

# Pilgrim Studio

Recording, Mixing & Mastering

Friedaustasse 17, 8003 Zürich

TEL: 079 744 61 19

WEB: [www.pilgrimstudio.ch](http://www.pilgrimstudio.ch)

E-MAIL: [Contact@pilgrimstudio.ch](mailto:Contact@pilgrimstudio.ch)

## Grundlagen der digitalen Audiotbearbeitung

### Teil 1

„Never turn your back on Digital!“

Bob Ludwig

## 1. Einleitung

Dieser Text liefert Grundwissen für das bewusstere Arbeiten mit digital gespeicherter Musik. Ich habe versucht den Text möglichst sinnvoll zu gliedern.

Eingerückte Textteile geben weiterführende Hintergrundinformationen.

Das Fettgedruckte ist die Minimal-Botschaft des Textes: es enthält Tips für die Verbesserung der Qualität Eurer digitalen Aufnahmen, liefert aber keine Erklärungen und hat deshalb oft die Form von „Du sollst!“.

### 1. 1. Vokabular

**ADC:** Abkürzung für Analog to Digital Converter: Ein Gerät das analoge Signale in digitale Signale umwandelt.

**Amplitude:** Die Intensität des Ausschlages einer Schwingung, was im Audiobereich der Lautstärke eines Signals entspricht.

**DAC:** Abkürzung für Digital to Analog Converter: Ein Gerät das digitale Signale in analoge Signale umwandelt.

**dB:** Dezibel ist die Einheit des Lautstärke-Pegels. Die Einheit Dezibel ist rein relativ, hat also keinen Fixpunkt; dies bedeutet dass eine Angabe in dB nur einen Sinn macht wenn der Wert auf einen anderen Wert bezogen wird. Die Aussage „Die Lautstärke ist 80 dB“ ist deshalb unsinnig, nicht aber die Aussage „Das Signal wurde um 6 dB erhöht“. Wird eine Dezibel-Angabe auf einen Fixpunkt bezogen, dann erhält das dB ein Suffix, also z. B. dBu, dBv, dBFS.

Eine Pegelerhöhung von 6 dB entspricht einer Lautstärke-Verdoppelung.

**dBFS:** Pegelangaben in dBFS beziehen sich auf den maximal möglichen Pegel in digitalen Systemen, also auf „Full Scale“ (FS). Pegelangaben in dBFS sind deshalb (fast) immer negativ – es gibt aber Fälle wo Pegel über 0 dBFS vorkommen.

**Headroom:** Der Bereich zwischen den maximalen Signalspitzen und dem Clipping Point.

**Jitter:** Verzerrungen, welche bei der Wandlung von digitalen Audiodateien entstehen, hervorgerufen durch minime zeitliche Unregelmässigkeiten bei der Abtastung.

**Lautheit:** deutsches Wort für englisch ‚**Loudness**‘. Lautheit und Loudness bezeichnen die empfundene Lautheit und sind *nicht zu verwechseln* mit **Pegel** bzw. **Level**, welche beide einen absoluten Wert auf einer Skala bezeichnen. Dies ist insbesondere wichtig um zu verstehen dass der Maximalpegel einer Aufnahme nicht darüber entscheidet wie laut die Aufnahme insgesamt ist. Wenn ein Schlagzeuger an einer Stelle besonders stark auf die Snare schlägt erscheint dadurch nicht die ganze Aufnahme lauter, aber der *Maximalpegel* ist höher.

Kompression und Limiting erhöhen die *Lautheit* einer Aufnahme, ohne dass sie den *Pegel* erhöhen. Im Gegenteil – wenn ich ein Signal komprimiere und die komprimierte Version gleich laut einstelle wie die nicht komprimierte Version, dann ist der Maximalpegel der komprimierten Version kleiner (das ist ja der Zweck von Kompression).

**Gain** wiederum bezeichnet eine *Pegeländerung*; der Ausdruck Gain wird beispielsweise bei einem Verstärker benutzt um anzugeben um wie viel der Pegel vom Verstärker erhöht wird<sup>1</sup>. Es ist nützlich sich über diese Begriffe im klaren zu sein und sie richtig anzuwenden.

