Karlsruher Institut für Technologie

Elektronik für Physiker

Vorlesungsskript

Prof. Dr. Marc Weber

Institut für Prozessdatenverarbeitung und Elektronik

Version vom 21. Januar 2016

Vorwort

Die Vorlesung „Elektronik für Physiker“ hat eine jahrzehntelange Tradition am Karlsruher Institut für Technologie und der vormaligen Universität Karlsruhe. Die Vorlesung samt Praktikum wurde von Dr. Wolfgang Jüngst und Prof. Christian Weddigen begründet. Ihr Buch zur Vorlesung ist trotz des großen Fortschritts der Mikroelektronik immer noch äußerst lesenswert und ist über den Semesterapparat leicht zugänglich. Die Vorlesung wurde 1993 von Prof. Hartmut Gemmeke weitergeführt und über die Jahre kontinuierlich modernisiert, insbesondere im Bereich der programmierbaren Hardware. Ich orientiere mich an den bewährten Konzepten meiner Vorgänger.

In diesem Skript wird versucht, in die Grundlagen der Analog- und Digitalelektronik einzuführen. Dabei gehe ich von dem in physikalischen Experimenten typischen Szenario aus, bei dem ein in der Regel schwaches analoges Eingangssignal zunächst verstärkt, gefiltert und digitalisiert wird, bevor es digital weiterverarbeitet, zwischengespeichert und schließlich analysiert werden kann.

Es gibt hoffentlich viele Gründe, diese Vorlesung zu besuchen oder sich mit dem Skript zu befassen. Ich möchte hier die offensichtlichen ausführen:

- Experimentalphysiker brauchen elektronische Grundkenntnisse. Zum Beispiel, um Laborgeräte zu bedienen, um bessere Experimente zu entwerfen, um zu verstehen, wo physikalische und messtechnische Grenzen liegen, um mit Ingenieuren effektiv zusammenzuarbeiten. Auch als Physiker in der Finanzindustrie, als Unternehmensberater und erst recht in der Autoindustrie und im Maschinenbau kann man Elektronikverständnis nutzen.

- Elektronik beeinflusst unser Leben – und zwar unseren Lebensstil wie unsere Lebensqualität. Beispiele in loser historischer Reihenfolge sind: elektrisches Licht, Telefon, Radio, PC, E-Mail, Internet, Mobiltelefon, Laptops, digitale Kameras, GPS, Medizintechnik und vieles mehr. Ich denke, es ist interessant zu verstehen, was diesen Geräten und Lebensformen zugrunde liegt.

- Die Halbleiterindustrie im weitesten Sinne ist ein sehr großer Markt mit jährlichen Gewinnen um 350 Milliarden USD (Stand 2014). Rund 10 % des Weltbruttoprodukts wird in der Mikroelektronikindustrie erwirtschaftet. Die Elektronikindustrie ist der Motor für viele andere Industriezweige. Die Wahrscheinlichkeit, dass Sie direkt oder indirekt mit der Elektronikindustrie zu tun haben werden, ist also nicht ganz klein.

Ich persönlich schätze an Elektronikentwicklung und Detektorbau allerdings besonders die folgenden Aspekte:

- Der technologische Fortschritt in der Mikroelektronik ist besonders groß. Ein populärer Beleg dafür ist „Moore's Law“, also die Verdopplung der Transistordichte in integrierten Schaltkreisen alle 18 - 24 Monate.

- Elektronikentwicklung und Detektorbau findet in der Regel in multidisziplinären (und oft auch internationalen) Gruppen statt.

- Man kann sich mit Elektronikkenntnissen sehr nützlich machen. Fast jedes Experiment, das an der Spitze der Forschung steht, ist auf anspruchsvolle Elektronik angewiesen.

Das Erstellen dieses Skriptes, auch in dieser noch vorläufigen Form, wäre ohne Hilfe anderer nicht möglich gewesen. Besonderer Dank gebührt Francesca Rindone für ihr Geschick im Umgang mit Latex und ihre Akribie als Lektorin. Saskia Baier und Thomas Schuh haben unzählige Gleichungen und Bilder erstellt. Außerdem möchte ich mich bei Francesca Rindone, Julius Hartmann, Thomas Schuh und den Hörern der Vorlesung für zahlreiche didaktische und inhaltliche Hinweise bedanken.

Februar 2015

Inhaltsverzeichnis

1 Grundlagen

1.1 Elementarladung, Strom und Spannung

Die grundlegenden elektronischen Größen sind Strom und Spannung.

Der **Strom** entspricht bewegter Ladung, also

$$\boldsymbol{I = \frac{dQ}{dt}} \quad . \tag{1.1}$$

Dabei steht I für Strom, Q für Ladung und t für Zeit. Die Einheit des Stroms I ist das Ampere A. Ein Strom von 1 A entspricht einem Ladungsfluss von einem Coulomb pro Sekunde, also 1 A = 1 C/s. Ein Coulomb entspricht der Ladung von $6,241 \cdot 10^{18}$ Elektronen.

Das Coulomb ist also historisch bedingt eine eher „große" Einheit. Praktischer ist oft das Femtocoulomb; 1 fC entspricht z. B. einer Ladung von rund 6000 Elektronen. Entsprechend werden wir selten mit Strömen größer als 1 A zutun haben, dagegen sehr häufig mit Strömen im Bereich von mA oder μA.

Die Stromflussrichtung zeigt in elektrischen Schaltkreisen konventionellerweise vom größeren Potential (+) zum kleineren Potential (−).

In metallischen Leitern sind Elektronen die beweglichen Ladungsträger, das heißt, der tatsächliche Ladungsfluss findet in entgegengesetzter Richtung statt. Es gibt natürlich auch Ionenleitung in Gasen, Flüssigkeiten und in Zellen (Ionenkanälen). In Halbleitern sind Elektronen und sogenannte Elektronenlöcher die beweglichen Ladungsträger.

Die **Spannung** U oder die Potentialdifferenz zwischen zwei Punkten entspricht der Energie W, die man aufbringen muss, um eine elektrische Ladung q in einem elektrischem Feld $\vec{E}$ zu bewegen,

$$\boldsymbol{\frac{W}{q} = U = \int_a^b \vec{E}\, d\vec{s}} \quad . \tag{1.2}$$

Die Einheit der Spannung ist das Volt V. Wenn man salopp von *Spannung an dem Punkt a einer Schaltung* redet, bezieht man sich implizit auf einen weiteren Bezugspunkt, der oft Masse oder Erde genannt wird.

Was ist wichtiger? Strom oder Spannung? Die Spannung zwischen zwei Punkten erzeugt den Strom. Andererseits führt Stromfluss durch elektronische Bausteine zu einem Spannungsabfall. Zum Beispiel fällt zwischen den beiden Enden eines ohmschen Widerstands R bei Stromfluss die Spannung $U = R\,I$ ab. Es gibt sogenannte Stromquellen wie Spannungsquellen.

Strom und Spannung können also etwas verwirrend sein, aber man gewöhnt sich mit der Zeit daran.

1.2 Passive elektronische Bauelemente

Die elektronischen Grundbausteine jeder Schaltung sind Widerstände, Kondensatoren, Spulen, Dioden und Transistoren. Wir werden uns zuerst mit dem Widerstand, dem Kondensator und der Spule befassen, den sogenannten linearen passiven Bauelementen.

1.3 Ohmscher Widerstand

Der Widerstand ist, wie auch Kondensator, Spule und Diode, ein Zweipol; das heißt, er hat zwei Anschlüsse. Der Strom durch einen Widerstand und die Spannung zwischen den Enden a und b des Widerstands sind zueinander proportional.

Dies entspricht dem **Ohmschen Gesetz**

$$\boldsymbol{U = R\,I} \quad . \tag{1.3}$$

wobei U für den Wert der Spannung, R für den Widerstand und I für den Strom steht.

Ganz praktisch ausgedrückt impliziert das Ohmsche Gesetz folgendes:

1. Legt man an den Widerstand R eine Spannung U an, so fließt der Strom $I = \frac{U}{R}$ durch den Widerstand.

2. Lässt man durch den Widerstand R einen Strom I fließen, so fällt die Spannung $U = R\,I$ am Widerstand ab.

Die Einheit des Widerstands ist das Ohm mit $1\,\frac{\mathrm{V}}{\mathrm{A}} = 1\,\Omega$.

Der Stromfluß führt zur Aufwärmung des Widerstands. Die thermische Verlustleistung P des Widerstands beträgt

$$\boldsymbol{P = U\,I = \frac{U^2}{R} = I^2\,R} \quad . \tag{1.4}$$

Auch elektrische Leiterbahnen auf einer Platine haben einen nichtverschwindenden Widerstand. Eine gute direkte Verbindung zwischen zwei Punkten eines Schaltkreises sollte einen Widerstand deutlich kleiner als $1\,\Omega$ aufweisen, z. B. $1\,\mathrm{m}\Omega$ oder noch besser einige $\mu\Omega$.

Der Widerstand eines Leiters hängt von seiner Geometrie und seinen Materialeigenschaften ab. So ist der Widerstand eines Leiters der Länge l und des Querschnittes A gegeben durch

$$\boldsymbol{R = \rho\,\frac{l}{A}} \quad . \tag{1.5}$$

Dabei ist ρ der spezifische Widerstand des Leitermaterials.

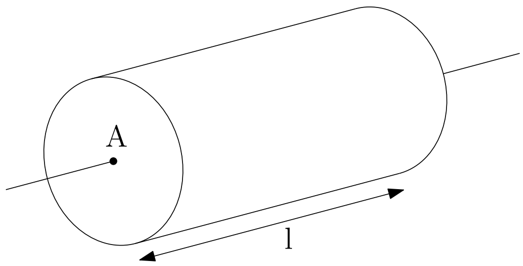


Abb. 1.1 Illustration des spezifischen Widerstands.

In der folgenden Tabelle sind beliebte Leitermaterialien und ihr spezifischer Widerstand aufgeführt.

Material	$\rho\,[10^{-8}\,\Omega\,\mathbf{m}]$	Z
Silber	1,59	47
Kupfer	1,68	29
Gold	2,21	79
Aluminium	2,65	13

Tab. 1.1 Spezifischer Widerstand bei 20 °C und Kernladungszahl Z verschiedener Metalle.

Silber ist ein sehr guter Leiter, aber für die meisten Anwendungen zu teuer. Daher wird meistens Kupfer verwendet. Kupfer ist ein gut formbares und verarbeitbares Material, das sich sehr gut löten lässt. Gold ist auch teuer, oxidiert aber nicht. Deshalb sind offen liegende Flächen einer Platine manchmal mit einer dünnen ($< 1\,\mu$m) Goldschicht überzogen, insbesondere für späteres Drahtbonden. Aluminium ist ein problematisches Material, das leicht oxidiert und sich deshalb schlecht löten lässt. Es ist aber leicht und hat eine kleine Atomzahl $Z = 13$. Daher wird es gerne in Detektoraufbauten verwendet, z. B. als Kabel oder Kühlrohr, wenn es auf geringe Vielfachstreuung ankommt[1]. Oft bereut man nachträglich diese Wahl.

Widerstände werden verwendet:

- zur Strombegrenzung bei fester Spannung
- zur Spannungsteilung (mit zwei oder mehr Widerständen in Reihe)
- zur Umwandlung eines gegebenen Stroms in eine definierte Spannung
- zum Entladen von Kondensatoren bzw. zur Einstellung von Zeitkonstanten
- zum Einstellen einer Verstärkung
- als Rückkopplungselement in Verstärkern.

Oft sind aber Transistoren die besseren „Widerstände“.

Man kann auch nichtlinearen Bauelementen einen sogenannten *dynamischen* Widerstand oder

[1] Die Strahlungslänge X_0 von Kupfer ist 14,4 mm und die von Aluminium 88,9 mm.

eine *dynamische Impedanz* zuordnen durch die Definition $R = dU/dI$. (In der Elektrotechnik werden dynamische Widerstände oft mit kleinen Buchstaben bezeichent, also $r = dU/dI$. Diese Konvention wird im Skript nicht verwendet.) Die dynamische Impedanz ist nicht konstant, sondern hängt vom gewählten Arbeitspunkt ab, das heißt vom Wert des Stroms, der Spannung oder – bei Wechselspannungen – der Frequenz.

Ein typisches Beispiel ist die Kennlinie einer Diode (Abb. 1.2). Der dynamische Widerstand entspricht dem differentiellen Widerstand dU/dI am Punkt (I_0, U_0) und ist klein, da die Kurve sehr flach verläuft. Der Wert von U_0/I_0 dagegen ist eher groß.

Der dynamische Widerstand bestimmt das Kleinsignalverhalten des Bausteins in einem Schaltkreis.

Anmerkung: Die Leistung P ist dagegen durch $P = U_0\, I_0$, die Großsignalgrößen, gegeben.

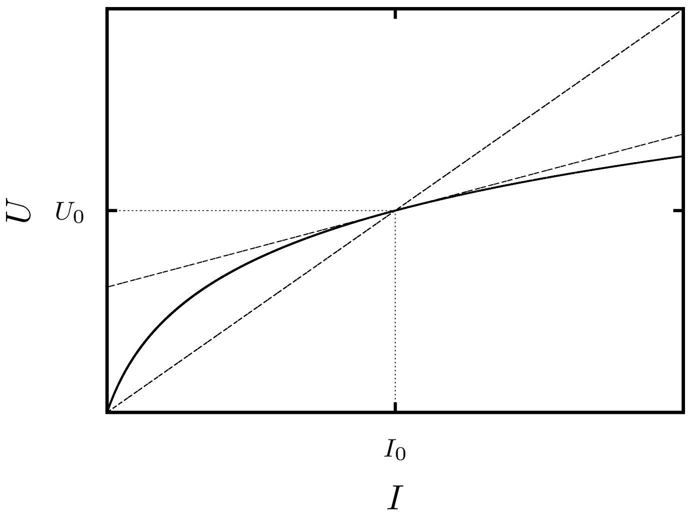


Abb. 1.2 Differentieller Widerstand und integraler Widerstand am Beispiel einer Diodenkennlinie.

Anmerkung: Die Achsenwahl in Abbildung 1.2 ist unüblich. Meistens werden die Spannung auf der Abszisse und der Strom auf der Ordinate aufgetragen, nicht umgekehrt.

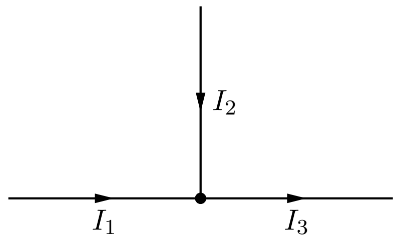
1.4 Die Kirchhoffschen Gesetze

Die Kirchhoffschen Gesetze beschreiben Beziehungen zwischen Strömen (Knotenregel) und zwischen Spannungen (Maschenregel) in elektrischen Netzwerken.

Knotenregel [2] (Ladungserhaltung):

Für jeden Knoten eines Netzes ist die Summe der zufließenden Ströme gleich der Summe der abfließenden Ströme (Abb. 1.3). Dabei ist die Zuordnung der Pfeilrichtungen willkürlich. Das korrekte Vorzeichen der Ströme ergibt sich automatisch als Ergebnis der Netzwerkanalyse.

[2] Engl. *Kirchhoff's Current Law (KCL)*

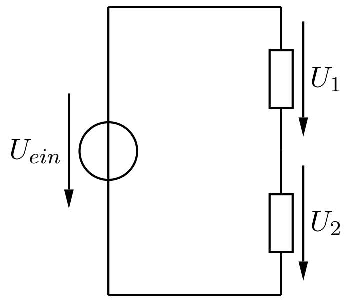


$$\boldsymbol{I_1 + I_2 = I_3} \tag{1.6}$$

Abb. 1.3 Knotenregel.

Maschenregel [3] (Energieerhaltung):

Die Summe der Spannungsabfälle entlang jeder geschlossenen Schleife eines Netzes ist null (Abb. 1.4).



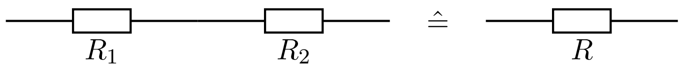
$$\boldsymbol{U_{ein} - U_2 - U_1 = 0} \tag{1.7}$$

Abb. 1.4 Maschenregel: Spannungsquelle mit zwei Widerständen.

Aus den Kirchhoffschen Gesetzen ergeben sich einfache Regeln zum Berechnen von Parallel- und Reihenschaltungen von Widerständen.

1.5 Reihenschaltung von Widerständen

Zwei Widerstände in Serie haben einen größeren Widerstand als jeder einzelne Widerstände (Abb. 1.5).



$$\boldsymbol{R = R_1 + R_2} \tag{1.8}$$

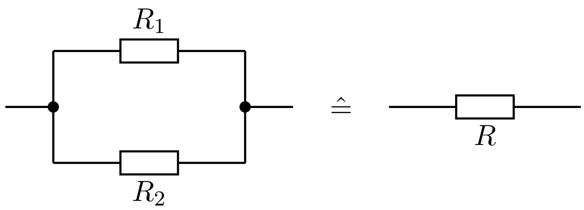
Abb. 1.5 Reihenschaltung von Widerständen.

[3]Engl. *Kirchhoff's Voltage Law (KVL)*

Anmerkung: Dies entspricht im Sinne von $R = \rho \frac{l}{A}$ einfach einem längeren Widerstand.

1.6 Parallelschaltung von Widerständen

Zwei parallele Widerstände haben einen geringeren Widerstand als die Einzelwiderstände (Abb. 1.6). Der kleinste Widerstand dominiert den Gesamtwiderstand R.



$$\boldsymbol{R = \frac{1}{\frac{1}{R_1} + \frac{1}{R_2}} = \frac{R_1\,R_2}{R_1 + R_2}} \tag{1.9}$$

Abb. 1.6 Parallelschaltung von Widerständen.

Diese Formel sollte man sich merken. Oft benutzt man die Schreibweise $R_1 \| R_2$ für den Gesamtwiderstand, ohne das Ergebnis explizit auszurechnen.

