

Die Zeitauflösung in der Musikübertragung

Ralf Koschnicke

Zur Renaissance analoger Mischkonsolen

Für viele mag die Markteinführung analoger Summierer- und Mischkonsolen zur Benutzung in Kombination mit digitalen Audioworkstations zu den ausgemachten Scurrilitäten der Audiobranche gehören, doch spätestens die mittlerweile sehr angewachsene Zahl von Herstellern – auch sehr renommierter –, die sich auf diesem Sektor engagieren, sollte zu genauerer Betrachtung der Hintergründe ermahnen. Denn offensichtlich gibt es auf Anwenderseite einen Bedarf in nennenswertem Umfang.

Während natürlich für manche notorisch unkritische Befürworter digitaler Audiotechnik – egal, ob Anwender oder Entwickler – sofort feststeht, dass dieser Gerätetypus nur völliger Unsinn sein kann und allenfalls ins Reich der Esoterik gehört, so lohnt sich doch die genauere Betrachtung der Vorgänge in solchen Geräten – genauso



wie in klassischen analogen Mischpulten – und dann der Vergleich mit den Eigenschaften digitaler Mischprozessoren. Denn, um es gleich vorwegzunehmen, es gibt eine eindeutige technische und theoretisch nachvollziehbare Erklärung für die klare Präferenz einer offensichtlich wachsenden Zahl von Anwendern für die 'alte' Analogtechnik. Gleich-

zeitig wird in diesem Artikel aber auch ein Ansatzpunkt deutlich, wie viele Diskussionen über scheinbar nicht zu erklärende Unterschiede zwischen analoger und digitaler Signalverarbeitung generell in eine ganz neue Richtung geleitet werden können.

Ralf Koschnicke

Jahrgang 1972, Studium der anwendungsorientierten Physik mit Wahlfächern Halbleitertechnik und Akustik an der TH Darmstadt; Diplom in Technologie Integrierter Schaltungen; private Instrumental- später Dirigierausbildung; schon während des Studiums Gründung einer Musikproduktionsfirma für Außenaufnahmen klassischer Musik und zahlreiche – zum Teil preisgekrönte (PdSK) – CD-Produktionen. Nun seit etwa 12 Jahren Inhaber der ACOUSENCE recordings. Inzwischen zunehmend Konzentrierung der Arbeit im eigenen Klassiklabel ACOUSENCE classics. Neueste Veröffentlichung, quasi bei konsequenter Nutzung der hier beschriebenen Erkenntnisse: Eine CD und 24Bit/96kHz DVD-V mit Werken von Hans Rott und Gustav Mahler kombiniert in einem Produkt.



Ausgangspunkt

Prinzipiell wird die digitale Audiotechnik viel zu sehr anhand statischer Modelle beziehungsweise anhand stationärer Schwingungen (periodische Sinusschwingungen) beschrieben. Bei dem in unserer Branche meist zu übertragenden Nutzsignal handelt es sich aber um Musik und Musiksignale spielen sich eigentlich nie im stationären Bereich einer Schwingung ab – jener Signaltyp für die Modellrechnung beziehungsweise die Messtechnik hat, um es mal ein wenig salopp auszudrücken, mit Musiksig

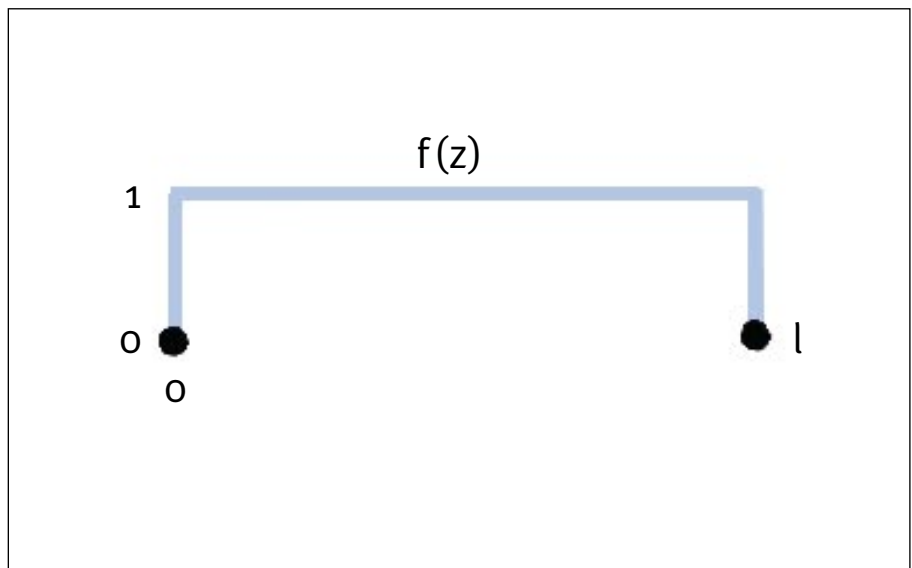
In der Analogtechnik wurde zwar auch nicht generell explizit das zeitliche Verhalten des Übertragungssystems betrachtet – wenngleich Begriffe wie Impulsverhalten und Anstiegsgeschwindigkeit noch vielen bekannt vorkommen dürften –, allerdings implizierten die gängigen Qualitätskriterien Frequenz- und Phasengang dort ein entsprechend gutes Verhalten auch bei nicht stationären Signalen.