**SRC:** Abkürzung für Sample Rate Converter: Ein Gerät das die Samplerate, d.h. die Frequenz des Abtasters, ändert.

## 1. 2. Grundsätzliches

Grundsätzlich gilt:

**Bei einem Mix kumulieren sich sämtliche Verschlechterungen des Klanges. Was vielleicht bei einer einzelnen Spur noch knapp durchgeht ist nicht mehr akzeptabel, wenn diese Verschlechterung auf allen Spuren, welche zusammengemischt werden, vorhanden ist. Deshalb ist es wichtig, dass auf eine saubere Signalführung geachtet wird und Fehler, welche sich kumulieren, so weit wie möglich vermieden werden.**

## 2. Bits

### 2.1. Was sind Bits?

Bit ist die Abkürzung für *binary digit* („zweiwertige Zahl“). Ein Bit ist eine binäre (zweiwertige) Informationseinheit – sein Wert ist entweder 1 oder 0 (Null).

Das Binärsystem wurde von Leibnitz eingeführt, und Leibnitz wiederum entdeckte das Binärsystem bei seiner Beschäftigung mit dem chinesischen Orakelbuch I Ging, das eine Wahrsage-Methode beschreibt, deren Grundlage das Binärsystem ist. Die Ursprünge dieser chinesischen Wahrsage-Methode verlieren sich in der Urgeschichte.

Auf dem Binärsystem beruht die ganze heutige Computertechnik. Auch in unserem Alltagsleben spielt die zweiwertige (binäre) Logik eine zentrale Rolle: Es ist die Logik von „Ja oder Nein“. Wichtig für die digitale Technik ist nun aber, dass man mit binären Zahlen sehr gut und einfach rechnen kann – und das geht weit über „Ja oder Nein“ hinaus.

Im Binärsystem wird jede Zahl durch eine Reihe von Einsen und Nullen dargestellt:

---

<sup>1</sup> Noch genauer betrachtet verstärkt aber ein Verstärker gar nicht, sondern er produziert ein Abbild mit grösserer Amplitude ;-)

0 ist 0  
1 ist 1  
2 ist 10  
3 ist 11  
4 ist 100  
5 ist 101  
6 ist 110  
7 ist 111  
8 ist 1000.

Auffällig ist, dass die Potenzen von Zwei immer eine Eins mit x Nullen ist, wobei x die Potenz ist (das ist nur ein Beispiel für die Eleganz des Binärsystems – Leibnitz hielt diese Eleganz für ein Zeichen, dass das Binärsystem grundlegende göttliche Prinzipien des Kosmos widerspiegelt bzw. aufzeigen kann).

Damit man die Binärzahlen von normalen Zahlen unterscheiden kann werden sie mit einem % davor geschrieben: %1001 ist demnach eine Binärzahl (und entspricht der Dezimalzahl 9).

Man kann mit Binärzahlen ganz normal rechnen: wenn ich beispielsweise  $3 + 5$  im Binärsystem rechne, dann zähle ich %11 und %101 nach folgenden Regeln zusammen: ich beginne bei der hintersten Stelle und zähle die beiden Zahlen zusammen. Ist das Resultat Zwei (was es ja im Binärsystem nicht gibt!), dann erfolgt ein Übertrag auf die nächste Stelle (ganz genauso wie bei  $9 + 1 = 10$ ). %1 plus %1 ergibt somit %10. Dann fahre ich mit der nächsten Stelle der Zahlen fort, indem ich diese und ggf. den Übertrag addiere – etc.  $3 + 5$  ist im Binärsystem also %11 + %101 und ergibt %1000:  
Hinterste Stelle: %1 plus %1 ergibt %10. die hinterste Stelle des Resultates ist also 0, und wir haben einen Übertrag 1.  
Zweitletzte Stelle: 1 plus 0 plus Übertrag 1 ergibt 0, Übertrag 1.  
Drittletzte Stelle: 1 plus Übertrag 1 ergibt 0, Übertrag 1.  
Viertletzte Stelle: Übertrag von der drittletzten Stelle: 1.  
Alles klar?

