

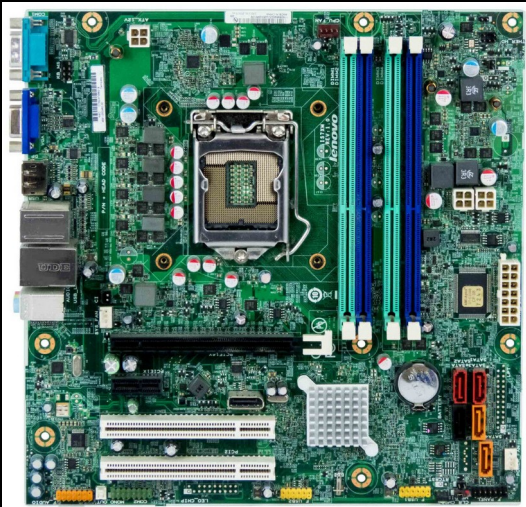
VL07: Rechnernutzung in der Physik

Rechnerarchitekturen & Softwareentwicklung

Günter Quast

Fakultät für Physik
Institut für Experimentelle Teilchenphysik

WS 2023/24

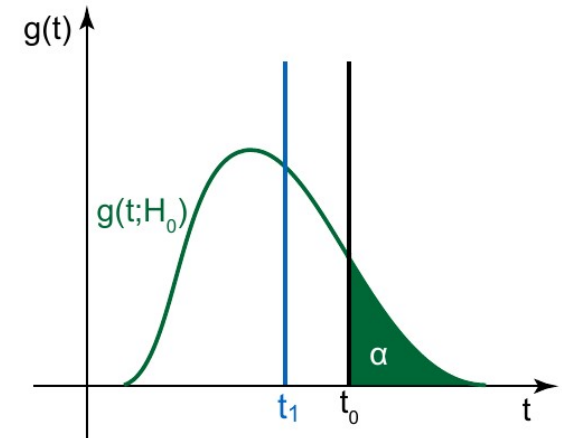


Zusammenfassung V06: Hypothesentest

Vergleich eines Modells (= Nullhypothese, $\mathcal{H}_0$) mit Stichprobe (Daten)

Zur Quantifizierung notwendig:

- Wahl einer Prüfgröße (engl. test-statistic) t
- Berechnung der Prüfgröße aus Daten
- ggf. Verwerfen der Hypothese mit (vor dem Test festgelegtem !) Konfidenzniveau α

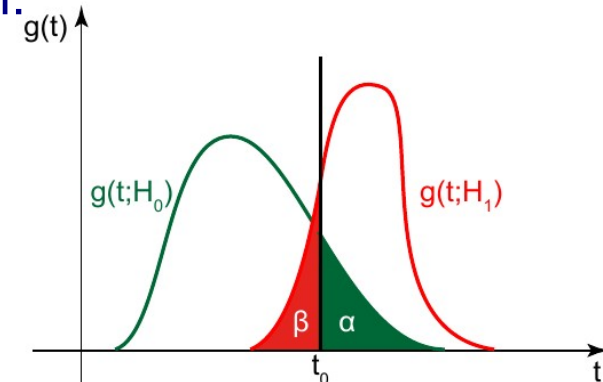


Optimale Prüfgröße bei Vergleich zweier Hypothesen:

Likelihood-Verhältnis

$$r(x) = \frac{\mathcal{L}(\mathcal{H}_0)}{\mathcal{L}(\mathcal{H}_1)} = \frac{\prod_{i=1}^n f(x_i|\mathcal{H}_0)}{\prod_{i=1}^n f(x_i|\mathcal{H}_1)} \leq \eta(\alpha)$$

(Neyman-Pearson Lemma)



α : Fehler 1. Art, fälschliches Verwerfen von $\mathcal{H}_0$ („false positive“)

β : Fehler 2. Art, fälschliches Akzeptieren von $\mathcal{H}_1$ („false negative“)

Gilt streng nur bei „einfachen Tests“ ohne freie Parameter;
Verallgemeinerung auf Profil-Likelihood in der Physik häufig verwendet:

$\mathcal{L} \rightarrow \mathcal{L}_{\text{prof}}$, $t = -2 \ln Q$, Q : Profil-Likelihood-Quotient

- Rechnerarchitekturen, Parallelisierung und Vektorisierung
- Kollaborative Softwareentwicklung

Das **fünfte Übungsblatt** ist on-line !

Rechnernutzung in der Physik

Institut für Experimentelle Teilchenphysik

Institut für Theoretische Teilchenphysik

Prof. G. Quast, Prof. M. Steinhauser

Dr. A. Mildenberger, Dr. Th. Chwalek

[ILIAS-Seite zum Kurs](#)

WS 2023/24 – Blatt 05

Abgabe: Montag 11.12.2023 bzw. Dienstag 12.12.2023

Abschlussprojekt: Ensemble-Test -

Überprüfung der Eigenschaften
eines Verfahrens zur Modellanpassung
auf Verzerrung und Coverage

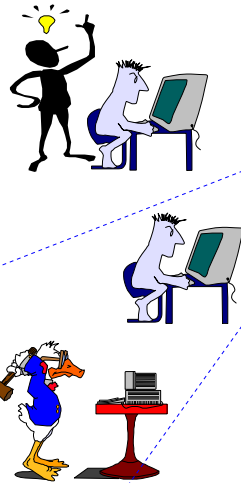
ggf. **Problem** mit nicht installiertem Paket *PhyPraKit*:

im Terminal des Jupyter-Notebooks eingeben: `pip install --user PhyPraKit`

Rechnernutzung in der Physik



Datenerfassung



*Einzelplatz-Rechner
als Schnittstelle zu
User (=Physiker:in)*



Server-Rack

Moderne Wissenschaft ohne Computer nicht denkbar



CERN Computer Centre



*Worldwide LHC
computing Grid*

Weltumspannende Netzwerke: Grid(s) und Cloud

Rechner in der Physik heute meist auf PC-Architektur (*PC=Personal Computer*) d.h. **Mikrocomputersysteme** bestehend aus:

- Mikroprozessor(en)
- (flüchtigem) Halbleiter-Speicher
- permanentem Magnet-Festplattenspeicher oder „SolidStateDisk“ (SSD)
- Ein-Ausgabeschnittstellen
- Tastatur, Bildschirm, Maus, Wechsellaufwerke, Drucker, Netzwerk, ...

Verfügbar als

- Einzelplatz-Desktop-PC
- vernetzter Cluster aus Desktop-PCs
- PC-Farm mit Fileservern
- Multi-Prozessor-Installationen für „High-Performance-Computing“
- vernetzte Rechenzentren („GRID“)
- Cloud-Computing:
 - virtueller Server („Platform as a Service“)



GridKa, KIT Campus Nord

Rechnerhardware

Rechnerhardware

Grundlage der Rechnertechnik ist die **Digitalelektronik**:

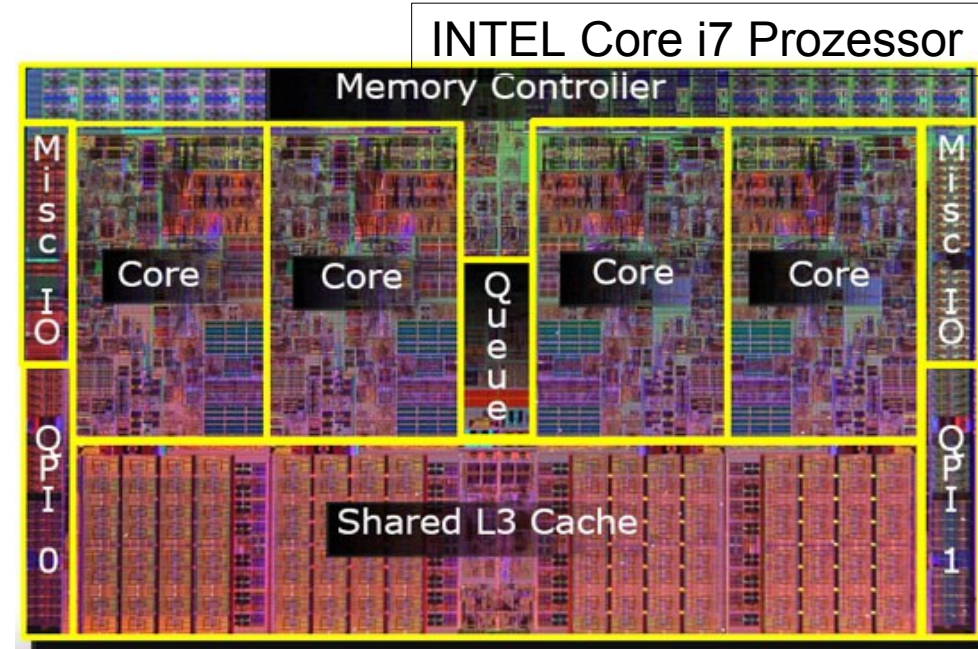
- Zustände 1 und 0 repräsentiert durch elektrische Signale

Boolesche Algebra mit Hilfe von elektronischen Bauteilen wie

- Gattern (AND, OR, XOR)
- Bistabilen Stufen („FlipFlops“) als Speicherelemente

daraus aufgebaut:

- Register, Addierer und weiter
- Rechenwerk, Steuerwerk, Speicher eines Mikroprozessors



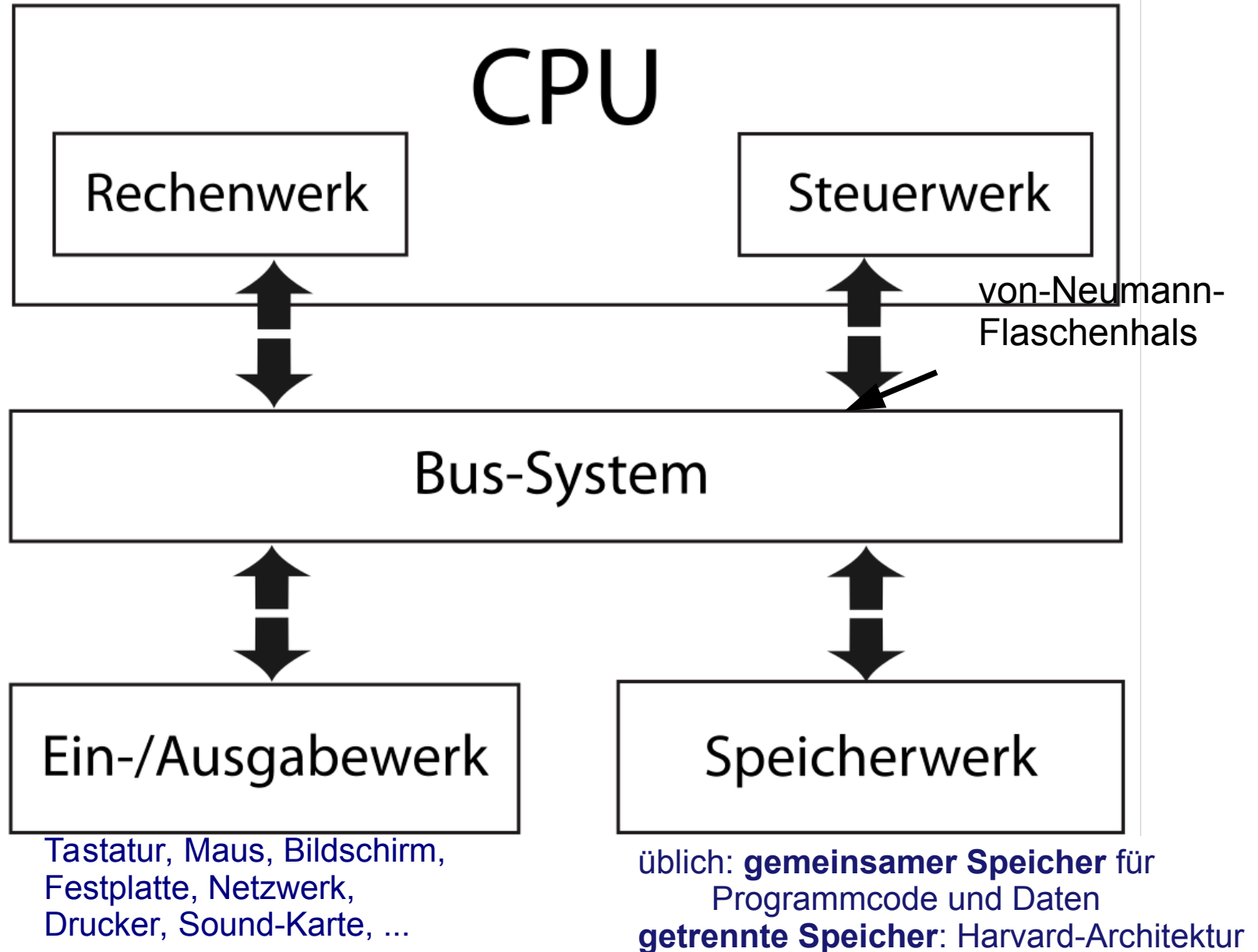
Zahlendarstellung im Binärsystem, d.h. z.B. 10 = 1010; meist 4 Binärstellen als eine „Hex“-Ziffer angegeben im Dezimalsystem, also 1010=\$A (oder 0xA)

Gesamte Rechnerhardware auf wenigen Chips integriert

Stand der Technik 2023:

- 3 - 5 nm Strukturgröße,
- ~ 10 Milliarden Transistoren / cm²
- Chips mit 20 Milliarden Transistoren und 64 Rechenkernen („cores“)

von Neumann-Architektur

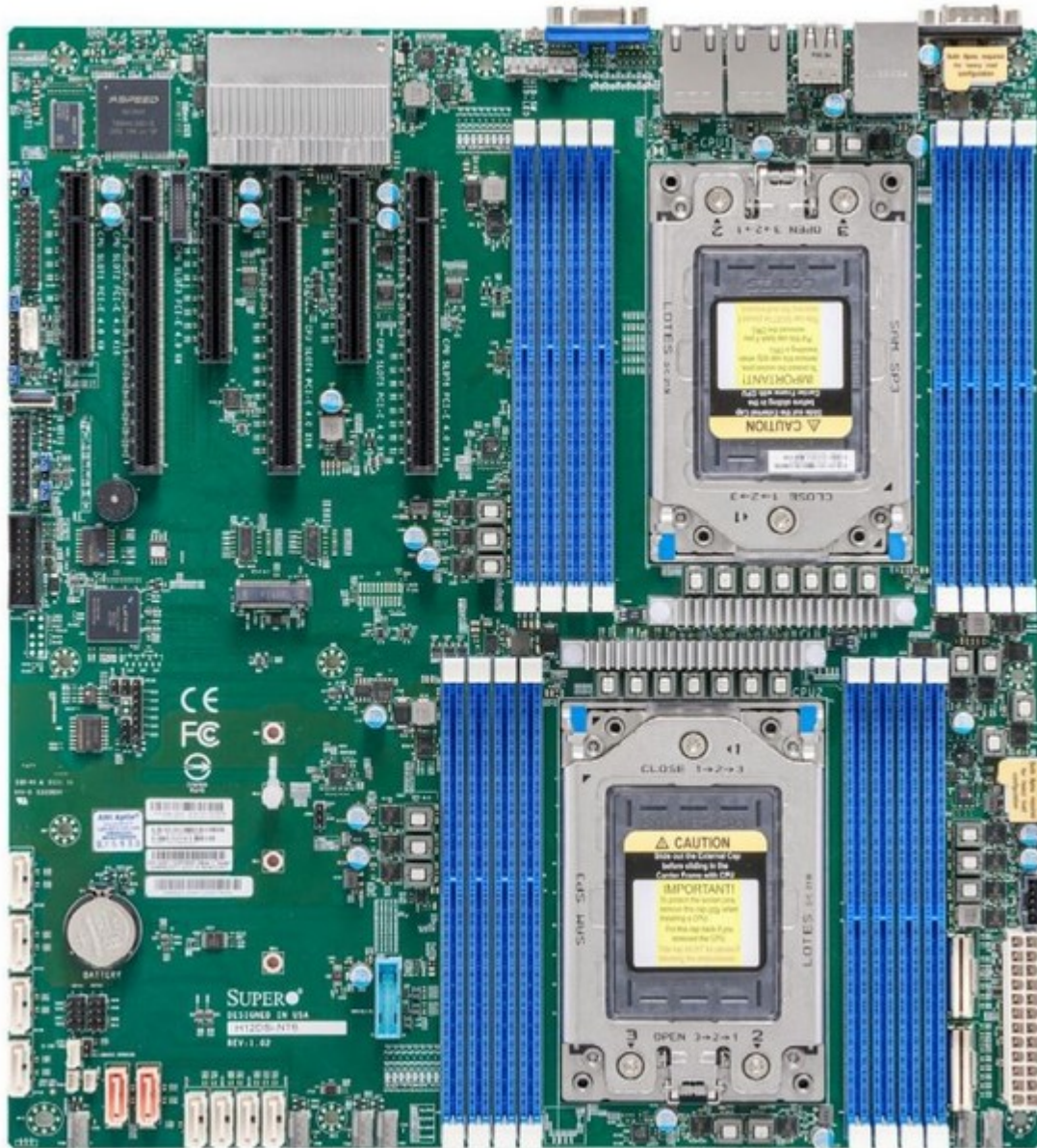


Hauptplatine

Modernes Mainboard für Server-System

Server Mainboard (2022) der Firma Supermicro

- Mainboard Formfaktor: E-ATX
- System on Chip
- CPU Sockel Typ: LGA 4094 (SP3)
- CPU Sockel Anzahl: 2
- max. RAM Kapazität (GB): 4096
- Netzwerk Port(s): 2x RJ45 - 10 Gbit/s



Mikroprozessor

Mikroprozessor ist die „CPU“ (**C**entral **P**rocessing **U**nit), die alle Systemkomponenten steuert und arithmetisch-logische Operationen ausführt.

