein Leerzeichen). So bedeutet `%6.3f` die 6-stellige Ausgabe von `db` mit 3 Stellen hinter dem Punkt: `6.021`. Übrigens, das `f` in `printf` steht für ‚formatiert‘.

17 Informatik 5 – Allgemeine Funktionen in C

Außer `main` und Standardfunktionen wie `printf`, `scanf`, `log` und `while` können zusätzliche benutzerdefinierte Funktionen im Programm aufgerufen werden; sie werden gewöhnlich vor der Funktion angebracht, in der sie aufgerufen werden.

L21 zeigt ein Programm, das Tabelle T02 erzeugt. Fangen wir mit dem Hauptteil an. Der Deklaration von `main` und der Variablen `Zaehler` folgt jene der mit einer [8] versehenen Variablen `Ziffern` als Typ `char` (f. engl. *character* = ‚Zeichen‘): es ist oft besser, mehrere gleichtypige Variablen unter einem Namen indiziert zu führen, als sogenanntes *Feld* oder *Array* (engl. f. ‚Reihe‘). `char Ziffern[8]` fasst bis zu acht Zeichen zusammen, angefangen mit dem nullten: `Ziffern[0]` (in C wird von 0 gezählt, deklariert wird die Anzahl der Elemente). `char` besetzt ein Byte; wird mit von 0 bis 255 reichenden Ganzzahlen gearbeitet, so benötigen diese höchstens ein Byte und können als `char` platzsparend deklariert werden, wogegen `int` gewöhnlich 4 Bytes (32 Bits) besetzt und von -2147483648 ($= -2^{31}$) bis $+2147483648$ ($= +2^{31}-1$) reicht.

Variablen des Typs `FILE *` (stets groß geschrieben, engl. f. ‚Datei‘) machen Dateien zugänglich. Der Dateiname – hier `Zahlensysteme.txt` – gefolgt von einem `"r"` (engl. *read* = ‚lesen‘) oder `"w"` (engl. *write* = ‚schreiben‘) werden der Standardfunktion `fopen` eingegeben, die dann der Dateivariablen zugewiesen wird: `FILE *Datei; Datei = fopen("Zahlensysteme.txt", "w");`. Nun lässt die Variable `Datei` auf die durch sie vertretene Datei zugreifen (s. `fprintf` u.).

Neben `while` ist ein weiterer Schleifenkonstrukt `for`. Aber während sich `while` nur mit Bedingungen befasst, initialisiert und verändert `for` einen Zähler, in L21 so: `for (Zaehler = 0; Zaehler<=Endzahl; Zaehler++)`; . Erst steht `Zaehler` auf 0. Die Schleifenbedingung: `Zaehler` darf `Endzahl` nicht übersteigen, eine in Zeile 3 auf 255 festgesetzte Konstante. `Zaehler++` erhöht `Zaehler` um 1 (dem Ausdruck `Zaehler = Zaehler + 1` gleich: zum Wert von `Zaehler` addiere 1 und weise das Ergebnis nach links `Zaehler` zu, den vorherigen Wert ersetzend). Jede Wiederholung des `for`-Anweisungsblocks erhöht `Zaehler` also um 1, bis 255 erreicht wird, wonach die Schleife (wegen der dann verfehlten Bedingung) aufhört.

`fprintf` (das erste *f* steht für *file*) schreibt dann den `Zaehler`-Wert in `Datei` nach dem Format `%3dd = :` die Zeichen `[%3d]` schreiben eine dreistellige Ganzzahl vor, von `[d =]` gefolgt, wie z.B. `186d =` auf der zweiten Seite von T02 unten links (dort wurden Schriftart und Größe von Hand gesetzt, nicht programmiert).

Jetzt sind wir bei der Beschreibung einer Funktion (`verwandle`) angelangt. In jedem ihrer drei Aufrufen 5-7 Zeilen von unten stehen drei Parameter in Klammern, eine Zahl (2, 8 bzw. 16) und die Variablen `Zaehler` und `Ziffern`. Die Funktion selbst ist ab Zeile 9 zu sehen: sie gibt kein Wert zurück (kein `return` am Ende), daher ist ihr Typ `void` (engl. f. ‚nichtig‘). In Klammern stehen die `int` Variablen `Basis` und `Zahl` und die `char` Variable `*Code` (das `*` gibt den Wert von `Code` an `main` zurück – s.u.).

Die eingeklammerten Parameter sind mit jenen im Aufruf direkt verbunden: Basis nimmt nacheinander 2, 8 und 16 und Zahl den jeweiligen Zaehler-Wert (0 bis 255) an. Die *lokalen* (nur in `verwandle` geltenden) Variablen `i`, `Potenz`, `Puffer`, `Rest` und `Teiler` werden deklariert. `Potenz` wird 1, `Puffer` 255. Während `Potenz` wiederholt mit `Basis` multipliziert wird, wird `Puffer` durch `Basis` geteilt (so knapp wie bei `++` bedeutet `a*=b` dasselbe wie `a=a*b` und `a/=b` dasselbe wie `a=a/b`): ist `Basis` 2, so wird `Potenz` in 8 `while`-Durchläufen 2, 4, 8, 16, 32, 64, 128 und 256, `Puffer` parallel dazu 127, 63, 31, 15, 7, 3, 1 und 0 (merke: $255/2 = \text{int } 127$, nicht $\text{float } 127,5$): (`Puffer>0`) ist nun falsch, `while` endet und `Potenz` bleibt 256.

`Rest` wird `Zahl` (sprich `Zaehler` im Aufruf) zugewiesen, `Teiler` der Wert von `Potenz/Basis` (nach obigem $256/2 = 128$) und `i` 0. Jetzt folgt eine zweite `while`-Schleife, in der wiederholtermaßen Folgendes geschieht:

1. `Rest` wird durch `Teiler` geteilt und das Ergebnis `Puffer` zugewiesen,
 2. der Rest nach der Teilung von `Rest` durch `Teiler` wird `Rest` zugewiesen (% ist hier ein *Modulo-Operator*, %= wird wie *= u.ä. gehandhabt),
 3. `Teiler` wird durch `Basis` geteilt und das Ergebnis `Teiler` zugewiesen.
- Folgende Wertveränderungen werden in 8 `while`-Durchläufen vollzogen – am Ende des 8. Durchlaufs wird `Teiler` 0 womit (`Teiler>0`) falsch ist und `while` endet:

`Teiler`: 128 →64 →32 →16 →8 →4 →2 →1 →0

Ist `Rest` (sprich `Zahl` im Aufruf) z.B. 186, ist sie in den selben 8 Durchläufen:

$$\underline{186\%128} = \underline{58\%64} = \underline{58\%32} = \underline{26\%16} = \underline{10\%8} = \underline{2\%4} = \underline{2\%2} = \underline{0}$$

`Puffer`: $186/128=1$, $58/64=0$, $58/32=1$, $26/16=1$, $10/8=1$, $2/4=0$, $2/2=1$, $0/1=0$,

d.h. 10111010 – eine Binärzahl, die auf der `Basis` 2 berechnet wurde.

In der zweiten Zeile dieses zweiten Schleifenblocks von `verwandle` wird eine weitere benutzerdefinierte `char` Funktion aufgerufen: `Hexzeichen`, in Zeilen 4-8 definiert: sie empfängt den `int` Parameter `num` (Wert: 0 - 15), und gibt ein ASCII-Zeichen `num+48` (d.h. 0 bis 9) wenn `num<10` sonst `num+55` (A bis F) zurück. `Hexzeichen` liefert die Hexadezimaldarstellung von `num`, z.B. aus 8 wird 8, aus 12 wird C. Die `if-else` (engl. für ‚wenn‘-, ‚sonst‘)-Kontrollstruktur ist somit dargelegt.

In `main` werden dann die Binär-, Oktal- (`Basis` 8) und Hexadezimal- (`Basis` 16) Zahlen als Zeichenketten gespeichert: in den 8 Durchläufe der 2. `while`-Schleife von `verwandle` hat der die 8 Code-Elemente (vgl. Ziffern) indizierende Zähler `i` die Werte 1 bis 8 angenommen (s. `i++`;) – die von `Hexzeichen` gelieferten Zeichen werden dem `i`-ten Code-Element zugewiesen. Nach dem Schleifenende erhält `Code` noch das Schlusszeichen `\0` und geht an `Ziffern` im `verwandle`-Aufruf zurück. Dann schreibt `fprintf` die Kette `Ziffern` in `Datei` (im Format `%s` für engl. *string* = ‚Kette‘), gefolgt von [`b =`] bzw. [`o =`] bzw. [`h =`] für `Basis` = 2 bzw. 8 bzw. 16.

Zum Schluss folgen `fclose(Datei)` ;, das Gegenstück zu `fopen` (hiermit wird die `Datei` ‚versiegelt‘), eine Erfolgsmeldung am Bildschirm (mit `printf`) sowie die Rückgabe einer Null ans Betriebssystem, womit das Programm endet.

1983 beschloss eine Konferenz namhafter Synthesizer-Hersteller die Einführung einer einheitlichen Norm für die externe digitale Steuerung ihrer Klangmodule: *MIDI*, die *Musical Instruments Digital Interface* (engl. f. ‚Digitale Musikinstrumenten-Schnittstelle‘). Diese Norm, die eine relativ einfache hardwaremäßige Schnittstelle voraussetzt, ermöglicht die Kommunikation zwischen verschiedenen Musikgeräten: die Tastatur eines mit MIDI ausgerüsteten Klangmoduls kann die Klanggeneratoren eines anderen ansprechen, auch ein Rechner kann eingreifen und diese in Echtzeit spielen. Nach außen hin besitzt die MIDI-Schnittstelle gewöhnlich drei fünfpolige, mit der Aufschrift MIDI-IN, -OUT und -THRU (= engl. f. ‚ein‘, ‚aus‘ und ‚durch‘) versehene DIN-Kupplungen – die ersten zwei sind für den Empfang bzw. Ausgabe von MIDI-Signalen, die letzte, eine Durchgangskupplung zur Weitergabe von für ein anderes Gerät bestimmten Signalen, erlaubt die Bildung eines Verbunds mehrerer Apparate (s. T22 für ein Beispiel). Die in den MIDI-Signalen einkodierte Kanalnummer bewirkt, dass alle folgenden Signale für den momentan auf diese Nummer eingestellten Klanggenerator bestimmt sind; wenn also die Signale mehrerer Kanäle durch alle Klangmodule durchrasen, spielt jeder Generator nur ‚seine‘ Stimme. Die Signale laufen über ein mit zwei fünfpoligen DIN-Steckern versehenes abgeschirmtes zweiadriges Kabel, und zwar seriell (d.h. als Bitfolge) mit einer Übertragungsrate von 31250 Bits, d.h. 3906 Bytes pro Sekunde.

Will man einen Ton einschalten, so schickt man dem Generator das *Note-Ein*-Signal, eine einbytige Hexadezimalzahl, deren zwei Halbbytes 9_h bzw. die Kanalnummer 0_h bis F_h sind – also besagt die Zahl 90_h , dass im ‚nullten Generator‘ ein Ton gleich gespielt wird. Dieses erste Byte wird das *Status-Byte* genannt, dessen erstes Halbbyte je nach Anweisungsart zwischen 8_h und F_h liegt (d.h. das erste, bedeutsamste Bit ist immer 1). Danach folgt unmittelbar die Tonhöhenangabe, das erste von zwei *Daten-Bytes*, deren erstes Bit immer auf 0 steht.. Alle verfügbaren Töne, 128 an der Zahl, gehören der zehneinhalboktavigen chromatischen Tonleiter vom (Sub-Subkontra-)C₃ bis zum g⁶ an, und zwar von 0 bis 127_d nummeriert. Der Ton c' entspricht 60_d; dadurch sind die Nummern aller gleichnamigen Noten C_n durch 12 teilbar. Bevor der Ton gespielt wird, muss auch seine Anschlagskraft angegeben werden, und zwar ebenfalls in einem Byte gepackt als 7-Bit-Zahl zwischen 0 und 127_d. Die Anweisung ‚spiele ein c' mittlerer Lautstärke in Kanal null‘ sieht also so aus (dezimal): 144 60 64.

Wie beim Anspielen eines Tons wird auch das Ausschalten oder Dämpfen des Tons mit drei Bytes bewerkstelligt – 1. das Signal Note-Aus, $8m_h$ (m = MIDI-Kanalnummer), 2. die *Notennummer* und 3. die *Dämpfungsgeschwindigkeit* – allerdings können nur wenige Klangmodule auf diese letzte Angabe reagieren. Die Bytefolge $8F_h$ 60_d 127_d bedeutet ‚Dämpfe raschestens das c' in Kanal F_h ‘. Da jeder Ton nach dieser Definition mindestens sechs Bytes zur Ein- und Ausschaltung braucht, kann MIDI theoretisch höchstens $3906/6 = 651$ Töne pro Sekunde übertragen, unbeschadet der Tatsache, dass gängige Klangmodule gerade einmal 200 pro Sekunde zu schaffen scheinen.

Ein Akkord wird durch das sequentielle Einschalten der Einzeltöne gespielt – die 1½ Millisekunden zwischen einem Note-Ein und dem nächsten ist bei einem kleineren Akkord kurz genug, damit dieser nicht als Arpeggio gehört wird. Durch die Folge Note-Ein, Notenummer, Null kann auch gedämpft werden – hier wird im sparsamen *Running Mode* (engl. f. *laufender Modus*) das Status-Byte nur zu Beginn eingesetzt, wonach paarweise nur noch Daten-Bytes folgen, um den Notenstrom zu beschleunigen.

Zwei wichtige MIDI-Anweisungen sind *Instrumentenwahl* (engl. *program change*) – 2 Bytes: Cm_h und i (die Klangfarbe als Zahl von 0 bis 127_d oder mehr, je nach Klangmodul) und *Tonhöhenradzustand* (engl. *pitch wheel*) zur mikrointervallischen Verschiebung aller Tonhöhen des Kanals – 3 Bytes: ϵm_h , 0 und v (die Verschiebung als Zahl zwischen 0 und 127_d , mit 64_d als Normalzustand; der Umfang wird durch eine firmenspezifische Systemanweisung auf 0 bis ± 12 Halbtöne gesetzt – bei ± 1 Halbton ist die kleinste durch das 2. Datenbyte ermöglichte Verschiebung $^{100}/_{64}$ oder unter 1,6 Ct).

Zuletzt sei hier die vielseitige *Steueränderungs-Anweisung* (engl. *control change*) erläutert: mit dem Status-Byte βm_h beginnend, gibt das erste Daten-Byte die Anwendungsart durch die *Kontrollnummer* an, wie z.B. 7_d für Gesamtlautstärke (engl. *volume* – dies erlaubt ein *crescendo/decrescendo* eines anhaltenen Klanges) oder 64_d für ein simuliertes Dämpferpedal; außer diesen beiden (die Werte laufen von 0 bis 127_d) sind manche standard, andere firmenabhängig – sie können im Handbuch zum Klangmodul nachgelesen werden. Das zweite Daten-Byte gibt den Wert der Steuerung an: bei Kontrollnummer 7_d bedeutet 0 ‚lautlos‘, bei Nr. 64_d ‚ohne Pedal‘; 127_d ist auch hier der Maximalwert. T23 listet die acht Hauptanweisungsarten auf (die mit A, D und F beginnenden sind etwas spezieller in ihrer Art und hier nicht kommentiert worden).