Da die Wichtigkeit des „vertikalen“ Auflösungsvermögens im Bereich der Signalamplitude, sprich Rauschabstand oder auch Quantisierung, wohl unumstritten ist und auch keine ausreichende Erklärung für das zu untersuchende Phänomen liefert, soll hier nur das zeitliche („horizontale“) Auflösungsvermögen des Übertragungssystems betrachtet werden.

Zentral für das weitere Verständnis ist nun zunächst einmal grundsätzlich festzuhalten, dass ein immer gültiges physikalisches Grundgesetz (ähnlich beispielsweise wie der Energieerhaltungssatz) eine feste Beziehung zwischen Übertragungsbereich (Frequenzumfang) und der zeitlichen Auflösung¹

Fouriertransformation

Fourier beschreibt mathematisch, dass jede periodische Signalform als Überlagerung diskreter periodischer Sinusschwingungen dargestellt werden kann. Das Ausgangssignal für unser Beispiel soll nebenstehende Rechteckschwingung sein.

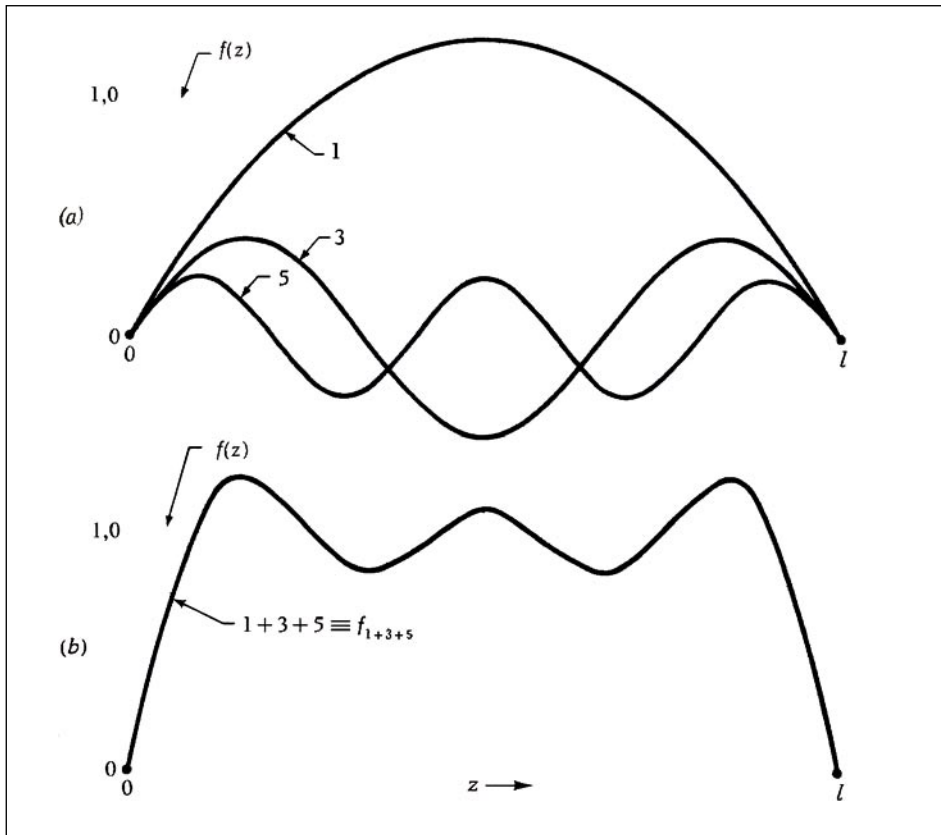


Rechteckwelle ($\lambda/2$ einer periodische Rechteckwelle)

vorgibt. Wegen der zentralen Bedeutung dieses Punktes nähern wir uns der Thematik nun zunächst gleich von zwei Seiten, aus einer eher anschaulichen Richtung und von einer eher theoretischen:

¹ Durch den Begriff „zeitliche Auflösung“ wird die Fähigkeit beschrieben, kleine Strukturen im Signal in der zeitlichen Ebene, in ihrem Verlauf, korrekt darstellen zu können).

Aus drei einzelnen Sinusschwingungen (Partialschwingungen) zusammengesetzt, sieht das Signal dann wie folgt aus:



oben die ersten drei Partialschwingungen separat,
unten das daraus resultierende Summensignal

Theoretisch setzt sich das Originalsignal gemäß folgender Gleichung aus unendlich vielen Partialschwingungen zusammen:

$$\begin{aligned}
 F(z) &= B_0 + \sum_{m=1}^{\infty} B_m \cos mk_1 z + \sum_{m=1}^{\infty} A_m \sin mk_1 z \\
 &= \frac{4}{\pi} \left\{ \sin k_1 z + \frac{1}{3} \sin 3 k_1 z + \frac{1}{5} \sin 5 k_1 z + \dots \right\} \\
 &= 1,273 \sin \frac{\pi z}{l} + 0,424 \sin \frac{3 \pi z}{l} + \\
 &\quad + 0,255 \sin \frac{5 \pi z}{l} + \dots
 \end{aligned}$$

Unsere auf drei Partialschwingungen reduzierte Darstellung wird beschrieben durch:

$$g(z) = 1,273 \sin \frac{\pi z}{l} + 0,424 \sin \frac{3 \pi z}{l} + 0,255 \sin \frac{5 \pi z}{l}$$

Zwei Dinge lassen sich nun anschaulich sehr leicht verstehen:

1.) Es kann in einem Übertragungssystem kein Signal dargestellt werden, welches kürzer ist, als die halbe Wellenlänge der im System maximal möglichen Schwingungsfrequenz.