Wichtig für unser Thema ist, dass Verrechnungen im Binärsystem sehr schnell zu mehr Stellen führen. Jede Stelle ist ein Bit. Wenn ich zwei 16-Bit-Werte zusammenzähle erhalte ich als Resultat eine Zahl, welche mehr als 16 Bit aufweist. Was mit den überzähligen Bits geschieht entscheidet massgeblich über die Klangqualität unserer Musik.

Mit einem einzigen Bit können zwei Werte wiedergegeben werden – eben 1 oder 0. Bei zwei Bits ergeben sich  $2^2 =$  vier verschiedene mögliche Werte: 00, 01, 10, 11. Bei drei Bits sind es  $2^3 = 8$  verschiedene mögliche Werte, bei 8 Bit bereits  $2^8 = 256$  verschiedene mögliche Werte. 16 Bit liefert  $65'536$  verschiedene mögliche Werte. 24 Bit liefert  $16'777'216$  verschiedene mögliche Werte.

Was bedeutet dies für die digitale Signalübertragung?

Entgegen einer verbreiteten Meinung führen mehr Bits *nicht* dazu dass ein Signal genauer abgebildet werden kann. Was zutrifft ist: Je mehr Bits verwendet werden desto grösser wird die mögliche Dynamik und der mögliche Rauschabstand. Die besten analogen Geräte erreichen heute einen Rauschabstand welcher etwa 22,5 Bits entspricht. Da sich Rauschen summiert und bei der üblichen Audiotbearbeitung mehr als eine analoge Stufe im Signalweg liegt (bereits der analoge Teil des Wandlers ist eine solche Stufe, dazu kommen Mikrofone, Mikrofon-Vorverstärker, EQs, Kompressoren, Mischpulte etc.) ist ersichtlich dass der Rauschabstand eines 24 Bit Wandlers bei weitem ausreicht.

Ein grösserer Dynamikbereich führt mit sich dass Signal-Anteile mit geringem Pegel (Fades, Hall- und Raumanteile, Ausklang etc.) klarer abgebildet werden. Dies führt dazu dass der Sound natürlicher ist, kann aber auch Mängel im Signal aufdecken. Ein 'natürlich klingendes' Rauschen wird von vielen Hörern als durchaus musikalisch empfunden. Trotzdem ist es ratsam die grösstmögliche Auflösung zu wählen, um möglichst viele Informationen des Signals zu bewahren.

Dazu kommt dass eine höhere Bit-Auflösung bei der digitalen Bearbeitung von Audio-Signalen von Vorteil ist.

Aus alledem folgt:

**Wählt wann immer möglich eine möglichst grosse Bit-Breite, sowohl bei der Aufnahme wie auch intern bei der Signalbearbeitung! Eine Auflösung von 16 Bit sollte nur das Endprodukt aufweisen – der CD-Premaster *nach dem Mastering*.**

### 3. Aufnahme

#### 3.1. Die Wandlung

Ein ADC (Analog to Digital Converter) wandelt analoge Signale, die als Wechselspannung am Eingang anliegen, in digitale Daten um.

Ein DAC (Digital to Analog Converter) wandelt digitale Daten in analoge Signale (in Form von Wechselspannungen) um.

Ein SRC (Sample Rate Converter) ändert die Abtastfrequenz, mit welcher digitale Signale verarbeitet werden.

Die Änderung der Abtastfrequenz fordert sehr hochwertige Wandler; sofern nicht ein wirklich erstklassiger SRC vorhanden ist empfiehlt es sich, Sampling Rate Conversion gänzlich zu unterlassen oder das Signal mit der ursprünglichen Sampling Rate analog auszuspielen (zum Beispiel auf eine Bandmaschine) und mit der Ziel-Sampling-Rate wieder neu digital aufzunehmen – die Qualität ist meistens besser als wenn ein mittelmässiger SRC verwendet wird.

Im folgenden wird die Wandlung der Sample Rate nicht weiter thematisiert, sondern nur die Wandlung von Analog zu Digital sowie von Digital zu Analog.

Die zur Zeit meistverbreitete Art der Wandlung ist die Pulse Code Modulation, kurz PCM.