Anmerkung: Parallelwiderstände entsprechen im Sinne von $R = \rho \frac{l}{A}$ einem Einzelwiderstand mit größerem Querschnitt.

Die Größe $\frac{1}{R} = G$ wird auch der **Leitwert** genannt. Die Einheit des Leitwertes ist das Siemens mit $\mathrm{S} = \frac{1}{\Omega}$.

Für parallele Widerstände gilt auch: Die Summe der Leitwerte der Einzelwiderstände ergibt den Leitwert des Gesamtwiderstands, also

$$\boldsymbol{G = \frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} = G_1 + G_2} \quad . \tag{1.10}$$

1.7 Spannungsquelle

Elektrische Schaltkreise brauchen eine Energiequelle zum Betrieb – entweder eine Spannungsquelle oder eine Stromquelle.

Die Spannungsquelle ist uns vertrauter. Die ideale Spannungsquelle ist ein Zweipol und liefert eine konstante Spannung unabhängig von dem Strombedarf [4] der Last. Beispiele dafür sind Batterien (Gleichspannung) und die Netzspannung an der Steckdose (Wechselspannung). Die reale Spannungsquelle kann allerdings keinen beliebig großen Stromfluss bewirken, sondern nur einen Maximalstrom. Die abgegebene Leistung ist begrenzt. Genauso ist die Spannung, falls sie überhaupt einstellbar ist, nach oben begrenzt. Die Spannung der realen Spannungsquelle nimmt mit zunehmendem Strombedarf der Last etwas ab (Abb. 1.7).

[4] Strombedarf, nicht Stromverbrauch!

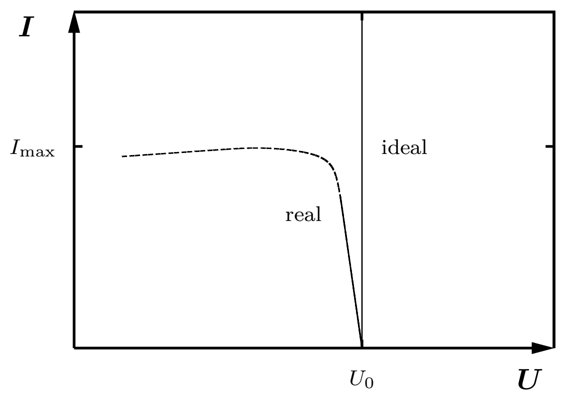


Abb. 1.7 Skizze der Strom-Spannungskennlinie einer idealen (durchgezogene Linie) und einer realen Spannungsquelle (gestrichelte Linie).

Es gilt

$$U_{aus} = U_0 - I\,R_i \quad .$$

Dabei ist U_{aus} die an der Last anliegende Ausgangsspannung (Klemmenspannung), U_0 die sogenannte Quellen- oder Leerlaufspannung, I der Strom und R_i der Innenwiderstand der Spannungsquelle. Eine reale Spannungsquelle lässt sich als eine ideale Spannungsquelle mit einem in Reihe geschalteten Innenwiderstand R_i darstellen (Abb. 1.8).

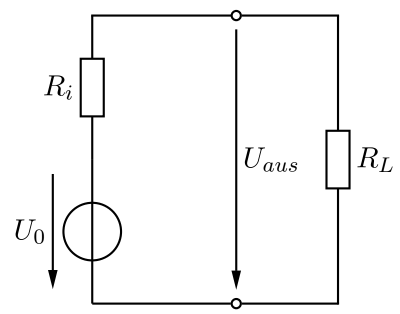


Abb. 1.8 Schaltkreis einer Spannungsquelle mit Innenwiderstand R_i und Lastwiderstand R_L.

Eine gute Spannungsquelle sollte einen kleinen Innenwiderstand $R_i < 1\,\Omega$ haben, damit die Abhängigkeit der Ausgangsspannung vom Laststrom klein ist. Typische Werte sind:

- Netzgerät: $R_i \sim 10^{-5}\,\Omega$
- Autobatterie: $R_i \sim 10^{-2}\,\Omega$
- Monozelle: $R_i \sim 0,1 - 1\,\Omega$

Spannungsquellen funktionieren am besten „ohne Last“ oder bei kleinen Lastströmen, am zweitbesten bei *konstantem* Strom. Für empfindliche analoge Schaltungen ist die Stabilität der Versorgungsspannung kritisch. Manchmal ist ein konstanter Strom besser als ein im Mittel geringerer, aber schwankender Strombedarf.

Fragen:

1. Wie misst man U_0 und R_i für eine gegebene Spannungsquelle?
2. Definieren Sie die Energieeffizienz einer Spannungsquelle in Abhängigkeit von R_L und berechnen Sie diese.

1.8 Stromquelle

Die ideale Stromquelle liefert einen konstanten Strom unabhängig von der Spannung, die sich an der Last einstellt. Eine reale Stromquelle dagegen kann den Strom nicht für beliebig hohe Spannungen konstant halten. Die Leistung ist begrenzt. Eine zu hohe Spannung würde auch die Komponenten der Stromquelle zerstören.

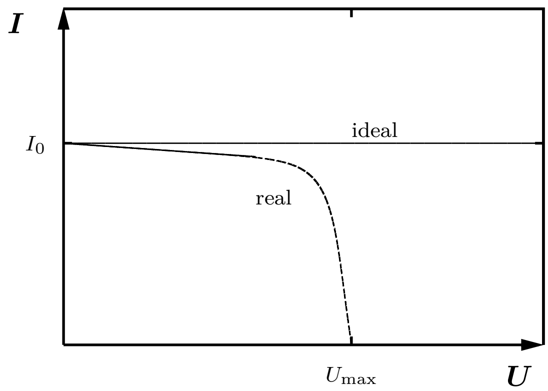


Abb. 1.9 Skizze der Strom-Spannungskennlinie einer idealen (durchgezogene Linie) und einer realen Stromquelle (gestrichelte Linie).

Eine reale Stromquelle hat ungefähr den in Abbildung 1.9 skizzierten Strom-Spannungsverlauf. Mit zunehmender Spannung U_{aus} an der Last nimmt der Strom I ab. Es gilt

$$I = I_0 - \frac{U_{aus}}{R_i} \quad . \tag{1.11}$$

Dabei ist I der durch die Last fließende Klemmenstrom, I_0 der Quellen- oder Leerlaufstrom und R_i der Innenwiderstand der Stromquelle. Die reale Stromquelle lässt sich als eine ideale Stromquelle mit einem parallel geschalteten Innenwiderstand R_i darstellen. Dies entspricht dem Schaltkreis in Abbildung 1.10.

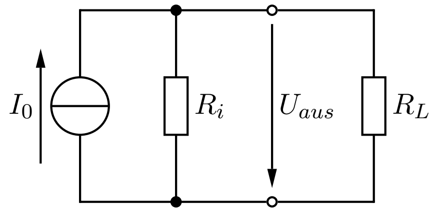


Abb. 1.10 Schaltkreis einer Stromquelle mit Innenwiderstand R_i und Lastwiderstand R_L.

Eine Stromquelle sollte einen möglichst großen Innenwiderstand R_i haben, ein guter Wert wäre $\sim 10^7\,\Omega$. Der Querbalken im Symbol der idealen Stromquelle deutet an, dass hier der Innenwiderstand unendlich ist. Die an der Last anliegende Spannung stellt sich nach

$$U_{aus} = U_L = I_0\,(R_i \| R_L)$$

ein.

Eine Stromquelle funktioniert am besten bei „Kurzschluss“. Leerlaufbetrieb ist keine gute Idee.

Frage: Wie bestimmt man den Innenwiderstand einer Stromquelle?

Die Messung des Innenwiderstands bei hohen Frequenzen ist schwierig.

Anmerkung: Die Variation der Lastspannung bei fester Frequenz f um $\Delta\,U$ führt ja nur zu kleinen $\Delta\,I$, da $R_i = dU/dI$ groß sein soll.

Stromquellen sind vielleicht weniger vertraut, aber wichtig und weit verbreitet.

1.9 Der Spannungsteiler

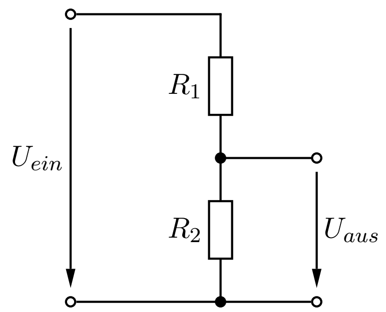


Abb. 1.11 Schaltkreis eines Spannungsteilers mit Lastwiderstand R_L.

Der Spannungsteiler (Abb. 1.11) ist einer der häufigsten Schaltkreisbestandteile. Es gilt $U_{ein} = U_1 + U_2 = (R_1 + R_2)\,I$ und $U_{aus} = R_2\,I$. Damit ergibt sich die Ausgangsspannung U_{aus} zu

$$U_{aus} = U_{ein}\,\frac{R_2}{R_1 + R_2} \quad . \tag{1.12}$$

Sie ist damit kleiner als die Eingangsspannung U_{ein}. Bei dieser Herleitung haben wir vorausgesetzt, dass durch die an U_{aus} angeschlossene Last kein oder nur ein sehr kleiner Strom fließt. Der Spannungsteiler ist keine gute Spannungsquelle, da durch R_1 und R_2 ein Verluststrom fließt und der Innenwiderstand R_i dieser „Spannungsquelle“ nicht klein ist.

1.10 Der Kondensator

Der Kondensator ist ein Zweipol. Ein einfaches Beispiel ist ein Plattenkondensator bestehend aus zwei Elektroden, die durch einen nichtleitenden Spalt voneinander getrennt sind. Entsprechend ist das Symbol für den Kondensator gewählt (siehe Abb. 1.12).

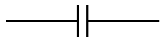


Abb. 1.12 Schaltsymbol für den Kondensator.

Die Spannung U zwischen den Elektroden ist proportional zu der auf ihnen gespeicherten Ladung Q [5]:

$$\boldsymbol{Q = C\,U} \tag{1.13}$$

mit der Kapazität C. Die Einheit der Kapazität ist das Farad mit $1\,\mathrm{F} = 1\,\frac{\mathrm{C}}{\mathrm{V}}$.

Typische Kondensatoren außerhalb der Leistungselektronik haben eine Kapazität von pF bis μF.

Kondensatoren werden verwendet:

- zur Blockierung des Gleichstromanteils in Schaltungen
- zur Siebung (Kurzschluss) eines Wechselstromanteils
- zum Filtern, wie z. B. mit Hochpass-, Tiefpass- und Schwingkreiskomponenten
- zur Ladungsspeicherung.

Die Kapazität eines Kondensators hängt von seiner Geometrie und der relativen Permittivität ϵ seines Dielektrikums im Spalt ab.

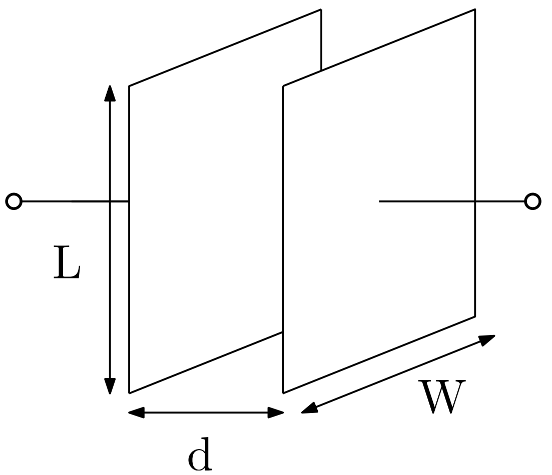


Abb. 1.13 Skizze eines Plattenkondensators.

$$\boldsymbol{C = \frac{\epsilon_0\,\epsilon_r\,A}{d}} \tag{1.14}$$

Dabei ist in Gleichung 1.14 $A = W\,L$ die Elektrodenfläche, d der Elektrodenabstand des jeweiligen Plattenkondensators, ϵ_0 die sogenannte Dielektrizitätskonstante, Permittivität oder

[5] $+\,Q$ auf der einen Elektrode und $-\,Q$ auf der anderen.

elektrische Feldkonstante und ϵ_r die materialabhängige relative Dielektrizitätszahl. Der Wert der Dielektrizitätskonstanten

$$\epsilon_0 = 8,859\,10^{-12} \quad [\mathrm{F/m}] \quad .$$

Die im Kondensator gespeicherte Energie E ist

$$\boldsymbol{E = \frac{1}{2}\frac{Q^2}{C} = \frac{1}{2}C\,U^2} \quad . \tag{1.15}$$

Die Strom-Spannung-Charakteristik eines Kondensators ergibt sich durch Differenzieren von Gleichung 1.13

$$\boldsymbol{I = \frac{dQ}{dt} = C\,\frac{dU}{dt}} \quad . \tag{1.16}$$

Was bedeutet das?

- Der Strom ist nicht proportional zur Spannung, sondern zur Änderung der Spannung.
- „Der Strom wird als Ladung Q gespeichert.“

Entladen eines Kondensators über einen Widerstand (Abb. 1.14)

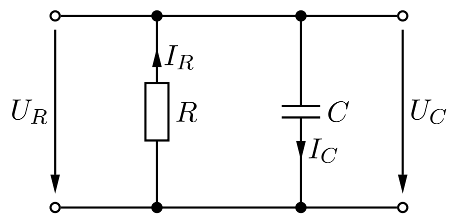


Abb. 1.14 Entladen eines Kondensators über einen Widerstand.

Mit KCL gilt

$$I_C = I_R \quad \text{und damit} \quad C\,\frac{dU_C}{dt} = -\,\frac{U_C}{R} \quad .$$

Die Lösung der Differentialgleichung ist

$$U_C(t) = U_0\,\mathrm{e}^{-\frac{t}{R\,C}} \quad \text{mit} \quad U_0 = U_C(t=0) \quad .$$

Die Kondensatorspannung fällt also exponentiell ab (Abb. 1.15).

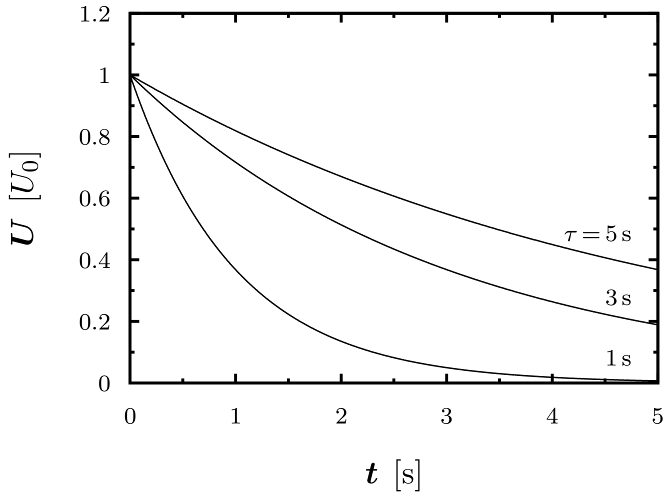


Abb. 1.15 Exponentielle Entladung eines Kondensators über einen Widerstand.

Die Größe

$$\boldsymbol{\tau = RC} \tag{1.17}$$

ist die Zeitkonstante der Schaltung und hat die Einheit Sekunden, $1\,\Omega\,\mathrm{F} = 1\,\mathrm{s}$. Nach der Zeit $\tau = RC$ ist die Spannung des Kondensators von U_0 auf $\frac{1}{e}\,U_0$, also auf $37\,\%$ von U_0 abgefallen, nach $5\,\tau$ auf $0,7\,\%$. Wegen $U_R = U_C = RI$ fällt auch der Entladestrom exponentiell ab.

Aufladen eines Kondensators über einen Widerstand (Abb. 1.16)

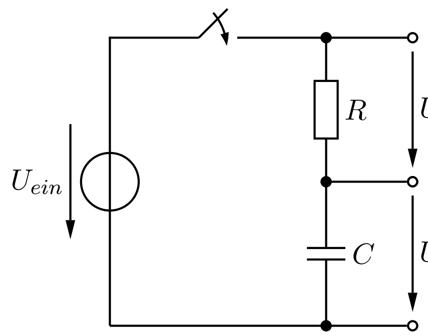


Abb. 1.16 Aufladen eines Kondensators über einen Widerstand mit der Eingangsspannung U_{ein}.

Anmerkung: Der Widerstand R mag klein sein, ist aber in jeder praktischen Schaltung ungleich Null wegen des Innenwiderstands der Spannungsquelle und wegen des Leitungswiderstands.

Der Kondensator sei für $t < 0$ entladen. Bei $t = 0$ wird der Schalter geschlossen. Dann gilt

$$I = C\,\frac{dU_C}{dt} = \frac{U_R}{R} = \frac{U_{ein} - U_C}{R} \;\Leftrightarrow\; \frac{dU_C}{dt} + \frac{U_C}{RC} = \frac{U_{ein}}{RC} \tag{1.18}$$

mit der Lösung $U_C = U_{ein} + A\,\mathrm{e}^{-\frac{t}{RC}}$.

Mit der Anfangsbedingung $U_C(t=0) = 0$ folgt $A = -U_{ein}$ und

$$U_C = U_{ein}\left(1 - \mathrm{e}^{-\frac{t}{RC}}\right) \quad . \tag{1.19}$$

Dieselbe Lösung für die Kondensatorspannung ergibt sich, wenn man in Abb. 1.16 die Position von Widerstand und Kondensator vertauscht.

Anmerkung: Einfacher ist das Aufladen eines Kondensators über eine Stromquelle (Abb. 1.17).

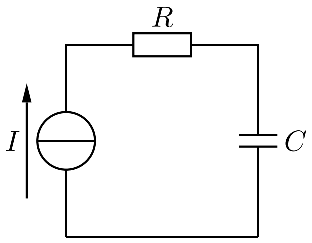


Abb. 1.17 Aufladen eines Kondensators über eine Stromquelle.