2.) Um dann bei entsprechend gegebener Strukturgröße noch einen Signalverlauf mit einer höheren Flankensteilheit als ein Sinus abbilden zu können, benötigt man zusätzlich noch ungeradzahlig Vielfache der Grundfrequenz. Wobei auch zu erkennen ist, dass der Beitrag der Partialschwingungen zum Gesamtsignal mit zunehmender Ordnung abnimmt.

Die Unschärferelation in der Audiotechnik

Dieser Abschnitt liefert nun eine wissenschaftlich umfassendere Erklärung für den im letzten Abschnitt dargestellten Zusammenhang:

Generell betrachtet man die Eigenschaften technischer Übertragungssysteme gerne anhand deren Verhalten bei der Durchleitung eines Impulses, dem so genannte Dirac-Stoß, durch das System (diesen Impuls kann man sich bildlich auch als halbe Periode eines ultrakurzen Rechtecksignals vorstellen). Der Impuls ist ein nichtperiodisches Signal und nach Fourier enthält er Partialschwingungen mit allen Frequenzen kontinuierlich bis ins Unendliche.

Nun beschreibt ein ganz allgemein gültiges physikalisches Grundgesetz unter anderem auch den Zusammenhang zwischen Übertragungsbandbreite und zeitlichem Auflösungsvermögen (siehe Heisenbergsche Unschärferelation in der Quantenphysik):

$$\Delta f \Delta t \approx 1$$

Wobei Δt die zeitliche Dauer des Impulses und Δf die Bandbreite der harmonischen Partialschwingungen bezeichnet, aus denen sich der Impuls zusammensetzt. Hieraus folgt nun, dass jedes Signal, dessen Bandbreite begrenzt ist, auch eine gewisse Breite besitzt. Ein realer Impuls besitzt somit immer eine bestimmte zeitliche Dauer, die unmittelbar mit der Bandbreite des Übertragungssystems korreliert.

Allgemein gilt: Je kleiner die Bandbreite, desto größer die zeitliche Ausdehnung des Impulses - und umgekehrt.

Technisch real existierende Übertragungssysteme sind nun immer in ihrer Bandbreite begrenzt und somit auch nicht in der Lage, unbegrenzt fein zeitliche Strukturen aufzulösen.

Nyquist-Shannon-Abtasttheorem

In der analogen Audiotechnik sind die limitierenden Faktoren zunächst die Mikrofone sowie (eher schon Vergangenheit) vor allem

die Bandaufzeichnung und die LP als ehemaliger Tonträger für den Endkonsumenten, wobei gleich wichtig anzumerken ist, dass bei diesen Gerätegruppen der in Datenblättern angegebene Übertragungsbereich, bezogen auf ein in der Amplitude lineares Übertragungsverhalten, und der tatsächlich übertragene Frequenzbereich stark differieren können. Denn es gibt keine Grenzfrequenz, oberhalb derer das Übertragungsmaß mit einer hohen Steilheit „schlagartig“ stark absinkt. Deshalb verfügen solche Geräte auch ohne großen linearen Übertragungsbereich noch über ein sehr viel besseres zeitliches Auflösungsvermögen; eine gedämpft übertragene Partialschwingung ist immer noch besser als gar keine.

In der digitalen Audiotechnik sieht dies nun ganz anders aus. Dort wird der Übertragungsbereich gezielt möglichst hart bei

tenen zeitdiskreten Signal das Ursprungssignal ohne Informationsverlust (aber mit unendlich großem Aufwand) rekonstruiert bzw. (mit endlichem Aufwand) beliebig genau approximiert werden kann.

In der Praxis bedeutet das Abtasttheorem nun, dass man vor der Abtastung die maximale Frequenz kennen oder herausfinden muss (zum Beispiel mit Hilfe der Fourier-Analyse eines hochfrequent abgetasteten Signals) und dass dann das Signal mit mehr als der doppelten Frequenz abgetastet werden muss, wenn man das Signal in guter Näherung rekonstruieren will.

Diagramm 1 zeigt nun das Frequenzspektrum eines Mikrofonsignals (Symphonieorchester mit Chor in einer großen Domkirche, das heißt viele Raumanteile, Hauptmikro DPA4006) aufgezeichnet bei einer Abtastrate von 96kHz. Diagramm 2 zeigt zum

folgt angewendet: Zu Beginn wurde erst gar nicht der Versuch unternommen, das real vorliegende Signal zu analysieren und danach die Abtastrate zu wählen, sondern die Regel wurde gemäß den eigenen Zielvorgaben, gewürzt mit etwas Zweckoptimismus und Geschäftsinteresse, umgestellt: Zuerst wird das Signal knapp oberhalb von 20 kHz hart beschnitten, um somit dann quasi erst in Konsequenz die Randbedingungen des Theorems zu erfüllen.

