


## 5. Ton und Klang

- 5.1 Ton: Physikalische und physiologische Aspekte
- 5.2 Kompression von Audio-Signalen: MPEG-Audio
- 5.3 Audio-Datenformate: Übersicht 
- 5.4 Klangerzeugung und MIDI

Weiterführende Literatur:

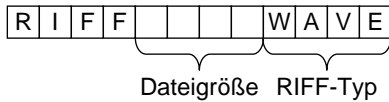
Arne Heyda, Marc Briede, Ulrich Schmidt: Datenformate im Medienbereich, Fachbuchverlag Leipzig 2003

## RIFF (Resource Interchange File Format)

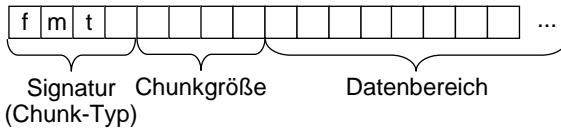
- **IFF:** 1985 von der Firma Electronic Arts eingeführt
  - Sehr einfaches Einheitsformat für verschiedene Arten von Multimedia-Daten, stark verbreitet auf AMIGA-Rechnern
  - Prinzip („Tagged File Format“):
    - » Header gibt Dateityp an
    - » Eigentliche Daten in einer Folge von ebenfalls (über Header) typisierten *chunks*
- **RIFF:**
  - Bestandteil der „Multimedia Programming Interface and Data Specifications“ von Microsoft und IBM, 1991
  - Basiert auf der Idee von IFF
  - Existiert prinzipiell in zwei Varianten:
    - » RIFF für Intel-Architektur („little-endian“)
    - » RIFX für Motorola-Architektur („big-endian“)(RIFX heutzutage auch auf Motorola-Prozessoren ungebräuchlich)

# Grundstruktur von RIFF-Dateien

RIFF-Header (in Bytes):



Chunk-Header (in Bytes):



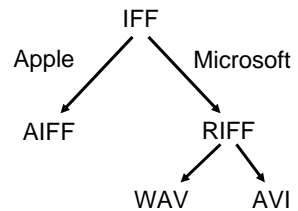
- Verbreitete RIFF-Datentypen (als eigenständige Dateiformate bekannt):
  - WAVE (oder .wav): Audio, unkomprimiert
  - AVI: Video (Audio/Video Interlaced), unkomprimiert
  - RMI: MIDI-Daten (sh. später)
  - BND: „Bündel“ von RIFF-Dateien

# Wave-Format

- Spezialfall des RIFF-Formats (RIFF-Typ „WAVE“)
- Zwei Arten von Chunks:
  - FMT-Chunk (Signatur „fmt “)
    - » Format-Typ (z.B. MS PCM, IBM ADPCM)
    - » Anzahl Kanäle
    - » Sampling-Rate (Hz)
    - » Datenrate (Bytes/s)
    - » Größe von Datenblöcken
    - » Formatspezifische Information (Z.B. bei MS PCM 2 Byte Sample-Größe (bits/Sample))
  - DATA-Chunk (Signatur „data“), meist nur ein solcher Chunk vorhanden
    - » Bei mehreren Kanälen „interleaving“, d.h. alle Kanäle für einen Zeitpunkt in Folge

## AIFF (Audio Interchange File Format)

- Herstellerspezifische Erweiterung von IFF durch Apple für unkomprimiertes Audio
  - Format-Chunk
  - Daten-Chunks, byteweise gepackt
- Audiodaten für bis zu 6 Kanäle (Surround Sound)
- Möglichkeit zur Einstreuung von MIDI-Chunks und Instrumenten-Chunks
- Spezialvariante AIFF-C für komprimierte Audiodaten (ca. 6:1)




## AU (Audio File Format)

- Bei NeXT entwickelt, weit verbreitet im UNIX-Bereich (z.B. Sun)
- Header:
  - Abtaste, Kanalzahl, Datenformat etc.
  - beliebig lange Textinformation
- Datenbereich:
  - Kanäle miteinander verschränkt
  - Viele Datenformate, z.B.:
    - » von 8 bis 32 Bit
    - »  $\mu$ -Law und linear
    - » Festkomma, Gleitkomma, doppelte Genauigkeit
- Unterstützung von Dateifragmentierung