Vorsicht: Man erreicht sehr schnell hohe Spannungen.

$$I = C\,\frac{dU_C}{dt} \quad \Leftrightarrow \quad U_C = \frac{1}{C}\int_0^T I\,dt = \frac{I\,T}{C}$$

Für $C = 1\,\mu$F und $I = 1\,$mA erreicht man nach einer Sekunde 1000 V! Der Wert des Widerstandes vor dem Kondensator ist irrelevant.

Beim Aufladen des Kondensators über Widerstände und Spannungsquellen ist dagegen U_{aus} durch U_{ein} begrenzt (Gl. 1.19).

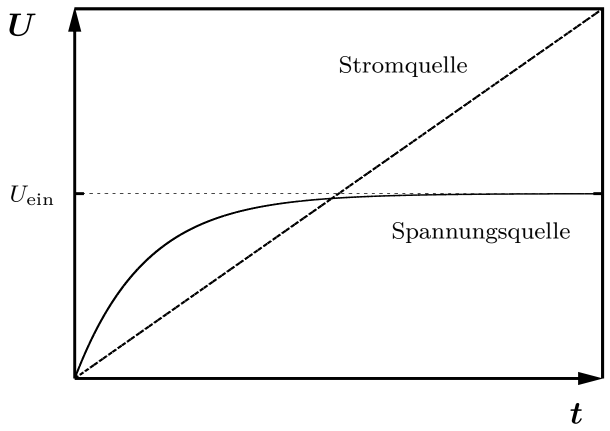


Abb. 1.18 Spannungsverlauf (Spannung gegen Zeit) beim Aufladen eines Kondensators über eine Strom- bzw. Spannungsquelle.

Übung: Überprüfen Sie, dass auf einem Kondensator die Energie

$$E = \frac{1}{2}CU^2$$

gespeichert ist.

Ansatz:

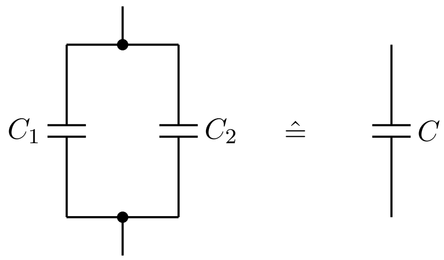
$$E = \int_0^\infty P\,dt$$

und Schaltung 1.14

Frage: Wieviel Energie braucht man um einen Kondensator über die Schaltung in Abb. 1.16 aufzuladen?

Frage: Warum hängt die Lösung nicht von R ab?

1.11 Parallelschaltung von Kondensatoren



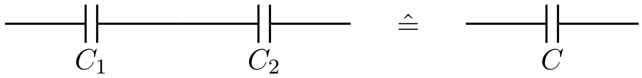
$$\boldsymbol{C = C_1 + C_2} \tag{1.20}$$

Abb. 1.19 Parallelschaltung von Kondensatoren.

Das Parallelschalten von Kondensatoren (Abb. 1.19) erhöht die Gesamtkapazität. Dies entspricht bei baugleichen Plattenkondensatoren einem einzelnen Kondensator mit einer größeren Plattenfläche A. Oft kommt es weniger auf den exakten Wert der Kapazität an, sondern auf das Hochfrequenzverhalten eines realen Kondensators. In diesem Fall ist es sinnvoll, zwei unterschiedliche Kondensatortypen parallel zu schalten.

1.12 Serienschaltung von Kondensatoren

Die Formel entspricht der Parallelschaltung von Widerständen (Abb. 1.20). Der kleinste Kondensator dominiert die Gesamtkapazität.



$$\boldsymbol{\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_2}} \quad \text{und damit} \quad \boldsymbol{C = \frac{C_1 C_2}{C_1 + C_2}} \tag{1.21}$$

Abb. 1.20 Reihenschaltung von Kondensatoren.

1.13 Integrator

Die Eingangsspannung U_{ein} einer Schaltung muss nicht zeitlich konstant sein. Interessanter ist es, wenn U_{ein} variiert und ein Signal darstellt. Mittels einer Integrator- oder Tiefpassschaltung (Abb. 1.21) kann man das Eingangssignal U_{ein} über die Zeit integrieren.

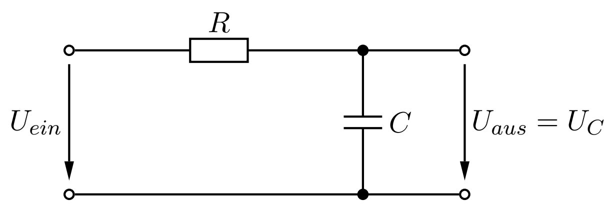


Abb. 1.21 Integrator oder Tiefpass.

Es gilt:

$$\frac{dU_C}{dt} + \frac{U_C}{RC} = \frac{U_{ein}}{RC}$$

(siehe Gleichung 1.18).

Für $t \ll RC$ gilt $U_C \ll U_{ein}$. Der Kondensator ist noch nicht aufgeladen. Dann vereinfacht sich die obige Gleichung zu

$$\frac{dU_C}{dt} = \frac{U_{ein}}{RC} \quad \text{mit} \quad U_C = U_{aus} = \frac{1}{RC} \int U_{ein}\, dt \quad . \tag{1.22}$$

Für große Zeiten gilt die obige Näherung nicht. Die Kondensatorspannung kann den Wert von U_{ein} nicht überschreiten. Der Wert von R und C muss an das Zeitverhalten des Eingangssignals angepasst werden. Viel bessere, aber aufwendigere Integratoren kann man mit Operationsverstärkern (OPVs) realisieren. Auch dann kann natürlich das Ausgangssignal nicht beliebig groß werden und den Wert der Versorgungsspannung des OPV nicht überschreiten.

Auf eine Folge von Reckeckpulsen ergibt sich die in Abb. 1.22 gezeigte Impulsantwort.

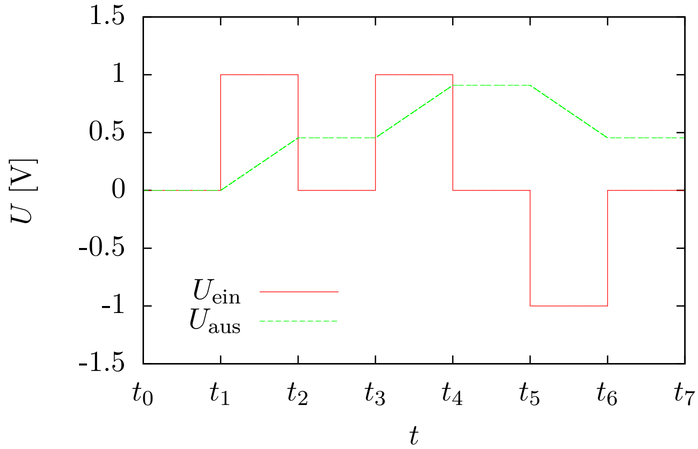


Abb. 1.22 Impulsantwort eines Tiefpasses.

Der lineare Anstieg ergibt sich aus

$$U_{aus} = U_C = U_{ein}\left(1 - \mathrm{e}^{-\frac{t}{RC}}\right) \approx \frac{U_{ein}}{RC}\,t \quad \text{für} \quad t \ll \tau \quad .$$

1.14 Differenzierer

Abbildung 1.23 zeigt ein Differenzierglied oder Hochpass. Im Vergleich zu Abbildung 1.21 sind die Positionen von R und C vertauscht.

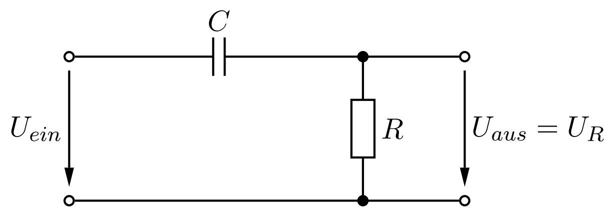


Abb. 1.23 Differenzierer oder Hochpass.

Es gilt

$$U_{ein} = U_C + U_R \quad \text{mit} \quad U_R = I\,R \quad \text{und} \quad I = C\,\frac{dU_C}{dt} \quad .$$

Daraus folgt

$$U_{ein} = U_C + R\,C\,\frac{dU_C}{dt} \quad \text{und} \quad U_{ein} - U_C = R\,C\,\frac{d(U_{ein} - U_{aus})}{dt} = U_{aus} \quad . \tag{1.23}$$

Für kleine Widerstände R fällt an R nur eine geringe Spannung ab und damit gilt $U_{ein} - U_{aus} \approx U_{ein}$. (Dies gilt auch für kleine Werte von C bzw. für langsame Änderungen der Eingangsspannung.) Gleichung 1.23 vereinfacht sich damit zu

$$U_{aus} = R\,C\,\frac{dU_{ein}}{dt} \quad . \tag{1.24}$$

Das Ausgangssignal ist also in Näherung proportional zur Ableitung des Eingangssignals.

Wie reagiert der Differenzierer auf eine Folge von Reckeckimpulsen am Eingang (siehe Abb. 1.24 und Tabelle 1.2)?

Der Kondensator sei zum Zeitpunkt t_0 entladen, $U_C = 0$ V. Wird er zu t_1 an $U_{ein} = 1$ V angeschlossen, so liegen zu diesem Zeitpunkt beide Elektroden auf demselben Potential (da $U_C = 0$ V), und damit gilt $U_R = 1$ V. Am Widerstand R liegt also eine von Null verschiedene Spannung an, und es fließt ein Strom I_R. Dadurch wird im Zeitinterval $t_1 - t_2$ der Kondensator aufgeladen. (Dies entspricht Abb. 1.16). Da die Eingangselektrode auf dem festen Potential U_{ein} liegt, geschieht das Aufladen durch Absenken des Potentials der Ausgangselektrode. Wir gehen zur Vereinfachung davon aus, dass U_C und U_R zum Zeitpunkt t_2 ihre asymptotischen Werte $U_C = 1$ V und $U_R = 0$ V erreicht haben. Interessant ist jetzt das Verhalten von U_{aus} zum Zeitpunkt t_2. Hier springt U_{aus} auf -1 V, da $U_{ein} = 0$ V und $U_C = 1$ V. Im Intervall $t_2 - t_3$ wird der Kondensator entladen und U_{aus} fällt exponentiell von -1 V auf 0 V ab. Dies entspricht der Schaltung aus Abb. 1.14, da der Innenwiderstand der Spannungsquelle sehr klein ist.

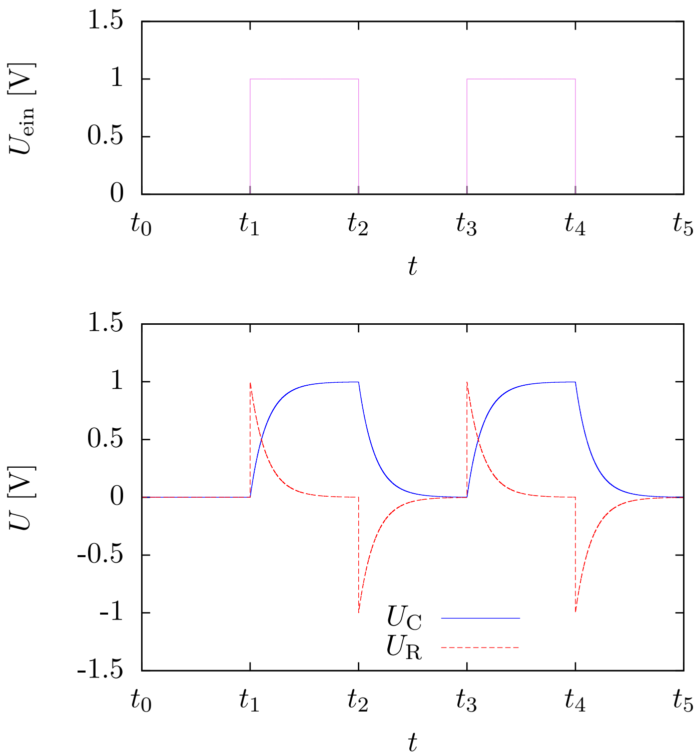


Abb. 1.24 Impulsantwort eines Hochpasses. Oben: Rechteckpuls als Eingangssignal U_{ein}. Unten: Das dazugehörige Ausgangssignal des Differenzierers $U_{aus} = U_R$ (rot, gestrichelt) und die entsprechende Kondensatorspannung U_C (blau, durchgezogen) als Funktion der Zeit t.

Zeit	U_{ein} [V]	U_R [V]	U_C [V]	**Kommentar**
$t_0 - t_1$	0	0	0	Kondensator ist entladen
t_1	1	1	0	
$t_1 - t_2$	1	$e^{-\frac{t}{\tau}}$	$1 - e^{-\frac{t}{\tau}}$	Kondensator wird geladen
t_2	0	-1	1	
$t_2 - t_3$	0	$-e^{-\frac{t}{\tau}}$	$e^{-\frac{t}{\tau}}$	Kondensator wird entladen

Tab. 1.2 Die funktionale Abhängigkeit der Kurven in Abbildung 1.24. Die Eingangsspannung wechselt zwischen 0 V und 1 V.

1.15 Die Spule

Die Spule, das dritte passive Bauelement, ist ein Zweipol. Das Strom-Spannungsverhalten einer Spule folgt der Gleichung

$$\boldsymbol{U = L\,\frac{dI}{dt}} \quad . \tag{1.25}$$

Dabei ist L die Induktivität der Spule mit der Einheit Henry [H].

Legt man an eine Spule eine konstante Spannung an, so nimmt der Strom also linear mit der Zeit zu. Die in der Spule gespeicherte magnetische Energie E ist

$$\boldsymbol{E = \frac{1}{2}\,L\,I^2} \quad . \tag{1.26}$$

Typische Spulen (außerhalb von Netzteilen) haben Induktivitäten im Bereich nH bis mH. Es gibt unterschiedliche Bauformen als Solenoid- oder Toroidwicklung mit oder ohne Eisenkern. Für Solenoid- oder Toroidstrukturen hängt die Induktivität quadratisch von der Zahl der Windungen n ab und ist proportional zur Permeabilität μ des Kernmaterials. Die Permeabilität μ ist allerdings nicht konstant und nimmt bei hohen Frequenzen und in äußeren Magnetfeldern stark ab. Für Luftspulen gilt $\mu = 1$. Für eine Eisenkernspule kann $\mu \approx 2000$ sein. Mit der Zahl der Windungen nimmt auch die Drahtlänge zu. Dies ist oft unerwünscht, da es den ohmschen Widerstand der Spule erhöht.

Spulen werden verwendet:

- zur Begrenzung von Stromschwankungen
- als Schwingkreiskomponenten in HF-Technik
- als effizienter Gleichspannungswandler, z. B. Buck Converter, Zündspule
- als Netzfilterkomponenten
- zur galvanischen Trennung in Transformatoren.

1.16 Parallelschaltung von Spulen

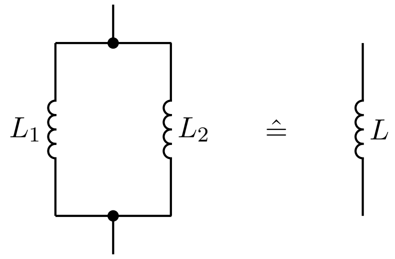


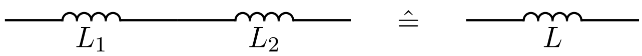
Abb. 1.25 Parallelschaltung von Spulen.

Für die Parallelschaltung von Spulen gilt

$$\boldsymbol{\frac{1}{L} = \frac{1}{L_1} + \frac{1}{L_2}} \quad \text{und damit} \quad \boldsymbol{L = \frac{L_1\, L_2}{L_1 + L_2}} \quad , \tag{1.27}$$

vorausgesetzt die Magnetfelder der Einzelspulen koppeln nicht aneinander.

1.17 Serienschaltung von Spulen



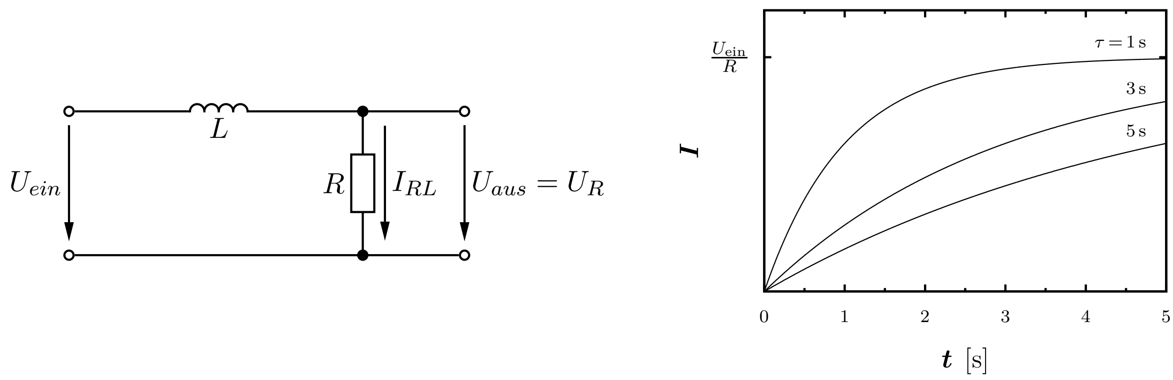
$$\boldsymbol{L = L_1 + L_2} \tag{1.28}$$

Abb. 1.26 Reihenschaltung von Spulen.

Für die Serienschaltung von Spulen gilt $L = L_1 + L_2$ vorausgesetzt die Magnetfelder der Einzelspulen koppeln nicht aneinander.

Magnetisieren einer Spule

Die Spule aus Abb. 1.27 wird zum Zeitpunkt $t = 0$ an eine konstante Spannung U_{ein} angeschlossen. Davor ($t < 0$) fließt durch die Spule kein Strom. Wie verhalten sich Spulenspannung und Spulenstrom für $t > 0$?



(a) (b)

Abb. 1.27 (a) Schaltkreis mit Spannungsquelle, Spule und Widerstand. (b) Spulenstrom nach Anschalten von U_{ein}.