Wissenschaftlich betrachtet ist dies aber nur zulässig, wenn man zweifelsfrei beweisen kann, dass diese Signalanteile über 20 kHz, die ja immerhin im Ausgangssignal deutlich vorhanden sind, andererseits nicht nötig sind. Wissenschaftlich ist dieser Umstand aber nicht belegt, stattdessen gibt es, wie gleich gezeigt wird, gegenteilige Erkenntnisse. Einziges Argument

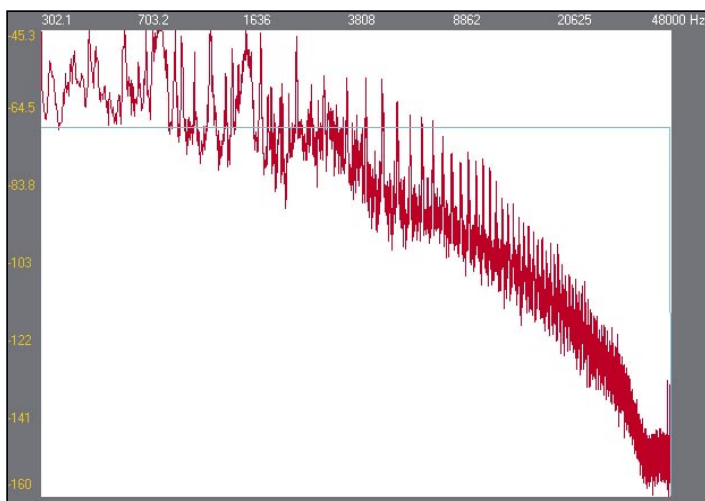


Diagramm 1 (Musik)

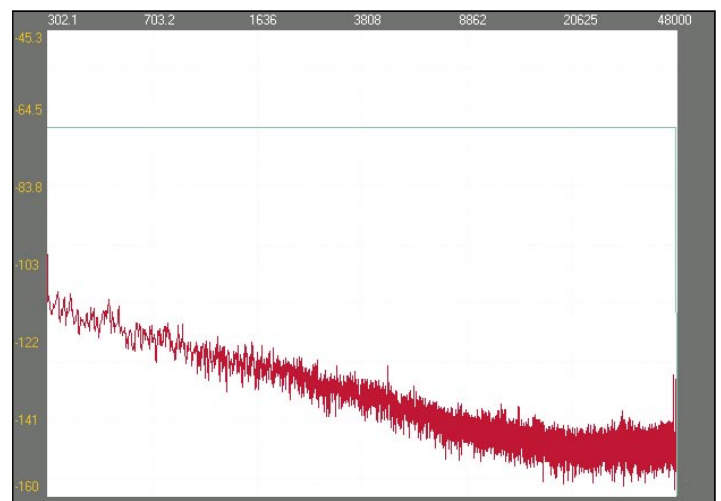


Diagramm 2 (Musikpause)

der Hälfte der Abtastfrequenz (f_s) bzw. schon knapp darunter beschnitten.

Somit gibt es garantiert keine Partialschwingung oberhalb $f_s/2$, wodurch wir dann auch gleich beim ersten ganz wesentlichen Fehler sind, mit dem jede heute erhältliche, mit Standard-Abtastern arbeitende digitale Audiotechnik behaftet ist: Die mathematische Grundlage für die Konstruktion eines digitalen Übertragungssystems liefert das berühmte *Nyquist-Shannon-Abtasttheorem*. Hier lohnt sich ein Blick auf die Interpretation seiner Urform:

Das Abtasttheorem besagt, dass ein kontinuierliches, bandbegrenzt Signal mit einer Minimalfrequenz von 0 Hz und einer Maximalfrequenz f_{max} mit einer Frequenz größer als $2 \cdot f_{max}$ abgetastet werden muss, damit man aus dem so erhal-

Vergleich das Signal des gleichen Mikrofons in einer Musikaufnahme.

Im Vergleich ist nun deutlich zu erkennen, dass bei Vorhandensein von Musik Signalanteile bis über 40kHz vorhanden sind, die sich deutlich vom allgemeinen Geräuschteppich abheben. (Möglicherweise sind auch noch darüber Signalanteile existent, denn bei der Abtastrate 96kHz ist hier ja bereits der Anti-Aliasingfilter des AD-Wandlers im Spiel.)

Rein formal richtig ausgelegt, verbietet das Abtasttheorem also für die Musikübertragung Abtastern unter 80 kHz. Bezieht man noch andere technische Notwendigkeiten mit ein, liegt man mit den heute standardisierten 96 kHz sicher richtig.

Bisher wurde das Theorem aber stets wie

war lediglich die Ansicht, der begrenzte Hörwahrnehmungsbereich eines menschlichen Ohres von maximal 20 kHz mache die technische Übertragung höherer Frequenzen überflüssig.

Das „Zurechtbiegen“ der Rahmenbedingungen ist nicht zulässig und, streng wissenschaftlich betrachtet, grober Unfug!

Das zeitliche Auflösungsvermögen beim menschlichen Hören

Die kleinste auf einer CD darstellbare Struktur kann nicht kürzer als 22 μ s sein. Das klingt zunächst noch sehr gut, aber wird den Fähigkeiten des menschlichen Hörsystems nicht gerecht.