## QuickTime

- Bibliothek von systemnahen Programmen für MacOS und Windows für die Bearbeitung von zeitbasierten Medien („movies“)
  - Entwickelt von Apple ca. 1991-heute
- Sehr allgemeines Konzept für Medienstrukturen
  - „Atom“: Allgemeiner Container für Mediendaten
  - Mehrere Tracks je Präsentation
  - Pro Track:
    - » Medienstruktur (Referenzen zu Medien verschiedenen Typs)
    - » „Edit List“ für Zeitsynchronisation
- QuickTime wurde als Basis für die MPEG-4 Dateistruktur gewählt.
- Viele verschiedene Dateitypen von QuickTime unterstützt
  - Wichtiges spezifisches QuickTime-Format: „Movie“ (MOV)

## 5. Ton und Klang

- 5.1 Ton: Physikalische und physiologische Aspekte
- 5.2 Kompression von Audio-Signalen: MPEG-Audio
- 5.3 Audio-Datenformate: Übersicht
- 5.4 Klangerzeugung und MIDI 

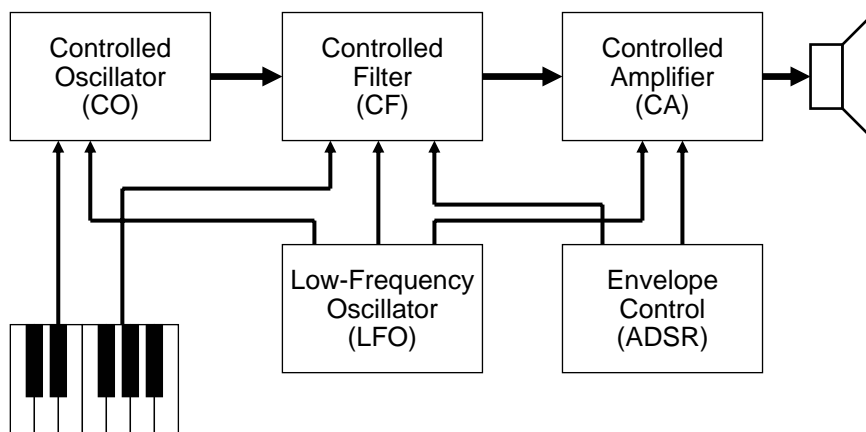
Literatur:

Hannes Raffaseder: Audiodesign, Fachbuchverlag Leipzig 2002

# Klangerzeugung

- Grundziel der Klangerzeugung:  
Erzeugung komplexer Klangereignisse mit elektronischen Mitteln
- Klänge für Musik oder Sprache können künstlich produziert werden
  - Tonhöhe, Lautstärke, Klangfarbe (*timbre*) einstellbar
- Klangerzeuger:
  - Einfache Klangerzeuger in Soundkarten enthalten  
(Frequenzmodulation einfacher Wellenformen)
  - Hochwertige Klangerzeuger z.B. in elektronischen Musikinstrumenten („Synthesizer“, MIDI-Keyboards)
    - » mehrstimmig (z.B. 128)
    - » multitimbral (d.h. verschiedene Klangfarben, z.B. 64)
- Historisch gesehen:
  - 1900 Dynamophone (Thaddeus Cahill), 1920 Termenvox (Lew Termen), 1930 Trautonium (Friedrich Trautwein)
  - Seit Anfang der 60er Jahre (Robert Moog): Moderne Synthesizer-Architektur
  - Musikalischer Durchbruch: 1968 Walter Carlos „Switched-on Bach“