- Befehlscodes für verschiedene Operationen im Speicher abgelegt („Programm“)
- Register als schnelle interne Datenspeicher
- Datentransfers
Speicher $\Leftrightarrow$ Register
- logische oder Arithmetische Operationen auf Registerinhalt oder externem Speicher

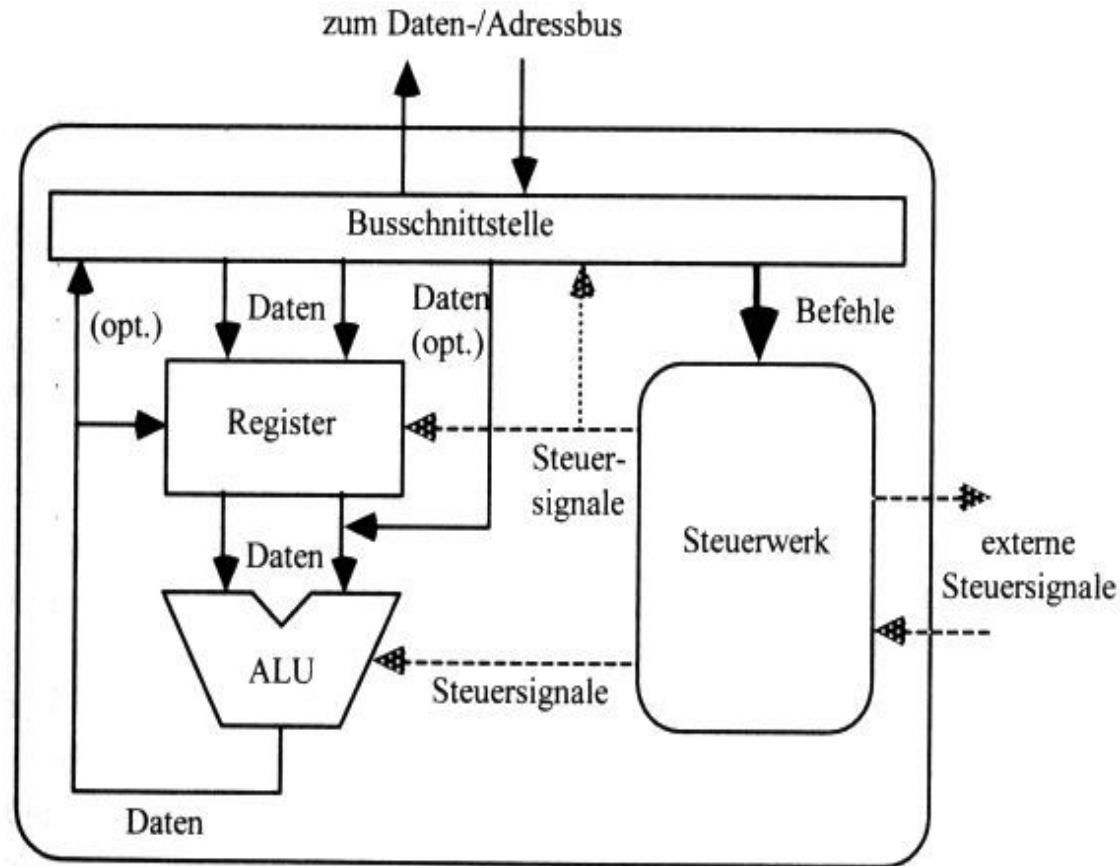


Abb. 2.17. Struktur eines einfachen Mikroprozessors

Zeitliche Entwicklung

Moore's Law: The number of transistors on microchips doubles every two years

Moore's law describes the empirical regularity that the number of transistors on integrated circuits doubles approximately every two years. This advancement is important for other aspects of technological progress in computing – such as processing speed or the price of computers.

Our World
in Data

Transistor count

50,000,000,000

10,000,000,000

5,000,000,000

1,000,000,000

500,000,000

100,000,000

50,000,000

10,000,000

5,000,000

1,000,000

500,000

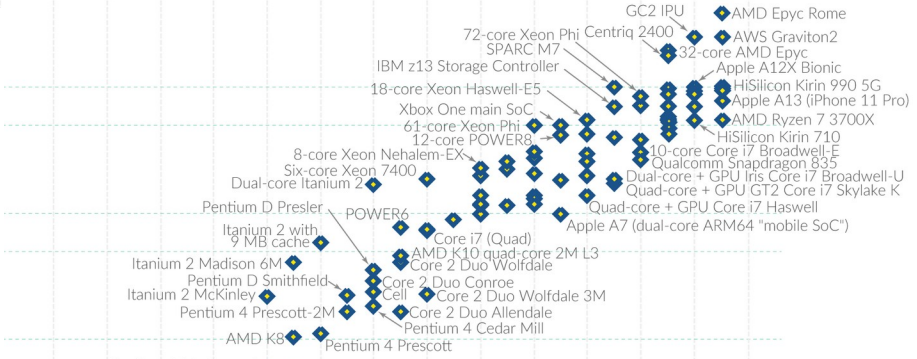
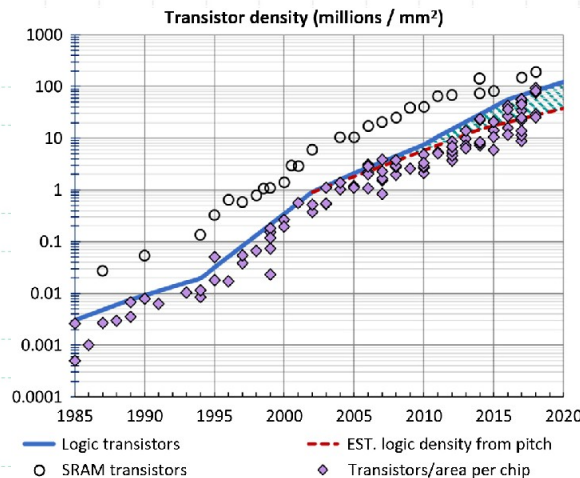
100,000

50,000

10,000

5,000

1,000



Moore'sches Gesetz:
Verdopplung der
Anzahl Transistoren
alle 1½ Jahre

Year in which the microchip was first introduced

Data source: Wikipedia (wikipedia.org/wiki/Transistor_count)

OurWorldinData.org – Research and data to make progress against the world's largest problems.

Licensed under CC-BY by the authors Hannah Ritchie and Max Roser.