Eine spannende, aber leider nicht sehr verbreitete Alternative ist das DSD (Direct Stream Digital)-Verfahren von Sony, das in der SACD und in den Wandlern von Pacific Intersonic (HDCD, heute im Besitz von Microsoft) eingesetzt wird. Diese Wandler sind sehr teuer (zwei Kanäle kosten um die 10'000 Franken) und mit den meisten heute gebräuchlichen digitalen Aufnahmesystemen nicht kompatibel – aber der Klang ist deutlich besser...

Aber wie gesagt ist die Pulse Code Modulation der heutige Standard (sie wird auch auf der CD eingesetzt).

Bei der Pulse Code Modulation wird der analog anliegende Spannungsverlauf des Signals in regelmässigen zeitlichen Abständen abgetastet.

Das Resultat jedes Abtastvorgangs ist der Wert der Spannung des Signals zum Zeitpunkt der Abtastung. Ein solches Resultat wird *Sample* genannt.

Wie oft pro Sekunde ein Sampling erfolgt ist durch die Sampling Rate festgelegt. Die Sampling Rate einer CD liegt beispielsweise bei 44,1 kHz, d.h. es erfolgen 44'100 Abtastvorgänge pro Sekunde.

Das Resultat dieses Abtastvorgangs ist eine Treppenkurve. Die Höhe der Treppenstufen entspricht dem gesampelten Wert zum jeweiligen Zeitpunkt.

Dieser Abtastvorgang findet im analogen Bereich statt, was bedeutet dass jedes einzelne Sample einen diskreten Spannungs-Wert aufweist, welcher dem tatsächlichen Wert der anliegenden Spannung zur Zeit des Samplings exakt entspricht.

Dieser Wert wird nun in eine digitale Zahl übersetzt – aber während das analoge Signal unendlich viele Wert-Stufen aufweisen kann, gibt es im digitalen Bereich nur eine begrenzte Anzahl möglicher Werte. Wie viele Werte möglich sind bestimmt die Bit-Auflösung des digitalen Systems.

Wir haben also eine analoge (und damit kontinuierliche) Spannungskurve, welche in einem ersten Schritt analog in diskrete Werte zerschnitten wird. Das digitale System kann aber nicht alle diese Werte abbilden. Was geschieht?

Für jeden am Eingang anliegenden Wert wird die nächstliegende Möglichkeit gesucht, den Wert digital abzubilden.

Je mehr digitale Wert-Stufen zur Verfügung stehen, desto genauer kann das System den analogen Wert digital abbilden. Die Anzahl möglicher digitaler Wert-Stufen wird in *Bits* angegeben (mehr dazu oben unter „Bits“).

Bei der Wandlung von *digital zu analog* wird der ganze Prozess umgekehrt: der digitale Wert wird in einen analogen Sample-Wert umgewandelt, und diese werden zu einer analogen Spannungskurve zusammengesetzt.

Bei der Wandlung von digital zu analog wird aber natürlich vom vorhandenen digitalen Material ausgegangen. Wenn bei der Wandlung von analog zu digital Fehler aufgetreten sind, so ist das Material mit diesen Fehlern gespeichert, und das Material wird folglich auch mit diesen Fehlern von digital nach analog gewandelt. So sind beispielsweise Rundungsfehler, welche bei der Digitalisierung auftraten, als Verzerrungen mit gespeichert und können hinterher nicht mehr entfernt werden.

**Zusammenfassung: Bei der Wandlung von analog zu digital wird das Signal in regelmässigen Abständen gemessen und in digitale Werte umgewandelt.**

**Die *Sampling Rate* bestimmt, wie oft pro Sekunde gemessen wird.**

**Die *Bit-Auflösung* des digitalen Systems bestimmt, wie genau das Signal in digitale Werte übersetzt werden kann.**

### 3.2. Aussteuerung der Wandler

Es sollte klar geworden sein, dass ein ADC und ein DAC keinesfalls blosse ‚digitale Kisten‘ sind, sondern dass diese Wandler einen analogen Teil beinhalten, welcher eine massgebliche Rolle bei der Wandlung spielt.