Allgemein wird der Hörbereich des Menschen auf 20 Hz-20 kHz festgelegt. Dar- aus folgerten dann die Erfinder der CD die

Abtastrate von 44,1 bzw. 48 kHz für digitale Audioübertragungssysteme. Könnte man den Zusammenhang zwischen Bandbreite und zeitlichem Auflösungsvermögen



eins zu eins auf den Menschen übertragen, wäre es auch tatsächlich widersinnig, von technischen Audioübertragungssystemen – gleich, ob analog oder digital – eine höhere zeitliche Auflösung zu fordern, da der Mensch diese dann ohnehin nicht wahrnehmen könnte. Diese seit Erfindung der Digitaltechnik proklamierte Festlegung muss aus heutiger Sicht jedoch deutlich in Frage gestellt werden.

Natürlich gilt die Physik auch bezogen auf das menschliche Hören. Die Bandbreite des menschlichen Gehörs wird anhand von Hörtests mit Sinusschwingungen auf maximal 20 kHz (Erwachsener höchstens 17 kHz) festgesetzt. Allerdings darf beim menschlichen Hören nie der „psychische Filter“, nämlich unser Gehirn, vernachlässigt werden. Das Gehirn ist Teil des menschlichen Übertragungssystems und bestimmt die Hörwahrnehmung immer mit.

Neuere wissenschaftliche Untersuchungen belegen beispielsweise durchaus die Fähigkeit, die Anwesenheit von eigentlich unhörbar hohen Frequenzen „spüren“ zu können. Hingegen kann das Ohr zwei zeitlich diskrete Signale nur mit einem zeitlichen Abstand größer 10 ms bewußt als zwei Signale voneinander trennen. Wiederum, wenn es um Richtungswahrnehmung geht, sind Winkelabweichungen von nur 3-5 Grad exakt wahrnehmbar. Das entspricht einer Laufzeitdifferenz zwischen beiden Ohren von etwa 10 μ s. Untersuchungen im Bereich Instrumentenakustik zeigen, dass der Mensch kleinste Signalstrukturen bis hinab zu einer Größenordnung von 5 μ s wahrnimmt und diese für einen natürlichen Höreindruck benötigt².

Rechnet man die 20 kHz der höchsten hörbaren Frequenz um in Zeitauflösung, ergibt

das 25 μ s. Rechnet man die 10 ms Zeitauflösung um in eine maximal hörbare Frequenz, so dürften wir keine Frequenzen über 50 Hz (das ist die Frequenz eines Netzbrum-

mens) hören. Gleichzeitig hören wir beim Bestimmen von Richtungen Zeitabstände von 10 μ s und nehmen in Klängen Signalstrukturen bis hinab zu 5 μ s wahr.

Insgesamt ist die Hörwahrnehmung wissenschaftlich noch nicht wirklich abschließend untersucht, wenn auch mittlerweile einige Hirnforscher auf diesem Gebiet tätig sind. Alle Beispiele zusammen genommen zeigen uns aber dennoch recht eindeutig, dass der physikalische Zusammenhang von Bandbreite und Zeitauflösung so nicht auf den Menschen angewendet werden darf. Der menschliche Hörbereich, die



Präzision in der Richtungswahrnehmung, die Empfindlichkeit bei der Bewertung von Klängen und das an sich begrenzte Selektionsvermögen bezüglich zeitdiskreter Ereignisse würden sonst in starkem Widerspruch zueinander stehen.

Evolutionswissenschaftler erklären übrigens das hohe zeitliche Auflösungsvermögen des Hörsystems dadurch, dass präzises Richtungshören für den Urmenschen als Jäger überlebenswichtig war. Auch die Art der Geräusche zu erkennen, war unter Umständen entscheidend über Leben und Tod. Für die bewusste Wahrnehmung extrem hoher Frequenzen zeigt uns die Evolutionsgeschichte allerdings keine Notwendigkeit.

² siehe Bericht zur 23. Tonmeistertagung 2004: Ernst-Joachim Völker – Direktschall von Musikinstrumenten

Die Konsequenzen für die Musikübertragung

Was in der theoretischen Betrachtung für den Impuls oder Rechtecksignale gilt, ist in der realen Audiotechnik genauso auf impulsähnliche Signale anzuwenden. Musik und Sprache gehören wegen ihrer komplexen Ein- und Ausschwingvorgänge generell zu dieser Kategorie. Stationäre Signale, wie reine Sinusschwingungen, gibt es in der Musik nicht.

Erst aus einer strukturiert über die Zeit ablaufenden Klangverteilung wird Musik. Jedem Musiker ist ganz natürlich klar, wie wichtig nicht nur ein in Intonation sicheres, sondern auch ein zeitlich präzises Spiel ist. Daneben bilden aber erst kleinste Variationen in Lautstärke und zeitlichem Ablauf, in Phrasierung und Artikulation, den Ausgangspunkt für ein musikalisch ansprechendes Spiel mit Seele; den Unterschied zwischen Mensch und Maschine. Überdies ist auch die zeitliche Struktur der Klänge für deren Wiedererkennung als natürlich entscheidend.

Beim natürlichen Hören sind wir gewohnt, diese zeitlichen Abläufe so wahrzunehmen, wie sie uns von den natürlichen Klangerzeugern angeboten werden. Unser Hörsystem besitzt dabei wie erwähnt prinzipiell ein sehr gutes Auflösungsvermögen.