Alle oben beschriebenen Signale werden von einem manuell gespielten Klangmodul direkt erzeugt; sie werden hier zwecks ihrer Ausgabe durch einen Rechner besprochen. L24a zeigt ein C-Programm, das für die Ausgabe der ersten Töne von Schumanns *Fröhlicher Landmann* über ein MIDI-Klangmodul gedacht ist; beispielsweise schicken `spiele-Aufrufe` in `main` Tonhöhe und Anschlagsstärke an die `spiele-Funktion`, wo sie `Tonhoehe` und `Anschlag` heißen; normalerweise müsste `spiele` über die Funktion `sende` drei Bytes zur MIDI-Steckdose schicken (Hexadezimalzahlen werden in C mit `0x` gekennzeichnet) – aber die genaue Form von `sende` ist vom Rechnertyp abhängig; also werden hier die Werte nur am Bildschirm angezeigt. `warte` prüft die Rechneruhr mittels (in `<time.h>` kodierte) `CLOCKS_PER_SEC` und `clock()` und wartet die in Sekunden angegebene Zeit ab. L24b ist kompakter: `Partitur = fopen(DATEINAME, "r");` öffnet die Datei `FROEHLICHER_LANDMANN.TXT` (durch die Präprozessor-Direktive `#define` als `DATEINAME` definiert) zum lesen. Die Schleife `while` mit der Funktion `Ausfuehren` läuft solange diese den Wert 1 (= ‚wahr‘) an `main` zurückgibt, d.h. solange die Eingabemenge (`s.fscanf`) 3 ist; ist sie es nicht, so wie beim Ende der Eingabedatei, so endet `Ausfuehren`. Das `,==‘` in `Ausfuehren` ist ein *Vergleichsoperator*, anders als der Zuweisungsoperator `,=‘`. Zeile 4 (mit `schreibe`) zeigt einen skurrilen typografischen Einsatz von `#define`.

19 Harmonik 4 – Zur quantitativen Erfassung von Harmonizität

Seit der Antike sagt man, dass zwei Töne, deren Frequenzen im einfachen Verhältnis zueinander stehen, ein ‚harmonisches‘ Intervall bilden. Doch ein ‚unharmonisches‘, das sich größtmäßig in der Nähe eines harmonischen Intervalls befindet, gerät sozusagen in den Sog des stärkeren und wirkt durch Zurechthören als dessen Annäherung.

Ist die Einfachheit eines Zahlenverhältnisses quantitativ auszudrücken? Die Verhältnisse 1:2 (Oktav), 2:3 (Quint), 3:4 (Quart), 4:5 (große Terz) oder 5:6 (kleine Terz) scheinen ‚einfacher‘ als 8:9 (großer Ganzton), 9:10 (kleiner Ganzton) und diese wiederum einfacher als 15:16 (kleine Sekunde), 32:45 (übermäßige Quart) oder 45:64 (verminderte Quint). Man merke: je harmonischer die Intervalle, desto kleiner sind die Zahlen, die sie bilden; aber ist die 3:5-große Sext harmonischer als die 4:5-große Terz?

Den oben fehlenden Verhältnissen 6:7 und 7:8 sowie 10:11, 11:12, 12:13, 13:14 und 14:15 sind keine klassischen Intervalle zuzuordnen – die Zahlen 8 und 9 wurden der kleineren 7 stets vorgezogen, die 15 und 16 den Zahlen 11, 13 und 14. Man merke: die historisch vorgezogenen Zahlen basieren auf den Primzahlen 2, 3 und 5, die anderen auf den höheren 7, 11 und 13. Zur Bildung eines harmonischen Intervalls ist sowohl die Kleinheit der Verhältniszahlen als auch ihre Teilbarkeit relevant. Zur Messung der *Harmonizität* vereinte ich 1978 diese beiden Eigenschaften in meiner sogenannten *Unverdaulichkeitsfunktion* $\xi(N)$; s.F25a. Die Potenz 2 in der Formel beeinflusst die Steilheit des Anstiegs der Unverdaulichkeit bei den Primzahlen, quasi als Maß der ‚Primfeindlichkeit‘. Praktisch ist, dass $\xi(a \cdot b) = \xi(a) + \xi(b)$ gilt, wie bei Logarithmen.

T26a zeigt die Unverdaulichkeiten der Zahlen 1 bis 100 – in steigender Reihenfolge ihrer Unverdaulichkeit aufgestellt, entsteht bei den Zahlen 1 bis 16 die Folge 1 2 4 3 8 6 16 12 9 5 10 15 7 14 11 13: die letzten vier sind besagte Aussätzige. Aus der umgekehrten Summe der Unverdaulichkeiten der teilerfremden Zahlen P und Q ließ sich eine Funktion für die Harmonizität des Intervalls Q/P bilden: je unverdaulicher P und Q, desto weniger harmonisch das Intervall – s. F25b.

1980 fragte ich zwölf Freunde nach der von ihnen einzeln empfundenen Reihenfolge der Schwierigkeit, eine Kreisform in 2 – 9 gleiche, vom Kreismittelpunkt ausgehende Segmente zu teilen; ihre Antworten im Durchschnitt: 2 4 (3 8) 6 (5 9) 7 (umklammerte Zahlen sind umtauschbar), mit der Unverdaulichkeit unübersehbar übereinstimmend. Mir unbekannt wurden 1975 Testpersonen an der Universität Stanford gebeten, die Ähnlichkeit der Ziffern 0 bis 9 nach mehreren Kriterien, darunter ‚abstrakte Qualität‘, zu bewerten (s. T27a); die Ergebnisse wurden dann *multidimensional skaliert*, eine Methode, die die Ziffern derart räumlich platziert, dass deren gegenseitige Abstände deren jeweiligen ‚Unähnlichkeiten‘ entsprechen. Man fand eine deutliche Trennung der geraden Zahlen von den ungeraden, der Primzahlen von den zusammengesetzten; auch wurden die Zahlen von links nach rechts größer. Und 1980 entdeckte ich entlang der Primzahlgrenze eine der Unverdaulichkeit verwandten Ordnung.

T26c zeigt eine Auflistung aller reingestimmten Intervalle innerhalb einer Oktav mit einer Harmonizität von 0,05 und mehr; manchmal auftretende Minuszeichen deuten eine Polarität des Intervalls zum oberen Ton an, der dann als Grundton wirkt. In der Tabelle steht neben jedem Intervallverhältnis dessen Primzerlegung in Potenzen der Primfaktoren; bei dieser *Mindestharmonizität* (MH) von 0,05 sind die für die Primzahlen 2, 3, 5 und 7 absoluten Grenzpotenzen 7, 4, 2 und 1, die *Höchstpotenzfolge* (HPF); höhere Primzahlen kommen nicht vor und haben daher die Potenz 0. Die Anzahl der in dieser einen Oktav zulässigen Intervalle beträgt 38. Bei einem Umfang von drei Oktaven und MH 0,04 sind bei einer 8 5 2 1 1-HPF 240 Intervalle zu verzeichnen, die T27b grafisch darstellt. T26b zeigt HPF und Intervallanzahl bei anderen MH-Werten beim willkürlichen Umfang von einer Oktav.

Nach dem Astronom Daniel Kirkwood (1814-1895) entstammen die Lücken im Asteroidengürtel einfachen Zahlenverhältnissen („Kommensurabilität“) zwischen dort zu erwartenden Umlaufzeiten und jener Jupiters um die Sonne: früher dort befindliches Material wurde aufgrund entstehender Resonanzen vom Jupiter angezogen und in eine langgestreckte Laufbahn gezwängt und schließlich von der Sonne zerstört. T27c zeigt einen Vergleich zwischen der Dichte des Gürtels und der Harmonizität besagter Umlaufzeitintervalle (bei der geringeren „Primfeindlichkeit“ 1,2) – eine allgemeine Übereinstimmung ist deutlich.

Wie zu erwarten, reduziert eine Anhebung der MH die Intervalldichte (die MH 0,1065 lässt nur noch eine mixolydische Skala durch!) und umgekehrt. Es gibt einen Zusammenhang zwischen der HPF und der entsprechenden MH – s. F25c; man bedenke, dass eine HPF Intervalle einschließt, deren Harmonizitäten auch unter dem entsprechenden Mindestwert liegen – sie garantiert nur, dass alle Intervalle harmonischer als der Mindestwert durch sie erfasst werden. Bei der einer MH von 0,03 entsprechenden HPF 12 8 3 2 1 1 gibt es immerhin 3.964 verschiedene Intervalle in einer Oktav(!), von denen nur 211 harmonischer als 0,03 sind. Die HPF der allgemein akzeptierten Intervallstimmung klassischer abendländischer (und indischer) Theoretiker ist 9 6 2 0 0..., während die Intervallbildung der Antike um Pythagoras von 9 6 0 0 0... ausging. Beide Folgen sehen interessanterweise wie zaghafte Ansätze zu einer auf einer 0,04-MH basierenden 9 6 2 1 1 1 0... aus. Höhere Primzahlen sind Harmonizitätshemmer; deswegen haben Theoretiker sie immer mit großer Vorsicht genossen – zur Intervalldeutung zogen sie früher monströse Konglomerationen kleinerer Primzahlen der vielleicht eleganteren Lösung kleinerer Produkte größerer Primzahlen vor (vgl. den 3-Limit-Tritonus 512:729 mit dem 5-Limit-Tritonus 32:45).

Auf die 3- und 5-Limit-Systeme folgend sind die nächstkomplexen die *7-Limit*-, *11-Limit-Systeme* usw., wobei es nicht plausibel ist, überzeugende Musik bei einem beliebig hoch gesetzten Prim-Limit zu erwarten ohne eine passende, wohlüberlegte Musikgrammatik für das daraus wachsende Tonsystem zu entwickeln.

Von einer Mindestharmonizität von 0,02 ausgehend, finden wir im Halbtonbereich 550 bis 650 Cent nicht weniger als 256 Intervalle vor (s. $\Gamma 28a_1$); darunter befinden sich als harmonischste die Verhältnisse 5:7 und 32:45, und (knapp schwächer) deren Umkehrungen 7:10 und 45:64, allesamt Intervalle, die als Tritonus gelten können. Es geht jetzt weiter um die Frage, welche dieser und anderer benachbarter Varianten verstanden wird, wenn ein Intervall von 600 Cent intoniert wird, eine Frage, die ohne eine genaue Inspizierung des musikalischen Zusammenhangs, in dem dies geschieht, nicht beantwortet werden kann.

600 Cent besitzen als Intervall keine eindeutige harmonische Bedeutung, genau wie ‚Summ-Pflicht‘, welches, je nach Zusammenhang, auch ‚Sumpf-Licht‘ heißen kann. Ein in Cent oder Halbtönen beschriebenes Intervall ist bloß eine Abstandsangabe. Erst durch das Zurechthören im Zusammenhang erhält das Intervall einen harmonischen Sinn. Um dies vorzuführen sei empfohlen, die Töne C D F E (Mozart) und Cis His E Dis (Bach) auf einem Klavier zu spielen und die Wirkung der großen Terz C-E mit jener der verminderten Quart His-E zu vergleichen (s. $\Gamma 28b_1$). Die deutlich unterschiedliche Spannung entspringt einem im Hörer vollzogenen, mehr oder weniger unbewussten Zurechtbiegen der Töne auf eine wunschmäßig vorgestellte Optimalstimmung.

Wenn, von einem C ausgehend, die um 3, 5 und 6 Halbtöne darüber liegenden Töne nacheinander erklingen, so fällt es nicht schwer, diese als Es, F und Ges zu deuten. Wenn aber die Töne in der Reihenfolge 0, 3, 7 und 6 Halbtöne über C vorkommen, suggeriert sich C, Es, G, Fis wie in $\Gamma 28b_2$ gezeigt. Die vier Töne der ersten Folge wären möglicherweise als 1:1, 5:6, 3:4, 45:64, die der zweiten als 1:1, 5:6, 2:3, 32:45 harmonisch verständlich, und zwar nach den in $\Gamma 28b_2$, als – bis auf jene zu Fis und Ges – relativ einfache Verhältnisse eingezeichneten Querverbindungen der vier Töne untereinander. Doch wirkt Fis (auf 32:45, 590 Ct berechnet) klanglich etwas schärfer, höher, als das rechnerisch höher vermutete Ges (45:64, 610 Ct), auch wenn die zwei Folgen auf einem gleichtemperierten Klavier gespielt werden – dies kommt wahrscheinlich dadurch, dass diese zwei Töne sogar ‚zeitlich rückwärts‘ eher als Leitdenn als Strukturtöne fungieren, das Ges zu F, das Fis zu G: sie werden als Tonabstände (‚melodisch‘) gehört und mikrotonal näher an ihre Verankerungen intoniert (das ‚Leittonphänomen‘). So ist in diesem Mischsystem die Stimmung von den melodisch gehörten Fis und Ges weniger zwingend relevant als die von den eher harmonisch gehörten C, Es, F, und G. Dennoch wird im Folgenden nur vom harmonischen Aspekt ausgegangen; die Relevanz der einzelnen Wertungen bleibt hier unberücksichtigt.

Zur Erlangung eines harmonischen Einblicks in eine irrationale Tonhöhenmenge, deren Intervalle größenmäßig bekannt aber verhältnismäßig unbekannt sind, müssen diese also *rationalisiert* werden, d.h. Centangaben müssen in Brüche umgewandelt werden. Zu kompositorischen Zwecken ging ich 1978 diese Aufgabe an.

Meine Lösung geht wie folgt vor: zuerst werden alle zulässigen *Stimmalternativen* der Töne ermittelt; die Auswahl hängt von zwei Hauptfaktoren ab, der *Mindestharmonizität* (MH) und der *Stimmtoleranz* (ST). Bei harmonisch komplexerer Musik ist die MH tiefer anzusetzen als bei harmonisch einfacherer; eine geringere Sensibilisierung hat eine höhere ST zur Folge. Aus einer gewählten MH wird nach der Formel in F25c eine Höchstpotenzfolge (HPF) errechnet, die zur Erzeugung einer (wie jener in T28a₁ abgebildeten) Intervallliste eingesetzt wird. Es ist möglich, die HPF willkürlich zu bestimmen, wie z.B. 0, 0, 0, 5, 3, 0, ..., eine Stimmung, die nur auf 7 und 11 basiert; eine solche Folge ist mit kompositorischer Überzeugung jedoch sehr schwer einsetzbar und wirkt unter Umständen als eine verstimmt Darbietung einer anderen Stimmung, die kleineren Primzahlen entsprungen ist.