Aus der Differentialgleichung

$$U_{ein} = L\,\frac{dI}{dt} + R\,I$$

folgt die Lösung (Abb. 1.27(b))

$$I(t) = \frac{U_{ein}}{R}\left(1 - \mathrm{e}^{-\frac{t}{\tau}}\right) \quad . \tag{1.29}$$

Hier steht die Zeitkonstante τ für

$$\boldsymbol{\tau = \frac{L}{R}} \quad . \tag{1.30}$$

Am Widerstand liegt die Spannung

$$U_R = U_{ein}\left(1 - \mathrm{e}^{-\frac{t}{\tau}}\right)$$

an, an der Spule die Spannung

$$U_L = U_{ein}\,\mathrm{e}^{-\frac{t}{\tau}} \quad .$$

Der Nominalstrom $I = \frac{U_{ein}}{R}$ wird erst nach Verzögerung ($t \gg \frac{L}{R}$) erreicht. Dann gilt $U_L = 0$, $\frac{dI}{dt} = 0$ und $U_R = U_{ein}$. Nach Erreichen des konstanten maximalen Stromes hat die Spule keinen Einfluss mehr auf den Stromkreis, speichert allerdings noch magnetische Energie.

Kurzschluss einer Spule über einen Widerstand

Verringert man U_{ein} zu einem späteren Zeitpunkt auf 0 V, so entspricht dies Abb. 1.28

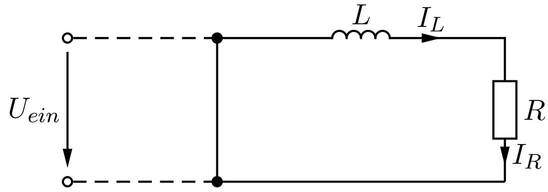


Abb. 1.28 Entladen einer Spule.

Die Lösung der Differentialgleichung

$$U_L + R\,I = 0 \Leftrightarrow L\,\frac{dI}{dt} = -R\,I \quad \text{ist} \quad I = I_0\,\mathrm{e}^{-\frac{t}{\tau}}$$

mit

$$I(t=0) = I_0 = \frac{U_{ein}}{R} \quad \text{und} \quad U_L = L\,\frac{dI}{dt} = -R\,I_0\,\mathrm{e}^{-\frac{t}{\tau}} \quad .$$

Die „Entladung“ einer magnetisierten Spule durch einen Widerstand erfolgt also exponentiell. Das heißt, der Strom nimmt exponentiell mit der Zeit ab. Die Polarität der Spulenspannung ändert sich, und U_L nimmt exponentiell ab.

Damit ergibt sich das folgende Verhalten der Schaltung 1.27 auf eine Folge von Rechteckpulsen am Eingang (Abb. 1.29 und Tabelle 1.3).

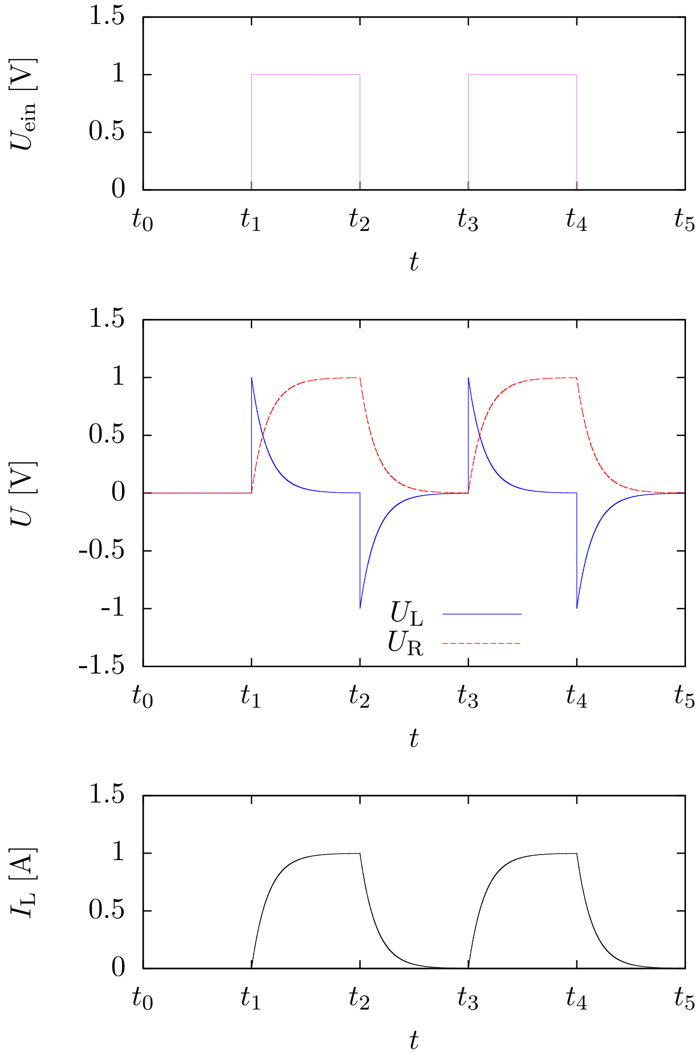


Abb. 1.29 Impulsantwort einer Serienschaltung von Spule und Widerstand auf Rechteckimpulse. Oben: Rechteckpuls als Eingangssignal U_{ein}. Mitte: Die erzeugten Spannungen an Widerstand U_R (rot, gestrichelt) und Spule U_L (blau, durchgezogen) als Funktion der Zeit t. Unten: Der Strom I_L als Funktion der Zeit t.

Zeit	U_{ein} [V]	U_R [V]	U_L [V]	I_L [A]	Kommentar
$t_0 - t_1$	0	0	0	0	
t_1	1	0	1	0	
$t_1 - t_2$	1	$1 - e^{-\frac{t}{\tau}}$	$e^{-\frac{t}{\tau}}$	$1 - e^{-\frac{t}{\tau}}$	Spulenstrom baut sich auf, U_L fällt exp. ab
t_2	0	0	-1	1	
$t_2 - t_3$	0	$e^{-\frac{t}{\tau}}$	$-e^{-\frac{t}{\tau}}$	$e^{-\frac{t}{\tau}}$	

Tab. 1.3 Die funktionale Abhängigkeit der Kurven in Abbildung 1.29. Die Eingangsspannung wechselt zwischen 0 V und 1 V.

Gleichspannungsfilter

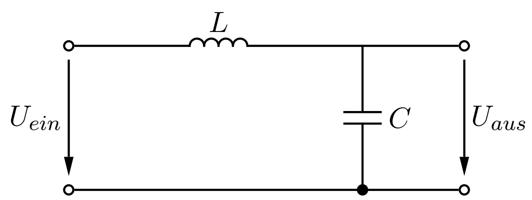


Abb. 1.30 Gleichspannungsfilter mit Spule und Kondensator.

Spannungsschwankungen der Eingangsgleichspannung kann man durch einen Gleichspannungsfilter, wie z. B. durch den *LC*-Filter (Abb. 1.30) glätten. Ein Vorteil der Spule ist ihr geringer Ohmscher Widerstand im Vergleich zu einem *RC*-Filter (Abb. 1.23). Dies ist bei nennenswerten Lastströmen relevant.

1.18 Analyse linearer Netzwerke

Lineare Netzwerke sind Schaltkreise aus Widerständen, Kondensatoren, Spulen und einer beliebigen Anzahl von und Strom- und Spannungsquellen.

1.19 Satz von Helmholtz (Thévenin's theorem)

Jedes lineare Netzwerk kann bezüglich zweier beliebiger Knotenpunkte als ideale Spannungsquelle und einem (Innen-)Widerstand in Serie beschrieben werden.

Beispiel: Spannungsteiler

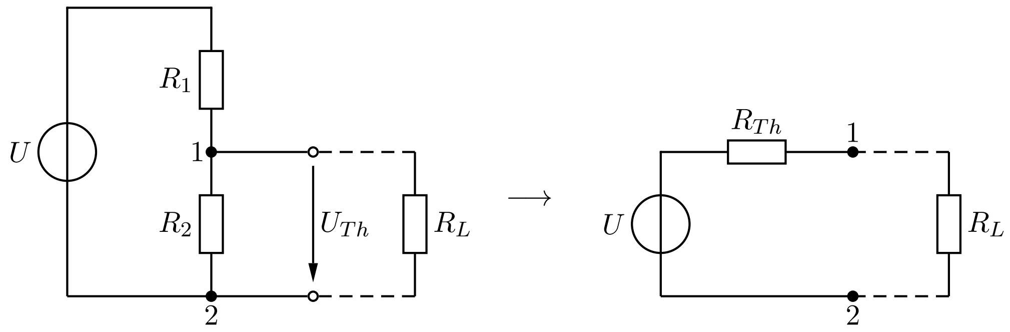


Abb. 1.31 Thévenins Ersatzschaltung für einen Spannungsteiler.

Wie berechnet man U_{Th} und R_{Th}?

Es gibt verschiedene Wege. Alle beruhen auf KCL und KVL sowie auf dem Ohmschem Gesetz:

a) U_{Th} entspricht der Ausgangsspannung des Ausgangsnetzwerks ohne Last ($R_L = \infty$).

b) $R_{Th} = \frac{U_{Th}}{I_{KS}}$, wobei I_{KS} der Ausgangsstrom des Ausgangsnetzwerkes bei Kurzschluss ($R_L = 0$) ist.

Diese Vorschrift ist plausibel:

a), da ohne Last kein Strom fließt und kein Spannungsabfall an R_{Th} auftritt. Den Wert von R_{Th} muss man also noch nicht kennen.

b), da beim Kurzschluss der Strom nur durch R_{Th} und U_{Th} bestimmt wird. Die Last ist irrelevant.

Für den Spannungsteiler ist die Berechnung einfach und ergibt

$$U_{Th} = \frac{U\,R_2}{R_1 + R_2} \quad , \quad I_{KS} = \frac{U}{R_1}$$

und damit

$$R_{Th} = \frac{R_2\,R_1}{R_1 + R_2} \quad .$$

Für kompliziertere Netze, z. B. ein T-Stück (Abb. 1.33), ist die Berechnung umständlich. Hier ist die folgende Alternativregel zu b) hilfreich:

b*) Alle Spannungsquellen werden durch Kurzschlüsse und alle Stromquellen durch offene Schaltungen ersetzt. Der Lastwiderstand wird entfernt. R_{Th} ist der Widerstand zwischen den Ausgangsklemmen des so modifizierten Schaltkreises.

1.20 Nortons Theorem

Jedes lineare Netzwerk kann bezüglich zweier beliebiger Knotenpunkte als ideale Stromquelle parallel zu einem einzelnen Widerstand dargestellt werden (Abb. 1.32). Insbesondere gilt: Jede Stromquelle ist elektrisch äquivalent zu einer Spannungsquelle mit $R_{Th} = R_{No}$.

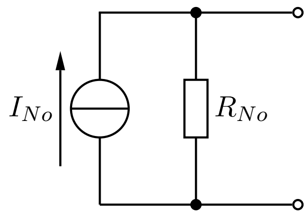


Abb. 1.32 Nortons Ersatzschaltung.

Wie berechnet man I_{No} und R_{No}?

Die Ableitung ist analog zum Helmholtz-Thévenin-Ersatzschaltkreis.

1. I_{No} entspricht dem Ausgangsstrom des Originalschaltkreises mit kurzgeschlossener Last.
2. Alle Spannungsquellen werden durch Kurzschlüsse und alle Stromquellen durch offene Schaltungen ersetzt. R_{No} ist der Widerstand zwischen den Ausgangsklemmen des so modifizierten Schaltkreises.

Es gibt zahlreiche formale Methoden lineare Netzwerke zu analysieren und zu vereinfachen. Neben den bereits angewandten Theorien von Thévenin und Norton kann man das Superpositionsprinzip ausnutzen, den Vielpolformalismus anwenden oder einfach lineare Gleichungssysteme lösen. Zur Illustration hier noch zwei weitere Lösungsverfahren am Beispiel des T-Stücks. Für uns ist es in der Praxis oft nur wichtig, diese Methoden prinzipiell zu kennen und die offensichtliche Lösung zu erraten.

1.21 Das Superpositionsverfahren

Lösungen für Ströme und Spannungen in linearen Netzwerken lassen sich als Überlagerung von Teillösungen von Netzwerken darstellen, bei denen alle passiven Elemente unverändert sind, aber die Strom- und Spannungsquellen bis auf eine Quelle durch ihre Innenwiderstände ersetzt werden.

Beispiel: T-Stück

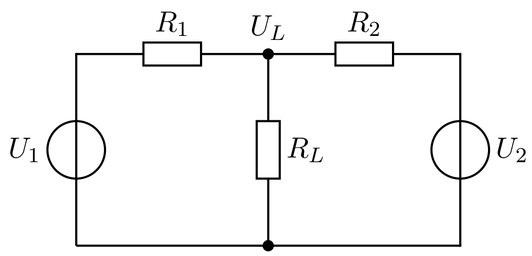


Abb. 1.33 Ausgangsschaltung für Superpositionsverfahren.

Wir wollen den Laststrom der Schaltung in Abbildung 1.33 berechnen. Der Laststrom ergibt sich als Supersposition der Lastströme $I_L = I_{L_1} + I_{L_2}$ der Schaltkreise in Abbildung 1.34.

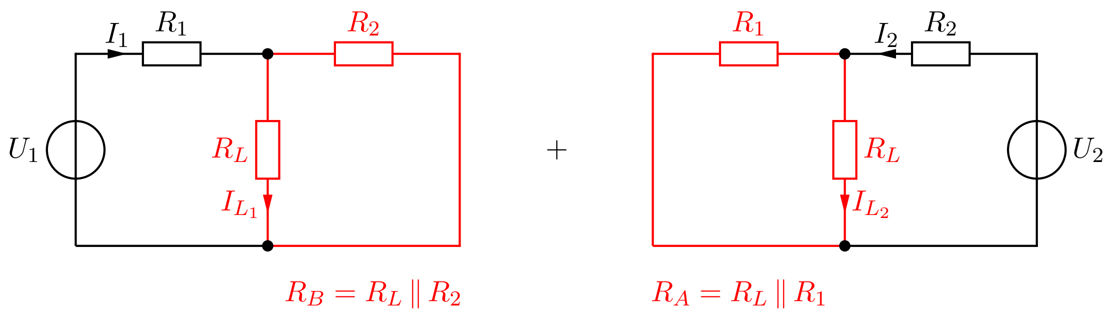


Abb. 1.34 Superpositionsverfahren: Ersatzschaltungen zu Abb. 1.33.

Hier wurden jeweils die Spannungsquelle U_1 bzw. U_2 durch den Innenwiderstand $0\,\Omega$ ersetzt. Die Widerstände R_L und R_2 im linken Bild bzw. R_L und R_1 im rechten Bild sind parallel geschaltet. Es ergibt sich also

$$U_{L_1} = \frac{U_1\,R_B}{R_1 + R_B} \quad , \quad U_{L_2} = \frac{U_2\,R_A}{R_2 + R_A} \quad \text{und} \quad U_L = U_{L_1} + U_{L_2} \quad .$$

Mit

$$I_{L_1} = \frac{U_{L_1}}{R_L} \quad \text{und} \quad I_{L_2} = \frac{U_{L_2}}{R_L} \quad \text{bzw. über} \quad I_L = \frac{U_L}{R_L}$$

folgt

$$I_L = I_{L_1} + I_{L_2} = \frac{U_1\,R_2 + U_2\,R_1}{R_1\,R_2 + R_L\,(R_1 + R_2)} \quad .$$

1.22 Lineare Gleichungssysteme

Uns interessiert wieder der Laststrom, der sich aber aus Kenntnis der Lastspannung U_L aus $I_L = \frac{U_L}{R_L}$ ergibt.

Zuerst stellen wir alle nicht trivialen Gleichungen, die sich aus KCL und KVL ergeben, zusammen. Das ist hier einfach. Bei Schaltungen mit mehr Elementen muss man systematisch vorgehen, die Knoten sauber nummerieren und gut buchhalten (Abb. 1.35).

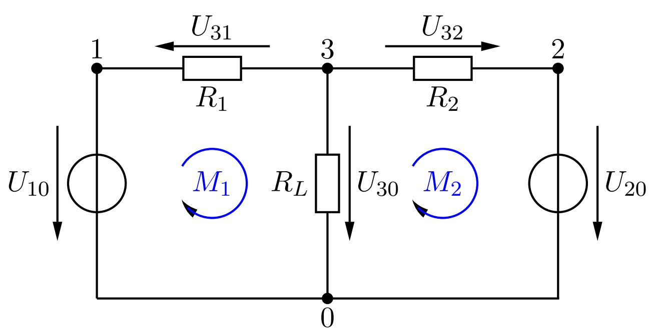


Abb. 1.35 Ausgangsschaltung fuer lineares Gleichungssystem.

Aus $\sum I_i = 0$ (alle Ströme I_i fließen *in* den Knoten 3) folgt

$$\frac{U_{31}}{R_1} + \frac{U_{32}}{R_2} + \frac{U_{30}}{R_L} = 0 \quad . \tag{1.31}$$

Aus $\sum U_{ij} = 0$ folgt

$$-\,U_{31} + U_{30} - U_{10} = 0 \quad (M_1) \tag{1.32}$$

und

$$-\,U_{30} + U_{32} + U_{20} = 0 \quad (M_2) \quad . \tag{1.33}$$

Mit $\frac{1}{R_i} = G_i$ gilt also in Matrixschreibweise

$$\begin{pmatrix} G_L & G_1 & G_2 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} U_{30} \\ U_{31} \\ U_{32} \end{pmatrix} = \begin{pmatrix} 0 \\ U_{10} \\ U_{20} \end{pmatrix} \quad .$$

Dieses Gleichungssystem lässt sich mit der Kramerschen Regel leicht lösen (Anhang). Die Lösung ist

$$U_{30} = \frac{G_1\, U_{10} + G_2\, U_{20}}{G_L + G_1 + G_2} \quad .$$

2 Schaltungen mit Widerständen, Kondensatoren und Spulen

2.1 Wechselspannung

Wir haben bis jetzt zeitlich konstante Ströme und Spannungen bzw. Ein- und Ausschaltvorgänge betrachtet. Fast alle elektronischen Schaltungen dieser Vorlesung werden mit Gleichspannung betrieben. Das heißt, die Versorgungsspannung ist (möglichst) konstant. Die Signale, die uns interessieren, sind dagegen naturgemäß variabel, sonst würde keine Information übertragen werden.