# Grundstruktur eines Synthesizers

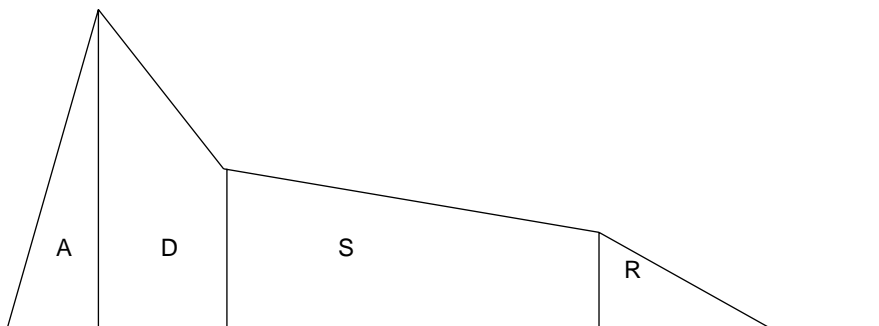


## Grundelemente bei der Klangerzeugung

- Oszillator
  - Erzeugt mehr oder weniger obertonreiches Signal, das die Grundfrequenz und auch wesentlich den Klangcharakter bestimmt
- Filter
  - Z.B. Hochpass, Tiefpass, Bandfilter
- Verstärker (*Amplifier*)
  - Kann über zeitabhängigen Pegelverlauf Klangempfindung wesentlich beeinflussen
- Hüllkurvengenerator (*Envelope Control*)
  - Zeitlicher Verlauf eines Klangereignisses auf ein einmaliges erzeugendes Ereignis hin (z.B. Tastendruck), meist ADSR (siehe nächste Folie)
- *Low Frequency Oscillator LFO*
  - Dient zur kontinuierlichen Veränderung eines klangbestimmenden Parameters innerhalb eines bestimmten Zeitintervalls
  - Frequenzen typischerweise zwischen 0 und 20 Hz

## ADSR-Modell

- Modulation nach dem ADSR-Modell
  - Attack (A), Decay (D), Sustain (S), Release (R)
  - Höhe und Breite der vier Parameter variabel
  - Verbreitet: Einstellung der A-, D-, S- und R-Zeiten über Regler



## Verfahren zur Klangsynthese

- Additive Klangsynthese
  - Umsetzung der Fourier-Analyse in die Praxis
  - Realisierung eines Klangs als Überlagerung von Sinustönen
- Subtraktive Klangsynthese
  - Erzeugung obertonreicher Grundsignale (z.B. Sägezahn, Dreieck, ...)
  - Steuerung der Spektren mit Filtern und der Amplitude mit Verstärkern
- Wavetable-Synthese
  - Vordefinierte, gespeicherte Wellenformen
  - Oszillator durchläuft Wavetable in programmierter Weise (z.B. LFO)
- Sampling
  - Wiedergabe digital aufgezeichneter akustischer Ereignisse
  - Multisampling: Viele Aufnahmen verschiedener Parameterwerte
- Granularsynthese
  - Zerlegung von Schallsignalen in *Grains* (wenige ms lange Abschnitte)
  - Entkopplung von Wiedergabegeschwindigkeit und Tonhöhe

## Akustische Modellierung

- Modellierung eines Klangerzeugers (physikalisch-akustisch)
  - Vereinfachtes physikalisches Modell der Klangerzeugung
  - Auflösung zu Wellenform: endlich viele gekoppelte nichtlineare Differentialgleichungen
  
- Modellierung eines Raums (architektonisch-akustisch)
  - Modifikation von Audiodaten gemäß Akustik eines speziellen Raums
  - Bestimmung der Raumakustik vor Ort mit Impulsschall verschiedener Frequenzen
  - Faltung des Audiosignals mit Akustik
  - Sinnvoll auch in Gebäudeplanung und -optimierung

## MIDI: Geschichte und Überblick

- Synthesizer: Revolutionäres Musikinstrument in den 70er Jahren
  - Beatles (White Album), Carlos (Switched-on Bach), ...
  - Technische Probleme:
    - Polyphonie, Kombination verschiedener Geräte, Synchronisation
- 1983: Erste Interoperabilitäts-Vorführung
- MIDI (Musical Instrument Digital Interface) Standard
  - International MIDI Association (IMA)
  - MIDI Manufacturers Association (MMA)
- Bedeutung für Multimedia:
  - Standardisierte Sprache für
    - » Übernahme von Daten aus Endgeräten, die Musikinstrumenten entsprechen (insb. Keyboard)
    - » Ansteuerung von Peripheriegeräten (Synthesizer, Beleuchtung)
    - » Abstrahierte Darstellung von gespielter Musik