Zur Anwendung der ST wird eine Gauss'sche Normalverteilungsglocke auf die zu stimmende Stelle gesetzt: diese nach deren Erfinder Johann Karl Friedrich Gauss (1777-1855) genannte Kurve, deren Breite (Varianz) zur angegebenen Toleranz proportional ist, dämpft – tonhöhenmäßig nach oben und unten zunehmend – die Harmonizitätswerte wie in T28a₂. Daraus entfallen als Stimmalternativen alle Intervalle, die von der Zentralstelle zu weit oder harmonisch zu schwach sind, um sich zu behaupten. Als *Nenntoleranz* (NT) gilt willkürlich der Tonhöhenabstand nach oben und unten, an dem die Dämpfung 20 beträgt. T28c zeigt diesen Dämpfungsvorgang bei der Stimmung einer in Cent angegebenen Durskala; die NT wurde hier auf 50 Ct festgesetzt, der Hälfte des kleinsten Abstands zwischen zwei benachbarten Skalenstufen.

Der nächste Schritt ist die Bestimmung der *Anzahl der Stimmalternativen* für jeden Ton; dann wird die Summe der *harmonischen Intensitäten* (Absolutwert der Harmonizität) aller intraskalaren Intervalle unter Einbeziehung aller Stimmalternativen ermittelt – dabei sind Stimmalternativen eines jeden Tones von einem Vergleich untereinander ausgeschlossen: die Wahl fällt auf jene Stimmkonstellation, bei der die Harmonizitätssumme am höchsten ist. Ist die Anzahl der zu stimmenden Tonhöhen beispielsweise 8 und die Anzahl deren jeweiligen Stimmalternativen 3, so ist die Anzahl der Stimmkonstellationen $3^8=6561$; eine 8-Ton-Skala hat 28 intraskalare Intervalle inne (die Formel lautet $n(n-1)/2$, wobei n die Tonanzahl ist), also müssen hierbei insgesamt $6561 \times 28 = 183708$ Harmonizitätswerte addiert werden, um die Optimalstimmung zu ermitteln, eine Aufgabe für einen Computer! Dabei stellte es sich als programmiertechnisch effizienter heraus, nicht die Harmonizitäten zu addieren, sondern deren Umkehrungen (jeweils die Summe zweier Unverdaulichkeiten) wobei dann die kleinste Gesamtsumme „prämiert“ wird. Die Gesamtzahl der intraskalaren Intervalle $n(n-1)/2$ (s.o.) geteilt durch die Kleinstsumme der Unverdaulichkeiten nenne ich die *spezifische Harmonizität* der Optimalstimmung der Tonhöhenmenge.

21 Harmonik 6 – Tonhöhenrationalisierung an praktischen Beispielen

Bei der Stimmung der einoktavigen Durskala ergibt sich bei einer Mindestharmonizität (MH) von z.B. 0,04 und einer Nenntoleranz (NT) von z.B. 50 Ct folgende vier für am harmonischsten befundene Stimmalternativen (SA) für jede der 8 Tonhöhen

1/1	9/8	5/4	4/3	3/2	5/3	15/8	2/1
80/81	10/9	81/64	27/20	40/27	27/16	243/128	81/40
81/80	8/7	56/45	21/16	243/160	128/75	256/135	160/81
63/64	35/32	32/25	48/35	32/21	12/7	40/21	63/32

Nach 6561maliger Prüfung aller 28 Intervallverbindungen der 3 besten Stufenkandidaten ($3^8=6561$) wählt der Rechner die Stimmkombination 1/1 9/8 5/4 4/3 3/2 5/3 15/8 2/1, eine seit sehr langem bekannte und allgemein akzeptierte Lösung – s. T13a. Die spezifische Harmonizität dieser Stimmung beträgt 0,2252.

Bei der harmonischen Mollskala wird unter gleichen Bedingungen folgende, ebenfalls geläufige Stimmkonstellation für optimal befunden: 1/1 9/8 6/5 4/3 3/2 8/5 15/8 2/1. Hier beträgt die spezifische Harmonizität 0,2032.

Die einoktavige pentatonische Skala (C D E G A c von C aus) wird wie folgt festgelegt: 1:1 9:10 4:5 45:64 5:8 9:16 1:2 (spezifische Harmonizität 0,1615).

Interessant ist die sogenannte Bohlen-Pierce-Skala (BP-Skala), bei der die reine Duodezime (1:3 oder 1902 Cent) in 13 gleiche Intervalle geteilt wird. Die Centwerte der einzelnen Stufen sind also 0 146 293 439 585 732 878 1024 1170 1317 1463 1609 1756 1902. Eine Rationalisierung dieser Skala ergibt – je nach MH und NT – verschiedene Stimmungen (es wurden aus Rechnerzeitgründen nur zwei Stimmalternativen untersucht). MH 0,03–0,04 und NT 40–50 ergab 1:1 32:35 5:6 7:9 32:45 2:3 3:5 5:9 1:2 7:15 32:75 2:5 9:25 3:1 bei einer spezifischen Harmonizität von 0,12. Die Abweichungen von den Sollwerten betragen hierbei 0 +9 +23 -4 +5 -30 +6 -6 +30 +2 +12 -23 +13 und 0 Cent. Die MH 0,05 und NT 0,01–0,03 zeigt Schwierigkeiten bei der Erteilung der Brüche: auf Grund des – durch den Konflikt zwischen hoher Harmonizitätsschwelle und niedriger Toleranz entstehenden – Mangels an Möglichkeiten werden die 1. und 2. Stufe gleichgestimmt, wie auch die 8. und 9. Erst bei NT 0,04 werden eindeutige Rationalisierungen ermittelt, wobei die Centabweichungen erwartungsgemäß im Schnitt noch höher liegen. Die gängige, aus 3:5:7-Ketten zusammengesetzte Reinstimmung der BP-Skala ist 1:1 25:27 21:25 7:9 5:7 49:75 3:5 5:9 25:49 7:15 3:7 25:63 9:25 1:3. Obwohl hier die Centabweichungen allgemein deutlich niedriger liegen als bei der obigen Rationalisierung, wurde jene wegen der intrastufigen Stimmigkeit gewählt.

Die bisherigen Ergebnisse können in größerem Detail in T29a ersehen werden. In T29b werden die Zahlenverhältnisse sowie Harmonizitäten der 78 Intervallverbindungen zwischen den 13 Stufen einer zwölfstemperierten Oktav gezeigt; diese Beststimmung entspringt folgenden Voraussetzungen – MH: 0,04, NT: 40 Cent, SA: 2. Diese Methode wird auch an den 14 Stufen einer in dreizehn gleiche Intervalle geteilten Oktav angewandt. Gleiche MH (0,04), NT (40 Cent – s. auch den letzten Absatz unten) und SA (2) ergeben ein Stimmungsnetz, das sich auch in T29b befindet.

Vergleiche unter den Ergebnissen unterschiedlicher MH und NT wurden auch an diesen zwei gleichstemperierten Skalen vorgenommen, wie in T30 veranschaulicht. Angestrebt war eine Art Landschaft, in der die Relevanz einer Stimmlösung durch deren Verbreitung überschaubar gemacht werden könnte. Zu diesem Zwecke entwickelte ich eine kompakte graphische Darstellungsmethode für Rationalisierung sowie Centgröße, *Ratioglyphen*, die das Schema in T30a (anhand von 2:3, 15:16 und 32:45 sowie 16:21, 128:243 und 24:35) erklärt: die obere Hälfte stellt die Primpotenzen an einem ‚Stamm‘ dar, an dem die Primzahlen aufwärts angeknüpft sind, deren Potenzen als horizontale ‚Astlänge‘ nach links (für negative Potenzen) und umgekehrt angegeben sind. Bei Überlänge darf der Ast platzsparend in die Vertikale ausweichen. Die untere Hälfte gibt die Intervallgröße in Cent an, und zwar erst in Hunderten und dann – als Abweichungen von diesen – in Zehnern und Einern. Dadurch erhält jedes gestimmte Intervall durch seine Ratioglyphe eine prägnante, erlernbare Gestalt.

Dazu entwickelte ich mit gängigen Zeichen eine Schreibweise für Notennamen, aus der die intervallische Herkunft der daran haftenden Töne ersichtlich wird: eine Zahl rechts gibt die Quintenanzahl (z.B. ³ oder ₃ für nach oben oder unten), ein Akzent die Terzen (˘ nach oben, ˘ nach unten), ein Fragezeichen die Septim an (? nach oben, ¿ nach unten – s. Folgendes). Der Ton ¿f¹ ist daher ()+Q-T-S (oder $2/1 \times 3/2 \times 4/5 \times 4/7 = 48/35$ oder 547 Ct) über C (S meint die *Naturseptim* 4:7). Bei mehreren Terzen bzw. Septimen wird das Zeichen entsprechend oft wiederholt, z.B. ˘˘˘ oder ??.

Zurück zu den Ratioglyphen: bei beiden Skalen ragt jeweils eine besondere Stimmlösung hervor, die sich durch ihre Verbreitung als die plausibelste empfiehlt: in T30 eingerahmt, gilt sie bei der 12-Ton-Skala von NT=40 bis 60, bei der 13-Ton-Skala um NT=40 (hier scheint eine höhere NT unangebracht; letztendlich würden bei genügender NT alle Töne zur Oktav und Quint bestimmt!). Die multidimensionale Skalierung (MDS) dieser beiden Optimalstimmungen zeigt jeweils ein Feld, in dem die Töne nach Harmonizität eingeteilt sind – je näher zwei Noten zueinander sind, desto harmonischer ist das zwischen ihnen liegende Intervall. Diese MDS-Felder können als ‚Landkarten zur harmonischen Modulation‘ betrachtet und für n-Limit-Systeme angelegt werden (merke z.B. die familiäre Gruppierung der ?- bzw. ¿-Töne in der MDS von T30b).

22 Metrik 1 – Zur Quantifizierung von Metrik

Die rhythmische Organisation von Musik zeigt im Allgemeinen eine mehr oder minder starke Anlehnung an ein innerlich hierarchisches *Metrum*, ein Raster von regelmäßigen aber akustisch unterschiedlich aktivierbaren Zeitpunkten. Die *Schichtung* eines Metrums, dessen innere akustische Hierarchie multiplikativer Art ist, wird am Beispiel eines **12/16**-Takts in T31a veranschaulicht: diese Taktart weist die Schichtung $2 \times 2 \times 3$ auf. Untersucht man sie in weiterer Tiefe über 32tel, 64tel usw., so ließe sich die angegebene geometrische Folge mit einer stetigen Zweiteilung fortsetzen: $\dots \times 2 \times 2 \times 2 \times 2 \dots$; die Stelle der letzten Zahl der Folge, die nicht 2 ist, bestimmt die *Ordnung* des Metrums – demzufolge ist **12/16** ein Metrum dritter, **6/8** [= 2×3] zweiter, **3/4** [= $3(\times 2 \times 2 \dots)$] erster und **4/4** nullter Ordnung.

Unabhängig von der Ordnung des Metrums zeigen dessen einzelne aufeinanderfolgende Pulse auf jeder Ebene der Schichtung eine unterschiedliche metrische Relevanz auf. Auf seiner ersten Ebene weist ein **12/16**-Takt zwei Schläge als punktierte Viertelnoten auf, einen metrisch stärkeren gefolgt von einem metrisch schwächeren. Bei den folgenden Ebenen mangelt es an sprachlichen Deskriptoren genügender Differenziertheit – die Auswertung der Pulsstärke der schon auf der zweiten Ebene vorhandenen vier punktierten Achtelnoten macht die Einführung von Zahlen nötig (hier bedeutet ‚Schlag‘ dasselbe wie die allgemeine Bezeichnung ‚Puls‘ nur auf den obersten zwei Ebenen).

Die unterschiedliche akustische Aktivität der Pulse eines Metrums – im einfachsten Falle werden sie angeschlagen oder ausgelassen – erzeugt einen das Metrum mehr oder weniger unterstützenden Rhythmus, d.h. die *metrische Feldstärke* kann dadurch erhöht oder gesenkt werden, wobei der Rhythmus zwischen metrisch (mit klar erkennbarem Grundschlag) und ametrisch (ohne erkennbaren Grundschlag) variiert. Dies kann durch die Erfassung und Einsetzung der Pulsstärke derart erfolgen, dass zur Unterstützung des Metrums die stärkeren Pulse häufiger als die schwächeren vorkommen; in einem ametrischen Rhythmus wären Pulse aller Stärken gleich häufig. Analogerweise kann in einer Tonhöhenkala die *tonale Feldstärke* oder ‚Tonalitätsgrad‘ beeinflusst werden, und zwar durch die Häufigkeit der Skalentöne, basierend auf der Harmonizität der Intervalle zwischen den Tönen und einem beliebig zu benennenden Grundton oder ‚Tonika‘. T31c stellt die Relation Feldstärke ↔ Häufigkeit der Elemente (Pulse eines Metrums, Töne einer Skala) als Gerade variabler Steigung dar.

Es bleibt noch, auf die – genügend differenzierte – Erfassung der Pulsstärke einzugehen: dies impliziert die Ermittlung einer einmaligen Wertung für jeden Puls des Metrums, so dass keine zwei Pulse als gleich stark empfunden werden. Diese Differenziertheit vermisste ich bei diversen Ende der 1970er Jahre vorliegenden Systemen und nahm mir daher vor, eine eigene passende Methode zu suchen.

1978 konnte ich eine Formel entwickeln, die für jeden Puls auf jeder Ebene eines multiplikativen Metrums jeglicher Ordnung eine Pulswertung (*Unverzichtbarkeit*, wie ich sie nenne) errechnen lässt; die Werte nach dieser Formel für die sechs Achtelnoten eines **3/4**-Takts sind 5 0 3 1 4 2, die für einen **6/8**-Takt sind 5 0 2 4 1 3. Beide Zahlenfolgen zeigen die stärkste Betonung auf der ersten Achtelnote; die zweitstärkste liegt bei **6/8** auf der vierten, bei **3/4** auf der fünften Achtelnote. Hier sind die Unverzichtbarkeiten der zwölf Sechzehntelnoten der dritten Schichtungsebene der Takte **3/4**, **6/8** und **12/16**:

3/4	(3×2×2):	11	0	6	3	9	1	7	4	10	2	8	5
6/8	(2×3×2):	11	0	6	2	8	4	10	1	7	3	9	5
12/16	(2×2×3):	11	0	4	8	2	6	10	1	5	9	3	7

Auch die Werte für die erste und zweite Ebene sind hierin enthalten: um sie ersichtlich werden zu lassen, subtrahiere man von der Unverzichtbarkeit die Differenz zwischen der Pulsanzahl auf der hier gezeigten und jener auf der erwünschten Ebene und behalte nur die positiven Zahlen, z.B. bei **3/4** (Pulsanzahl der gezeigten 3. Ebene = 12):

- 1. Ebene (Pulsanzahl 3: subtrahiere 12-3 = 9): 2 - - - 0 - - - 1 - - -
- 2. Ebene (Pulsanzahl 6: subtrahiere 12-6 = 6): 5 - 0 - 3 - 1 - 4 - 2 -

Man merkt: auf allen Ebenen ist die Unverzichtbarkeit des ersten Pulses stets eins weniger als die Pulsanzahl, die des zweiten stets null. Die Relevanz der Unverzichtbarkeitswerte kann durch einen sogenannten Verdünnungsprozess veranschaulicht werden: wenn die ‚verzichtbareren‘ Anschläge auf der anfangs anschlagsmäßig gesättigten α -ten Ebene eines Taktes nach und nach entfernt werden, so bleibt der Eindruck der Taktart immer gut erhalten, wie z.B. in der vorletzten Zeile der in F31b gezeigten Takte **3/4** und **6/8**.