Wichtige Signalformen sind periodische harmonische Schwingungen, periodische Rechtecksignale (*square waves*), Sägezahnsignale (die man auch beim Betrieb eines Oszilloskops oder anderen Bildschirmen verwendet), Rechteckpulse, Pulse die beim Durchgang eines geladenen Teilchens durch Detektoren entstehen und viele mehr (Abb. 2.1).

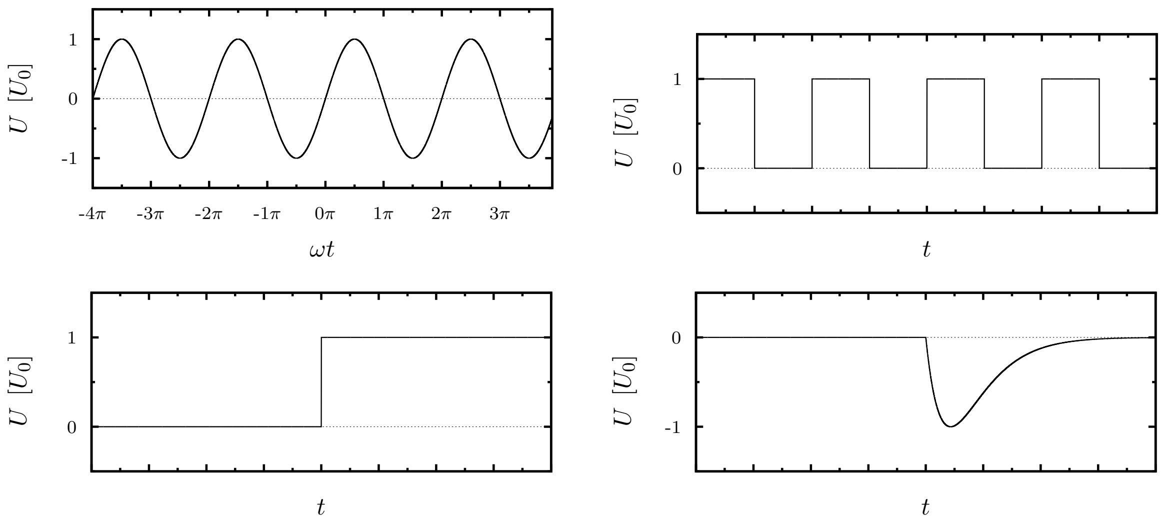


Abb. 2.1 Darstellung häufiger Signalformen (in offensichtlicher Reihenfolge): harmonische Schwingung, Rechteckpulse, Stufenpuls, Detektorsignal nach Pulsformung durch einen RC-CR - Filter.

Alle periodischen Signale lassen sich nach Fourier als Summe von harmonischen Schwingungen darstellen; auch für nichtperiodische Signale ist die Fourierzerlegung ein zentrales Hilfsmittel. Es ist deshalb wichtig, die Wirkung von Widerstand, Kondensator und Spule auf Wechselspannungen und Wechselströme einer festen Frequenz f zu verstehen.

Sei $U(t)$ eine harmonische periodische Schwingung $U(t) = U_0\ \sin(\omega\, t{+}\varphi)$. Dabei sind $\omega = 2\,\pi\, f$ die Kreisfrequenz, f die Frequenz der Schwingung in Hertz [Hz], U_0 die maximale Amplitude bzw. der Scheitelwert in Volt [V] und φ der Phasenwinkel in Bogenmaß.

Betrachtet man die Beziehung von Spannung und Strom an einem Widerstand, so gilt wegen $U = R\, I$ auch $\boldsymbol{U(t) = R\, I(t)}$. Die Spannung zwischen den Polen des Widerstands und der Strom durch den Widerstand sind immer in Phase (Abb. 2.2).

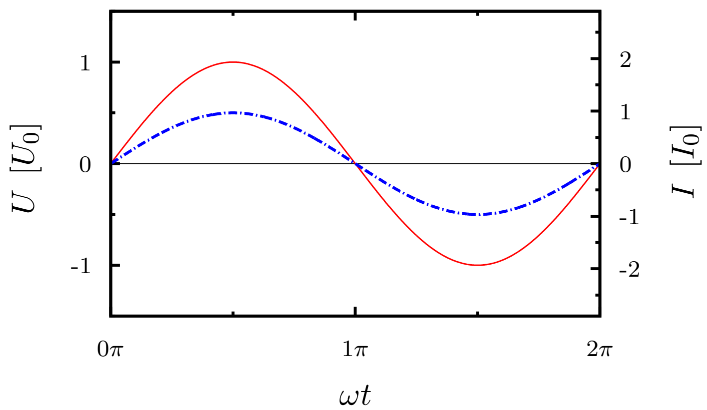


Abb. 2.2 Widerstand: Wechselspannung (rot, durchgezogen) und Wechselstrom (blau, gestrichtelt) als Funktion der Zeit t.

Dies ist bei Kondensator und Spule nicht so. Für einen Kondensator gilt wegen

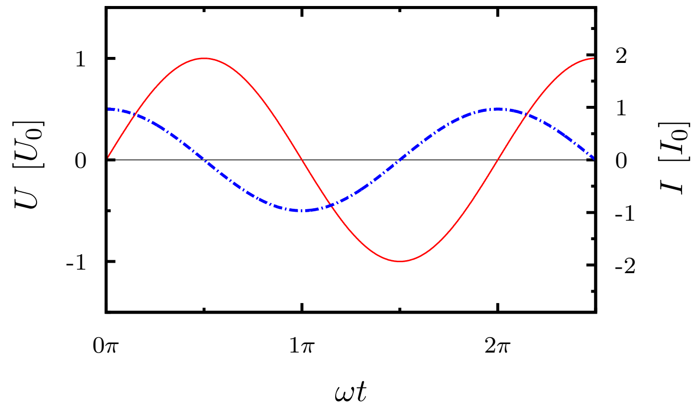
$$I = C\,\frac{dU}{dt} \quad \text{mit} \quad U(t) = U_0\ \sin(\omega\, t) \overset{1}{=} U_0\ \cos\left(\omega\, t - \frac{\pi}{2}\right)$$

für den Wechselstrom

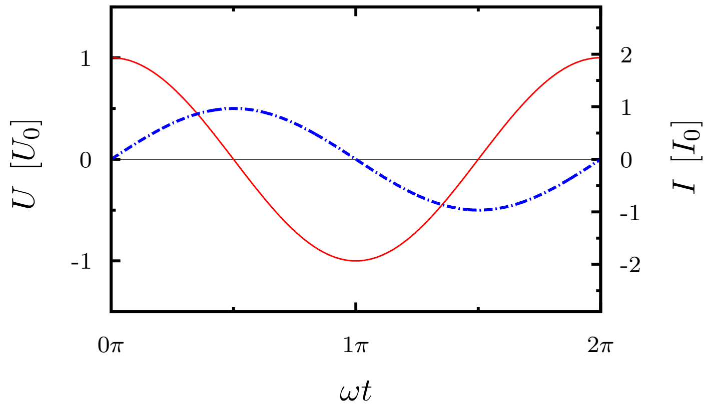
$$I(t) = U_0\,\omega\, C\ \cos(\omega\, t) \quad .$$

Strom- und Spannungsamplitude stehen im Verhältnis $\frac{U_0}{I_0} = \frac{1}{\omega\, C}$. Gleichzeitig sind die Schwingungen von Strom und Spannung um 90° verschoben. Der Strom eilt der Spannung um 90° voraus (Abb. 2.3(a)).

[1]Wir setzen ohne Einschränkung der Allgemeinheit $\varphi = 0$, und wählen eine Sinusspannung bzw. einen Sinusstrom statt einer Linearkombination von Kosinus und Sinus.



(a)



(b)

Abb. 2.3 (a) Kondensator und (b) Spule: Wechselspannung (rot, durchgezogen)und Wechselstrom (blau, gestrichtelt) als Funktion der Zeit t.

Für eine Spule gilt wegen

$$U = L\,\frac{dI}{dt} \quad \text{mit} \quad I(t) = I_0\,\sin(\omega\,t) = I_0\,\cos(\omega\,t - \frac{\pi}{2})$$

für die Wechselspannung

$$U(t) = I_0\,\omega\,L\,\cos(\omega\,t) \quad .$$

Strom- und Spannungsamplitude stehen im Verhältnis $\frac{U_0}{I_0} = \omega\,L$. Hier verspätet sich der Strom gegenüber der Spannung um 90° (Abb. 2.3(b)).

Dasselbe Ergebnis erhält man natürlich bei Wahl eines Stromes als $I = I_0\,\cos(\omega\,t)$ oder wenn man statt vom Strom von der Spannung ausgeht.

Im Folgenden werden wir mit komplexen Größen arbeiten, dann kann man Phasen und Amplituden in einer kompakten Schreibweise erfassen.

2.2 Komplexe Zahlen

Komplexe Zahlen sind hier unterstrichen. Die Größe $\underline{z}$ ist eine komplexe Zahl. Statt $\mathrm{i} = \sqrt{-1}$ bzw. $\mathrm{i}^2 = -1$ schreiben wir $\mathrm{j} = \sqrt{-1}$, da „i" in der Elektrotechnik den Strom symbolisiert. Jede komplexe Zahl hat einen Realteil und einen Imaginärteil und lässt sich schreiben als

$$\underline{z} = \mathrm{Re}(\underline{z}) + \mathrm{j}\,\mathrm{Im}(\underline{z}) = a + \mathrm{j}\,b \quad . \tag{2.1}$$

Dabei sind a und b reelle Zahlen.

$\underline{z}^*$ ist die konjugiert komplexe Zahl

$$\underline{z}^* = a - \mathrm{j}\,b \quad . \tag{2.2}$$

Es gilt

$$(\underline{z}_1^*\,\underline{z}_2^*) = (\underline{z}_1\underline{z}_2)^* \quad . \tag{2.3}$$

Insbesondere gilt damit auch

$$\left[\frac{a + \mathrm{j}\,b}{c + \mathrm{j}\,d}\right]^* = \frac{a - \mathrm{j}\,b}{c - \mathrm{j}\,d} \quad . \tag{2.4}$$

z ist der Betrag von $\underline{z}$ mit

$$z = \sqrt{\mathrm{Re}(\underline{z})^2 + \mathrm{Im}(\underline{z})^2} = \sqrt{\underline{z}\,\underline{z}^*} \quad . \tag{2.5}$$

$\underline{z}$ ist auch in der Form $z\,\mathrm{e}^{\mathrm{j}\varphi}$ darstellbar; dabei ist z die reelle Amplitude und φ ist die Phase mit $\tan\varphi = \frac{b}{a}$.

Das Ganze wird sehr anschaulich in der komplexen Ebene.

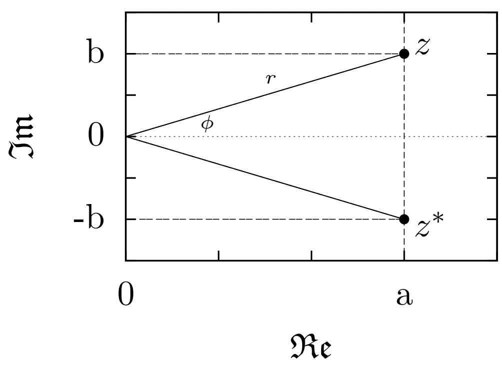


Abb. 2.4 Komplexe Ebene. Die Abzisse entspricht dem Realteil von $\underline{z}$, die Ordinate dem Imaginärteil.

Für die Exponentialfunktion gilt der **Eulersche Satz**

$$\mathrm{e}^{\mathrm{j}\,\varphi} = \cos\varphi + \mathrm{j}\,\sin\varphi \quad . \tag{2.6}$$

Hilfreich sind auch die Beziehungen

$$\begin{aligned} \mathrm{e}^{\mathrm{j}\,\frac{\pi}{2}} &= \cos\left(\frac{\pi}{2}\right) + \mathrm{j}\,\sin\left(\frac{\pi}{2}\right) = \mathrm{j} \quad , \\ \mathrm{e}^{\mathrm{j}\,\frac{3\pi}{2}} = \mathrm{e}^{\text{-}\mathrm{j}\,\frac{\pi}{2}} &= \cos\left(\frac{-\pi}{2}\right) + \mathrm{j}\,\sin\left(\frac{-\pi}{2}\right) = -\mathrm{j} \quad , \\ \mathrm{e}^{\mathrm{j}\,\pi} &= \cos(\pi) + \mathrm{j}\,\sin(\pi) = -1 \quad , \\ \mathrm{e}^{\mathrm{j}\;2\pi} &= 1 \quad . \end{aligned} \tag{2.7}$$

Sind die bisherigen Definitionen konsistent?

Für den Betrag von $\underline{z}$ gilt, wie oben angegeben,

$$|\underline{z}| = \left|z\,\mathrm{e}^{\mathrm{j}\,\varphi}\right| = z\;|\cos\varphi + \mathrm{j}\sin\varphi|$$

$$= z\,\sqrt{(\cos\varphi + \mathrm{j}\,\sin\varphi)\,(\cos\varphi - \mathrm{j}\,\sin\varphi)} = z\,\sqrt{(\cos^2\varphi + \sin^2\varphi)} = z$$

wie erwartet.

Für das Rechnen mit komplexen Zahlen ergeben sich die folgenden einfachen Regeln:

Addition:

$$\underline{z}_1 + \underline{z}_2 = a_1 + a_2 + \mathrm{j}\,(b_1 + b_2) \tag{2.8}$$

Multiplikation:

$$\underline{z}_1\,\underline{z}_2 = a_1\,a_2 - b_1\,b_2 + \mathrm{j}\,(a_1\,b_2 + a_2\,b_1) = z_1\,z_2\,\mathrm{e}^{\mathrm{j}\,(\varphi_1+\varphi_2)} \tag{2.9}$$

Potenzieren:

$$\underline{z}^n = z^n\,\mathrm{e}^{\mathrm{j}\,(n\,\varphi)} \tag{2.10}$$

Im Komplexen ergeben sich die **n-ten Einheitswurzeln**, also die Lösung der Gleichung $\underline{z}^n = z^n\,\mathrm{e}^{\mathrm{j}\,(n\,\varphi)} = 1$, zu $\mathrm{e}^{\mathrm{j}\,\frac{2\pi \mathrm{k}}{n}}$ mit $k = 1, ..., n$.

Lösungen für $n = 2$ sind -1 und 1 (Abb. 2.5).

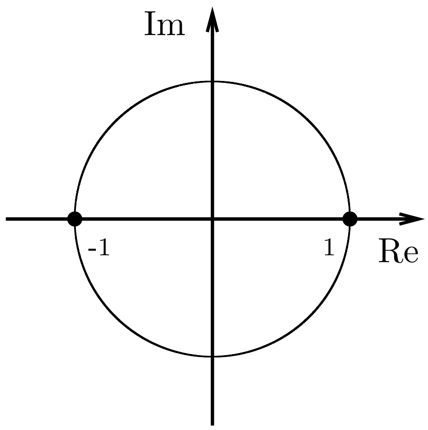


Abb. 2.5 Illustration der Lösungen der Gleichung $\underline{z}^2 = 1$ in der komplexen Ebene.

Lösungen für $n = 3$ sind $\mathrm{e}^{\mathrm{j}\,\frac{2\pi}{3}}$, $\mathrm{e}^{\mathrm{j}\,\frac{4\pi}{3}}$, $\mathrm{e}^{\mathrm{j}\,2\pi}$ bzw. $-\frac{1}{2} + j\frac{\sqrt{3}}{2}$, $-\frac{1}{2} - j\frac{\sqrt{3}}{2}$, 1 (Abb. 2.6).

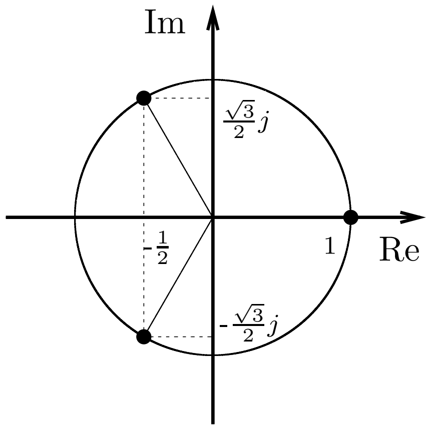


Abb. 2.6 Illustration der Lösungen der Gleichung $\underline{z}^2 = 1$ in der komplexen Ebene.

Lösungen für $n = 4$ sind -1, +1, -j, +j (Abb. 2.7).

Beispiel: $\left(-\frac{1}{2} - j\frac{\sqrt{3}}{2}\right)^3 = \left(\frac{1}{4} + j\frac{\sqrt{3}}{2} - \frac{3}{4}\right)\left(-\frac{1}{2} - j\frac{\sqrt{3}}{2}\right) = \left(-\frac{1}{2} + j\frac{\sqrt{3}}{2}\right)\left(-\frac{1}{2} - j\frac{\sqrt{3}}{2}\right) = \frac{1}{4} + \frac{3}{4} = 1$.

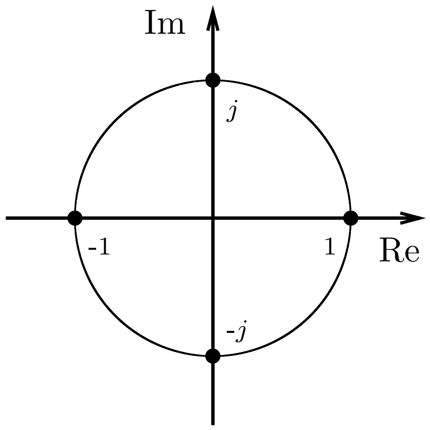


Abb. 2.7 Illustration der Lösungen der Gleichung $\underline{z}^4 = 1$ in der komplexen Ebene.