Quelle: Wikipedia

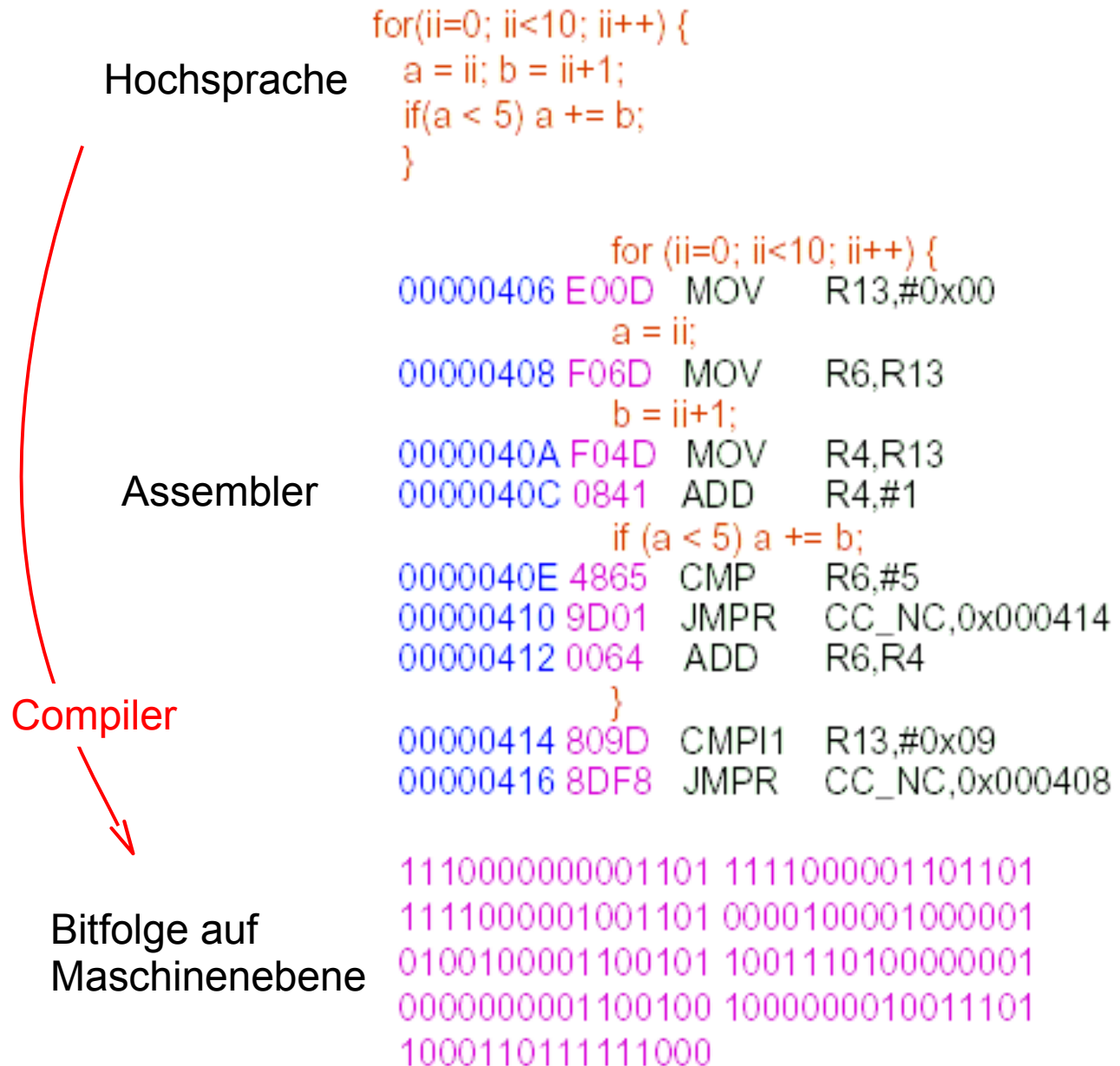
Hardware – was ein Prozessor tut

**Mikroprozessoren
verarbeiten intern
Folgen von Bits**

Bitfolge kann als

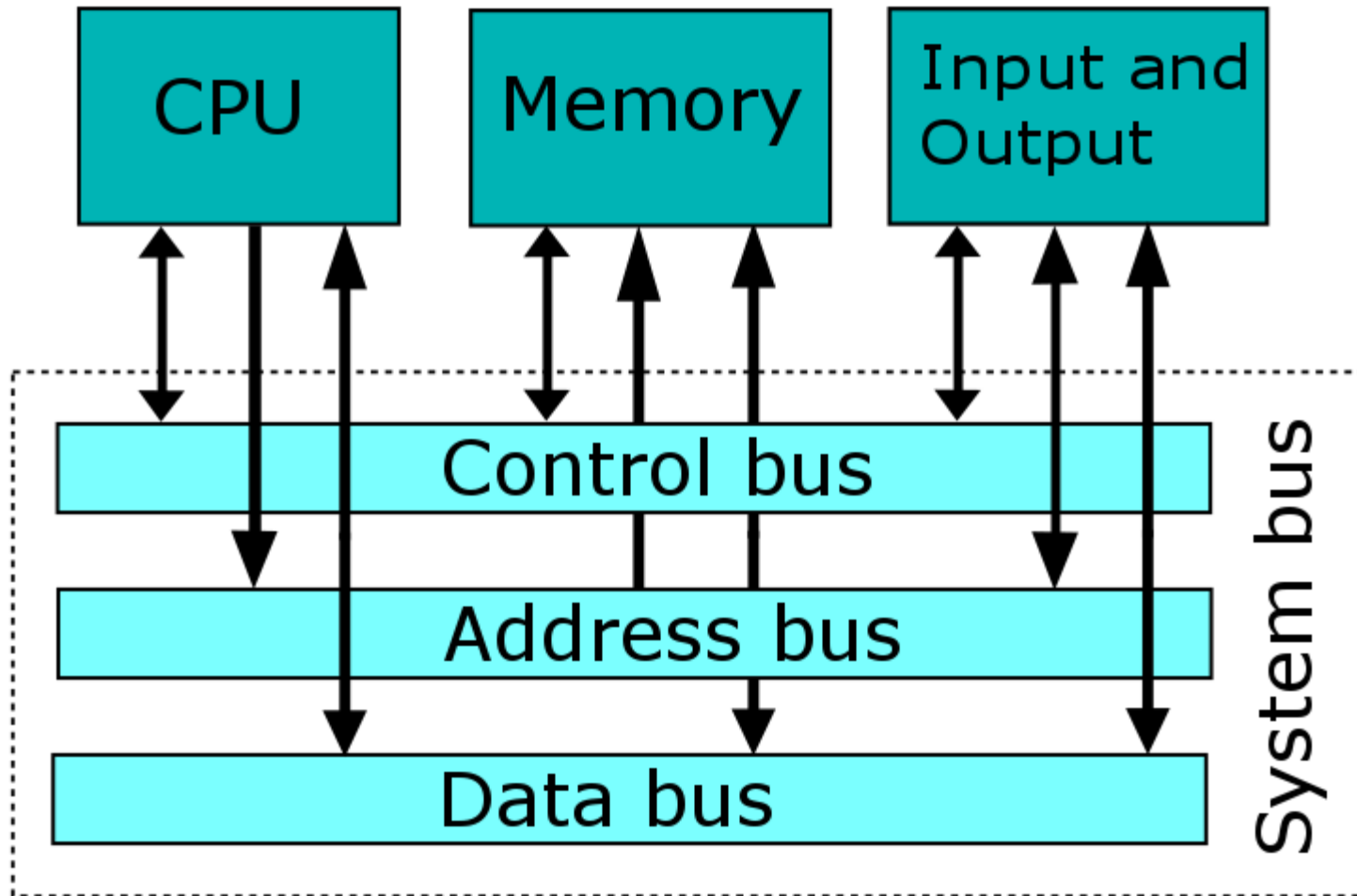
- Befehl
- Datum
- Adresse

interpretiert werden
(***kontextabhängig !***)



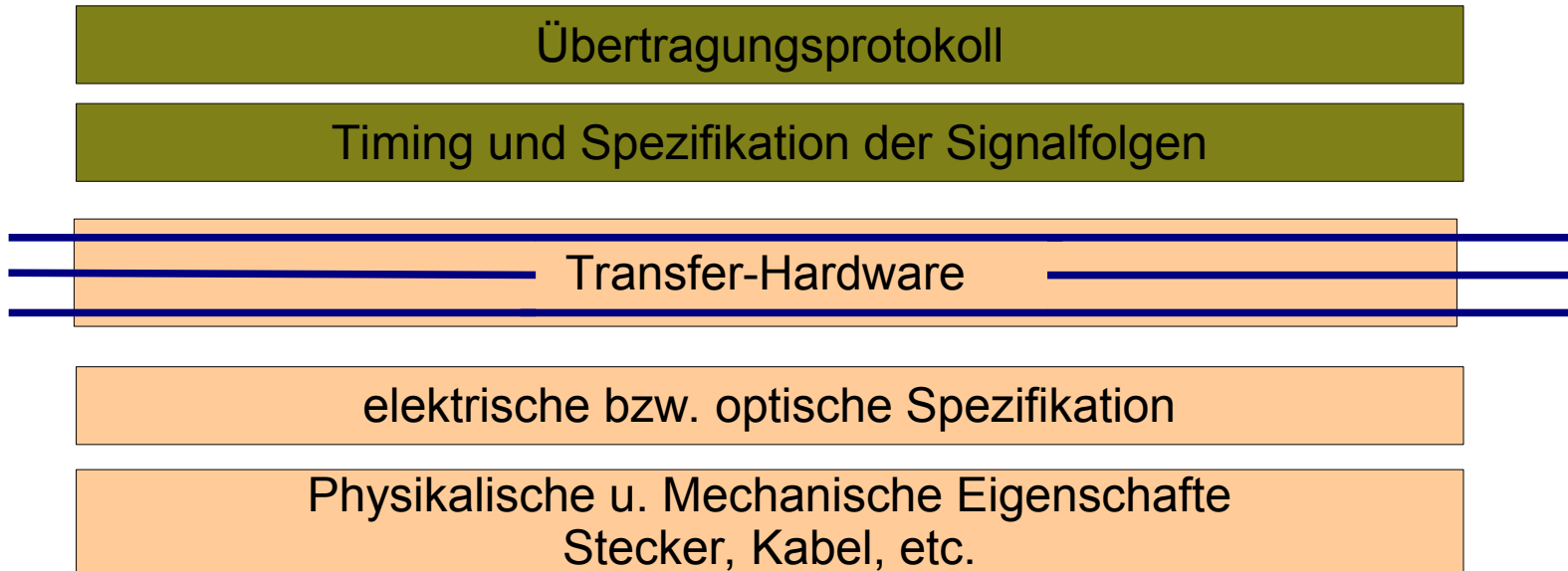
Bussystem

Bus-System verbindet CPU, Speicher und externe Geräte



Bussystem (2)

Verschiedene Ebenen der Datenübertragung:



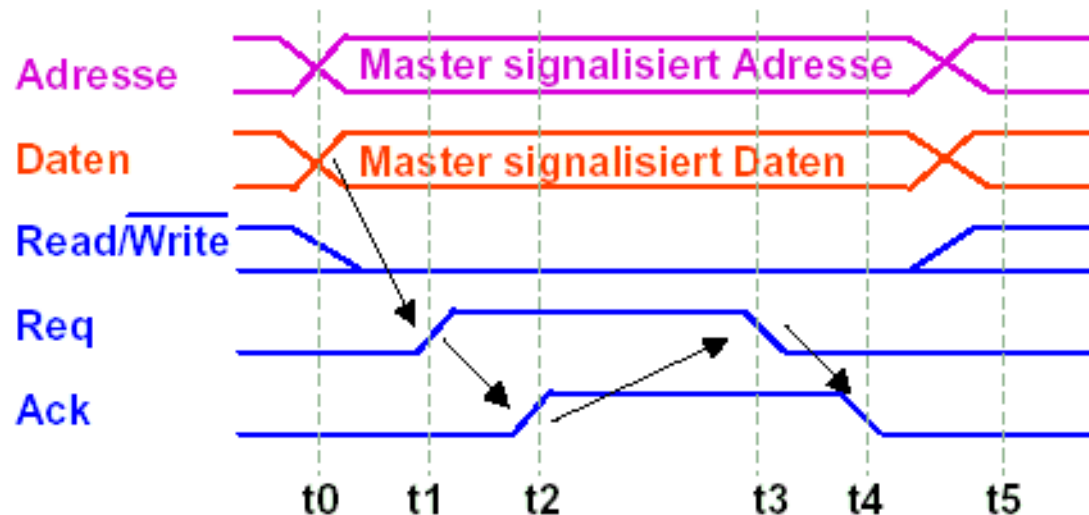
Daten- und Signaltypen auf Bussen:

Multiplexing: mehrere
Typen auf einem Bus



Bus-Timing (Prinzip)

Einfachstes Protokoll für (asynchrone) Datenübertragung :



Busankopplung über „Tristate-Gatter“ mit Zuständen
L (low)
H (high)
Z (hochohmig)

- t_0 : Master hat Buskontrolle und signalisiert Adresse, Richtung, Daten, wartet eine spezifizierte Zeit bis Daten am Bus stabil getrieben
- t_1 : Master signalisiert Anforderung Signal (Req)
- t_2 : Slave signalisiert Antwort (Ack) später, wodurch der Datenempfang bestätigt wird
- t_3 : Master löscht Anforderungssignal (Req)
- t_4 : Slave löscht Antwortsignal (Ack)

Aber: echte Bus-Systeme sind viel komplizierter !

Bussystem (3)

Historisch: Bus und CPU hatten gleiche Taktfrequenz → nur ein Bus notwendig

heute: eine Reihe von hierarchisch verbundenen Bus-Systemen mit unterschiedlichen Geschwindigkeiten

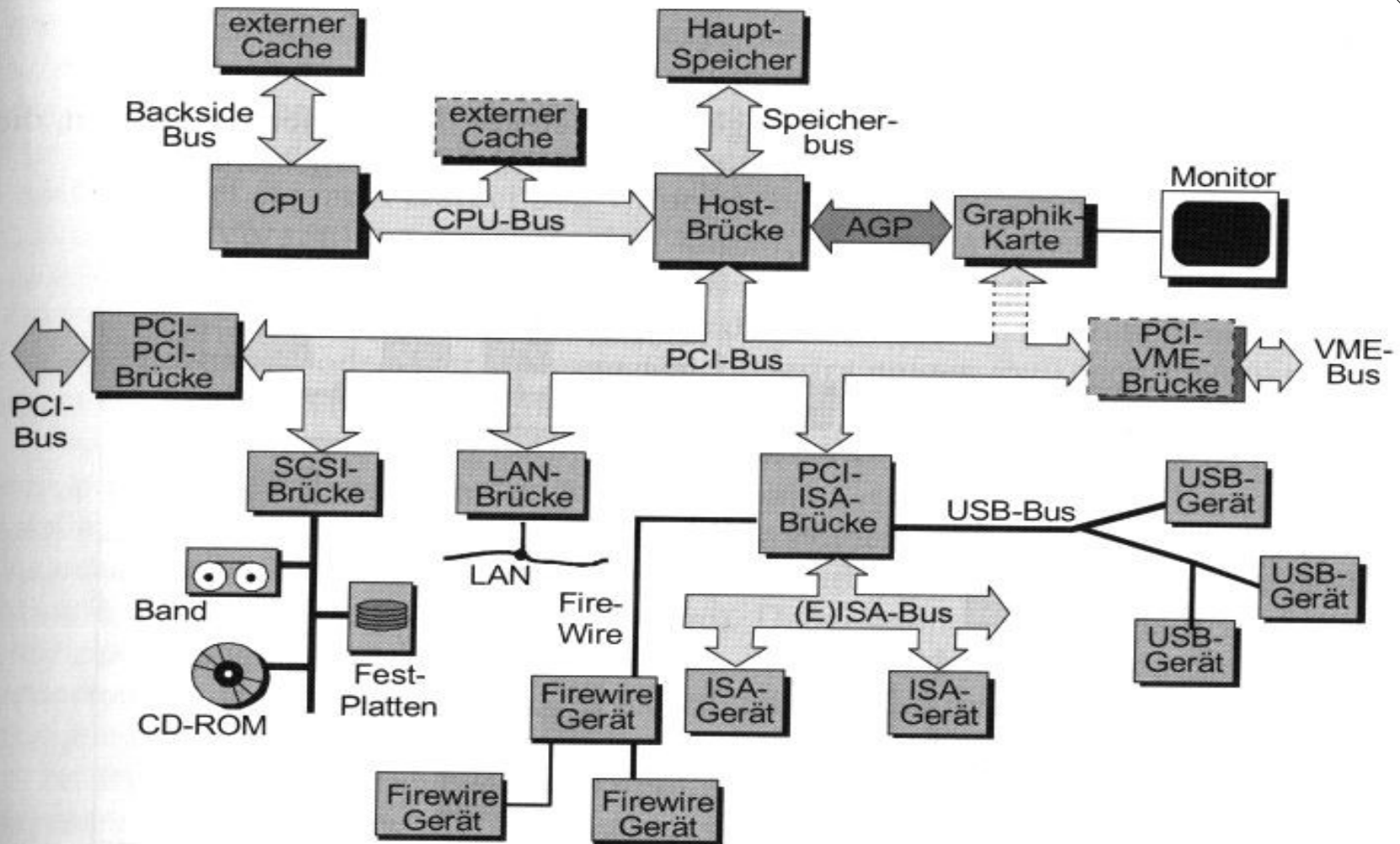


Bild 1.2-2: Typisches hierarchisches Bussystem

Datenaustausch zwischen

- Rechner und Peripheriegeräten bzw.
- mehreren Rechnern

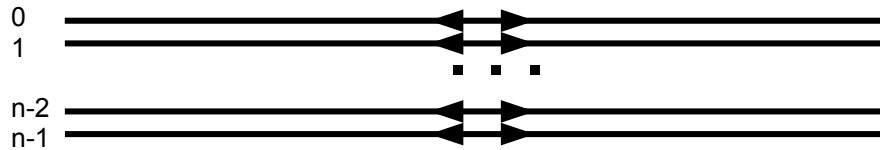
erfolgt über **Schnittstellen**.

Schnittstellen müssen standardisiert sein:

- elektrische Signale
- Steckermechanik und Belegung
- Timing
- Übertragungsprotokoll

Übertragung parallel ↔ seriell

parallel: gleichzeitige Übertragung von n Bits,
Leitungen typisch bidirektional schaltbar



beschränkt auf kurze Entfernungen

Beispiele:

Parallelport, ISA, PCI,
ATA , SCSI ...

seriell: Übertragung der Bits nacheinander einer Leitung
meist eine Übertragungsstrecke pro Richtung



typisch ist die „Taktrückgewinnung“ aus dem empfangenen Signal

Beispiele:

Serieller Port, USB,
SATA, Ethernet,
iSCSI ...

*Anm.: Anders als intuitiv erwartet sind die schnellsten
Übertragungen heute seriell !*

Probleme der parallelen Übertragung

- unterschiedliche Laufzeiten auf den einzelnen Signalwegen
- Übersprechen auf langen Leitungen

Schnittstellen (2)

Klassisch:

- die serielle Schnittstelle (RS232 oder V.24 Standard)
(früher) gebräuchlich für Maus, Tastatur, Modem, div. Messgeräte
- die parallele Schnittstelle

Moderne Standards:

- USB A („universal serial bus“) bzw. USB C
die aktuelle Schnittstelle (Maus, Tastatur, PC-Erweiterungen, Messgeräte)
es existieren Adapter seriell $\Leftrightarrow$ USB oder Parallel $\Leftrightarrow$ USB



USB 1.0/2.0 TYP A



USB 2.0 MINI-A



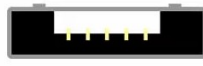
USB 1.0/2.0 TYP-B



USB Typ-C



USB 3.0 TYP A



USB 2.0 MICRO-A



USB 3.0 TYP-B

Server

Client

Server/Client automatisch

- FireWire

- besonders im Multi-Media-Bereich (DigiCams, aber auch Festplatten etc.)
- Zugriff auf PC-Bus mittels Steckkarten im PCI-Standard
- Zugang über Netzwerk (Ethernet-Schnittstelle, 10/100/1000/10000 Mbit/sec)

der aktuelle Standard: Universal Serial Bus(USB)