## MIDI-Grundbegriffe

- Ereignis (*event*):
  - Musikalische Aktion, z.B. Musiker drückt Taste auf Keyboard mit bestimmter Anschlagsstärke (*velocity*)
    - » etwa: „NOTE ON C3 velocity 100“
  - Jedes Ereignis findet zu einem bestimmten Zeitpunkt statt (Zeitstempel)
- Nachricht (*message*):
  - Binäre Codierung der in einem Ereignis enthaltenen Information
  - Kann gespeichert, weitergegeben, vervielfältigt, modifiziert werden
- Befehl (*command*):
  - Anweisung an ein externes Gerät, bestimmte musikalische Aktionen auszuführen
- Klangfarbe (*timbre*):
  - Charakteristik eines bestimmten wiederzugebenden Instruments
  - „Multitimbral“
- Kanal (*channel*):
  - Identifikator für bestimmten Empfänger (traditionell 16 Kanäle)



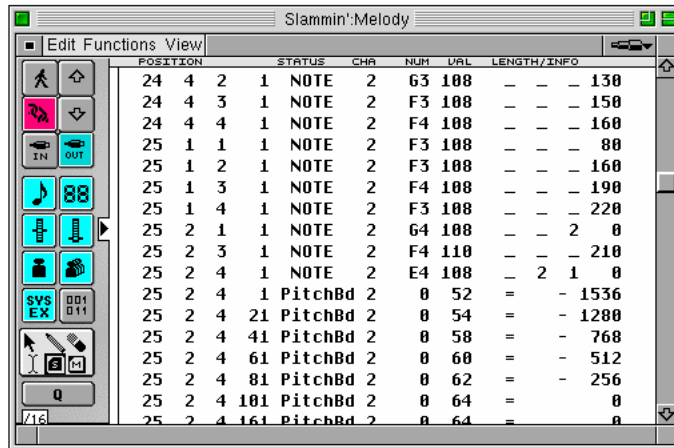
## MIDI-Nachrichten

- Channel Voice Messages
  - Eigentliche Musikdaten (sh. nächste Folie)
- Channel Mode Messages
  - Steuerung des Synthesizers
    - » Ein-/Ausschalten der eigenen Tastatur (z.B. bei Keyboard/Synthesizer)
    - » Testmodus
    - » Polyphonie-Steuerung
- System Real-Time Messages
  - Synchronisationstakt
  - Synchronisierte Sequenzen
  - Überprüfung der Verfügbarkeit von Geräten
- System Exclusive Messages (SysEx)
  - Weitergabe herstellerepezifischer Information an individuelle Geräte

## Inhalt einer MIDI-Datei: MIDI-Ereignisse

- Header-Information
- Track-Information
  - *Track* = Separat abspielbare und bearbeitbare Musikspur
- Track-Information Teil 1: Metainformation
  - Track-Nummer, -Name
  - Angaben zum Instrument (z.B. aus *General Midi*-Instrumenten)
  - Zeitbasis
- Track-Information Teil 2: Melodie
  - Folge von Channel Voice Messages, jeweils mit Zeitstempel relativ zur Zeitbasis
  - Note On (Parameter Notenwert, Anschlagstärke)
  - Note Off (Parameter Notenwert, Anschlagstärke)
  - Polyphonic Key Pressure (Parameter Notenwert, Anschlagstärke)  
(Änderung der Anschlagstärke über die Zeit)
  - Pitch Bend Change (Parameter Verschiebung)  
(Tonhöhenverstellung)

## MIDI Ereignisse: Beispiel



The screenshot shows a MIDI editor window with a table of MIDI events. The table has columns for POSITION, STATUS, CHA, NUM, VRL, and LENGTH/INFO. The events include NOTE and PitchBd (Pitch Bend) messages.