Die Vorteile der komplexen Darstellung werden später in den kommenden Unterkapiteln deutlich werden:

- Die Schreibweise ist kompakt und erspart vielfache Anwendung der trigonometrischen Formeln.
- Die Phasen- und Amplitudeninformation ist schnell überschaubar.

Es ist einfach, von den komplexen Größen durch Bilden des Realteils zu den realen Strömen und Spannungen überzugehen.

Außer den gezeigten elementaren Beziehungen braucht man in der Elektrotechnik gelegentlich noch den **Integralsatz von Cauchy** und den **Residuensatz**. Informationen dazu finden sich im Appendix.

2.3 Verallgemeinertes Ohmsches Gesetz

Durch Einführen des komplexen Widerstandes für Wechselströme und -spannungen der festen Frequenz $f = \omega/2\pi$, der sogenannten Impedanz $\underline{Z}$, besteht eine Beziehung

$$\boldsymbol{\underline{U} = \underline{Z}\,\underline{I}} \quad \text{mit} \tag{2.11}$$

Widerstand:

$$\boldsymbol{\underline{Z}_R = R}$$

Kondensator:

$$\boldsymbol{\underline{Z}_K = \frac{1}{\mathrm{j}\,\omega\,C} = -\,\mathrm{j}\,\frac{1}{\omega\,C}}$$

Spule:

$$\boldsymbol{\underline{Z}_S = \mathrm{j}\,\omega\,L}$$

Dies ist das verallgemeinerte Ohmsche Gesetz. Die Impedanz $\underline{Z}$ ist für ohmsche Widerstände reell. Für Kondensatoren und Spulen ist sie imaginär.

Die letzten beiden Beziehungen ergeben sich aus

$$\underline{U} = U_0\,\mathrm{e}^{\mathrm{j}\,\omega\,t} \quad \text{und} \quad \underline{I} = C\,\frac{d\underline{U}}{dt} \quad \Leftrightarrow \quad \underline{I} = \mathrm{j}\,\omega\,C\,U_0\,\mathrm{e}^{\mathrm{j}\,\omega\,t}$$

und damit

$$\frac{\underline{U}}{\underline{I}} = \frac{1}{\mathrm{j}\,\omega\,C}$$

bzw. aus

$$\underline{U} = L\,\frac{d\underline{I}}{dt} \quad \Leftrightarrow \quad L\,\underline{I} = \int \underline{U}\,dt = \frac{1}{\mathrm{j}\,\omega}\,U_0\,\mathrm{e}^{\mathrm{j}\,\omega\,t}$$

und damit

$$\frac{\underline{U}}{\underline{I}} = \mathrm{j}\,\omega\,L \quad .$$

In der komplexen Ebene stellen sich $\underline{Z}_R$, $\underline{Z}_K$, $\underline{Z}_S$ so dar:

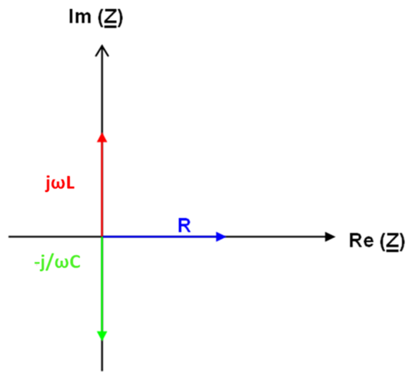


Abb. 2.8 Darstellung des komplexen Widerstands Z in der komplexen Ebene.

Alle bisher genannten Regeln, z. B. KCL, KVL, die Theoreme von Helmholtz-Thévenin und Norton, gelten unverändert für die komplexen Größen. Der Gesamtwiderstand $\underline{Z}$ zweier komplexer Widerstände in Reihe ist

$$\underline{Z} = \underline{Z}_1 + \underline{Z}_2 \quad .$$

Der Gesamtwiderstand $\underline{Z}$ zweier paralleler komplexer Widerstände ist

$$\underline{Z} = \left(\frac{1}{\underline{Z}_1} + \frac{1}{\underline{Z}_2}\right)^{-1} \quad .$$

Durch Bildung des Realteils kann man nach Abschluss der Schaltungsanalyse von den komplexen Größen wieder auf reale Ströme und Spannungen zurückkommen. Dabei bleiben die korrekten Amplituden und Phasenbeziehungen erhalten, z. B. die Phase zwischen der Ein- und der Ausgangsspannung.

2.4 Hochpass

Ein Vorteil der komplexen Schreibweise ist, dass wir den Hochpass (Abb. 2.9) als einen Spannungsteiler auffassen können. Es gilt

$$\underline{U}_{aus} = \underline{U}_{ein} \frac{\underline{Z}_2}{\underline{Z}_1 + \underline{Z}_2}$$

mit

$$\underline{Z} = \frac{\underline{Z}_2}{\underline{Z}_1 + \underline{Z}_2} = \frac{R}{\frac{1}{\mathrm{j}\,\omega\, C} + R} \stackrel{1}{=} \frac{\mathrm{j}\,\omega\, R\, C}{1 + \mathrm{j}\,\omega\, R\, C} \stackrel{2}{=} \frac{\mathrm{j}\,\omega\, R\, C + (\omega\, R\, C)^2}{1 + (\omega\, R\, C)^2} \quad .$$

Die Größe $\frac{\underline{U}U_{aus}}{\underline{U}_{ein}}$ wird auch komplexe **Übertragungsfunktion** [3] genannt. Wir werden später die Übertragungsfunktion des Hochpasses genauer betrachten. Hier sind wir nur an den Amplitudenverhältnissen interessiert. Wir schreiben deshalb $\underline{Z}$ als $Z\,\mathrm{e}^{\mathrm{j}\,\varphi}$ mit $\tan\varphi = \frac{\mathrm{Im}(\underline{Z})}{\mathrm{Re}(\underline{Z})}$.

[1] Multiplikation des Nenners und Zählers mit $\mathrm{j}\,\omega\, C$.
[2] Multiplikation des Nenners und Zählers der vorherigen Gleichung mit $1 - \mathrm{j}\,\omega\, R C$.
[3] Engl. *transfer function*.

$$\text{Es ergibt sich} \quad Z \overset{4}{=} \frac{\omega\, R\, C}{\sqrt{1 + (\omega\, R\, C)^2}} \quad \text{und} \quad \tan\varphi = \frac{1}{\omega\, R\, C} \quad . \tag{2.12}$$

Für hohe Frequenzen $\omega \gg \frac{1}{R\,C}$ gilt, $Z \approx 1$ und damit $U_{aus} \approx U_{ein}$. Für kleine ω dagegen gilt $Z \approx 0$ und $U_{aus} \approx 0$.

Der Kondensator ist für hohe Frequenzen durchlässig, für kleine Frequenzen hat er einen hohen Widerstand. Gleichspannung wird blockiert.

Achtung: Sehr hohe Gleichspannungen können zum Durchbruch des Kondensators und damit zum Kurzschluss führen.

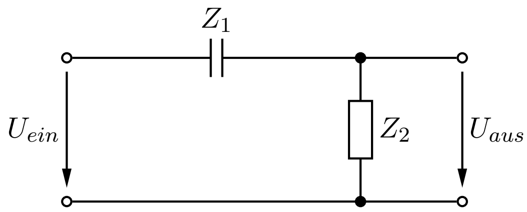


Abb. 2.9 Differenzierer oder Hochpass.

2.5 Tiefpass

Auch der Tiefpass (Abb. 2.10) ist in komplexer Schreibweise ein einfacher Spannungsteiler mit

$$\frac{\underline{U}_{aus}}{\underline{U}_{ein}} = \underline{Z} = \frac{\underline{Z}_2}{\underline{Z}_1 + \underline{Z}_2} = \frac{\frac{1}{\mathrm{j}\,\omega\, C}}{\frac{1}{\mathrm{j}\,\omega\, C} + R} \overset{5}{=} \frac{1}{1 + \mathrm{j}\,\omega\, R\, C} \overset{6}{=} \frac{1 - \mathrm{j}\,\omega\, R\, C}{1 + (\omega\, R\, C)^2} \quad .$$

Wieder schreiben wir $\underline{Z}$ als $Z\,\mathrm{e}^{\mathrm{j}\,\varphi}$ mit

$$Z = \frac{1}{\sqrt{1 + (\omega\, R\, C)^2}} \quad \text{und} \quad \tan\varphi = \frac{\mathrm{Im}(\underline{Z})}{\mathrm{Re}(\underline{Z})} = -\,\omega\, R\, C \quad . \tag{2.13}$$

Hier gilt $Z \approx 1$ für kleine Frequenzen. Kleine Frequenzen können unabgeschwächt passieren. Für hohe Frequenzen $\omega \gg \frac{1}{R\,C}$ gilt $Z \approx 0$.

Der Kondensator ist für hohe Frequenzen durchlässig und führt so zum Kurzschluss der Ausgangsspannung. Diese Schaltung ist auch als Gleichspannungsfilter geeignet. Wechselspannungsstörungen auf einer Gleichspannung werden abhängig von ihrer Frequenz abgeschwächt. Außerdem ist die Ausgangsspannung phasenverschoben und läuft der Eingangsspannung hinterher. Die Phasenverschiebung nimmt mit ansteigender Frequenz zu (siehe Abschnitt 2.10).

[4] $Z = \sqrt{\underline{Z}\,\underline{Z}^*}$ also $Z = \sqrt{\left(\frac{\mathrm{j}\,\omega\, R\, C}{1+\mathrm{j}\,\omega\, R\, C}\right)\left(-\frac{\mathrm{j}\,\omega\, R\, C}{1-\mathrm{j}\,\omega\, R\, C}\right)}$

[5] Multiplikation des Nenners und Zählers mit $\mathrm{j}\,\omega\, C$.

[6] Multiplikation des Nenners und Zählers mit $1 - \mathrm{j}\,\omega\, C$.

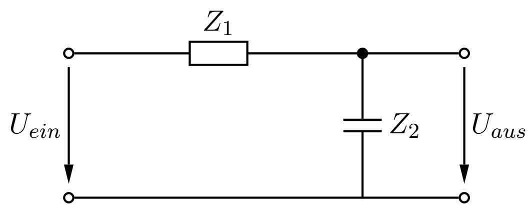


Abb. 2.10 Integrierer oder Tiefpass.

In Abbildung 2.11 ist die Wirkung eines Tiefpasses auf drei exemplarische Eingangssignale unterschiedlicher Frequenz dargestellt. Wie erwartet ist die Abschwächung für die hohen Frequenzen stärker.

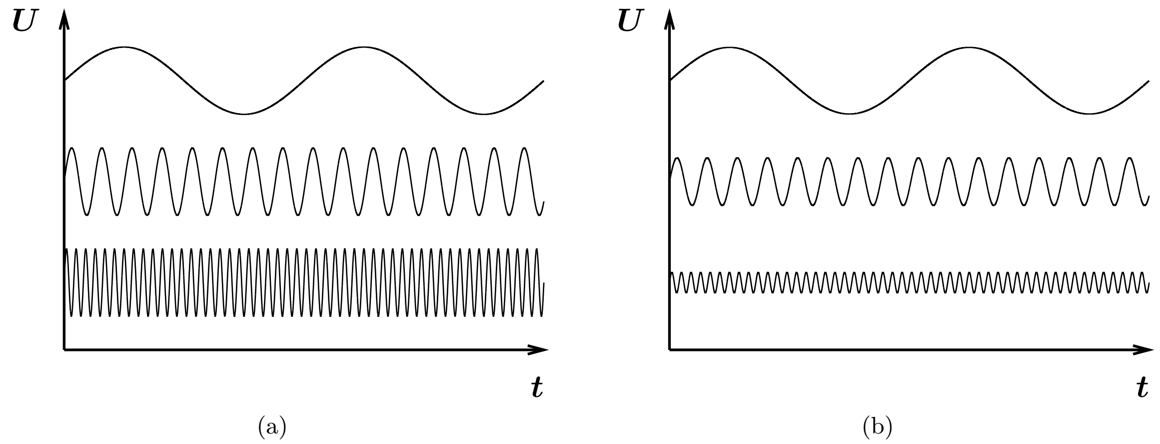


Abb. 2.11 (a) Drei Eingangssignale gleicher Amplitude aber unterschiedlicher Frequenz und (b) die abgeschwächten Ausgangsignale nach Durchlaufen des Tiefpasses. Zur besseren Sichtbarkeit sind die Kurven im Schaubild vertikal gegeneinander verschoben. Die Phasenbeziehung zwischen Ein- und Ausgangssignal ist im Bild nicht berücksichtigt.

Das Verhältnis der Ausgangsamplitude zur Eingagsamplitude ist in Abb. 2.12 dargestellt.

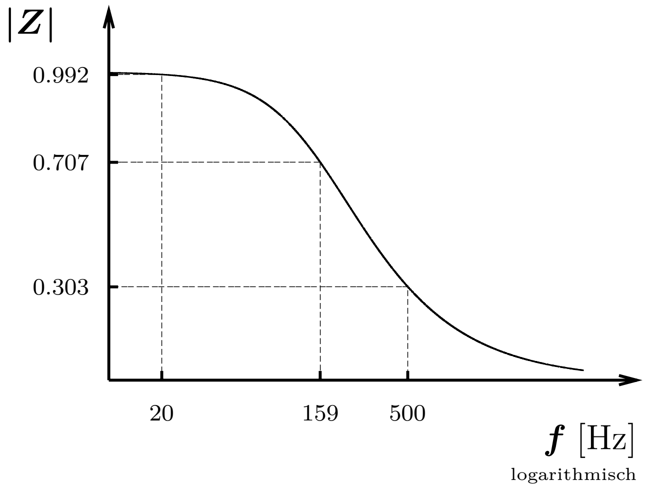


Abb. 2.12 Übertragungsfunktion eines Tiefpasses als Funktion der Frequenz. Die Übertragungsfunktion ist linear, die Frequenz logarihtmisch aufgetragen.

Anmerkung: Die lineare Darstellung der vertikalen Achse ist recht unüblich.

2.6 Reihenschaltung von Widerstand und Spule

Wir betrachten die Schaltung in Abbildung 2.13 Welche Spannung liegt an dem Widerstand und welche an der Spule an?

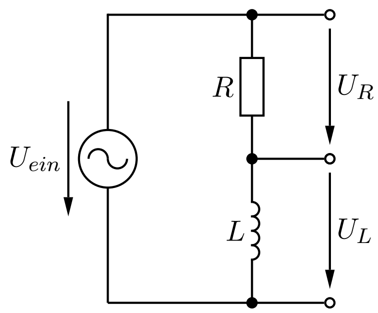


Abb. 2.13 Reihenschaltung von Widerstand und Spule [1].

Die Schaltung ist ein Spannungsteiler mit den komplexen Impedanzen $\underline{Z}_R = R$ und $\underline{Z}_L = \mathrm{j}\,\omega\,L$. Es gilt also

$$\underline{U}_L = \underline{U}_{ein}\,\frac{\underline{Z}_L}{\underline{Z}_R + \underline{Z}_L} \quad \text{und} \quad \underline{U}_R = \underline{U}_{ein}\,\frac{\underline{Z}_R}{\underline{Z}_R + \underline{Z}_L}$$

Die anliegende Wechselspannung $\underline{U}$ setzen wir mit $\underline{U}_{ein} = U_0\,\mathrm{e}^{\mathrm{j}\,\omega\,t}$ an. Die Phase der Eingangsspannung sei gleich 0.

Wir definieren

$$\underline{Z} = \underline{Z}_R + \underline{Z}_L = R + j\omega\,L = \sqrt{R^2 + \omega^2\,L^2}\;\mathrm{e}^{\mathrm{j}\arctan\left(\frac{\omega\,L}{R}\right)} \quad .$$

Im Folgenden kürzen wir $\arctan\left(\frac{\omega L}{R}\right)$ mit φ ab. Der Betrag von $\underline{Z}$, der auch Scheinwiderstand genannt wird, ist $Z = \sqrt{R^2 + \omega^2 L^2}$. Die Größen $\underline{Z}$, $\underline{Z}_R$ und $\underline{Z}_L$ sind in Abbildung 2.14 dargestellt.

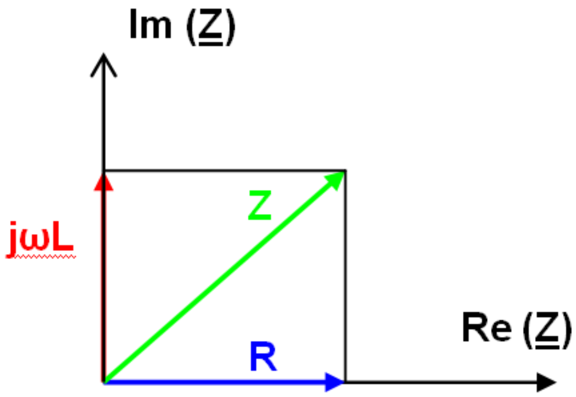


Abb. 2.14 Darstellung von $\underline{Z} = \underline{Z}_R + \underline{Z}_L$ (grün) mit $\underline{Z}_R = R$ (blau) und $\underline{Z}_L = \mathrm{j}\,\omega\, L$ (rot) in der komplexen Ebene.

Für die komplexen Spannungen $\underline{U}_R$ und $\underline{U}_L$ ergibt sich

$$\underline{U}_R = U_0\,\frac{R}{Z}\,\mathrm{e}^{\mathrm{j}\,\omega\, t - \varphi} \quad \text{und} \quad \underline{U}_L = \frac{\mathrm{j}\,\omega\, L}{Z}\,U_0\,\mathrm{e}^{\mathrm{j}\,\omega\, t - \varphi} = \frac{\omega\, L}{Z}\,U_0\,\mathrm{e}^{\mathrm{j}\,\omega\, t - \varphi + \frac{\pi}{2}} \quad .$$

Die Phasenverschiebung zwischen den beiden Spannungen beträgt also wie erwartet 90°.