Die Unverzichtbarkeitsformel ist für sukzessive Unterteilungen eines Taktes durch beliebige Primzahlfaktoren bestimmt. Voraussetzung bei jedem dieser Faktoren ist eine Folge von *Grundunverzichtbarkeiten* für eine Taktebene mit der entsprechenden Primanzahl von Pulsen, z.B. für 2 die Folge 1-0, für 3 die Folge 2-0-1, für 5 die Folge 4-0-3-1-2 usw. Bei einem Takt mit der Schichtung $p_1 \times p_2 \times p_3 \times \dots \times p_n$ ist die Unverzichtbarkeit $\Psi(n)$ des n -ten Pulses durch die Formel in F32a gegeben (die Folge der Primzahl-Grundunverzichtbarkeit ist durch die Funktion $\Psi(x)$ angezeigt, wobei x dem in runden Klammern gefassten Teil der Formel entspricht).

Die Primzahl-Grundunverzichtbarkeiten Ψ können nach eigenem Ermessen oder durch die Formel in F32b erstellt werden; Grundlage für ihre Bemessung ist hier die Folge der Unverzichtbarkeit für eine Taktebene mit einem Puls weniger, in der die obersten Teiler die größten sind (z.B.: um Ψ für einen Zyklus von 23 Pulsen zu ermitteln, werden zuerst die Werte nach der Ψ -Formel für 22 Pulse errechnet, wobei 22 als 11×2 dargelegt wird; die Werte für 11 Pulse basieren auf denjenigen für 10, als 5×2 zu verstehen, usw.).

23 Metrik 2 – Zur Messung metrischer Kohärenz

Bei der Suche nach einer Methode zur Bestimmung der Ähnlichkeit zwischen zwei Metren unterschiedlicher Schichtung und Geschwindigkeit konnte ich die Ermittlung der Unverzichtbarkeit befriedigend einsetzen. Zum Beispiel wird ein $2 \times 2 \times 3$ -Metrum im Ganztakt-Tempo \dot{a} 20 mit einem 3×5 -Metrum im Ganztakt-Tempo \dot{a} 16 verglichen: die Kleinstpuls-Tempi sind mit \dot{a} 240 gleich; auch sind 5 Takte des ersten Metrums gleich lang wie 4 Takte des zweiten (jeweils 60 Kleinstpulse). Also lege man, Puls für Puls, eine fünfmal durchgenommene Unverzichtbarkeitsfolge für $2 \times 2 \times 3$ Pulse einer viermal durchgenommenen Folge für 3×5 gegenüber, wie in T33a abgebildet. Wenn die beiden Metren keine gleichschnellen Kleinstpulse aufweisen, so müssen deren Schichtungen durch weitere Teiler solange in die Tiefe fortgesetzt werden, bis dies der Fall ist; z.B. werden ein 2×5 -Metrum bei Takttempo \dot{a} 50 (Kleinstpulstempo \dot{a} 500) und ein 3×2 -Metrum bei \dot{a} 60 (Kleinstpulstempo \dot{a} 360) auf $2 \times 5 \times 3 \times 3 \times 3 \times 2$ bzw. $3 \times 2 \times 5 \times 5$ erweitert, weil das Kleinstpulstempo beider Metren nun gleichermaßen \dot{a} 9000 beträgt. Auch hier werden den größeren Teilern die oberen Ebenen zugeteilt und die kleineren nach hinten geschoben (z.B. die Erweiterung des 2×5 -Metrum ist $\times 3 \times 3 \times 3 \times 2$, nicht $\times 2 \times 3 \times 3 \times 3$ oder $\times 3 \times 3 \times 2 \times 3$).

Die beste Methode, die ich zur Ermittlung metrischer Ähnlichkeit finden konnte, war, für jeden Puls die nach einem der Metren bemessene *relative Unverzichtbarkeit* (Nennwert geteilt durch Höchstwert, d.h. durch Pulsanzahl minus eins) mit derjenigen, die nach dem anderen bemessen wird, zu multiplizieren und die Summe aller Produkte festzustellen; Zusammenfälle stärkerer Pulse wirken begünstigend auf die Gesamtsumme, und zwar noch besser, wenn die Produkte der relativen Unverzichtbarkeit noch vor der Addition quadriert werden. Es stellt sich heraus, dass das durchschnittliche Produktquadrat (nennen wir es ρ) größer ist für offensichtlich verwandtere Metren (z.B. für 2×3 und 3×2 in gleichem Takttempo: 0,3245) als für unverwandtere (z.B. für $2 \times 2 \times 3 / \dot{a}$ 20 und $3 \times 5 / \dot{a}$ 16: 0,1573). Bei diesem Verfahren passiert es, dass bei den unverwandtesten Metren das ρ gegen $1/9$ tendiert (es stellt sich heraus, dass das ρ aller Paare von Zahlen kleiner oder gleich N ist $N^4/9 + N^3/3 + 13N^2/36 + N/6 + 1$); also tendiert $9\rho - 1$ im genannten Fall gegen null. Mit der halben negativen Umkehrung des natürlichen Logarithmus von $(9\rho - 1)/3,5$ (3,5 ist die höchstmögliche $9\rho - 1$ zwischen zwei 2×2 Metren gleichen Tempos) habe ich schließlich eine Skalierung gefunden, die mir am übersichtlichsten und am überzeugendsten wirkt; dieser Koeffizient ist in der Formel in F34 definiert.

T35a ist eine Tabelle der errechneten metrischen Ähnlichkeiten der vier dreischichtigen Metren bis zu dritter Ordnung, deren Pulsanzahl 12 nicht übersteigt ($2 \times 2 \times 2$, $2 \times 2 \times 3$, $2 \times 3 \times 2$, $3 \times 2 \times 2$); die Tempoverhältnisse sind hier durch Kombinationen der Ganzzahlen 1 bis 3 bestimmt. Ein Blick auf den ersten Wert in dieser Tabelle – 0,46382 als metrische Ähnlichkeit zwischen einem $2 \times 2 \times 2$ -Metrum und sich selbst – zeigt, dass das Wort ‚Ähnlichkeit‘ ein Problem aufwirft; hier sollte sich sicherlich eine Identität oder Gleichheit vorfinden, die mit dem Wert 1,0 plausibler auszudrücken wäre. Die Identität

zwischen $[3 \times 2 \times 2]$ und sich selbst wird (in derselben Tabelle) jedoch mit 0,41454 benotet – man sieht hierdurch, dass diese Bewertung im Allgemeinen auch die metrische Einfachheit berücksichtigt – $[2 \times 2 \times 2]$ ist ja einfacher als $[3 \times 2 \times 2]$. Aus diesem Grund habe ich beschlossen, diese die Zusammengehörigkeit zweier Metren und jene interne Zusammengehörigkeit von deren Pulsen verbindende Eigenschaft *metrische Kohärenz* zu nennen. Die eigentliche ‚Ähnlichkeit‘ ist nur bei einem Reihenversuch zu ermitteln, wobei eines der Metren allen Vergleichen gemeinsam bleibt: die bei der Gegenüberstellung dieses festen Metrums mit einem anderen Metrum errechnete Kohärenz geteilt durch die Kohärenz des festen Metrums mit sich selbst (die *Autokohärenz*) ergibt die metrische Ähnlichkeit; so beträgt diese bei zwei gleichen Metren 1, der Wert der *Identität*.

Zu der Zeit, als ich die Formel für metrische Unverzichtbarkeit entwickelte, habe ich auch eine Formel für die Harmonizität von Tonhöhenintervallen erstellt, die jedem rationalen Frequenzverhältnis $P:Q$ einen Koeffizienten für dessen Harmonizität zuordnet. Bei harmonisch ‚starken‘ Intervallen wie der Oktav 1:2 und Quint 2:3 ist dieser Wert höher als bei ‚schwächeren‘ wie dem Tritonus 32:45. Würde man eine hörbare Tonhöhe als eine äußerst rapide Folge von Pulsen betrachten, deren Tempo durch die Frequenz angegeben wird (ich nenne dies die *Frhythquenz*, die ‚Rhythmus-Frequenz‘), so wäre die Harmonizität eine Art ‚mikrometrische Kohärenz‘! Die Tabelle in T35b zeigt die metrische Kohärenz von 32 verschiedenen Taktpaaren gleichen Takttempos; zum Vergleich sind die Harmonizitäten der entsprechenden Tonhöhenintervalle angeführt. Die zwischen ihnen befindliche Parallelität wird in grafischer Form noch ersichtlicher – s. T33b. So wirft sich die Frage auf: ‚Ist die Harmonik ein Spezialfall der Polymetrik?‘.

Diese Bewandnis regte mich zur Anwendung der Methode der Tonhöhenrationalisierung auch auf die rhythmische an. Ein dem Rechner aus einer vorliegenden Partitur per Hand durch eine MIDI-Tastatur eingespielter, in Millisekunden gespeicherter Rhythmus (bedacht waren Anfangszeitpunkte der angeschlagenen Noten) wurde in Centwerte wie folgt umgerechnet: eine willkürlich gewählte Zeiteinheit wurde einem willkürlichen Centwert zugeordnet; nach der Formel zur Umrechnung von Frequenzquotienten in Cent wurden jede gemessene *Deltzeit* (die Zeitdifferenz zwischen einem Ereignis und dem folgenden) in Cent umwandelt – so wurde eine Zeit, die doppelt so lang war wie die Grundzeiteinheit durch -1200 Ct, eine halb so lange durch $+1200$ Ct vertreten. Die so entstandene Tonhöhenmenge wurde dann nach der Harmonizitätsmethode (mit den Parametern MH, NT sowie Stimmalternativen) rationalisiert und dann in Rhythmus zurückumwandelt. Die Ergebnisse waren durchaus zufriedenstellend: sogenannte N-tolen (wie Triolen und Quintolen) sowie punktierte Rhythmen wurden an deren von der Partitur vorgesehenen Stellen gebracht, obwohl die Einspielung alles andere als präzise war.

24 Mathematik 6 – Stochastische Analyse und Synthese

Als häufigster Buchstabe der deutschen Sprache gilt E – durchschnittlich soll es 18% aller Buchstaben ausmachen. Danach folgen N mit etwa 11% sowie l, R, S, A und T mit von 6–8% Häufigkeit. Die verbleibenden Buchstaben sind seltener, mit J, Y, X und Q als Schlusslichter (Umlaute, Eszet und Leerzeichen werden hier nicht berücksichtigt). Andere Sprachen weisen andere Häufigkeiten auf. Würde ein Text zufällig, bei Einhaltung nur dieser Eigenschaft der Buchstabenhäufung erzeugt, so könnte er so aussehen:

...DEASSBLANELTEDRDAFNEIEENNGTNDHEHEDGANNDREAIN...

Das sieht nicht sehr nach deutsch aus; die relative Häufigkeit der Buchstaben ist zufriedenstellend, nicht aber deren Reihenfolge. Es sollten daher nicht isolierte Buchstaben, sondern deren Kombinationen gezählt werden, z.B. Bigramme, d.h. alle Buchstabenpaare wie AA, AB, AC, ... BA, BB, BC usw., wobei am häufigsten EN, ER, CH, DE, ND, IE und EI vorzufinden sind. Längere Ketten, Trigramme (wovon die häufigsten interessanterweise EIN, ICH, NDE, DIE, UND und DER sind), führen zu noch getreueren Synthesen. T36a zeigt die Statistik sämtlicher Zeichen sowie der häufigsten Bigramme des vorliegenden Kapitels; die Anzahl der Erscheinungen sind links vor dem ‚x‘ angegeben.

Ketten von Elementen, wie hier Buchstaben, oder sonst Zeichen allgemeiner Art, wurden zu Beginn des 20. Jahrhunderts vom Mathematiker Andrei Andrejewitsch Markoff (1856-1922) wie oben geschildert statistisch untersucht; in dem Zusammenhang heißen sie daher *Markoff-Ketten*. Je länger die Kette, d.h. je höher deren Ordnung, desto mehr Information liefert sie über das allgemeine Verhalten der Elemente und desto näher befindet sich eine Resynthese der Originalquelle. T36b listet zu oberst einen Ausschnitt aus einem Satz über Markoff. Darunter befinden sich vorerst acht Resynthesen des Textes, den Markoff-Ordnungen 1 bis 8 entsprechend; der langsame Übergang vom Kauderwelsch der Ordnung 1 über das schon Sprechbare der Ordnung 2 zum deutschsprachigen Unsinn der Ordnung 3 ist deutlich. Danach folgen bloß sprachliche Verbesserungen.

An sich ganzzahlig, sind reellzahlige Ordnungen wie folgt auch realisierbar: man nehme bei einer reellen Ordnung $G+B$ (wobei G eine ganze Zahl und B ein Bruch ist) $100(1-B)\%$ der Bewertungen auf der Ordnung G und $100B\%$ auf der Ordnung $G+1$ vor, z.B. bei Ordnung 2,2 ($G=2$, $B=0,2$) sind 80% der Synthesen auf Ordnung 2 und 20% auf Ordnung 3. T36b zeigt zum Schluss vier weitere Synthesen bei Ordnungen zwischen 2 und 3 – auch hier ist ein langsamer sprachlicher Übergang wahrnehmbar.

Für die oben beschriebenen Synthesen wurden im Originaltext alle Ketten der Länge 1 bis 8 ermittelt, die – ähnlich T36a – in T36c zu sehen sind.

Musikelemente, z.B. Noten, können auf Markoffsche Weise analysiert und resynthetisiert werden. ¶37a zeigt ein Fragment einer Bach-Invention im Original sowie in Markoffscher Resynthese der Ordnung 1-8: auch hier ist ein allmählich abnehmendes ‚Chaos‘ vernehmbar. Die Phrase wurde in 32 einzelnen 16.-Noten aufgelistet, und zwar staccato (im Notenbeispiel zu legato ausgebügelt), wobei ‚-‘ die Pausen angibt:

--C-D-E-F-D-E-C-G---C'---HAH-C'---

Zahlenmäßig sind die Töne C-C' 2 2 2 1 1 1 2 2 plus 19 Pausen. Von diesem ausgehend verläuft die Synthese 1. Ordnung so:

D-HA-E-H-EHC'---EE-D-EEHF--H--F--

Hier zählen die selben acht Töne 0 2 6 2 0 1 5 1 plus 15 Pausen; die Häufigkeits-Abweichungen vom Original werden im weiteren Verlauf der Resynthese ausgeglichen. Bei der 7. Ordnung ist das Original wiedererlangt.