Inbesondere sollten einige Dinge bezüglich Pegel beachtet werden:

0 dBFS (FS bedeutet ‚Full Scale‘ – 0 dBFS ist der höchstmögliche Pegel, welcher ein digitales System erreichen kann) entsprechen im analogen Bereich +24 dBu. +24 dBu ist für analoge Geräte ein sehr hoher Pegel – nur wirklich hochwertig gebaute Geräte arbeiten bei diesem Pegel gänzlich verzerrungsfrei.

Die Hersteller von Mittelklasse-Wandlern kümmern sich meistens mehr um einen imposanten Wert beim Rauschabstand etc. – gerade der Rauschabstand ist für die Qualität eines Wandlers nahezu irrelevant! – als dass sie einen wirklich hochwertigen analogen Teil in den Wandler einbauen. Dies hat zur Folge dass der analoge Teil des Wandlers bei Pegeln von annähernd 0 dBFS bereits im Verzerrungsbereich arbeitet. Dasselbe gilt für alle analogen Geräte vor dem

Wandler, welche dem Wandler diesen hohen Pegel liefern müssen (Mikrofon-Vorverstärker, Ausgänge von Kompressoren oder Mischpulten etc.).

**Deshalb empfiehlt es sich, bei 24-Bit-Aufnahmen das Signal so auszusteuern, dass der analoge Teil des Wandlers einen Maximalpegel von höchstens –6 dBFS erhält – dies entspricht +18 dBu, was immer noch ein sehr hoher Pegel ist.**

Das Resultat dieser Pegelverminderung wird sein, dass sämtliche Geräte im analogen Bereich sauberer arbeiten, während die digitale Auflösung kaum hörbar verringert wird. Der Klang wird offener, klarer, detaillierter.

Dasselbe gilt für die Wandlung von digital zu analog: Maximalpegel bei –6 dBFS oder sogar –10 dBFS, und der Sound wird klarer...

Bei *16-Bit-Aufnahmen* fällt aber die geringe digitale Auflösung schon so stark ins Gewicht dass hier auf eine möglichst hohe Aussteuerung geachtet werden muss – was ein weiterer Grund dafür ist nicht mit 16 Bits zu arbeiten...

### **3.3. Word Clock bei der Wandlung**

**Bei jeder Wandlung (ADC, DAC sowie asynchrones SRC) muss immer der *Wandler* der Clock-Master sein. Dadurch wird Jitter auf ein Minimum beschränkt. Alle anderen Geräte arbeiten als Slave des Wandlers.**

Jitter ist eines der am wenigsten verstandenen Phänomene der digitalen Audiotbearbeitung und schwierig zu erklären. Bob Katz schrieb mir bezüglich des Kapitels über Jitter in seinem (sehr empfehlenswerten) Buch "Audio Mastering. The art and the science": "It took me 20 revisions before I was happy with the language of the chapter. Jitter was the hardest thing I have ever written about because the language is so slippery."

Jitter entsteht, wenn die Abtastung der Daten bei der Wandlung nicht regelmässig erfolgt, sondern zeitlichen Schwankungen unterliegt. Forschungen haben ergeben, dass Schwankungen von mehr als 10 Picosekunden hörbar sein können (eine Picosekunde ist ein Millionstel eines Millionstels einer Sekunde. Zum Vergleich: eine Periode der Abtastrate 44,1 kHz dauert 22,7 Mikrosekunden, also 22'700'000 Picosekunden).

Nehmen wir zur Veranschaulichung eine einfache Sinus-Schwingung, welche in digitale Daten umgewandelt und danach wieder in ein analoges Signal umgewandelt wird:

Der ADC folgt dem Amplitudenverlauf der Sinuskurve, indem er in regelmässigen Abständen eine Probe (englisch ‚Sample‘) der am analogen Eingang anliegenden Spannung nimmt (sogenannte Puls-Amplitudenmodulation, PAM). Während die Spannung am analogen Eingang also kontinuierlich steigt und dann wieder kontinuierlich fällt, hält der ADC nur momentane Punkte dieser Welle fest.

Die Werte dieser Punkte werden als Treppenspannung festgehalten (wobei die Höhe einer Treppenstufe der Höhe der Sinus-Spannung zum Zeitpunkt der Abtastung entspricht), und die Werte dieser Treppenspannung werden dann (mit Hilfe der sogenannten Pulse Code Modulation, PCM) in digitale Daten umgewandelt (das alles ist oben unter ‚Die Wandlung‘ beschrieben).