Vergleich „analog“ – „digital“

Analoge Bandmaschinen erreichen zwar meist keine 25 kHz Bandbreite mit im Pegel linearem Frequenzgang. Der Übertragungsbereich wird aber nicht rapide begrenzt, wie in jedem digitalen Gerät, sondern der Übertragungspegel sinkt erst ganz allmählich. Die Impulsübertragung ist damit zwar nicht ideal, aber immer noch deutlich besser als in der Digitaltechnik mit herkömmlichen Abtastraten; gleiches gilt wieder auch für das zeitliche Auflösungsvermögen. Ein Beispiel:

Die analoge Studiotonbandmaschine Otari MTR12 beispielsweise besaß eine Phasenkompensationsschaltung zur Verbesserung der Impulsübertragung. Diese Schaltung war mit Hilfe eines 10 kHz-Rechteckprüfsignals zu justieren.

Schickt man dieses Prüfsignal aber durch eine digitale Übertragungstrecke bei 44,1 kHz oder 48 kHz, bleibt nur noch ein reiner Sinus übrig (!), weil schon das erste ungeradzahlige Vielfache der Grundfrequenz (d.h.

30 kHz) nicht mehr übertragen werden kann. Die Entwickler dieser Maschine haben also viel Aufwand darauf verwendet, die Maschine in einem Punkt zu verbessern, in dem herkömmliche Digitaltechnik von vornherein völlig versagt.



Erst gute hochauflösende Digitaltechnik mit mindestens 96 kHz schafft ein zeitliches Auflösungsvermögen, das mit dem einer analogen Bandmaschine in etwa vergleichbar ist.

Spezialfall Mischpult

Ganz unabhängig von der Diskussion, welche Abtastrate für die Tonaufzeichnung zu benutzen ist, muss das Bearbeitungswerkzeug *Mischpult* betrachtet werden. Ein Mischpult ist wie kein anderes Gerät in der Audiotechnik vorrangig dazu konstruiert, einzelne Signale zu einem Ganzen zu summieren. Dabei ist selbstverständlich nicht nur die Genauigkeit in der Darstellung der Einzelsignale und deren Addition bezogen auf die Lautstärkeverhältnisse wichtig, sondern auch deren korrekte zeitliche Zusammenführung liefert erst das genaue akustische Abbild des realen Geschehens und sollte deshalb auch möglichst exakt erfolgen.

Beispielsweise werden die Signale vieler Instrumente, die ein Hörer im Schallfeld eines Orchesters im Raum hört, zusammen mit den Raum- und Hallsignalen ohne zeitliches Raster am Ohr des Hörers wahrgenommen (abgesehen von den Grenzen, die die Luftmolekülbewegungen setzen). Folglich sollten verschiedene Mikrofonsignale, die diese Instrumente und den Raum aufzeichnen, wie selbstverständlich möglichst auch ohne zeitliches Raster addiert werden.

Üblicherweise wird die Funktionsweise digitaler Mischpulte in der Technik nur statisch betrachtet oder deren Verhalten wird theoretisch nur anhand stationärer Signale untersucht. Hauptsächlich wird der Blick auf eine einwandfreie Funktion bezüglich der Pegeladdition und die Verzerrungsarmut der Einzelsignale gelenkt. Die zeitliche Präzision glaubt man über die Summierung der Phasenlage zu erreichen.

Das zeitliche Auflösungsvermögen wird nun aber von der Übertragungsbandbreite begrenzt. Und die ist in einem digitalen Mischpult heutiger Technologie sehr begrenzt. Selbst die aufwendigsten Mischpulte im derzeitigen Industrie- und Rundfunkstan-

dard mit Abtastraten von 48 kHz liefern nur eine Bandbreite von max. 24 kHz.

Analoge Verstärkertechnik – eben auch Mischpulte – liefert leicht Bandbreiten von einigen hundert Kilohertz, was nach obiger Gleichung gleichzusetzen ist mit einer zeitlichen Auflösung von Strukturen weitaus kleiner als 5 μ s. Beispielsweise überträgt ein gutes Analogpult ein 80 kHz Rechtecksignal noch deutlich als solches erkennbar (zwar mit abgerundeten Flanken aber noch insgesamt sehr hoher Flankensteilheit). Ein Digitalpult bei 48 kHz Abtastrate (egal welcher Güte und Preisklasse) überträgt noch nicht einmal ein 8 kHz Rechteck. Stattdessen sieht man am Ausgang nur einen 8 kHz Sinus! Erst Abtastraten von über 500 kHz würden hier mit analoger Audiotechnik vergleichbare Werte liefern.

Ein Erklärungsversuch

Wie kam es nun dazu, dass trotz solcher deutlicher Unterschiede in den Geräteeigenschaften die zeitliche Auflösung nie thematisiert wurde. Allgemein anerkanntes Entwicklungsziel bei analoger Audiotechnik war neben

einem linearen Amplitudenfrequenzgang im Hörbereich auch ein möglichst stabiler Phasengang, das heißt im Idealfall konnte das System alle Frequenzen zwischen 20 Hz und 20 kHz mit der selben Amplitude und Phasenlage übertragen. Die Phasenlage ist aber wiederum auf physikalischem Wege direkt verknüpft mit dem Übertragungsbereich, denn nur ein großer Übertragungsbereich kann eine stabile Phasenlage liefern. Gleichzeitig folgt, wie wir oben gezeigt haben, aus einem großen Übertragungsbe-

reich wiederum auch ein hohes zeitliches Auflösungsvermögen. Hochwertige Analogtechnik mit gutem Phasengang hatte somit immer auch eine gute Zeitauflösung beziehungsweise Impulstreue.