Die reellen Spannungen ergeben sich mit der korrekten Phasenbeziehung durch Bilden des Realteils der komplexen Größen, also

$$U_R = U_0\,\frac{R}{Z}\,\cos\left[\omega\, t - \arctan\left(\frac{\omega\, L}{R}\right)\right] \quad \text{und} \quad U_L = U_0\,\omega\,\frac{L}{Z}\,\cos\left[\omega\, t - \arctan\left(\frac{\omega\, L}{R}\right) + \frac{\pi}{2}\right] \quad .$$

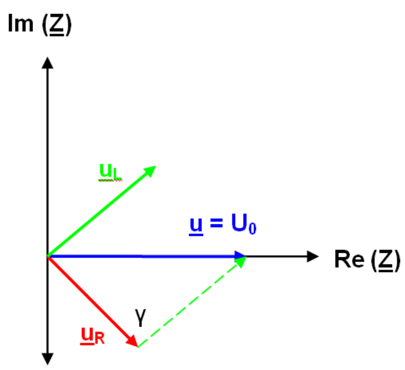


Abb. 2.15 Darstellung der komplexen Spannungen $\underline{U}, \underline{U}_R$ und $\underline{U}_L$ in der komplexen Ebene für $t = 0$ und $\frac{\omega\, L}{R} = 1$ (z. B. bei der willkürlichen Wahl von $\omega = 10\,\mathrm{MHz}$, $L = 100\,\mu\mathrm{H}$, $R = 1\,\mathrm{k}\Omega$ und damit $\phi = \frac{\pi}{4}$).

2.7 Frequenzkompensierter Spannungsteiler

Die Schaltung in Abbildung 2.16 ist bei der Nutzung des Oszilloskops von Bedeutung und ist Teil eines Praktikumsversuchs. Der parallelgeschaltete Widerstand R_1 und Kondensator C_1 des Tastkopfes liegt in Reihe mit der Eingangsimpedanz des Oszilloskops, die aus dem parallelgeschalteten Widerstand R_2 und Kondensator C_2 besteht. C_1 kann man adjustieren. Auch diese Schaltung ist wieder ein Spannungsteiler. Es gilt

$$\frac{\underline{U}_{aus}}{\underline{U}_{ein}} = \frac{R_2 \parallel \underline{Z}_2}{(R_1 \parallel \underline{Z}_1) + (R_2 \parallel \underline{Z}_2)} \quad ,$$

wobei $\underline{Z}_1$ bzw. $\underline{Z}_2$ die Impedanz von C_1 bzw. C_2 ist. Der Eingangswiderstand R_2 des Oszilloskops ist entweder $1\,\mathrm{M\Omega}$ oder für schnelle Signale, bei denen es auf eine korrekte Terminierung ankommt, $50\,\Omega$. Die Kapazität C_2 wird unter anderem durch die Kabelkapazität bestimmt und liegt im Bereich von $100\,\mathrm{pF}$. C_2 wirkt zusammen mit R_1 wie ein Tiefpass und verfälscht die Messung hoher Frequenzen. Durch die korrekte Wahl des Kondensators C_1 kann man das vermeiden. C_1 bildet mit dem Eingangswiderstand R_2 des Oszilloskops einen Hochpass.

Wann ist das Spannungsverhältnis $\frac{\underline{U}_{aus}}{\underline{U}_{ein}}$ frequenzunabhängig und welchen Wert hat es dann?

Es gilt

$$\frac{\underline{U}_{aus}}{\underline{U}_{ein}} = \frac{\frac{R_2}{1+\mathrm{j}\,\omega\,R_2\,C_2}}{\frac{R_1}{1+\mathrm{j}\,\omega\,R_1\,C_1} + \frac{R_2}{1+\mathrm{j}\,\omega\,R_2\,C_2}} \quad .$$

Nach Multiplikation mit $(1 + \mathrm{j}\,\omega\,R_1\,C_1)\,(1 + \mathrm{j}\,\omega\,R_2\,C_2)$ in Zähler und Nenner ergibt sich

$$\frac{\underline{U}_{aus}}{\underline{U}_{ein}} = R_2 \, \frac{1 + \mathrm{j}\,\omega\,R_1\,C_1}{R_1\,(1 + \mathrm{j}\,\omega\,R_2\,C_2) + R_2\,(1 + \mathrm{j}\,\omega\,R_1\,C_1)} \quad .$$

Es ist bequemer, das invertierte Verhältnis $\frac{\underline{U}_{ein}}{\underline{U}_{aus}}$ zu betrachten mit

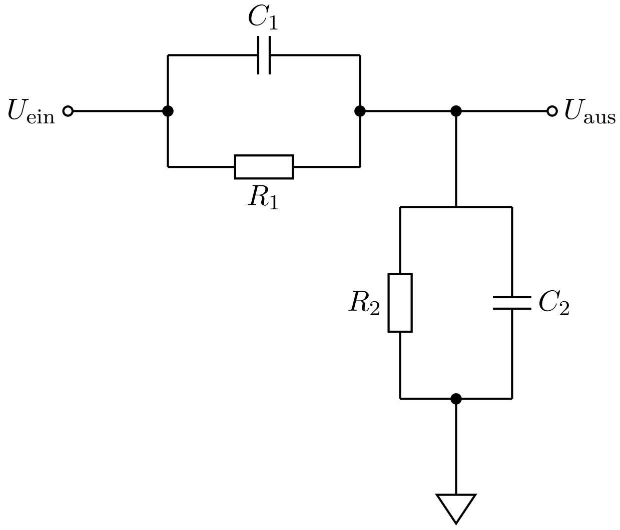
$$\frac{\underline{U}_{ein}}{\underline{U}_{aus}} = 1 + \frac{R_1\,(1 + \mathrm{j}\,\omega\,R_2\,C_2)}{R_2\,(1 + \mathrm{j}\,\omega\,R_1\,C_1)} \quad .$$

Für $R_1\,C_1 = R_2\,C_2$ gilt

$$\frac{\underline{U}_{ein}}{\underline{U}_{aus}} = \frac{U_{ein}}{U_{aus}} = 1 + \frac{R_1}{R_2} \quad .$$

Das Spannungsverhältnis ist dann wie gewünscht frequenzunabhängig, man sagt kompensiert, und entspricht dem des rein ohmschen Spannungsteilers.

(a)



(b)

Abb. 2.16 Frequenzkompensierter Spannungsteiler am Oszilloskop. (a) Foto eines Tastkopfes. (b) Schaltkreis des Tastkopfes mit Eingangswiderstand des Oszilloskops [2].

Es gibt sehr unterschiedliche Tastköpfe je nach Art der Schaltung, die untersucht werden soll. Oft ist $R_1 = 9\,\text{M}\Omega$, was bei der Einstellung von $R_2 = 1\,\text{M}\Omega$ einer Abschwächung des Eingangssignals von 1:10 entspricht. Die Abschwächung verringert die ohmsche und kapazitive Last, die die zu untersuchende Schaltung „sieht“.

Hier handelte es sich um einen passiven Tastkopf. Weitere Typen sind differentielle Tastköpfe oder aktive Tastköpfe mit einem integrierten Feldeffekttransistor zur Verstärkung des Signals direkt an der Schaltung und nicht erst im Oszilloskop.

2.8 Kettenschaltung von RC-Gliedern*

Die Schaltung in Abbildung 2.17, die im Praktikum untersucht wird, kann man als eine Folge von Spannungsteilern auffassen.

$$\text{Es gilt} \quad \underline{U}_{aus} = \underline{U}_{R_1} = \underline{U}_{R_2} \frac{R_1}{Z_1 + R_1} \quad \text{mit} \quad Z_i = \frac{1}{\mathrm{j}\,\omega\, C_i} \quad \text{für} \quad i = 1, 2, 3 \quad .$$

Dies ist die bekannte Spannungsteilerformel. Es spielt keine Rolle, dass im Schaltkreis in Abbildung 2.17 Strom sowohl durch R_1 wie auch durch R_2 fließt.

Frage: Warum nicht?

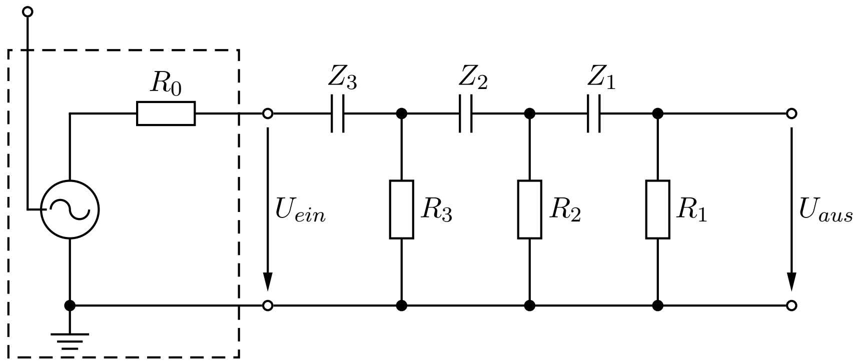


Abb. 2.17 Kettenschaltung von RC-Gliedern [1].

Außerdem gilt

$$\underline{U}_{R_2} = \underline{U}_{R_3} \frac{R_2 \parallel (\underline{Z}_1 + R_1)}{\underline{Z}_2 + R_2 \parallel (\underline{Z}_1 + R_1)} \tag{2.14}$$

und

$$\underline{U}_{R_3} = \underline{U}_{ein} \frac{R_3 \parallel (\underline{Z}_2 + R_2 \parallel (\underline{Z}_1 + R_1))}{\underline{Z}_3 + R_3 \parallel (\underline{Z}_2 + R_2 \parallel (\underline{Z}_1 + R_1))} \quad . \tag{2.15}$$

Für die Praktikumsaufgabe wählen wir $R_1 = R_2 = R_3 = R$ und $\underline{Z}_1 = \underline{Z}_2 = \underline{Z}_3 = \underline{Z}$. Dann ergibt sich (siehe Appendix)

$$\frac{\underline{U}_{ein}}{\underline{U}_{aus}} = 1 + \left(\frac{\underline{Z}}{R}\right)^3 + 5\left(\frac{\underline{Z}}{R}\right)^2 + 6\,\frac{\underline{Z}}{R} \tag{2.16}$$

$$= 1 - \frac{5}{(\omega\, R\, C)^2} + \mathrm{j}\, \frac{\frac{1}{(\omega\, R\, C)^2} - 6}{\omega\, R\, C} \tag{2.17}$$

$$= 1 - \frac{5}{(\omega\, R\, C)^2} + \mathrm{j}\,\varphi \quad \text{mit} \quad \varphi = \frac{\frac{1}{(\omega\, R\, C)^2} - 6}{\omega\, R\, C} \quad . \tag{2.18}$$

Die Phasenverschiebung zwischen Ausgangs- und Eingangsspannung hängt von der Frequenz ab. Für $\omega = \frac{1}{R\,C\,\sqrt{6}}$ wird der Imaginärteil gleich 0. Das heißt die Phasenverschiebung ist entweder 0° oder 180°. Durch Betrachten des Realteils können wir beide Fälle unterscheiden. Da der Realteil für $\omega = \frac{1}{R\,C\,\sqrt{6}}$ den Wert $\frac{U_{ein}}{U_{aus}} = -29$ annimmt, also negativ ist, ist die Phasenverschiebung 180° und nicht 0°.

2.9 Dezibel

Das Dezibel dB wird gerne benutzt um das Verhältnis von Leistungen, Spannungen oder Strömen auszudrücken. So gibt man die Verstärkung eines Verstärkers oder die Abschwächung eines Filters in dB an. Für Leistungen ist das Dezibel definiert als

$$\mathbf{10\ log}\ \frac{\boldsymbol{P_{aus}}}{\boldsymbol{P_{ein}}} \quad ,$$

für Spannungen oder Ströme als

$$\mathbf{20\ log}\ \frac{\boldsymbol{U_{aus}}}{\boldsymbol{U_{ein}}} \quad \text{bzw.} \quad \mathbf{20\ log}\ \frac{\boldsymbol{I_{aus}}}{\boldsymbol{I_{ein}}} \quad .$$

log steht für den Zehnerlogarithmus. Häufig auftretende Werte sind in Tabelle 2.1 zusammengestellt. [7]

dB	**Spannungsverhältnis** $20\ \log \frac{U_{aus}}{U_{ein}}$	**Leistungsverhältnis** $10\ \log \frac{P_{aus}}{P_{ein}}$
100	10^5	10^{10}
40	100	10^4
20	10	100
$6,02$	2	4
$3,01$	$\sqrt{2}$	2
0	1	1
$-3,01$	$\frac{1}{\sqrt{2}}$	$\frac{1}{2}$
$-6,02$	$\frac{1}{2}$	$\frac{1}{4}$
-20	$0,1$	10^{-2}
-40	$0,01$	10^{-4}
-100	10^{-5}	10^{-10}

Tab. 2.1 Übersicht der gängigsten Dezibelwerte.

Durch den Logarithmus in der Definition des Dezibels kann man auch sehr große oder sehr kleine Verhältnisse kompakt darstellen. Ansonsten ist das Dezibel eher verwirrend – zum einen, da nicht immer klar ist, ob es sich bei der in Dezibel ausgedrückten Größe um Leistungen oder Spannungen bzw. Ströme handelt, zum anderen, da es zahlreiche Sonderdefinitionen gibt, mit denen man statt relativen absolute Leistungs- oder Spannungswerte ausdrückt.

So ist der Wert der Leistung P ausgedrückt in dBm definiert als $10\ \log \frac{P[\mathrm{W}]}{1\,\mathrm{mW}}$. Eine Leistung von $P = 1\,\mathrm{W}$ entspricht damit $30\,\mathrm{dBm}$.

Entsprechend ist das dBmA definiert als $20\ \log \frac{I[\mathrm{A}]}{1\,\mathrm{mA}}$. Ein Strom von $1\,\mu\mathrm{A}$ entspricht also

[7] Hilfreich ist es auch folgende Werte zu kennen: $\log 2 \approx 0,3010$, $\frac{1}{\sqrt{2}} \approx 0,707$.

$-60\,\text{dBmA}$.

Die Verstärkung oder Dämpfung zweier hintereinander geschalteter Verstärker, Filter etc. in dB addieren sich, da $\log(a\,b) = \log a + \log b$.

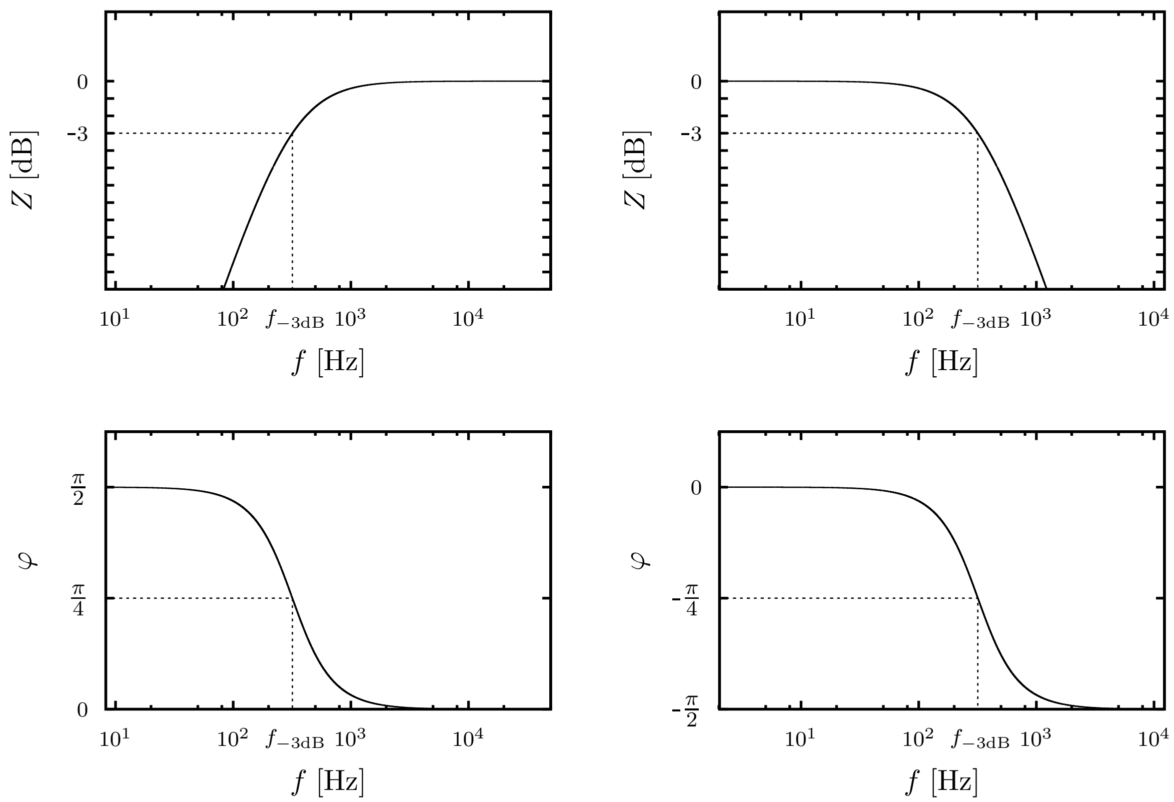
2.10 Bodediagramm

Das Bodediagramm ist eine beliebte graphische Darstellung der komplexen Übertragungsfunktion $Z\,\mathrm{e}^{\mathrm{j}\varphi}$. Dabei werden der Betrag Z der Übertragungsfunktion in dB und die Phase φ linear gegen den Zehnerlogarithmus der Frequenz $f = \frac{\omega}{2\pi}$ aufgetragen (Abb. 2.18).

Das Bodediagramm ist nicht nur hilfreich zur Veranschaulichung des Frequenzverhaltens von Hoch- und Tiefpässen und von Filtern im Allgemeinen. Es dient auch der Darstellung der Verstärkung und der Stabilität von Verstärkern.