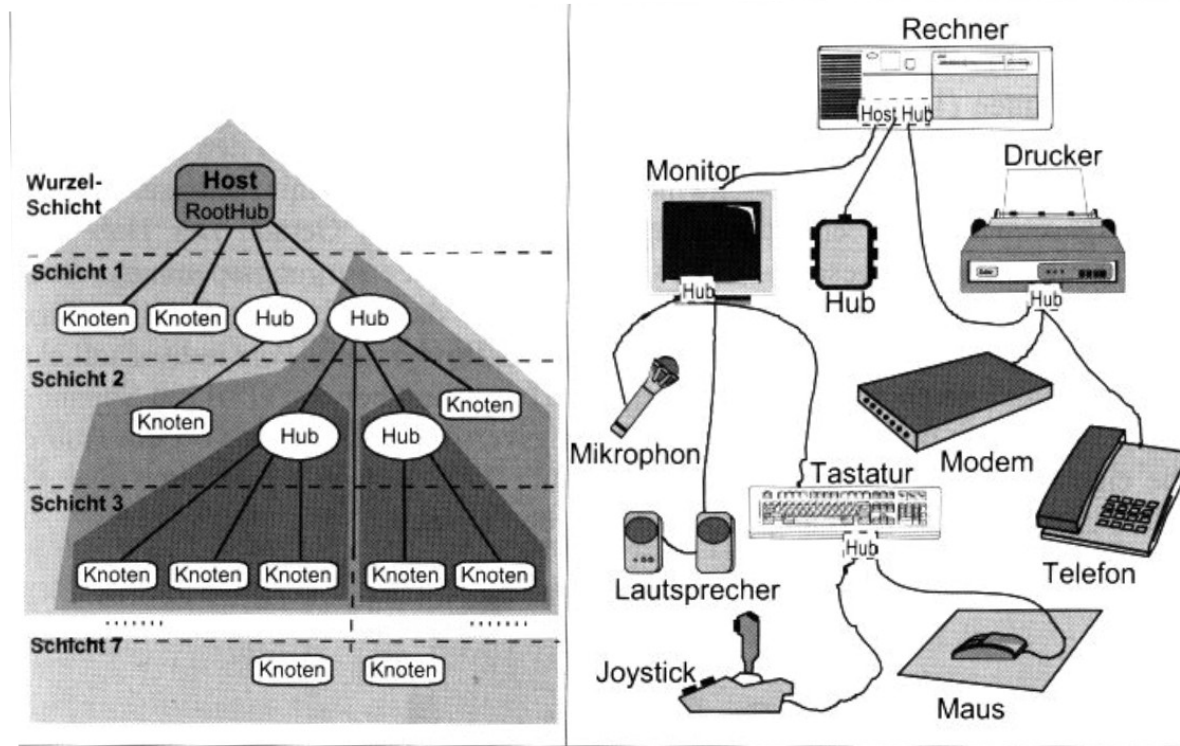


Bild 1.7-2: USB-Topologie und Systembeispiel

Preiswerte Stecker und Kabel,
flexibel konfigurierbar,
„plug-and-play“

Geschwindigkeiten:
USB 1: 1,5 und 12 Mbit/sec, USB 2: 480 Mbit/sec,
kompatibel zu USB 1, USB 3: 5000 Mbit/sec,
kompatibel zu USB 2

Praktisch alle Peripherie-Geräte nutzen heute USB 2/3 als
physikalische Kommunikationsschicht, auch Messgeräte !

Ethernet (IEEE 802.3)

Ethernet ist **DER** Netzwerk-Standard im LAN (=Local Area Network)

Serielle Übertragung mit „embedded clock“

Protokoll: **carrier sense multiple access collision detect (CSMA/CD)**

- 10 Mbps (10Base-T Ethernet) (**Koaxial-Kabel, „twisted pair“**)
- 100 Mbps (Fast Ethernet) (**twisted pair**)
- 1000 Mbps (Gigabit Ethernet) (**twisted pair, optisch**)
- 10-Gigabit Ethernet (**optisch, kurze Strecken auch twisted pair**)

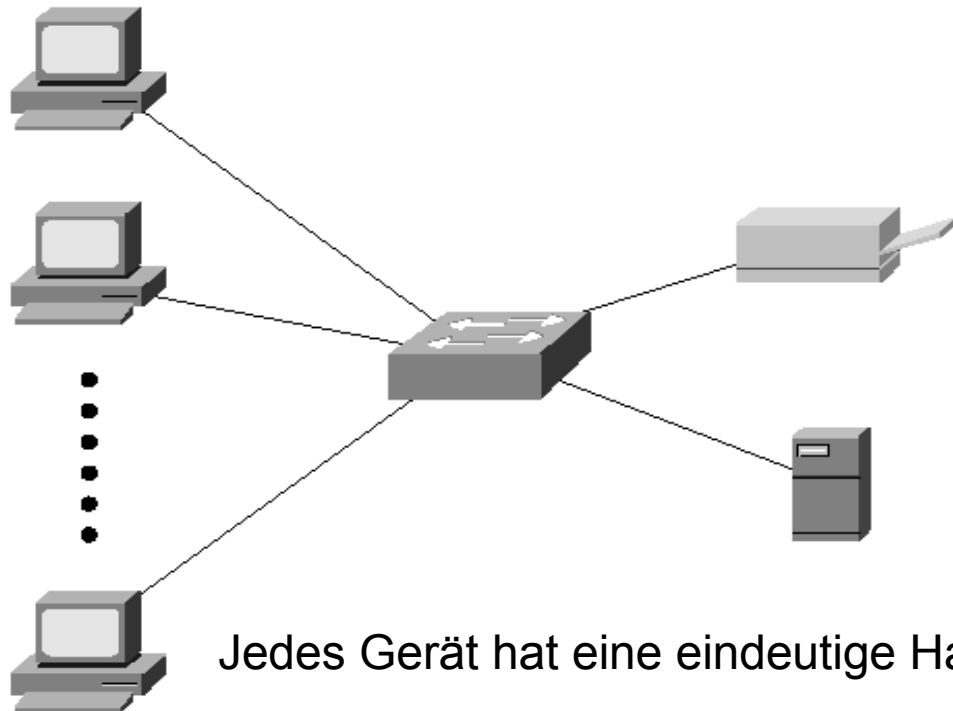
Hardware-Schnittstelle am PCI-Bus (Einsteckkarte oder auf MB integriert)

Verbindung von Rechnern, und – zunehmend - auch (intelligenter) Peripherie
über Hubs, Switches, Router (verbinden Sub-Netze)

Direkte Verbindung von zwei Computern im voll-Duplex-Modus
(**twisted pair**) durch „cross-over“-Kabel

Preiswerte Kabel, vielseitige Netz-Topologien

Hardware - Netzwerk

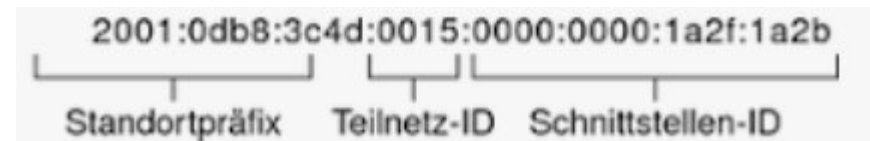


Stern-Topologie
über Hub (simpler Verstärker)
oder Switch (mit „Intelligenz“)

Einzelne Kabel bis max. 100m

Jedes Gerät hat eine eindeutige Hardware-Adresse: die MAC-Adresse;

Auf „Protokoll-Ebene“ (IP-Protokoll) gibt es eine logische Adresse
veraltet („IP4“) 4 Bytes (z.B. 192.168.0.1), aktuell 16 Bytes („Ipv6“)



fast immer auch mit einem Namen verknüpft, z.B.
`fphctssh.physik.kit.edu`

Weitere Bussysteme ...

Es gibt noch eine Vielzahl anderer Bussysteme

- CAMAC (alter Industrie-Standard)
- VME (**Industriestandard, basierend auf Motorola 68000 Busprotokoll**)
- FASTBUS (10 Mbit, 32-bit parallel, HEP Eigenentwicklung)
- Fibre-Channel (z.Zt. 2 Gbit seriell, optisch, für Storage Area Networks)
- iSCSI (SCSI über Ethernet)
- InfiniBand
- ...

Heute ist es meist nicht mehr nötig, eigene Bus-Interfaces zu entwickeln.

Es gibt

- fertige Chips mit
- entsprechenden Treibern für die gängigen Betriebssysteme

Steuerung von Datentransfers

Datenübertragung zwischen Peripherie und Speicher muss gesteuert werden:

1. „Polling“:

vollständig unter CPU-Kontrolle; CPU fragt Status ab und transferiert Daten über die internen CPU-Register in den Speicher

hohe CPU-Belastung während Datentransfer, während der Abfrage ist die CPU blockiert

2. Interrupt-gesteuert:

Schnittstelle unterbricht die CPU nach dem aktuellen

Befehl über spezielle Interrupt-Leitungen; Datentransfer über CPU;

Rücksprung zu vorheriger Tätigkeit

keine Blockierung der CPU durch Status-Abfragen, CPU-Belastung durch Datentransfers

3. DMA (Direct Memory Access):

Interrupt-gesteuert, CPU initiiert Datentransfer, Controller überträgt Daten und signalisiert Ende der Übertragung per Interrupt.