POSITION	STATUS	CHA	NUM	VRL	LENGTH/INFO
24 4 2	1 NOTE	2	63	108	- - - 130
24 4 3	1 NOTE	2	F3	108	- - - 150
24 4 4	1 NOTE	2	F4	108	- - - 160
25 1 1	1 NOTE	2	F3	108	- - - 80
25 1 2	1 NOTE	2	F3	108	- - - 160
25 1 3	1 NOTE	2	F4	108	- - - 190
25 1 4	1 NOTE	2	F3	108	- - - 220
25 2 1	1 NOTE	2	64	108	- - 2 0
25 2 3	1 NOTE	2	F4	110	- - - 210
25 2 4	1 NOTE	2	E4	108	- 2 1 0
25 2 4	1 PitchBd	2	0	52	= - 1536
25 2 4	21 PitchBd	2	0	54	= - 1280
25 2 4	41 PitchBd	2	0	58	= - 768
25 2 4	61 PitchBd	2	0	60	= - 512
25 2 4	81 PitchBd	2	0	62	= - 256
25 2 4	101 PitchBd	2	0	64	= 0
25 2 4	161 PitchBd	2	0	64	= 0

- MIDI-Dateien sind extrem kompakt.
- MIDI-Aufzeichnungen sind genauer als normale Notenschrift!

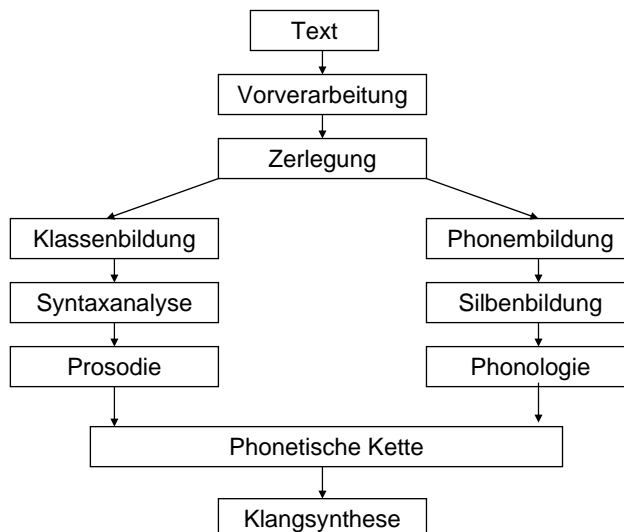
## Typische Funktionen von MIDI-Sequenzern

- Aufnehmen und Wiedergeben von Tonspuren
- Verschiedene Ansichten der gleichen Information:
  - Partitur, Keyboard-Matrix
  - Zeitgenaue Liniendarstellung
  - Darstellung von Zusatzinformation (z.B. *velocity*)
- Musik-Editor:
  - Komponieren (Noten einsetzen und verschieben, Längen verändern, Transponieren, ...)
  - Instrumente variieren
  - Effekte einfügen
  - Synchronisieren von Spuren und Abmischen
  - Oft integriert mit klassischer Mischpult-Funktionalität
  - Oft integriert mit Notensatz-Funktionalität

# Sprachanalyse und Sprachsynthese

- Ein- und Ausgabe in natürlicher Sprache
  - Alter Traum der Informatik
  - Grenzgebiet zu Computerlinguistik, KI
- Sprachausgabe:
  - relativ stabile Technologie
  - Bestandteil vieler Standard-Betriebssysteme
- Spracheingabe:
  - immer noch relativ wenig beherrscht
  - Trainingsfreie Systeme noch störanfällig
  - Trainingsgebundene Systeme existieren mit akzeptabler Leistung

## Sprachsynthese: Grob Ablauf



## Weiterentwicklung im Bereich Klangerzeugung

- MPEG-4 Standard:
  - *Structured Audio Format* ermöglicht Spezifikation von Klangerzeugern
  - *SAOL (Structured Audio Orchestral Language)* zur Beschreibung von elektronischen Instrumenten und Audioeffekten
  - *SASL (Structured Audio Source Language)* erlaubt differenzierte Formulierung von Spielanweisungen (über MIDI hinaus)
- Anwendungsfeld Interaktion:
  - Akustische Signale in Spielen und Softwaresystemen tendieren dazu, den Benutzer durch stupide Wiederholung zu ermüden
  - Softwaresynthese von Klängen eröffnet die Möglichkeit, situationsabhängig neue Klänge zu generieren, wo erwünscht