Das Entwicklungsziel Phasenlage war kein Selbstzweck, denn in realen Signalen gibt es keine Phasenlage und das menschliche Ohr kann keine Phase auswerten, sondern erst die damit verknüpfte Eigenschaft Zeitauflösung liefert die Übertragungsgüte. Andererseits lässt sich die Phasenlage messtechnisch sehr leicht überprüfen und wurde deshalb Qualitätskriterium für analoge Audiotechnik. In der digitalen Audiotechnik ist die Phasenlage nun eigentlich kein Qualitätskriterium, weil hiermit nur Signale überprüft werden, mit denen Digitaltechnik grundsätzlich nie Probleme hat, die aber andererseits mit realen Musik- und Sprachsignalen wenig gemein haben. Dennoch wurden die beiden Kriterien *menschlicher Hörbereich = Übertragungsbereich des technischen Systems und stabiler Phasengang innerhalb des Übertragungsbereichs* nun bei Einführung der digita-



len Audiotechnik eins zu eins, als hinreichend die Übertragungsqualität beschreibend, übernommen. Aus dem menschlichen Hörbereich folgten die Abtastraten von 44,1

hier nur ein allgemeiner, vom Signal modulierter Klangteppich ohne präzise Struktur übertragen. Diese Art von Verfälschungen sind nicht unbedingt für jeden sehr of-

den, sehe ich übrigens auch den Hauptgrund für die Renaissance hochkarätiger alter Aufnahmen auf modernen Tonträgern, die in Atmosphäre und „Spaß an der Musik“ von modernen Aufnahmen nur selten zu übertreffen sind. Manche Musikliebhaber auf Konsumentenseite handeln also schon längst.

Eigentlich einen eigenen Artikel wert, will ich hier dennoch kurz darauf aufmerksam machen: Es lohnt sich ebenfalls auch eine Bewertung datenreduzierter Formate unter den in diesem Artikel angesprochenen Gesichtspunkten. Denn hier passieren mit der Musik noch ungleich gravierende Dinge, auch wenn sich oberflächlich betrachtet manches klanglich gar nicht so schlecht anhört.



bzw. 48 kHz. Die Phasenstabilität innerhalb dieses Hörbereichs war für digitale Übertragungssysteme systembedingt leicht zu erfüllen, betrachtet man wie üblich Sinusschwingungen, allerdings ohne gleichzeitig noch nicht einmal annähernd auch die gleiche Zeitauflösung und Impulsübertragung zu gewährleisten.

Die Bewertung der damals neuen Digitaltechnik anhand von Qualitätskriterien der alten Analogtechnik – die so bis heute noch anhält – täuscht völlig falsche Tatsachen vor und demonstriert damit die scheinbare Überlegenheit. Auf dem für die Musikübertragung wichtigen Sektor Zeitauflösung bewegen wir uns aber tatsächlich technologisch auf dem Stand der 1950er Jahre.

Die Auswirkungen

Neben unbestreitbaren Fortschritten, vor allem in Handhabung und Funktionalität, gab es also mit der Entwicklung der digitalen Audiotechnik im letzten Vierteljahrhundert auch einige schwerwiegende Rückschritte, gemessen an der „gefühlten“ Audioqualität und Musikalität.

Die digitale Aufzeichnung arbeitet überhaupt nur so gut wie wir sie kennen, weil für Frequenzanteile unterhalb etwa 10 kHz die Theorie ausreichend gut auch in der Praxis funktioniert. Der größte Teil realer Klänge spielt sich in diesem Bereich ab und auch das menschliche Ohr ist in diesem Bereich am empfindlichsten. Hauptsächlich verfälscht wird der Bereich der Ober-töne und der feinen räumlichen Wahrnehmung, denn bei genauer Betrachtung wird

fensichtlich hörbar; insbesondere ohne direkten Vergleich zum Ursprungssignal. Der Klang leidet nur im Detail, die Musik und Atmosphäre einer Aufnahme leiden aber meist deutlich.

Meiner Erfahrung nach gehen, zumindest bei akustisch instrumentierter Musik, sehr leicht Anteile in beträchtlichen Ausmaß verloren, die eigentlich zwingend zum Gesamtbild der künstlerischen Aussage des Interpreten gehören. Auch verliert man reichlich tiefgründige Charakteristika der Instrumentenklänge. Und leider vermisst man

Der eigentliche Vorteil eines hochauflösenden Digitalsignals mit hohen Abtastraten liegt nicht primär in der Erweiterung des Übertragungsbereichs, sondern erst ein sekundärer Effekt bringt den entscheidenden Vorteil. Je höher die Übertragungsbandbreite eines technischen Übertragungssystems, desto besser sein zeitliches Auflösungsvermögen. Erst eine Abtastfrequenz von 96 kHz erfüllt formal das Nyquist-Shannon-Abtasttheorem und ist in punkto Zeitauflösung in etwa vergleichbar mit einer professionellen analogen Tonaufzeichnung. Damit erfüllt erst



diese Anteile nicht zwangsläufig, solange man es nicht besser kennt oder mal wieder ein Instrument „von Angesicht zu Angesicht“ hört.