Lohnt sich bei der Übertragung großer Datenmengen

geringst-mögliche CPU-Belastung

Software-Emulation von Hardware auf Wirts-System

- seit langem bei Großrechnern
- alte Rechner auf moderner Hardware/OS, z.B. Atari, V24, Apple
- Windows- oder Linux-Betriebssystem auf Linux und/oder Windows (bisher kommerziell)
- neue Entwicklung:
 - Open Source-Virtualisierung XEN oder KVM unter Linux (schon enthalten in modernen Distributionen)
 - freier Client für VM-Ware virtuelle Maschinen
DOS, Windows oder Linux auf Linux oder Windows
siehe <http://www.vmware.com>
 - VirtualBox, freie Virtualisierungslösung Windows, Linux u. Mac, siehe <http://www.virtualbox.org>

Erlaubt Standard-Umgebung auf unterschiedlichster Hardware bzw. OS

Virtuelle Hardware (2)

Software-Emulation von Hardware auf Wirts-System

- seit langem bei Großrechnern
- alte Rechner auf moderner Hardware/OS, z.B. Atari, V24, Apple II
- Windows- oder Linux-Betriebssystem auf Linux und/oder Windows

aktuell empfohlen:

VirtualBox, freie Virtualisierungslösung für Linux, Windows, Mac
siehe <http://www.virtualbox.org>

Virtualisierung erlaubt

Standard-Umgebung auf unterschiedlichster Hardware bzw. OS

Virtuelle Hardware: Container

Aktuell haben sich sogenannte „**Container**“ als Virtualisierungslösung durchgesetzt.

Nutzung von großen Teilen der Software-Infrastruktur des Wirtssystems - nur notwendige, unterschiedliche Bibliotheken zusätzlich zum Anwendercode

Beispiel: **Docker**

für Linux, Windows und MAC OS



Basierend auf einem Basis-Container (z.B. Ubuntu 22.04) werden in einem „Dockerfile“ die Befehle zur Installation von Zusatz-Paketen angegeben, siehe

<https://etpwww.etp.kit.edu/~quast/Docker/DockerContainers.html>

```
FROM ubuntu:22.04
USER root
# general update and packages
ENV DEBIAN_FRONTEND=noninteractive
RUN apt-get update && apt-get -y upgrade && \
    apt-get install -y sudo binutils git && \
    apt-get install -y python3-pip && \
    apt-get clean && rm -rf /var/lib/apt/lists/* && \
    pip3 install --no-cache-dir --upgrade pip
# python packages for data analysis
RUN pip3 install --no-cache-dir numpy scipy matplotlib \
    pandas seaborn \
    jupyterlab \
    uncertainties \
    iminuit \
    kafe2 \
    PhyPraKit \
    vector
# Create cgda user, include as sudoer
ARG USR=dausr
ARG UID=12345
ARG GID=100
RUN adduser --disabled-password --gecos "" -u ${UID} --gid ${GID} ${USR}
RUN echo "${USR} ALL=(ALL) NOPASSWD:ALL" > /etc/sudoers
WORKDIR /home/${USR}
ENV HOME /home/${USR}
COPY dot-bashrc $HOME/.bashrc
COPY . $HOME
RUN chown --recursive $USR $HOME
RUN chgrp --recursive users $HOME
EXPOSE 8888
USER $USR
ENV SHELL /bin/bash
CMD jupyter-lab --ip 0.0.0.0 --no-browser
## to execute, start with
# docker run -p 8888:8888 cgda:22.04
```

*Docker-Container incl. Jupyter-Server
für diesen Kurs (incl. Beschreibung)*

Analog ↔ Digital

Literatur

(für den Physiker) :

Rohe, Elektronik für Physiker

Renk, Meßdatenerfassung in der Kern- u. Teilchenphysik

(beide als Teubner Studienbücher)

Analog-Digitalwandlung

Wird immer dann benötigt, wenn ein Rechner Ausgangssignale zur Steuerung und Regelung liefern soll, z.B.

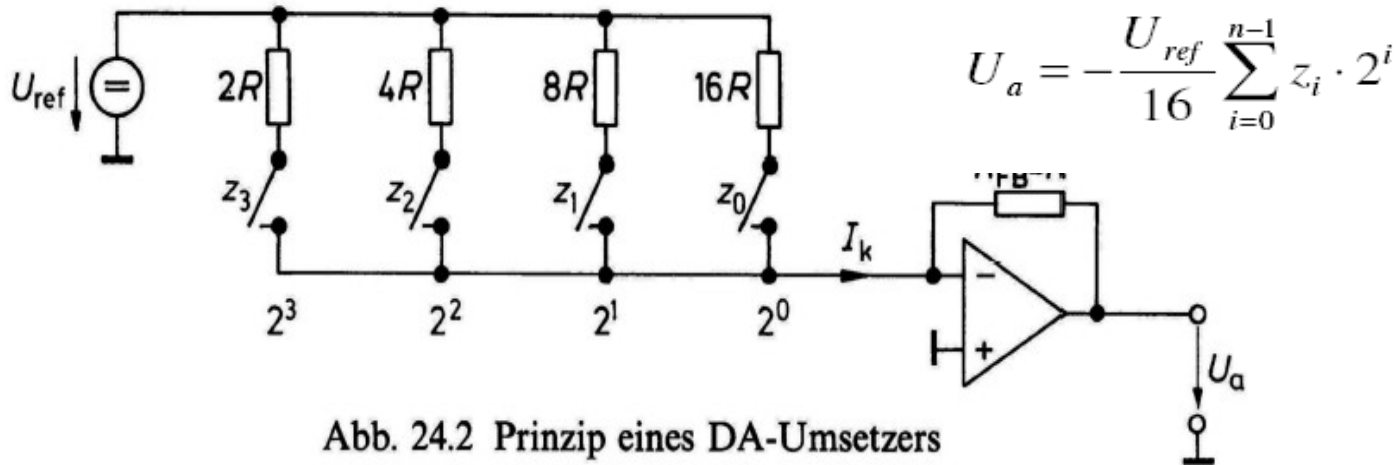
Ansteuerung von

- Lautsprechern
- analogen Video-Geräten
- Steuerspannungen für alle Arten von Reglern
- ...

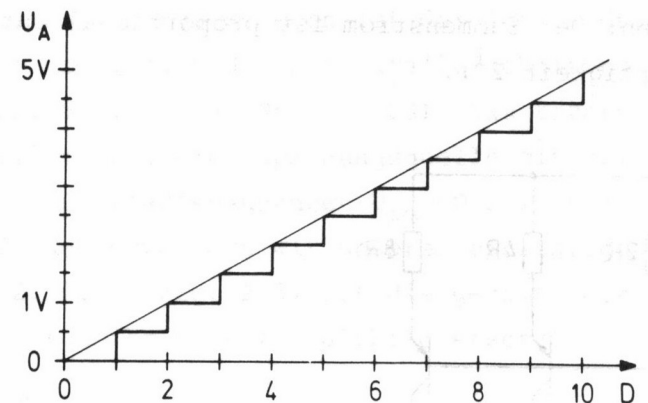
Digital-Analog-Wandler setzen einen digitalen Wert D in eine entsprechende Spannung um U .

d.h. Abbildung von $0 \leq D < 2^n \rightarrow U_{\min} \leq U < U_{\max}$

Prinzip Digital-Analog-Wandlung



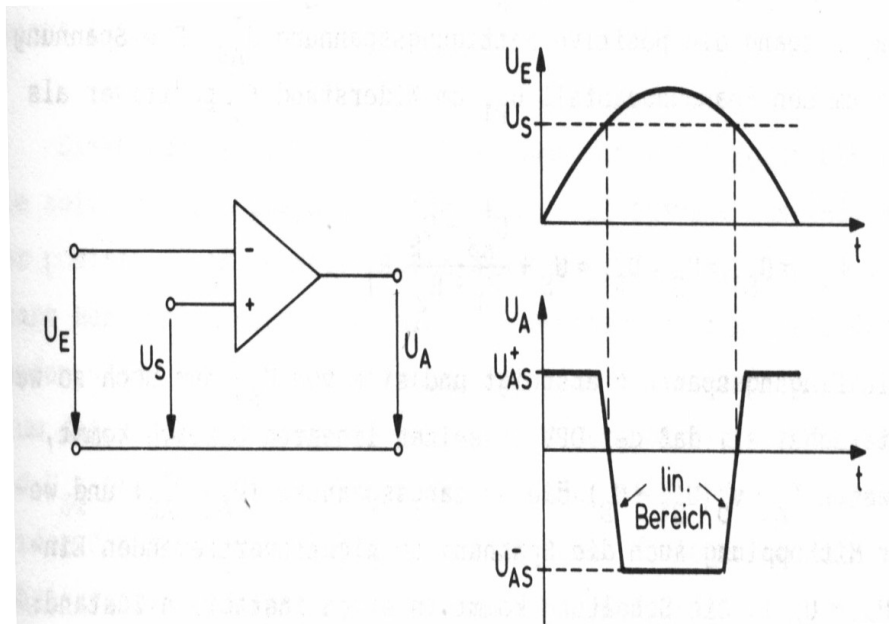
Prinzip: den Werten der einzelnen Bits einer Binärzahl entsprechende Ströme werden aufaddiert.



Stufencharakter der Ausgangsspannung unvermeidlich !