Die *Stochastik* (vom griech. *stochastikos* = ‚geschicktes Zielen‘ im Sinne von ‚gekonnt erraten‘) ist das Studium der Wahrscheinlichkeit zufallsabhängiger Ereignisse. 1957 führte der Komponist Iannis Xenakis (1922-2001) bei statistischen Klangberechnungen den Begriff *Stochastische Musik* ein.

Die zur Komposition einsetzbaren Wahrscheinlichkeitswerte sind nicht nur aus schon bestehender Musik zu gewinnen; häufiger ist der Ausgangspunkt ein musiktheoretischer oder ein kompositorisch erwogener. ¶37b schildert einen Fall, in dem eine Folge von 9 Wahrscheinlichkeitsvorgaben in Noten übersetzt werden. Grafik 1 oben links zeigt einen Strich in voller Länge bei MIDI-Note 76 (e'', x-Achse), einer 100%igen Wahrscheinlichkeit für diese Tonhöhe entsprechend, welches im Takt 1 des Notenbilds unten links bestätigt wird. Grafik 9 zeigt Gleiches für MIDI-Note 66 (fi'): – s. Takt 9. Dazwischen zeigen die Grafiken variierende Wahrscheinlichkeiten für die Noten vom e'' bis fi': eine glockenförmige Woge erscheint von rechts, sich dabei nach links verschiebend, wo sie verschwindet. Den dort angezeigten Wahrscheinlichkeiten entsprechend (die Summe aller Werte in jedem Bild wird stets auf 100% gerechnet) wurden die unten stehenden Noten erzeugt; je höher die Wahrscheinlichkeit, desto häufiger ist generell die entsprechende Tonhöhe. Allerdings müssten die einzelnen Takte viel länger sein, um die tatsächlich vorgeschriebenen Wahrscheinlichkeiten in den Noten gut einzuhalten; jedenfalls zeigt das Notenbild eine von e'' bis fi' – den Grafiken entsprechend – langsam sinkende Tonfolge.

Verschiedene Methoden können mittels eines Computers zur Erzeugung stochastischer Musik eingesetzt werden; am einfachsten ist es, den Bereich von 0 bis 1 in gleich viele Zonen wie zu verteilende Elemente zu teilen, wobei die Breite einer Zone zur entsprechenden Wahrscheinlichkeit proportional ist (für 50% stünde eine Zone von z.B. 0,23 bis 0,73) – danach werfe man zwischen 0 und 1 erzeugte Zufallszahlen in den so aufgeteilten Bereich und notiere die Zone, in die sie fallen – s. die Pentatonik in ¶37c.

Entsteht eine Schallwelle durch mehrere gleichzeitig überlagerte Sinusschwingungen, so nimmt sie die Form der Summe dieser Sinuskurven an. $\Gamma 38a$ zeigt zwei Kurven im Oktavenverhältnis, an den Periodenlängen ersichtlich. Auch die Amplituden sind im 1:2-Verhältnis – die langsamere Kurve schlägt doppelt so weit aus wie die schnellere (das Spektrum unten zeigt teiltonbezogen die Amplitude bzw. den Schalldruck an). Die fette Summenkurve ist durch die Periodizität der Komponenten selbst periodisch. Spektren können auch die Schallintensität oder den dB-Pegel angeben.

$\Gamma 38b$ stellt ein etwas reichhaltigeres Spektrum dar – hier sind fünf Teiltöne vorhanden, die frequenzmäßig als 1:2:3:4:5, amplitudenmäßig $1/1: 1/2: 1/3: 1/4: 1/5$ zueinanderstehen. $\Gamma 38c$ geht mit 25 Teiltönen weiter; auch hier weist der n -te Teilton die Frequenz nF und die Amplitude A/n auf, wobei F bzw. A die Frequenz bzw. Amplitude des ersten Teiltönen, des Grundtons sind. Die auffälligen Stufen sind im Vergleich zu $\Gamma 38a$ und $\Gamma 38b$ kleiner aber zahlreicher geworden, ein Prozess, der bei weiterer Fortsetzung in die Kurvenform in $\Gamma 38e$ übergeht (aufgrund ihrer Gestalt eine *Sägezahnkurve* genannt).

Die *Fourier-Analyse* – nach ihrem Erfinder, dem Mathematiker Jean-Baptiste Joseph Fourier (1768-1830) genannt – entknäuelte die einzelnen Sinuskomponenten einer wie oben zusammengesetzten periodischen Kurve: als Eingabe dient eine Kurvenperiode, als Ausgabe erhält man u.a die Amplitude der Komponenten. Das Spektrum in $\Gamma 38e$ zeigt, dass die abgebildete Sägezahnkurve der Summe der ersten 500 Teiltöne mit Amplitude A/n entspricht, dass die in $\Gamma 38f$ befindliche *Rechteckschwingung* die Summe nur ungeradzahlgiger Teiltöne mit Amplitude A/n darstellt; die *Dreieckschwingung* in $\Gamma 38g$ besteht ebenfalls nur aus ungeradzahlgigen Teiltönen – jedoch sind hier die Amplituden quadriert (sie betragen $1/1, 1/9, 1/25$ usw. mal die Grundamplitude) und: die Phasen von Teiltönen Nr. 3, 7, 11 usw. sind umgekehrt! $\Gamma 38h$ zeigt Spektrum und Schwingung aus der Amplitudenfunktion $1/\xi$, wobei ξ der von mir zur Messung von Harmonizität (s. Kapitel 19) eingesetzten Unverdaulichkeitsfunktion entspricht – die ‚verdaulichen‘ Teiltöne sind hier lauter.

So lassen sich *harmonische* (frequenzmäßig in ganzzahligen Verhältnissen) Spektren aller – sogar nichtperiodischen (s.u.) – Schallwellen errechnen. Unter den Varianten der Fourier-Analyse gibt es die sehr gebräuchliche *Diskrete Fourier-Transform* (DFT), die in Zeitfenstern, einer hypothetischen Periodenlänge entsprechend, erfolgt (es gilt $N=\mathcal{R}/f$, wobei N die pro Zeitfenster analysierten Samples, \mathcal{R} die Abtastrate und f die spektrale Grundfrequenz sind). Schallwellen aller Art können mittels eines willkürlich gewählten Grundtons analysiert werden, dessen Spektrum von einem Zeitfenster zum nächsten errechnet wird. Die Höhe dieses Grundtons bestimmt die Dichte der Information im zu untersuchenden Frequenzbereich: ist der Grundton nur knapp darunter, so könnten dessen Teiltöne im besagten Frequenzbereich zu weit auseinander liegen. Wichtig ist vor allem zu wissen, dass das ermittelte Spektrum amplitudenmäßig eine absolute Tonhöhengültigkeit besitzt, mehr oder weniger unabhängig vom Grundton.

Die in T38d abgebildete Kurve besitzt ein Sägezahnspektrum, obwohl sie optisch nicht so anmutet – im Gegensatz zu allen bisherigen Beispielen sind die Teiltöne hier phasenungleich, d.h. die Sinuskomponenten setzen zu Beginn der Periode bei unterschiedlichen Winkeln an, obwohl das klingende Resultat dem phasengleichen Sägezahnklang auch klanggleich ist; auch hier ermittelt eine DFT das Spektrum. Beim Anhören von Klängen verfährt das menschliche Ohr ähnlich; einzelne Teiltöne können – unabhängig von der Phase – durch sorgfältiges Zuhören herausgehört werden.

Auf diese Weise ist der von einem Fagott gespielte Ton g' (392 Hz) analysiert und in T38i-l dargestellt worden; zu sehen – mit Splines verbunden – sind die relativen Amplituden (%) bzw. die Lautstärkepegeln (dBu). Nach dem 12. Teilton (4704 Hz) sinkt der Pegel unter -30 dBu bzw. 3,2%-ige Amplitude und wird hier nicht mehr angegeben.

Die *Fourier-Synthese*, auch *additive Synthese* genannt, ist die umgekehrte, hier zu Beginn geschilderte Methode zum Aufbau von Schallwellen; sie eignet sich nicht für Geräusche. Wird z.B. ein synthetischer Klavierklang angestrebt, so werden Spektren sukzessiver Perioden der Schallwelle eines bestimmten Klaviertons fourieranalytisch ermittelt und digital gespeichert. Da der Klang (vor allem direkt nach Anschlag) kein zeitlich statischer ist, werden mehrere Perioden bis zu einer Dauer von etwa 0,3–0,4 Sekunden (etwa 100 Perioden beim c') analysiert; dabei müssen nicht die Spektren aller Perioden einzeln gespeichert, sondern können in Zeitfenstern von z.B. 10 ms additiv zusammengefasst sein. Nach diesem simulierten Anschlag kann der Ausklang durch die lautstärkenmäßig allmählich ausgeblendete schleifenmäßige Wiederholung der letzten Perioden des Tonanfangs realisiert werden. Um möglichst naturgetreue Übergänge der Klangfarbe herzustellen, wird dieses Verfahren bei ausreichend vielen Grundtonhöhen wiederholt, und zwar angemessenerweise im Abstand von bis zu einer großen Terz. Bei 22 Grundtonhöhen (7 Oktaven) × 30 Zeitfenster (300 ms) × (sagen wir) 50 Teiltönen kämen 33000 Zahlen zustande, die bei beispielsweise je 16 Bit eine Speicherkapazität von etwa 66 Kilobyte benötigen. Aus diesem Material können dann aus komponierten Tonhöhen und Lautstärken Schallwellen längerer Dauer berechnet und von digital zu analog umgewandelt werden; diesem (übersimplifiziert hergestellten) ‚Klavierklang‘ fehlen jedoch die typischen, vornehmlich vom Hammeranschlag geprägten Geräuschanteile.

T38m stellt eine per Hand gezeichnete Wellenperiode dar, deren Fourieranalyse bis zum 99. Teilton in T39 gezeigt wird. Diese Analyse bildet die Basis einer Resynthese, mit 6, 16 und 75 Teiltönen in T38n-p gezeigt; es ist nicht zu übersehen, wie die ursprüngliche Gestalt mit zunehmender Teiltonanzahl nach und nach wiederhergestellt wird.

In der Frühzeit Elektronischer Musik, um 1952 maßgeblich von den Komponisten Herbert Eimert (1897-1972) und Karlheinz Stockhausen (1928-2007) eingeleitet, wurde die analoge Klangerzeugung vornehmlich mit einfachen Tongeneratoren bewerkstelligt (meistens boten diese nur einfache Wellenformen wie Sinus, Sägezahn und Rechteck); dazu kamen u.a. Rauschgeneratoren und diverse klangverfremdende (z.B. Ring-) Modulatoren. Auf digitalem Wege – durch Max Mathews (geboren 1926) – konnte ab 1957 eine größere Vielfalt an vorzuerrechnenden Klängen und Geräuschen erzeugt werden, eine allerdings wegen der damals noch geringeren Rechengeschwindigkeit sowie Speicherkapazität recht aufwändige Alternative. Heute – nicht zuletzt durch die Einführung von MIDI – umfasst das Geräteangebot eine große Vielzahl von Synthesizern und Samplern.

Die Samplertechnik – heute zur bedeutendsten Methode der Klangerzeugung avanciert – ist denkbar einfach: für die befriedigende Wiedergabe z.B. einer bekannten Klangart wie der Klavierklang über den kompletten benötigten Umfang müssen etwa alle drei oder vier Halbtöne die über Mikrophon aufgenommene, AD-umgewandelte Originalschallwelle als Zahlenfolge (Samples) gespeichert werden. Wellen dazwischenliegender Grundtonhöhen werden durch Interpolation simuliert. Um die Speicherkapazität bei lang anhaltenden Klängen sparsamer zu nützen, braucht nur der zur Klangwiedererkennung wichtige Anfang (bis ca. 300–400 ms) abgespult zu werden; der Rest kann dann aus ein paar typischen Perioden vom Ende dieses Anfangs geformt werden, und zwar als eine sich wiederholende, von einer passenden Hüllkurve amplitudenmäßig gesteuerte Schleife.

Wurden vor 1980 neue, bisher unbekannte Klangarten erwünscht, so musste allgemein auf Technologien wie errechnete, durch DA-Synthese realisierte Schallwellen gegriffen werden, eine damals sehr zeitaufwändige Angelegenheit. Auch setzte man die Fourier-Synthese mit mehreren gleichzeitig klingenden Sinustongeneratoren in Echtzeit ein – wegen der vielen, für die Teiltöne eines Klanges benötigten Klangerzeuger nicht unkompliziert – oder man war auf analoge Synthesizer angewiesen, die nach festgelegten elektronischen Schaltungen ziemlich spezifische, doch zahlenmäßig begrenzte Klangarten zu erzeugen vermochten. Als relativ unaufwändig dagegen erwiesen sich in den 1980er Jahren entwickelte Methoden wie *Frequenzmodulation* (FM) oder *Phasenverzerrung* (PD vom engl. *Phase Distortion*), bei denen aus geschicktem Umgang mit relativ wenigen Tongeneratoren eine spektrale Reichhaltigkeit erfolgte.

Die 1973 von John Chowning (geboren 1934) veröffentlichte FM-Methode wird schon seit längerem bei der Rundfunkausstrahlung eingesetzt: eine zeitlich leicht schwankende Sinuswellenfrequenz kann als Frequenz/Zeit-Kurve analytisch ermittelt werden. Sind diese Schwankungen den Windungen der tragenden Welle gegenüber relativ langsam, ist die Dekodierung präziser zu vollziehen: so funktioniert FM- (zu deutsch UKW-) Radio, bei dem eine hohe Trägerfrequenz (88–108 MHz) durch eine viel langsamer vibrierende Schallwelle (20–20000 Hz) moduliert wird. Sind aber Träger und Modulation frequenznah, so sind sie weniger deutlich auseinanderzuhalten – durch ihre gegenseitige Beeinflussung entstehen spektrale Seitenbänder, die die Klangfarbe bereichern. So konnte Chowning aus nur zwei Frequenzen mehrere weitere erzeugen. Die Grundformel für eine FM unterworfenen Kurve hat die Form $e = A \sin(\alpha t + I \sin \beta t)$, wobei e die Wellenabweichung von der Mittellage, t die Zeit, A die maximale Abweichung (Amplitude), α und β die Perioden des Trägers bzw. der Modulation sind; I ist der (die Amplitude beeinflussende) Modulationsindex. $\Gamma 40b$ zeigt eine mit 10 Hz modulierte Trägerwelle von 100 Hz bei Modulationsindex 0,2. Die resultierende Welle (bei $\Gamma 40a$ noch unmoduliert), zur Mitte hin ‚ausgebeult‘, an den Rändern ‚zusammengedrückt‘ – welches erst bei höheren Indizes sichtbar wird (bis $\Gamma 40h$: Index 15) – zeigt bei Fourieranalyse anfangs symmetrisch bewegliche Seitenbänder (die Teiltöne sind als Vielfache von 10 Hz abgebildet); ab $\Gamma 40f$ erscheinen in den negativen Frequenzbereich verschobene Bänder im Positiven wieder, aber dadurch mit dämpfender Wirkung auf andere auch vorhandene Frequenzen. Der Amplitudenverlauf der Seitenbänder werden auch mittels der nach dem Mathematiker Friedrich Wilhelm Bessel (1784–1846) benannten *Bessel-Funktionen* berechnet, deren Erklärung den Rahmen dieses Buches sprengen würde. Bei gleicher Träger- und Modulatorfrequenz entstehen bei Index 1 bzw. 10 die in $\Gamma 40i$ bzw. k ersichtlichen Kurven. FM kann kaskadiert werden: die Modulation kann selbst moduliert werden usw. – der erste bekannte FM-Synthesizer, der Yamaha DX-7, setzte sechs auf 32 verschiedene Weisen miteinandergekoppelte Tongeneratoren ein.