Wenn nun die Abtastfrequenz des ADC bei der Puls-Amplitudenmodulation nicht völlig regelmässig ist, dann misst er ‚etwas zu spät‘ oder ‚etwas zu früh‘, und der Wert, welcher weiter verarbeitet wird, entspricht nicht exakt dem Wert der Sinuskurve, der tatsächlich (d.h. zum richtigen Zeitpunkt) hätte gemessen werden sollen – der Wert, den die Puls-Amplitudenmodulation festhält, ist so entweder zu hoch oder zu niedrig, entspricht jedenfalls nicht dem Wert, der bei exakter Abtastfrequenz gemessen worden wäre.

Der DAC, der diese Daten wieder in analoge Spannungen umwandeln muss, weiss nichts von diesen Schwankungen der Abtastfrequenz des ADC. Der Abtast-Vorgang des DACs erfolgt mit derselben Frequenz wie derjenige des ADC, aber der DAC wiederholt nicht die Schwankungen der Abtastfrequenz des ADC – im Gegenteil, der DAC weist selber Schwankungen in der Abtastfrequenz auf...

Wenn deshalb die so verfälschten Daten nun von einem DAC wiedergegeben werden, dann liegt am Ausgang nicht eine schöne Sinus-Spannung vor, sondern eine, welche Dellen und Buckel hat; diese entstehen eben dadurch, dass nicht jeder Punkt der Sinuskurve an den richtigen Platz gesetzt wird. Diese Dellen und Buckel in der Sinuskurve sind nichtlineare Verzerrungen und können sich als hörbare Klangveränderungen bemerkbar machen.

Und wenn auch noch der DAC eine unregelmässige Abtastfrequenz aufweist wird die Sinus-Schwingung weiter verfremdet – das hört sich dann eben an wie ‚typisch digital‘ – hohl, metallisch, flach.

Gemäss der Firma RME produziert ein Wandler, welcher als Slave nach einem externen Clock arbeitet, 15mal mehr Jitter als wenn er vom internen Clock gesteuert wird. Die Erörterung der Gründe dafür würde hier zu weit führen – es genügt sich an die Regel zu halten: Jeder Wandler produziert am wenigsten Jitter, wenn er während der Zeit der Wandlung der Master ist.

Durch Jitter bei der Wandlung von Analog zu Digital und umgekehrt hervorgerufene Verzerrungen sind *nicht* (im nachhinein) *korrigierbar*, sondern nur durch sehr stabile Clocks *vermeidbar*.

Hingegen sind durch Jitter hervorgerufene Verzerrungen irrelevant, wenn sie beispielsweise beim *Abhören* von *rein digitalem* Material auftreten. Denn diese Verzerrungen entstehen bei der für das Abhören benötigten Wandlung – das digital gespeicherte Material ist von dieser Wandlung jedoch nicht betroffen, und somit stellen die durch Jitter hervorgerufenen Verzerrungen auch keine Beeinträchtigung des Materials dar. Das Material ist in besserem Zustand als es sich anhört, und bei einem Wandler mit besserer Clock würden diese durch Jitter hervorgerufenen Verzerrungen nicht auftreten.

(Zu ergänzen bleibt: Die Leute welche Eure Musik dann zuhause auf der HiFi-Anlage anhören haben ziemlich sicher in ihrem CD-Player einen DAC, welcher sehr viel mehr Jitter produziert als Euer Studio-Wandler...)

Wichtig festzuhalten ist, dass Jitter nur bei der Wandlung eine Rolle spielt, *nicht* aber bei *reinem Daten-Transfer*. Wenn also ein Signal digital von einem Gerät in ein anderes geladen wird spielt die Regelmässigkeit des Word Clocks keine Rolle, da es bei Daten nur darum geht, dass die Daten *in der richtigen Reihenfolge* übermittelt werden. Es spielt jedoch überhaupt keine Rolle, wie regelmässig die Datenpakete übermittelt werden – solange das empfangende Gerät in der Lage ist, die Datenpakete in der richtigen Reihenfolge aneinander zu reihen, bleibt die Audioqualität unbeeinflusst.