Analog-Digital-Wandlung

Erfassung von Messdaten durch Rechner erfordert
Wandlung der analogen Werte in digitale Daten



Schlüsselement:
Komparator

z.B. mit einfachem OP:

$$U_E < U_S : U_A = U^+$$

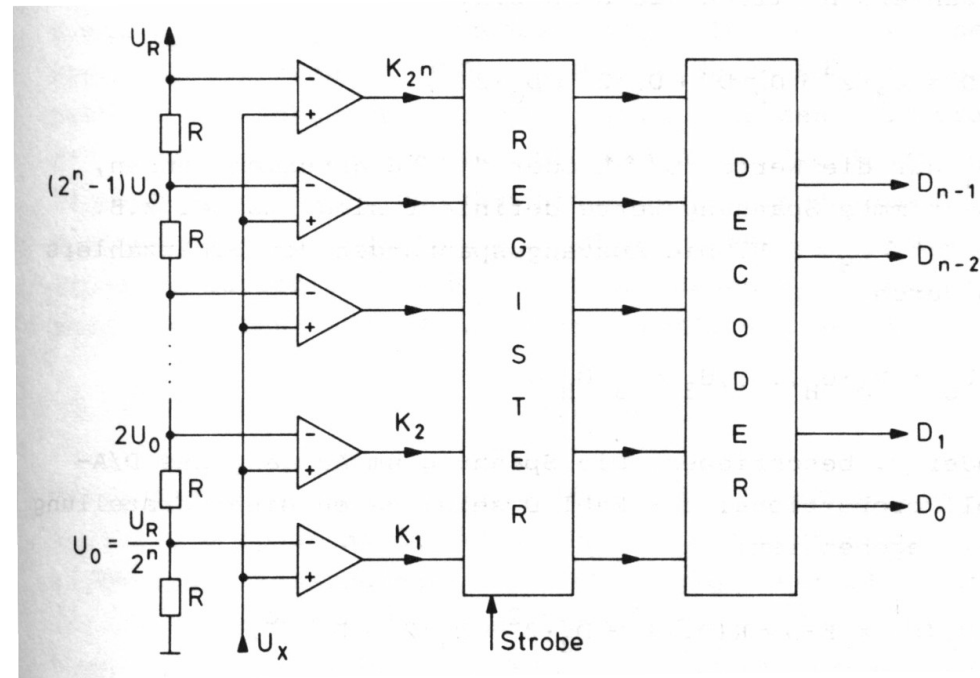
$$U_E > U_S : U_A = U^-$$

d.h. **digitale Ausgabe !**

Grundprinzip einer Analog-Digitalwandlung mittels Komparator

Analog-Digital-Wandlung

schnellstes, aber aufwändiges Verfahren: **Flash-Wandlung**



Prinzip: Simultaner Vergleich mit vielen Referenzspannungen

sehr schnell, **aber aufwändig** (1024 Komparatoren und ein
1024 → 10 Encoder für 10 bit Auflösung)

Zur Zeit sind Wandlungsraten bis zu einigen GHz möglich

Analog-Digital-Wandlung

Digitalisiertes Signal entspricht nur näherungsweise dem Original,
„Quantisierungsfehler“ sind unvermeidlich !

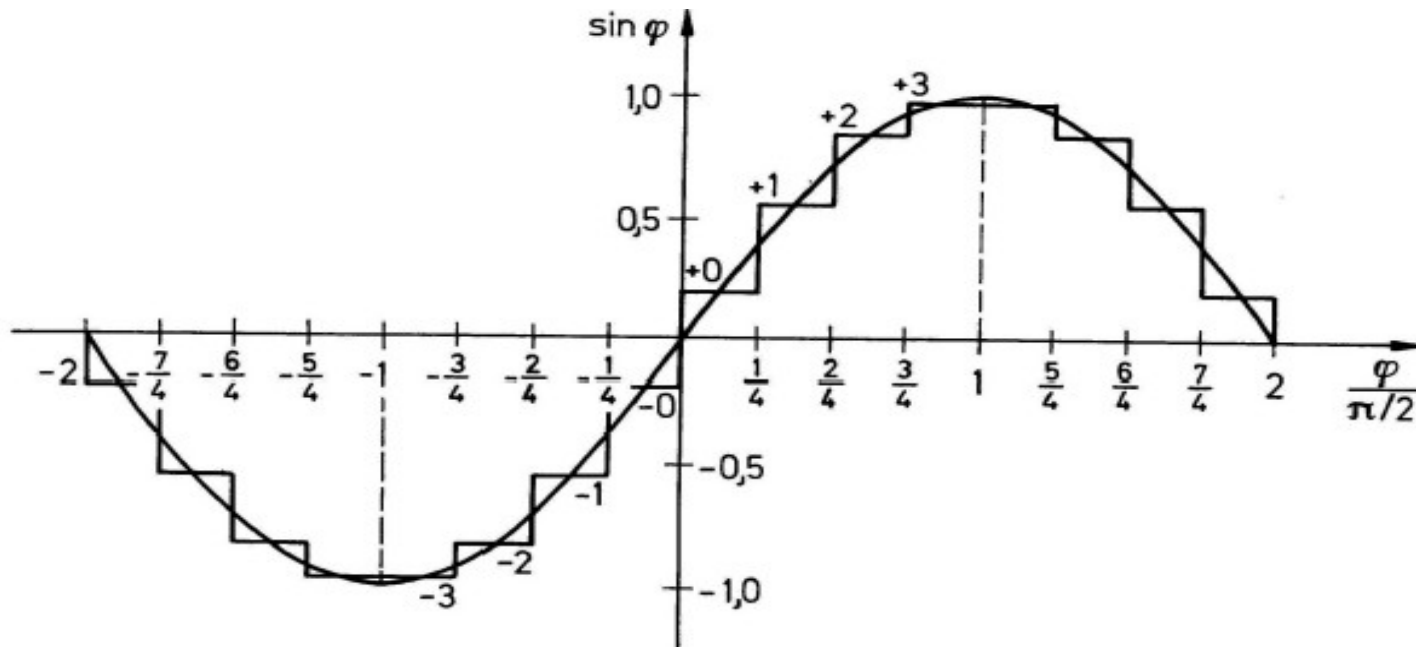


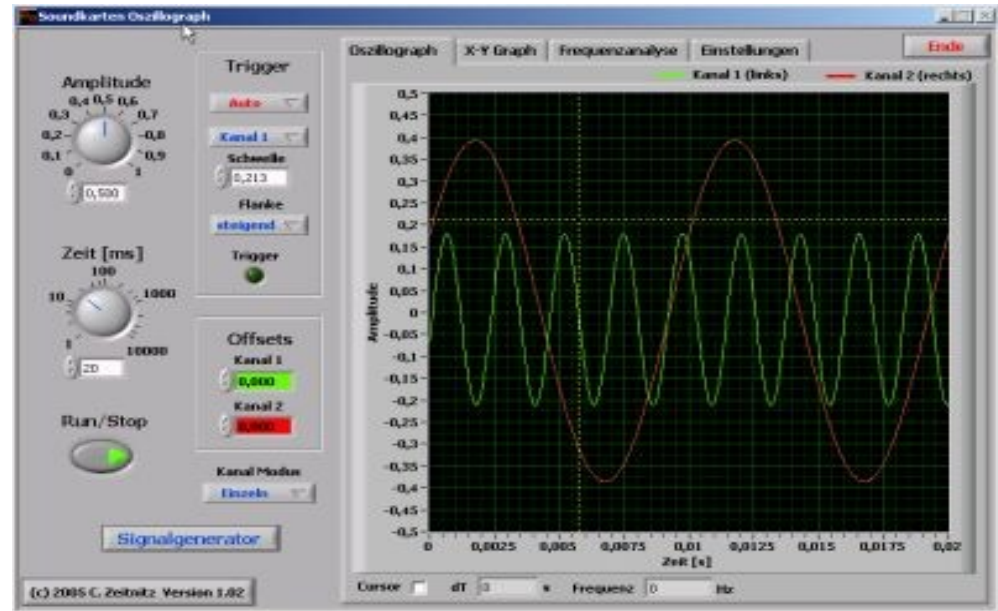
Abb. 24.17 Approximation einer Sinusschwingung mit 16 Stufen

„Abtastrate“ muss mindestens dem zweifachen der höchsten im
Signal vorhandenen Frequenz entsprechen (Nyquist-Theorem)

Oszilloskop für jedermann

Moderne Soundkarten haben

- (mindestens) 2 Eingangskanäle
- (mindestens) 2 Ausgänge
- mit (mindestens) 44,1 kHz Abtastfrequenz
- und 8 – 16 Bit Auflösung



Nachteil:

- Ein- und Ausgangsspannungen im Audio-Bereich ($< 0.7 \text{ V}$)
- nur Wechselspannungen oberhalb $\sim 20 \text{ Hz}$

Einen Oszillographen mit Frequenzanalyse für die Soundkarte gibt es (leider nur für Windows) unter

http://www.zeitnitz.de/Christian/Scope/Scope_ger.html

entwickelt mit  National Instruments LabVIEW

(leider fällt dieser Vorlesungsteil aus)

- Rechnerarchitekturen, Parallelisierung und Vektorisierung
- Kollaborative Softwareentwicklung