Die Phasenverzerrung (PD) wurde ungefähr zeitparallel zur Entstehung der FM-Klangsynthese von mehreren Forschern entwickelt – sehr einfach ausgedrückt, wird hiermit eine Eingangskurve durch eine Tabelle geschickt, in der für jeden möglichen y -Wert ein Ersatzwert angegeben wird. $\Gamma 40m$ zeigt fünf verschiedene Ergebnisse: die links oben angegebene Ursprungskurve (hier immer eine Sinuswelle) wird der eins-zu-eins-Korrelation in den grauen Feldern in der Mitte unterworfen (‚Alt‘-Wert auf der x -Achse gegen ‚Neu‘-Wert auf der y -Achse): das veränderte Resultat rechts ist in zwei der Beispiele deutlich komplizierter, obertonreicher als das Eingangssignal. Eine grundlegendere Form der PD-Synthese ist die Anwendung einer variablen Abtastrate auf die Schallwelle (hier ist die Bedeutung der Bezeichnung Phasenverzerrung eher zu suchen): auch hier ist es möglich, mit einfachen Mitteln komplexe Klänge zu erlangen.

Γ38i-l zeigen eine Anzahl paralleler senkrechter Linien, die auf der x -Achse die Teiltonfrequenzen eines Fagott-Spektrums, auf der y -Achse deren jeweiligen Amplituden bzw. Laustärkepegel angibt. Alle Teilton enthaltende Spektren, sowohl harmonische, bei denen die Teiltöne ganzzahlige Frequenzverhältnisse bilden und daher einer Obertonreihe zugewiesen werden können, als auch inharmonische, deren Frequenzen von dieser abweichen (z.B. bei vielen Glockenarten), können so dargestellt werden. Eine derart aus Teiltönen bestehende akustische Erscheinung wird ein Klang genannt: dessen gezeichnetes Spektrum besteht nur aus parallelen Linien; bei einem harmonischen Spektrum ist die Schallwelle periodisch.

Vergleichen wir dies mit dem in Γ41a gezeigten Spektrum eines kräftig geschlagenen Tamtams. Anstelle von parallelen Linien ist hier eine Hüllkurve zu sehen; die Klangfarbe dieses Instruments, ein Geräusch, enthält keine wahrnehmbaren Sinuskomponenten. Was für Frequenzen enthält ein solches Spektrum?

Es wird allgemein behauptet, dass ein von zwei Eckfrequenzen begrenztes Rauschband ‚alle dazwischen liegende Frequenzen‘ enthalte, und dass ein jeder Punkt auf der Hüllkurve die Amplitude der dort auf der x -Achse angezeigten Frequenz angebe. Jedes auch so schmale Rauschband muss in dieser Vorstellung aus unendlich vielen Sinustönen bestehen (denn für jedes Paar benachbarter Frequenzen gibt es noch eine dazwischen), die sich entweder zu einer Schallwelle von unendlich großer Amplitude summieren lassen oder individuell einen unendlich niedrigen Lautstärkepegel (null) aufweisen müssen.

Mit dieser Vorstellung kommt man nicht weiter. Hier wird gezeigt, dass Schallwellen von Geräuschen stochastische Eigenschaften haben, d.h., dass sie mit Zufall und Wahrscheinlichkeitsrechnung verquickt sind. Γ41b zeigt in der Mitte 882 willkürlich erzeugte Zufallszahlen in graphischer Form – von links bis rechts wurden 16-bit-Ganzzahlen zwischen einem gewissen Minimum und einem gewissen Maximum (hier -32768 bzw. +32767) durch einen Zufallsgenerator erzeugt. Als Samples durch einen DA-Umwandler geschickt, erklingt diese Folge als *Weißes Rauschen*, bei dem alle (hörbaren) Frequenzen gleichmäßig wahrzunehmen sind (bei einer Abtastrate von 44100 Hz dauert das Bild genau 20 ms). Die Fourieranalyse zeigt ein gut verteiltes Spektrum: in Γ41b oben werden neun, jeweils von einem Neuntel (d.h. je 98) der Zufallszahlen erstellten DFTs gezeigt, die zwar unterschiedlich, aber sich allesamt von 0 bis 20000 Hz erstrecken (zusätzlich zeigt jedes Kästchen die anderen acht Spektren in grau). Geräusche eignen sich ebenso wie Klänge zur Fourieranalyse: hier werden die Amplitude eines fiktiven Grundtons und seiner ebensofiktiven Teiltöne errechnet, zwischen denen dann interpoliert werden kann – je tiefer der ‚Grundton‘, desto dichter liegen dann die ‚Teiltöne‘ der Mittellage zusammen.

Betrachten wir die – in $\Gamma 41b$ im grauen Kästchen links befindlichen – 98 Zufallszahlen unter der Lupe. Horizontal vergrößert, sieht man sie wieder im länglichen Kästchen darunter: jedes Samplepaar ist als Eckpunkte eines Segments einer Sinuskurve maximaler Amplitude (in Zahlen -32768 bis $+32767$ bei 16 Bits) zu sehen – die bewegliche Frequenz dieser lautstärkenmäßig unveränderlichen Welle konnte anhand der Abtastrate (hier 44100 Hz) und Steilheit des Samplepaares errechnet werden – wie im Bild unten (mit der Bezeichnung ‚SIS‘) ersichtlich – und erlaubte die Deutung des Rauschens als umherfliegenden ‚Sinuston‘ zufälliger momentaner Frequenz: das unten gezeichnete Tonhöhen-Histogramm (‚Hgm‘) lässt eine besonders deutliche Präsenz im Bereich -856 bis $+510$ Cent um 44100 Hz wahrnehmen, welches den DFTs oben zu widersprechen scheint, zu denen jedoch eine direkte Verbindung nicht zwingend ist.

Diese Überlegungen bilden die Grundlage meines Schallwellen-Systems ISIS (‚Intra-Samplare Interpolation mit Sinuswellen‘): in diesem wird jedem Samplepaar einer beliebigen Schallwelle ein dazwischen liegendes Sinussegment eindeutiger Frequenz und maximaler Amplitude zugeordnet. Die Formel für diese Frequenz ist

$$f = \varrho(\hat{U}(s_2) - \hat{U}(s_1)) / 2\pi,$$

wobei f die ISIS-Frequenz, ϱ die gegebene Abtastrate und s_1 und s_2 zwei aufeinander folgende, im Bereich ± 1 befindliche Samples sind. Aus Gründen der Eindeutigkeit und Konsequenz fand ich es sinnvoll, den Wert von f durch die Addition von $+\varrho$ oder $-\varrho$ in den Bereich $(\varrho \pm 1)/2$ zu legen, d.h. Abtastrate \pm Nyquist-Limit ($\varrho \pm N$). Diese Form von ISIS beinhaltet zusätzlich *Kontigualphasen*, d.h. zwei benachbarte Sinussegmente treffen sich – ähnlich wie bei Splines – immer mit gleicher Phase.

$\Gamma 41c$ zeigt ein Beispiel der *ISIS-Analyse* mit 12 zufälligen Samples (vgl. $\Gamma 03g$) – den 11 Samplepaaren werden je eine Frequenz mit der obigen Formel abgewonnen (z.B. 43195 , 42685 , 49460 Hz, oben im Diagramm gezeigt – sehr hohe Frequenzen!); sie können aber transponiert oder als Tonhöhen anderweitig bearbeitet werden, und dann in eine Schallwelle (rück-)umgewandelt werden, und zwar mittels der Formel

$$s_2 = \beta(\hat{U}(s_1) + (2\pi f / \varrho)),$$

wobei f , ϱ , s_1 und s_2 die gleiche Bedeutung wie oben haben.

Für $\Gamma 41d$ wurde dieser umgekehrte Weg der *ISIS-Synthese* gewählt – die Beispiele zeigen 1) ein wiederholtes a' , 2) ein a' - a'' -Tremolo, 3) eine Zufallstonfolge zwischen a' und a'' sowie 4) eine stochastische Tonfolge mit Maxima der Wahrscheinlichkeit bei a' und a'' . Eine auf $73\frac{1}{3}$ Hz basierende DFT zeigt bei 1) den 6. Teilton (440 Hz), 2) den 9. Teilton (660 Hz, zu welchem die Oktave $440+880$ Hz ‚verschmolzen‘ ist!), 3) ebenfalls den 9. Teilton mit leisen Begleitfrequenzen und 4) ein sehr reichhaltiges Spektrum. Die je oben im Kästchen angezeigten, reanalysierten ISIS-Frequenzen und deren unten befindlichen Histogramme zeigen die Richtigkeit der Durchführung der Vorgaben. Es ist also plausibel, ein Geräusch als einen tonhöhenmäßig schnell umherfliegenden Sinuston zu betrachten, dessen Frequenzverteilungs-Histogramm einen besonderen, mir noch unbekanntem Bezug zum Spektrum hat.

Die Wahrnehmung von Tonhöhen zeigt sich in zwei Hauptformen. Zum einen bilden Frequenzen hörbare intervallische Verbindungen, wodurch die Welt der Harmonik entsteht (rational-intervallisches Hören). Viel häufiger ist jedoch der Einsatz des Gehörs um hoch und tief zu unterscheiden, um Sprachformanten und Sprachmelodien zu erkennen, ohne dass das intervallische Hören eintritt (Tonlagenwahrnehmung). Ein geflüstertes s. klingt höher als ein geflüstertes sch.; es ist jedoch in der Regel nicht nötig, das Intervall zwischen diesen zwei Rauschbändern zu erkennen oder zu merken.

Ein Experiment: Zwei relativ schmale Rauschbänder mit Zentralfrequenz 300 bzw. 2000 Hz werden Testpersonen vorgestellt – sie werden gebeten, ein drittes Rauschband auf eine Tonhöhe in die Mitte zwischen den anderen beiden zu bringen. Intervallische Überlegungen suggerieren hierfür eine Zentralfrequenz von der geometrischen Mitte 775 Hz; tatsächlich aber wird die ausgewählte Zentralfrequenz bei 920 Hz liegen, etwa eine kleine Terz höher. Das ist daher zu erklären, dass wenn das intervallische Hören ausgeschaltet ist, eine andere, von der physikalischen Centskala abweichende subjektive Einteilung des Hörbereichs ins Spiel kommt – s. $\Gamma 42a$.

Ein zweites Experiment: Die Frequenz eines Sinustons wird stochastisch zu einem Rauschband ‚geschmiert‘; bleibt der Schalldruck konstant, so ändert sich die Lautheit nicht, auch wenn sich die Bandbreite des Rauschens bei gleichbleibender Zentralfrequenz ausweitet – das sollte auch nicht überraschen, denn bei konstanter Tonhöhe und konstantem Schalldruck soll die Lautheit auch konstant bleiben. Plötzlich aber, ab einer gewissen Bandbreite, beginnt die Lautheit bei weiterer Ausdehnung doch deutlich zuzunehmen – diese *kritische Bandbreite* ist relativ klar ermittelbar und besitzt eine für jede Zentralfrequenz spezifische Größe – s. $\Gamma 42b$.

$\Gamma 42c$ zeigt den Verlauf der kritischen Bandbreite in Bereich von 50 Hz bis 18 kHz; der Zentralfrequenz eines Bandes auf der x -Achse entsprechend, ist dessen Breite an der y -Achse abzulesen. Ist wie in $\Gamma 42e$ die untere Grenze eines Bands 100 Hz, so ist die obere Grenze dieses Bands bei etwa 200 Hz (verfolge die y -Achse); das bei 200 Hz beginnende Band endet bei 300 Hz, das bei 300 Hz beginnende bei 400 Hz. Die Bändergrenzen scheinen also eine Obertonreihe zu bilden – 100, 200, 300, 400 Hz...; aber statt der zu erwartenden 500 Hz ist die nächste Grenzfrequenz 510 Hz, von 630, 770, 920, 1080, 1270, 1480, 1720, 2000, 2320, 2700, 3150, 3700, 4400, 5300, 6400, 7700, 9500, 12000 sowie 15500 Hz gefolgt – diese Werte sind in $\Gamma 42e$ an der y -axis über die in der Mitte des Diagramms diagonal laufenden, von 1 bis 24 durchnummerierten Zahlen zu sehen. In Cent gemessen, nehmen die Bandbreiten zuerst rapide ab, dann pendeln sie sich bei etwa 2000 Hz auf eine Breite von 260 Ct ein, wonach sie sich wieder bis auf etwas über 400 Ct wieder ausdehnen, wie wieder in $\Gamma 42c$ zu sehen.

Zurück zum ersten Experiment – ein Blick auf die Bändergrenzen zeigt 920 Hz auf dem halben Weg zwischen 300 Hz und 2000 Hz – s. Striche 3, 8, 13 in T42e. Diese Bändergrenz-Frequenzfolge stellt die oben erwähnte subjektive Tonhöhenkala dar; die Einheit dafür ist *Bark* (abgekürzt ‚Bk‘) nach dem Akustiker Heinrich Barkhausen (1881-1956), dem Erfinder der Einheit Phon, genannt. Eberhard Zwicker (1924-1990), der sich sehr früh mit der kritischen Bandbreite beschäftigte, nennt diese Skala jene der ‚Tonheit‘, wohl analog zu subjektiven ‚Lautheit‘-skala; es findet jedoch das von mir vorgezogene Wort ‚Tonlage‘ hier Verwendung. T43 gibt detaillierten Angaben über die Entsprechung Bark→Hertz; sie wurde durch Formeln von Ernst Terhardt (1979) und Hartmut Traunmüller (1990) als Annäherung an die empirisch gewonnenen Werte erstellt – der Übergang von der einen zur anderen Formel ist bei 219,5 Hz oder 2,16 Bk:

Ist $b \leq 2,16$: $f = (4000/3) \hat{=} (b/13.3)$ (Terhardt)

Ist $b \geq 2,16$: $f = 1960(b+0,53)/(26,28 b)$ (Traunmüller)

Ist $b \geq 20,1$, verbessert Traunmüller seine Formel durch folgende vorherige Operation:
 $b \leftarrow (b+4,422)/1,22.$

Die umgekehrten Formeln:

Ist $f \leq 219,5$: $b = 13,3 \cup (3f/4000)$ (Terhardt)

Ist $f \geq 219,5$: $b = ((26,81f)/(1960+f)) - 0,53$ (Traunmüller)

Danach, ist $b \geq 20,1$: $b \leftarrow b + 0,22(b - 20,1)$ (Traunmüller-Korrektur)

(f ist die Frequenz in Hz, b ist die Tonlage in Bk).