In der weitgehend analogen Produktion auf einem Qualitätsniveau, bei dem diese Details schon sehr exakt abgebildet wer-

mindestens diese Abtastfrequenz alle Bedingungen, die man an professionelle Audiotechnik im Bereich Schallaufzeichnung stellen muss und bis zur Erfindung der digitalen Audiotechnik Anfang der 1980er Jahre auch gestellt hat. Im Bereich Signalverarbeitung einzelner Signale³ sind zum Teil noch hö-

Fazit

here Abtastraten zu empfehlen.

Bei Mischpulten würden erst Abtastraten von über 500 kHz mit Analogtechnik vergleichbare Ergebnisse liefern. Darin liegt dann auch die Erklärung für die doch deutliche Überlegenheit analoger Mischungen gegenüber digitalen; überall dort, wo es noch auf eine facettenreiche naturgetreue Musikalität und einen möglichst farbenreichen Klang ankommt. Denn auch für den

In der Musikproduktion beziehungsweise überall, wo Kunst gestaltet wird, muss sich die Wahl des Werkzeugs – ob analog oder digital – dann vielleicht nur noch alleine anhand rein ästhetischer Aspekte beziehungsweise am persönlichem Geschmack des Nutzers orientieren.



dies zur Zeit nicht realisierbar ist, stattdessen Mischung analog mit einer klassischen Konsole, oder eben – als Kompromiss zwischen beiden Welten, wenn ein höherer Grad an Funktionalität und ein schonender Umgang mit dem Budget gefordert ist – eine digitale Workstation mit hohen Abtastraten in Verbindung mit einem guten externen Analogsummierer.

Übrigens lohnt sich dieses Vorgehen auch für die Herstellung von CDs. Zwar besitzt der Tonträger dann die genannte Einschränkung, aber



„gefühlten“ Klang ist die möglichst genaue Abbildung des originären Schallverlaufs in seiner komplexen spektralen Zusammensetzung wichtig.

Ein Ausblick

Etwas versöhnlich mit der Entwicklung kann man dennoch umgehen, wenn man berücksichtigt, dass auch die „alte“ Analogtechnik alles andere als perfekt ist, dort aber prinzipiell nur noch wenig Entwicklungspotential schlummert. Der Weg hin zur Digitalisierung stimmt grundsätzlich, nur das momentan übliche Niveau schließt nicht an den Entwicklungsstand der analogen Audioteknik der 1980er Jahre an. Nimmt man die hier beschriebenen Erkenntnisse auf, so gibt es auf der digitalen Ebene noch ein enormes Entwicklungspotential. Denn die zur Verfügung stehende Prozessorleistung wächst stetig und somit werden auch irgendwann – demnächst? – komplett digitale Signalstrecken möglich, die rein technisch betrachtet der analogen Technik überlegen sind. Im Bereich der reinen Aufzeichnung ist dies im Falle hoher Abtastraten ja bereits heute der Fall.

Auch wenn die Massen bekanntermaßen einen anderen Weg eingeschlagen haben (datenreduzierte Formate: mp3 beim Konsumenten und mpeg, das mittlerweile schon als „CD-Qualität“ verkauft wird, bei Funk und Fernsehen), so möchte ich abschließend den Weg für den weiterhin qualitätsbewussten Teil unserer Branche zu prognostizieren versuchen: Aufzeichnung digital, grundsätzlich bei höheren Abtastraten (96 kHz oder 192 kHz)⁴ und damit höheren Audiobandbreiten. In der weiteren Signalverarbeitung sind nochmals höhere Abtastraten je nach Anwendung sinnvoll und in digitalen Mischpulten sogar notwendig. Da

zeichnet man bei 96 kHz auf und mischt analog, so hat man in der Stereosumme, wegen der höheren zeitlichen Präzision und den Misch- und Interferenzprodukten aus den eigentlich nicht direkt hörbaren höherfrequenten Anteilen, ein insgesamt natürlicheres Produkt. Man verliert später (nur) den Signalanteil, der auf der CD selbst nicht darstellbar ist. Und anhand eigener Experimente kann ich sicher beurteilen, dass der Verlust an Musikalität und Klang so weitaus geringer ausfällt, als bei einer digitalen Mischung bei Standardabtastraten. Dies kann jeder mal gerne selbst ausprobieren... ■

³ Um den Rahmen dieses Artikels nicht gänzlich zu sprengen, wurde das Thema Signalverarbeitung ausgeklammert. Betrachtet man aber das zeitliche Verhalten von digitalen Bearbeitungsgeräten, so wird man auch einige interessante Phänomene beobachten, zum Beispiel gilt Heisenberg auch ganz besonders bei so genannten phasenlinearen FFT-EQ's; phasenstarr für Sinussignale „Ja“, aber „zeitlich korrekt“?

⁴ für noch höhere Abtastraten bei der Aufzeichnung gibt es technisch kein Argument, denn spätestens die Lautsprecher setzen Grenzen und AD-Wandler arbeiten somit spätestens bei 192 kHz zeitlich gesehen exakt genug (siehe Beitrag von Mike Story und der Firma DCS Ltd.: A Suggested Explanation For (Some Of) The Audible Differences Between High Sample Rate And Conventional Sample Rate Audio Material – http://www.dcsLtd.co.uk/technical_papers/aes97ny.pdf)