T42d zeigt die Ergebnisse dieser Formeln sowie vier anderer Annäherungen – die senkrechten grauen Linien zeigen eine von 100 Hz ausgehende Obertonreihe.

Die kritische Bandbreite und die Barkskala zeigen sich in einer Reihe von anderen Phänomenen: z.B. das kleinste noch hörbare Intervall zwischen zwei alternierenden Sinustönen beträgt etwa 0,02 Bk im ganzen Hörbereich, also ca. 35 Ct bei 100 Hz, 6 Ct bei 1000 Hz, 8 Ct bei 10000 Hz. Die recht grobe Intervallwahrnehmung bei tiefen Frequenzen erklärt die bekannte Praxis, einen Kontrabass oder eine Pauke zu stimmen durch Flageolets (höhere Teiltöne) auf den Kontrabass-Leersaiten bzw. am heller klingenden Rand der Pauke.

Die Wissenschaftler Reinier Plomp und Willem Levelt (1965) und andere haben auch gezeigt, dass die zwischen benachbarten Tönen in Akkorden klassischer Musik typische Intervallgröße durchaus mit der Barkskala zu tun hat: tiefere Intervalle, in Cent gemessen, sind generell größer als höher liegende. Verschiedene Versuche zeigen eine Tendenz zu etwa 1,3 Bark als beinahe lagenunabhängiger Favorit.

Auch das dissonanteste Intervall zwischen zwei simultanen Sinustönen kann – nach einer legendären Arbeit von Plomp und Levelt in Bark angegeben werden (0,2), unabhängig von der Frequenz – s. Kapitel 30.

Komponisten, die eine Musik anstreben, bei der das intervallische Hören irrelevant sein soll (z.B. weitreichende Glissandi oder punktuelle Texturen) sei hier empfohlen, sich an der Barkskala zu halten – die Verteilung der Tonhöhen wirkt viel gleichmäßiger.

29 Akustik 9 – Lautstärkensummierung und Maskierung

Erklingen zwei Sinuswellen gleichzeitig, so addieren sich ihre Kurven zu einer zusammengesetzten Form, wie die in $\Gamma 44$ grau gezeigten: hier sind zwei Frequenzen im Oktavverhältnis bei a) gleicher Anfangsphase, b) einer Phasenverschiebung von einer Achtelperiode der höheren Frequenz. Die physikalische Gesamtlautstärke in dB(SPL) hängt u.a. vom RMS (der Quadratwurzel des Durchschnittquadrates) der Kurve ab. Hierbei spielen die eigentlich klangirrelevanten Phasen eine Rolle, wie im RMS-Unterschied zwischen diesen beiden Bildteilen ersichtlich. Eine Fourieranalyse der zwei Summenkurven gibt die Originalfrequenzen, -amplituden und -phasen wieder – siehe diese rechts im Bild (bei ‚DFT‘): sie sind hier als 8. und 16. Teiltöne einer willkürlich gesetzten Grundfrequenz zu sehen, von der eine Periode die komplett angezeigte Schallwelle umfasst; die Phase wird durch die Neigung der Spektrallinien angegeben. Im Extremfall der Addition zweier frequenz- und amplitudenmäßig identischen Sinuskurven umgekehrter Phase beträgt die Gesamtlautstärke null, Stille(!): die Kurven löschen sich gegenseitig aus. Hier kann die Fourieranalyse die Originalkurven natürlich nicht rekonstruieren.

$\Gamma 44c$ zeigt zwei Sinuskurven umgekehrter Anfangsphase im Frequenzverhältnis 8:9 – die Summe ist eine lautstärkenmäßig schwebende Kurve: die Frequenz der Schwebung entspricht der Differenz der beiden Sinusfrequenzen, hier $9-8=1$, d.h. ein Achtel der tieferen, ein Neuntel der höheren Komponente – hiervon zeigt das Bild eine Periode. Wieder vermag die Fourieranalyse die Ausgangssituation plausibel zu rekonstruieren (die entgegengesetzte Phase spiegelt sich in den Richtungen der Spektrallinien wieder).

$\Gamma 44d$ zeigt zwei Sinuskurven im irrationalen, also aperiodischen Verhältnis des Goldenen Schnitts ($1:\frac{1}{2}(1+\sqrt{5})$) oder 1:1,618034, nahe bei 8:13); eine Untersuchung des RMS, Periode für Periode der unteren Frequenz, ergibt unregelmäßige Schwankungen. Dies kann als irregulär verlaufende Lautstärke ausgelegt werden; die tatsächliche Frequenz und die Trägheit des sie wahrnehmenden Gehörsinns stellt die sinnvolle Größe des Zeitfensters und damit die Konstanz der angeblichen Lautstärke fest. In der Tat könnte deswegen selbst ein einziger Sinuston bei ausreichend großem Zeitfenster als mit konstanter, bei kleinem Zeitfenster als mit (bis zu im Sampletakt) schnell schwankender Lautstärke dargestellt werden.

Die subjektiv empfundene Gesamtlautheit lässt sich ebenfalls ermitteln, und zwar wie folgt:

Ist die Kurve aperiodisch, so dass die Anfangsphasen der Sinusbestandteile keine Rolle spielen, so kann die Gesamtlautstärke bzw. die Gesamtlautheit durch die Addition der einzelnen Schallintensitäten in W/qm bzw. der Lautheiten in Sone ermittelt werden. Als Beispiel soll hier ein Sägezahnnton mit Grundfrequenz 100 Hz dienen, wie in $\Gamma 44e$ abgebildet. Die Phasen der Teiltöne werden hier außer Acht gelassen. Per Definition sind die Amplituden eines Sägezahnspektrums proportional zur Umkehrung der Teiltonnummer, d.h. A/n , wobei A die Amplitude des Grundtons ist. So sind die Intensitäten dementsprechend i/n^2 , wobei i (proportional zu A^2) die Intensität des Grundtons ist. Daher tendiert bei praktisch unendlichen vielen Teiltönen die Gesamtintensität gegen $(1+1/4+1/9+1/16\dots)\times i = 1,694\times i$. Jetzt sind die Einzelintensitäten ermittelbar, und zwar nach i/n^2 als 59% (=1/1,694), 14,8%, 6,5%, 3,6% usw. der absoluten Gesamtintensität (beträgt der Gesamtpegel z.B. 90 dB(SPL) so ist nach $\Gamma 08$ die Gesamtintensität 1 mW/qm und damit die des Grundtons 0,59 mW/qm). Um die Gesamtlautheit zu ermitteln, muss das Spektrum als Funktion der Amplitude abhängig von der Teiltonnummer bzw. der -frequenz in eine physiologisch relevante umskaliert werden, nämlich in Sone als Funktion von Bark, wie in $\Gamma 44f$ ersichtlich. Dies erfolgt vornehmlich durch das Absuchen der Fletcher-Munson-Tabellen (s. $\Gamma 12$): ist die Frequenz (Hz) und die Lautstärke (als Intensität oder dB(SPL)-Pegel) bekannt, so kann die Lautheit in Phon und dadurch in Sone abgelesen werden. Jetzt wäre durch die Summierung der Sonewerte die Gesamtlautheit ermittelt, wenn nicht das Phänomen der Maskierung wäre.

Es hat sich experimentell herausgestellt, dass wenn zwei Sinustöne gleichzeitig erklingen, der untere Ton eine dämpfende Wirkung auf die Lautstärke des oberen hat, vor allem wenn sich beide in einem Abstand von weniger als einem Bark befinden. Das heißt: sind sich die zwei Töne in der Höhe nah genug und deren Lautstärkenabstand zu Gunsten des unteren groß genug, könnte es sogar sein, dass der obere völlig verdeckt wird und unhörbar bleibt. $\Gamma 44g$ zeigt das Sägezahnontonspektrum als Phänomen der sogenannten *Lautheitsdichte* in Sone pro Bark als Funktion der subjektiven Tonhöhe in Bark. Statt durch einen vertikalen Strich wird die Lautheit eines jeden Tones durch eine schweifende Kurve dargestellt, die links steil steigt und rechts langsam fällt: bei einem einzelnen Sinuston ist es die darunter bis zur x -Achse enthaltene Fläche, die der Lautheit des Tones in Sone entspricht; die Flächeneinheit ist das Produkt jener der x -Achse (Bark) mit jener der y -Achse (Sone/Bark), also Sone. Bei einem Spektrum der Lautheitsdichte (wie in $\Gamma 44g$) formen die Maximalwerte aller Schweife zusammen eine gemeinsame Hüllkurve, die mit der x -Achse eine die Gesamtlautheit angegebende Fläche einschließt. Diese sich überlappenden Einzelflächen verdeutlichen das Phänomen der Maskierung: jede gemeinsam geteilte Lautheitsfläche zählt nur einmal – allein die Gesamtfläche aller Teiltöne entspricht der tatsächlichen Lautheit in Sone.

Eine vom Ohr wahrgenommene Schallwelle setzt einiges in Bewegung. Zuerst kommt das zwischen dem Außen- und Mittelohr gelegene Trommelfell, welches die Bewegung an drei im Mittelohr befindlichen, winzigen beweglichen Knochen (,Hammer‘, ,Amboss‘ und ,Steigbügel‘) weitergibt. Letzterer ist mit der kleinen Membran verbunden, die das ,ovale Fenster‘ zur im Innenohr gelegenen ,Schnecke‘ verschließt, einem beim Menschen zweieinhalbmal gerollter, sich zuspitzender Schlauch voll Flüssigkeit etwa 35 mm lang und 3 bis 9 mm im Durchmesser – s. $\Gamma 45a_1$ – rechts im Bild zur Veranschaulichung ,ausgerollt‘, dann um 90° um ihre Längsachse gedreht – $\Gamma 45a_2$ – dann in der Horizontalachse seitlich um 90° gedreht, um eine Ansicht ,von vorne‘ bzw. von typischen Querschnitten zu gewähren – $\Gamma 45a_3$. Dieser Schlauch ist längs in drei Kämmerchen geteilt, und zwar durch zwei Membrane – s. die Vergrößerung des Querschnitts $\Gamma 45a_4$; eine von diesen, die Basilarmembran – im Bild sichtbar innerhalb der (aus)gerollten Schnecke – ist eine Art ,Gehör-Tastatur‘, und ist mit etwa 3500 Haaren bestückt (s. das Cortische Organ in der weiteren Vergrößerung $\Gamma 45a_5$), die durch Berührung unterschiedliche Tonhöhenindrücke dem Gehirn vermitteln. Eine Teilung der Basilarmembran in 24 gleich lange Abschnitte von je 1,5 mm (150 Haaren) entspricht der Barkskala.

Die Berührung der Haare erfolgt so: wenn eine ankommende Schallwelle das ovale Fenster rüttelt, wird die Basilarmembran wellig, ähnlich wie beim Ausschütteln eines Tuches – innerhalb von Millisekunden läuft eine Welle vom ovalen Fenster zur Schneckenspitze hin, wobei die Amplitude der Welle zu- und dann abnimmt; die Stelle des Maximums hängt, durch die unterschiedliche Membrandicke bedingt, von der Frequenz des Eingangsignals ab – zum ovalen Fenster hin ist die Membran dünner und fester und reagiert stärker auf höhere Frequenzen (in $\Gamma 45a_2$ ist die Basilarmembran – unterschiedlich dick – von der Seite zu sehen). An den Maxima werden die Haare gebogen und teilen diesen Zustand durch Nervenstränge dem Gehirn mit. Das kleinste noch wahrnehmbare Intervall zwischen zwei alternierenden Sinustönen beträgt 0,02 Bk; diese müssen die Basilarmembran also mindestens $150 \times 0,02 = 3$ Haare bzw. 30 μm auseinander erregen, um voneinander unterschieden zu werden.

Die Hörnerven zwischen Basilarmembran und Gehirn enthalten Neuronen, Zellen, die bei absoluter Stille unregelmäßig und bis zu 150 mal in der Sekunde ,feuern‘ und Stromimpulse von etwa 1 ms Dauer und wenigen Mikrovolt Spannung zum Gehirn schicken. Werden die Haare der Basilarmembran im Rhythmus der Schallwellenperioden gedrückt, so koordinieren sich die Neuronen zeitlich – die trägeren unter ihnen feuern nicht bei jeder Schallperiode, aber der Gesamteindruck ist periodisch. So erhält das Gehirn eine doppelte Information: 1. örtlich – durch die mit den Erregungsstellen der Basilarmembran verbundenen Hörnervenden entsteht eine grobe Analyse in Bark für das tonlagenmäßige, nicht-intervallische Hören; 2. zeitlich – der Rhythmus der Neuronen-Stromstöße lässt Frequenzen und dadurch Intervalle rezipieren, notwendig für den Musikgenuss.

1965 stellten Reinier Plomp und Willem Levelt fest, dass zwei gleichzeitig erklingende Sinustöne am dissonantesten (= ‚am wenigsten wohlklingend‘) bei einem Abstand von etwa 0,2 Bk wirken – $\Gamma 45b$ zeigt eine stufenlose Darstellung des Dissonanzgrads als Funktion des Tonhöhenabstands in Bark. Das mit 0,2 Bk dissonanteste Intervall variiert von 350 Ct bei 100 Hz über 60 Ct bei 1000 Hz bis zu 80 Ct bei 10000 Hz.

Es könnte verwundern, dass Tritonus, große Septim (ca. 1,1 Bk bzw. 1,8 Bk bei 300 Hz) und andere für ihre Dissonanz bekannte Intervalle keine Schwankungen in dieser Kurve verursachen. ‚Konsonanz‘/ ‚Dissonanz‘ und ‚Harmonizität‘/ ‚Inharmonizität‘ werden hier deutlich unterschieden: Harmonizität (‚intervallische Klarheit‘) wird durch die Einfachheit des Verhältnisses zwischen den zwei Intervall-Frequenzen bedingt. Die Klangfarbe spielt dabei kaum eine Rolle – Musik für Streicher kann für Bläser uninstrumentiert bzw. transponiert werden ohne ihren harmonischen Sinn zu verlieren. Dagegen ist Konsonanz die ‚Glätte‘ (Dissonanz die ‚Rauheit‘) eines Klanges. In der tiefsten Klavieroktav sind – durch den dort physikalisch großen Umfang der kritischen Bandbreite bedingt – reine Quart und Ganzton dissonanter als Tritonus bzw. Halbton. Das psychologische Phänomen Harmonizität entspringt der zeitbedingten Rezeption der Neuronenfeuerungen im Gehirn; die physiologische Konsonanz entspricht Abschnitten auf der Basilarmembran (die Dissonanz 0,2 Bk entspricht ca. 0,3 mm oder 30 Haaren).

Ein Sinustonpaar im Abstand einer übermäßigen Quart oder großen Septim wirkt in der Tat nicht besonders dissonant; das werden erst durch die Reibung der kleinen Intervalle zwischen den Obertönen komplexer Klänge. $\Gamma 45c$ zeigt das Konsonanz-Verhalten des Tritonus h-f' (250–354 Hz) anhand der Spektren der zwei Töne; je sechs Teiltöne mitsamt aller Reibungsstellen sind abgebildet – die Summe der Dissonanz beträgt nach Plomp und Levelt 2,8. $\Gamma 45d$ zeigt diese Dissonanzsumme als stufenlose Kurve über zwei Oktaven (die Kurve der unteren Oktav ist mit der in Plomp und Levelts 1965er Veröffentlichung praktisch identisch) – siehe auch zum Vergleich die unten grau gezeichnete Kurve der Harmonizität. Meines Erachtens spielt auch die Teilton-Lautheit eine Rolle: wird diese (in Sone) mit der Dissonanz vor deren Summierung multipliziert, fallen die leiseren Teiltöne weniger ins Gewicht wie in $\Gamma 45e$ ersichtlich.

$\Gamma 45f$ zeigt die Ermittlung der Dissonanzen in einer Oktav nach derselben Methode. Sägezähntöne mit 6, 12 bzw. 24 Teiltönen wurden in drei Tonlagen untersucht, nämlich die jeweils von 27,5 Hz, 220 Hz sowie 440 Hz aufwärts laufende Oktav. Es stellte sich heraus, dass für die zwei oberen Bereiche die klassischen Starkdissonanzen – kleine Sekund, Tritonus und große Septim – wie erwartet deutlich dissonanter als die große Sekund, reine Quart und kleine Septim sind. Jedoch ist dieses Verhalten in der 27,5 Hz-Oktav gar nicht so eindeutig: die große Sekund, kleine Septim und selbst die Quart sind hier vergleichbar dissonant wie die vorhin als dissonanter erwähnten Intervalle. In dieser tiefen Lage wirkt der Halbton beinahe wie eine verstimmte Prim, der Tritonus wie eine verstimmte Quint, die große Septim wie eine verstimmte Oktav. $\Gamma 45f$ zeigt auch – in grau – die Ergebnisse der Berechnungen nach Plomp und Levelt mit sechs gleich lauten Teiltöne für die drei Tonlagen.

31 Phonetik 1 – Physiologische Phonetik: zur Erzeugung von Sprachlauten

Im Körper der Klangfarben gibt es eine besondere Untermenge, die trotz der einfachen Art, der geringen Zahl sowie der genauen Einhaltung der sie bestimmenden Gesetze durch den hohen Grad der Ausschöpfung der zulässigen Möglichkeiten eine reichhaltige Vielfalt und dennoch ausgeprägte Differenziertheit aufweist – die Laute menschlicher Sprache. Die Phonetik ist das Studium der Gesetze dieser Klangfarben.

Auf physiologischer Ebene betrachtet, entstehen Sprachlaute in erster Linie durch die Spektralfilterung von Klängen (z.B. Stimmbänderschwingungen) und/oder Geräuschen (z.B. das Zischen des Atemstroms) durch die verstellbare Größe und damit variierbaren Resonanzen des Mund-, Nasen- und Rachenraums. Fließt Luft durch den offenen Mundraum, entstehen *Vokale* (wie a, e, i); ist der Mundraum durch das Anheben der Zunge oder das Drücken einer Lippe an die Zähne eingeengt, entstehen *Reibelaute* (z.B. s, sch, f); wird er von Zunge oder Lippen geschlossen, entstehen bei gleichzeitig offenem Nasenraum *Nasallaute* (z.B. m, n, ng), und bei geschlossenem Nasenraum unter erhöhtem internen Luftdruck *Verschlusslaute* (z.B. p, b, d), die durch die Wiederöffnung des Mundes erklingen. Das Drücken der Zungenspitze an den Gaumen mit Freiraum an den Zungenseiten erzeugt *Seitenlaute* (z.B. l wie in ‚laut‘).

In der Deutschen Sprache gibt es etwa 22 Konsonanten sowie 13 klar unterscheidbare Vokale, letztere z.B. in den Silben ‚biet(e) Bitt(e) Beet Bett Bad Bott(ich) Boot Butt Bud(e) Bött(cher) böt(e) Bütt Büd(chen)‘, dazu 3 Doppellaute (‚mein Heuhaus‘; dazu könnte z.B. auch ‚nur‘ hinzugezählt werden). Diese drei Dutzend Grundlaute werden durch 30 Buchstaben in nicht immer eindeutiger Kombination dargestellt – oft weist nur der Inhalt auf die richtige Aussprache: z.B. klingt ‚Weg‘ in ‚Weg von hier!‘ anders als in ‚Der Weg ist lang‘. Von der Internationalen Phonetischen Assoziation (IPA) werden über 60 in allen Weltsprachen vorkommende Konsonanten sowie über 20 Vokale anerkannt; in der von der IPA entwickelten Internationalen Lautschrift werden die dreizehn oben zur Veranschaulichung der Vokale angegebenen Silben wie folgt geschrieben: bi t bl t bet bEt bat bWt bot bUt bud b#t bOt bYt byd. T46 zeigt die meisten von der IPA anerkannten Laute tabellarisch, nach biologisch-artikulatorischen Gesichtspunkten unterteilt (s.u.).

Doch zurück zu den deutschen Vokalen: vier zeigen Lang- und Kurzcharakter (bi t/bl t, bet/bEt), andere (bWt/bot, bUt/bud, b#t/bOt, bYt/byd) kurz/lang. Wer die ersten neun Vokale hintereinander – i l e E a W o U u – ausspricht, wird merken, dass die Zunge, zuerst hoch im vorderen Mund sich zum a hin senkt, wobei deren höchste Stelle sich etwas nach hinten schiebt; danach erhebt sie sich wieder bei einer weiteren Verschiebung nach hinten. Diese Bewegung verursacht eine Größenveränderung des Mundraums vor der Zunge sowie des restlichen Mund- und Rachenraums dahinter, welches eine direkte Auswirkung auf die Lautbildung hat: die Höhe der Zunge ist hierfür eine entscheidende Variable, die Vorne-hinten-Position im Mund ihrer höchsten Stelle eine zweite.

T47a zeigt einen Querschnitt der Mund- und Nasenhöhle sowie des Rachens. Wenn nun die Lippen gerundet werden, entsteht ein vierter Raum zwischen Lippen und Zähnen, der die Vokalbildung weiter beeinflusst: wird ein i mit gerundeten Lippen (,labial‘) gesprochen, so wird es zu einem y (Ü) umgeformt, ein illabiales e wird zum labialen O (Ö). Umgekehrt: ein normal labiales u wird mit gespreizten Lippen zum ɜ, das z.B. im Portugiesischen vorkommt, ein W wird ʌ, wie im englischen (*but*, mit deutschem Akzent bət ausgesprochen). Das # findet man im französischen bʃf (*bœuf*), auch regional im Deutschen k#p̂ (Körper). Das ɔ kommt im Englischen vor, z.B. kɔt (*cat*, mit deutschem Akzent kEt). Das dunkle britische ɑ in *guard* wird mit A wiedergegeben, das @ entspricht dem amerikanischen o im ähnlich klingenden *God* ausgesprochen. Die Zungenstellung bei den acht Vokalen i e E a A W o u wird in T47b abgebildet – der Phonetiker Daniel Jones [1881-1967] nannte diese acht die Kardinalvokale.

Die Zunge spielt eine eben so große Rolle bei der Erzeugung der Konsonanten, die nach drei Prinzipien klassifiziert werden – w (,Artikulationsort‘ in T46b – als Angabe dienen die in T47a gezeigten lateinischen Namen der verschiedenen Mund-/Rachenraumteile; ,retroflex‘ und ,glottal‘ bedürfen einer besonderen Erklärung – s.u.), w (die ,Artikulationsart‘) und ob mit schwingenden Stimmbändern (stimmhaft) oder nicht (stimmlos). Zu ,retroflex‘: die Zunge ist in dieser Position, wenn ihre Spitze zum Gaumen hin zurückgebogen wird – der bekannte ,indische Akzent‘ verdankt dieser Stellung ihre besondere Färbung. Das ,Glottale‘ kann durch den Unterschied zwischen der süddeutschen Aussprache des Wortes ,Verein‘ (fE'ra|n) und der norddeutschen (fE'/a|n) verdeutlicht werden (ein ' als phonetisches Symbol bedeutet eine Betonung auf der folgenden Silbe). T47c zeigt die Zungenstellung bei der Erzeugung der Reibelaute [fTʃSCX] wie in ,Riff^ʃ, engl. ,with^ʃ, ,Riss^ʃ, ,Tisch^ʃ, ,ich^ʃ und ,ach^ʃ (besonders tief im Hals artikuliert). T47d zeigt die Zungenstellung bei der Erzeugung der Nasalen [mnʃN] wie in ,fromm^ʃ, ,Bonn^ʃ, frz. ,Cologne^ʃ und ,Gong^ʃ

T46a listet die Vokale, T46b die Konsonanten in der Lautschrift der IPA nach ihrer physiologischen Entstehung auf.

Laute menschlicher Sprache – ‚Phoneme‘ genannt – werden vornehmlich durch das Pressen von Luft durch den Vokaltrakt im Hals und Mund erzeugt; dazu kommt die eventuelle gleichzeitige Betätigung der Stimmbänder bzw. anderer Effekte wie z.B. das Klicken der Zunge. Des Vokaltrakts variable Durchmesser führt zur Bildung darin von einzelnen, miteinander verbundenen Räumen, wie z.B. zwischen den Lippen und den Zähnen, den Zähnen und der Zungenwölbung, der Raum hinter der Zunge usw. Die unterschiedliche Größe dieser Räume verstärkt entsprechende Frequenzbereiche im Luftstromgeräusch und ggf. auch im Stimmbänderklang o.ä..

Ein jedes Phonem wird stets durch die etwa gleiche Mundstellung und damit die gleiche Frequenzbeschaffenheit erzeugt; daher kann jedem Laut ein für ihn typisches Spektrum zugeordnet werden. Die durch die Trakträume verstärkten, im Spektrum hervorgehobenen Frequenzbereiche heißen *Formanten*. $\Gamma 48a_1$ zeigt schematisch und zweimal das Spektrum des Vokals [i] – auf c (131 Hz) und es' (311 Hz) basiert. Die Obertonreihen beider Spektren sind mit ober- bzw. unterhalb liegenden Pünktchen markiert: in beiden sind die Frequenzbereiche um 250 Hz und 2500 Hz, Formanten von [i], lauter als die umliegenden Frequenzen und sind mit F1 und F2 gekennzeichnet. Wegen der Unabhängigkeit der Stimmbänder vom Vokaltrakt bleiben diese absoluten Bereiche beim gleichen Vokal auch bei anderen Grundtönen erhalten. $\Gamma 48a_2$ zeigt zwei Spektren vom [a], ebenfalls von c und es' ausgehend – die Formanten sind hier um F1=750 und F2=1500. $\Gamma 48b$ zeigt als Hüllkurven die ersten zwei Formanten der acht Vokale [i eEQ] und [uOWA].

Es hat sich herausgestellt, dass für die Erkennung aller Vokale die ersten zwei Formanten ausreichen, obwohl mehrere höhere gemessen werden können. Der erste Formant, geformt durch den Raum hinter der Zunge, rangiert von ca. 250 Hz bei hoch gewölbter Zunge (wie bei [i] und [u]) bis 800 Hz bei tief gelegener Zunge (s.[a]). Er hängt also von der Zungenhöhe ab. Der zweite Formant, geformt durch den Raum zwischen der Zunge und den Zähnen, rangiert von 800 Hz bei hinterer Zungenwölbung (s.[u]) bis 2500 Hz bei vorderer Wölbung (s.[i]). Dieser Formant ist also das Ergebnis der Bewegung der Zunge auf der horizontalen Längsachse.

Diese Zweidimensionalität kann genutzt werden, um sämtliche Vokale auf einer Ebene darzustellen, wie in $\Gamma 48c$: die zwei Formanten sind x- und y-Koordinaten zugeteilt. Gezeigt werden jeweils im einem Feld gefassten Messungen aus sechs über dem Diagramm angegebenen Quellen. Dass die Felder sich praktisch nie überlagern ist bemerkenswert. In (fast) jedem Feld ist der Mittelpunkt durch die Angabe des Vokals markiert. Man beachte auch die vier geraden Linien, die die Felder [i eEQ], [yO#A], [3F^O] und [uOW@] verbinden – die Orientierungspunkte aller sechzehn Vokale befinden sich hier auf einem Netz von Tritoni (c', fi', c'' und fi'') für F1 und Ganztönen (f'', g'', a'', h'', cis''' usw.) für F2 – s. auch $\Gamma 48e$. Von diesen Positionen ausgehende Klangsynthesen der sechzehn Vokale wirken überzeugend.

Γ48d stellt – das Netz der sechzehn soeben vorgestellten Formantenpositionen beibehaltend – die Formanten der Vokale [i eE0a@Wou3yi] dar. Diese Trapezform stellt die äußersten Grenzen der biologischen Vokalbildung dar; alle menschlich erzeugbaren Vokale können einem darin liegenden Punkt zugeordnet werden. Die links im Bild (mit erfunden Symbolen) angezeigten, F1-gespiegelten Formantenpaare (auch ‚Vokale‘) sind technisch realisierbar aber sprengen die Möglichkeiten des menschlichen Mundes.

Vokale (vom lat. *vōx* = ‚Stimme‘ u.a.) werden wegen ihres selbstständigen Charakters auch Selbstlaute genannt; dagegen sind die ‚Konsonanten‘ (= ‚Mitlaute‘, vom lat. *con+sonare*, weil sie die Vokale meist begleiten) eine diffuse Gruppe – von (eigentlich selbstständigen) Nasalen, (z.B. *n<N7Mm*), den Lateralen (l ! \;) und den rauschenden Reibelauten oder Frikativen (fvTDSzSZCH usw.) zu den wirklich kon-sonanten Explosivlauten (pbtdkg usw.). Während den Nasalen in ihrer stimmlosen Form keine besonderen allgemein alphabetischen oder phonetischen Zeichen zugeteilt wurden, ist dies bei den anderen Gruppen auffallend. Das stimmlose [1] z.B. im Walischen (dort ‚ll‘ geschrieben) ist eigentlich ein Reibelaut. Γ48f zeigt schematisch (als Gauss’sche Verteilungskurven) die Spektren von sechs Reibelauten, [xftSCs]. Hier geht es weniger um Formanten sondern um mehr oder weniger breite Geräuschbänder, weswegen die x- und y-Achsen in Bark und Phon kalibriert sind. Während die Spektren von [f] und [T] sich stark ähneln (und sie werden häufig unabsichtlich vertauscht), es ist deutlich, wie der Übergang [xftSCs] eine steigende Tonhöhe erzeugt. Diesen Übergang zeigt Γ48g in der Form eines *Sonagramms* (vom lat. *sonus* = ‚Schall‘ und griech. *grámma* = ‚Geschriebenes‘): die x-Achse stellt hier die Zeit, die y-Achse die (hier subjektive) Tonhöhe dar und die Lautstärke wird durch Grauwerte wiedergegeben. Der Gebrauch von Sonagrammen ist vor allem bei der Spektralanalyse von Gesprochenem, also mit einer Zeitkomponente versehen, sehr verbreitet.