Wir hatten in Kapitel 2.4 hergeleitet, dass für den Hochpass bei hohen Frequenzen gilt $Z \approx 1$. Dies entspricht 0 dB wie im Diagramm in Abbildung 2.18(a) links auch angezeigt. Ein Vorteil der logarithmischen Darstellung des Bodediagramms ist, dass man auch den Verlauf von Z für kleine Frequenzen gut darstellen kann. Der Abfall von Z mit abnehmender Frequenz beträgt 20 dB pro Frequenzdekade [8] bzw. 6 dB pro Oktave. D. h., einer zehnfach kleineren Signalfrequenz entspricht eine zehnfach kleinere Verstärkung. (Eine Oktave entspricht, wie in der Musik, der Verdopplung oder Halbierung der Frequenz.) In der logarithmischen Darstellung fällt Z proportional zu $\frac{1}{f}$ ab.

[8] Für kleine Frequenzen gilt: $Z = \frac{\omega\,R\,C}{\sqrt{1+(\omega\,R\,C)^2}} \approx \omega\,R\,C$. Außerdem gilt $20\,\log\,\frac{\omega\,R\,C}{10\,\omega\,R\,C} = 20\,\log\,\frac{1}{10} = -20$.



(a) (b)

Abb. 2.18 (a) Bodediagramm eines Hochpasses und (b) eines Tiefpasses.

Man definiert die Grenzfrequenz $f_{-3\,\mathrm{dB}}$ als die Frequenz, bei der Z auf den Wert $\frac{1}{\sqrt{2}} = -3\,\mathrm{dB}$ abgefallen ist. Die Grenzfrequenz ist für den Hochpass und den Tiefpass gleich groß und beträgt $f_{-3\,\mathrm{dB}} = \frac{1}{2\pi RC}$ bzw. $\omega_{-3\,\mathrm{dB}} = \frac{1}{RC}$. Dies überprüft man leicht durch Einsetzen in die Formel für Z in Kapitel 2.4 und 2.5.

Anmerkung: Die Beziehung $\omega_{-3\,\mathrm{dB}} = \frac{1}{RC}$ darf man sich merken. Sie wird oft benötigt.

Beispiel: Für R = 10 kΩ und C = 100 nF ergibt sich $\omega_{-3\,\mathrm{dB}} = 1$ MHz und $f_{-3\,\mathrm{dB}} \approx 159$ kHz.

Wichtig ist auch das Verhalten der Phase, das heißt die Phasendifferenz zwischen der harmonischen Eingangsspannung und der Ausgangspannung gleicher Frequenz. Für hohe Frequenzen ist die Phasenverschiebung des Hochpasses $\varphi = 0°$. Dies ergibt sich aus $\tan\varphi = \frac{1}{\omega RC} \approx 0°$ für hohe Frequenzen $\omega \gg \frac{1}{RC}$, und entsprechend gilt $\arctan\varphi \approx 0°$. Für kleine Frequenzen strebt $\tan\varphi$ gegen $+\infty$ und $\varphi \approx +90°$. Das heißt die Ausgangsspannung ist um 90° verschoben. Der Wert der Phase bei der Grenzfrequenz beträgt 45°.

Oft ist die Phasenbeziehung irrelevant, aber bei Verstärkern mit Rückkopplung ist die korrekte Phasenbeziehung entscheidend.

Für den Tiefpass gilt $Z \approx 0$ dB für kleine Frequenzen, da dann $Z = \frac{1}{\sqrt{1+(\omega RC)^2}} \approx 1$. Hier fällt Z mit 20 dB pro Dekade für hohe Frequenzen ab. Die Phasenverschiebung ist für kleine

Frequenzen $\varphi \approx 0°$ und für hohe Frequenzen $-90°$. Die Ausgangsspannung ist verzögert. Dies kann man sich damit erklären, dass es Zeit kostet, den Kondensator aufzuladen bzw. ihn zu entladen.

Man kann sich den Phasenverlauf eines Hoch- und Tiefpasses sehr leicht merken. Bei $|Z| = 1$ bzw. 0 db beträgt die Phasenverschiebung 0°, bei $|Z| << 1$ ist die Phase entweder um +90° (Hochpass) oder −90° (Tiefpass) verschoben. Die Phasenverschiebung bei $\omega_{-3\,\mathrm{dB}}$ beträgt ± 45°.

2.11 Filter zweiter Ordnung*

Der Hoch- und Tiefpass aus dem vorherigen Abschnitt sind Filter erster Ordnung. Es sind die einfachsten elementaren Filter. Es ist leicht, komplexere und bessere Filter aufzubauen, wenn man mehr Komponenten einsetzt. So kann es zum Beispiel vorteilhaft sein, wenn die Übertragungsfunktion Z des Tiefpasses schneller als $-20\,$db pro Frequenzdekade abnimmt. Im folgenden werden einige einfache Filter zweiter Ordnung vorgestellt, die aus jeweils einer Spule und einem Kondensator als frequenzabhängige Impedanzen aufgebaut sind.

2.11.1 Tiefpass

Abb. 2.19 zeigt einen Tiefpass zweiter Ordnung.

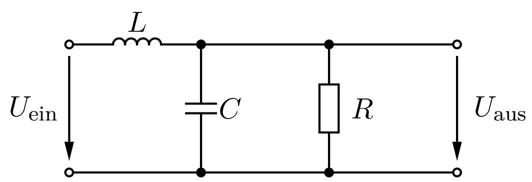


Abb. 2.19 Beispiel eines Tiefpasses zweiter Ordnung.

Die Übertragungsfunktion ergibt sich mit dem Widerstand R, den komplexen Impedanzen Z_C, Z_L und der Abkürzung $s = j\,\omega$ zu

$$Z = \frac{Z_C||R}{Z_L + Z_C||R} = \frac{\frac{R\frac{1}{sC}}{R+\frac{1}{sC}}}{sL + \frac{R\frac{1}{sC}}{R+\frac{1}{sC}}} = \frac{\frac{R}{sRC+1}}{sL + \frac{R}{sRC+1}} = \frac{R}{s^2RCL + sL + R} \quad . \tag{2.19}$$

Im Nenner tritt also ein Polynom zweiten Grades in s, also der Frequenz, auf. Dadurch fällt der Tiefpass für hohe Frequenzen mit $-40\,$db pro Frequenzdekade ab. Außerdem variiert die Phase nicht zwischen 0° und -90°, sondern zwischen 0° und -180°. Mit den drei unabhängigen Werten von R, C und L gibt es einen weiteren Freiheitsgrad im Vergleich zu Filtern erster Ordnung. Dies ist die Güte Q. Der Wert der Grenzfrequenz ω_0 und der Güte Q ist

$$\omega_0 = \frac{1}{\sqrt{LC}} \quad , \quad Q = \omega_0\,RC \quad .$$

Mit den Größen ω_0 und Q lässt sich die Übertragungsfunktion des Tiefpasses auch als

$$Z = \frac{\omega_0^2}{s^2 + \frac{\omega_0}{Q}s + \omega_0^2}$$

schreiben. Je größer Q, desto höher ist die Spitze der Übertragungsfunktion bei $\omega = \omega_0$. Das Bodediagramm des Tiefpasses ist in Abbildung 2.20 dargestellt.

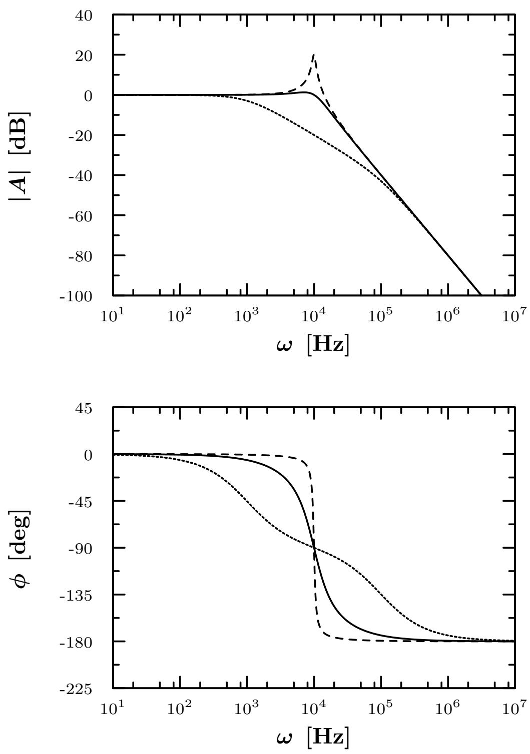


Abb. 2.20 Bodediagramm eines Tiefpasses zweiter Ordnung mit $\omega = 10$ kHz und $Q = 1$ (durchgezogen), $Q = 10$ (gestrichelt) bzw. $Q = 0,1$ (gepunktet).

2.11.2 Hochpass

Abb. 2.21 zeigt einen Hochpass zweiter Ordnung.

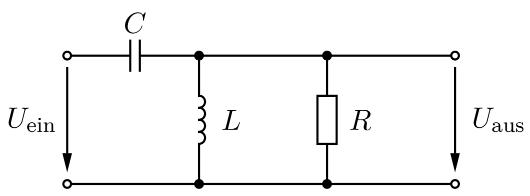


Abb. 2.21 Beispiel eines Hochpasses zweiter Ordnung.

Die Übertragungsfunktion ergibt sich zu

$$Z = \frac{sL||R}{\frac{1}{sC} + sL||R} = \frac{\frac{sLR}{sL+R}}{\frac{1}{sC} + \frac{sLR}{sL+R}} = \frac{sLR}{\frac{1}{sC}(sL+R) + sRL} \tag{2.20}$$

$$= \frac{s^2RCL}{sL + R + s^2RCL} = \frac{s^2}{s^2 + \frac{\omega_0}{Q}s + \omega_0^2} \quad . \tag{2.21}$$

Man sieht leicht, dass der Betrag der Übertragungsfunktion bei kleinen Frequenzen gegen den Wert Null strebt und bei sehr hohen Frequenzen gegen 1 konvergiert. Auch hier beträgt der Anstieg 40 db pro Dekade, und $\omega_0 = \frac{1}{\sqrt{LC}}$ und $Q = \omega_0 RC$.

Das Bodediagramm des Hochpasses ist in Abbildung 2.22 dargestellt.

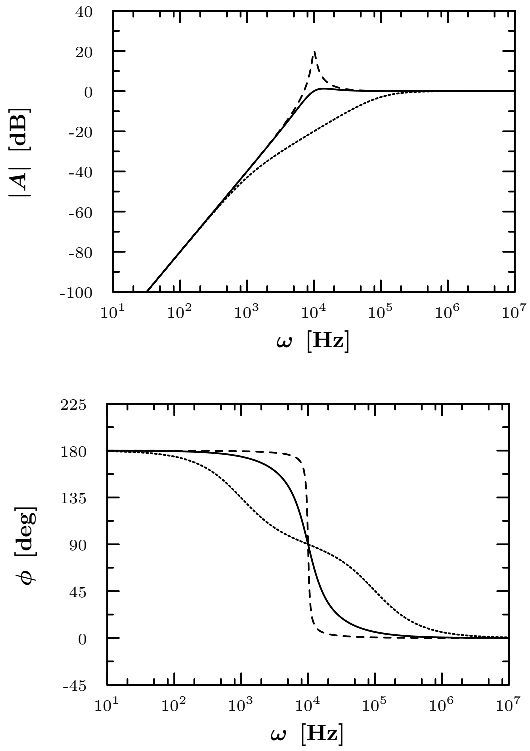


Abb. 2.22 Bodediagramm eines Hochpasses zweiter Ordnung mit $\omega = 10$ kHz und $Q = 1$ (durchgezogen), $Q = 10$ (gestrichelt) bzw. $Q = 0,1$ (gepunktet).

2.11.3 Bandpass

Abb. 2.23 zeigt einen Bandpass zweiter Ordnung.

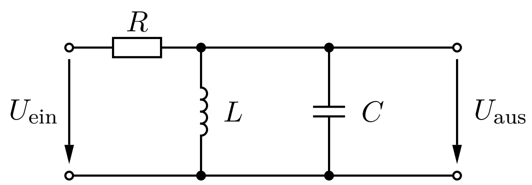


Abb. 2.23 Beispiel eines Bandpasses zweiter Ordnung.

Die Übertragungsfunktion ergibt sich zu

$$Z = \frac{sL||\frac{1}{sC}}{R + sL||\frac{1}{sC}} = \frac{\frac{sL}{s^2LC+1}}{R + \frac{sL}{s^2LC+1}} = \frac{sL}{s^2RLC + sL + R} = \frac{\frac{\omega_0}{Q}s}{s^2 + \frac{\omega_0}{Q}s + \omega_0^2} \quad . \tag{2.22}$$

Der Bandpass sperrt bei sehr kleinen und sehr großen Frequenzen, bei mittleren Frequenzen ist er durchlässig. Es gilt $\omega_0 = \frac{1}{\sqrt{LC}}$ und $Q = \omega_0 RC$.

Das Bodediagramm des Bandpasses ist in Abbildung 2.24 dargestellt.

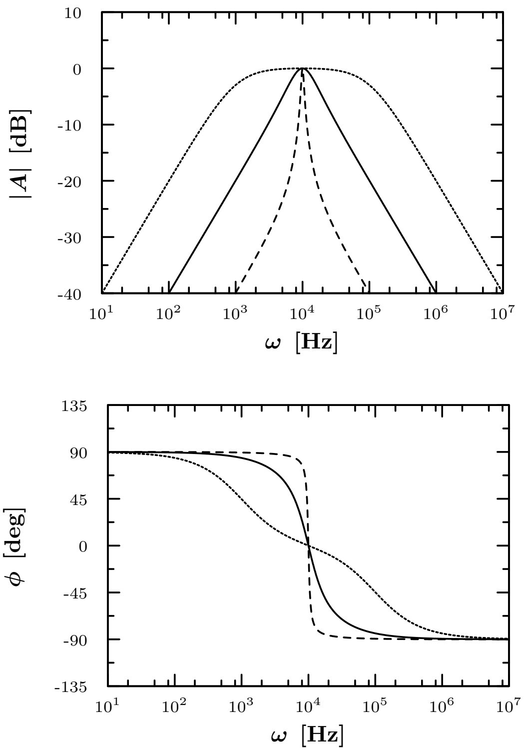


Abb. 2.24 Bodediagramm eines Bandpasses zweiter Ordnung mit $\omega = 10$ kHz und $Q = 1$ (durchgezogen), $Q = 10$ (gestrichelt) bzw. $Q = 0,1$ (gepunktet).

2.11.4 Bandsperre

Die Bandsperre (Abb. 2.25) verhält sich gerade umgekehrt wie der Bandpass.

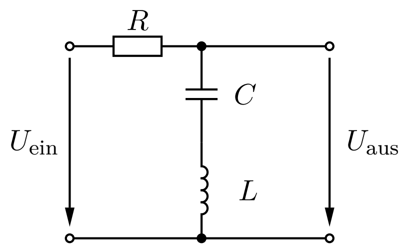


Abb. 2.25 Beispiel einer Bandsperre zweiter Ordnung.

Die Übertragungsfunktion ergibt sich zu

$$Z = \frac{sL + \frac{1}{sC}}{R + sL + \frac{1}{sC}} = \frac{s^2LC + 1}{sRC + s^2LC + 1} = \frac{s^2 + \omega_0^2}{s^2 + \frac{\omega_0}{Q}s + \omega_0^2} \quad . \tag{2.23}$$

mit $\omega_0 = \frac{1}{\sqrt{LC}}$ und $Q = \frac{L}{R}\,\omega_0$. Hier ist Q ein Maß für die Breite des abgeschwächten Bereichs. Das Bodediagramm einer Bandsperre ist in Abbildung 2.26 dargestellt.

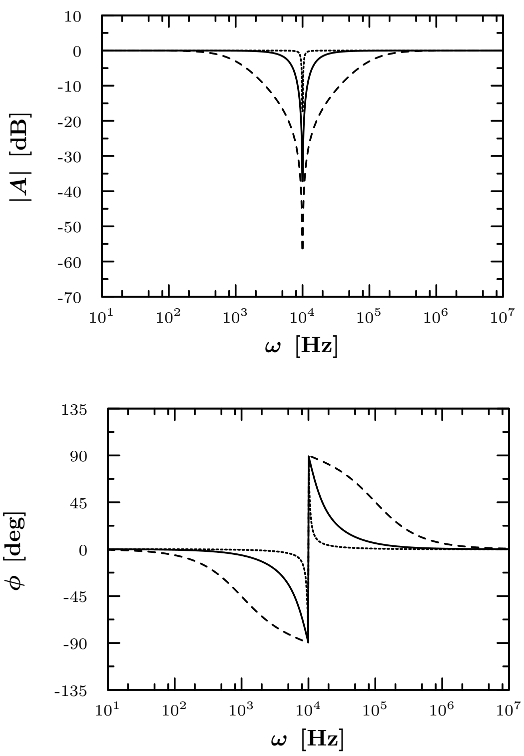


Abb. 2.26 Bodediagramm einer Bandsperre zweiter Ordnung mit $\omega = 10$ kHz und $Q = 1$ (durchgezogen), $Q = 0,1$ (gestrichelt) bzw. $Q = 10$ (gepunktet).

Anmerkung: Man kann Filter höherer Ordnung auch durch das Hintereinanderschalten von Hoch- und Tiefpässen erster Ordnung aufbauen (siehe Abb. 2.27). Welche Nachteile hat dies?

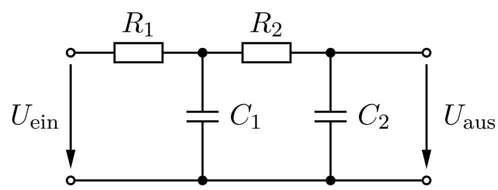


Abb. 2.27 Filter aus zwei hintereinandergeschalteten passiven Tiefpässen.

2.12 Allpassfilter*

Ein Allpass überträgt alle Frequenzanteile des Eingangsignals ungedämpft auf den Ausgang. Der Betrag der Übertragungsfunktion $|\mathrm{Z}(\omega)| = 1$ bzw. eine andere frequenzunabhängige Konstante wie wir unten sehen werden. Wozu ist der Allpass dann gut? Er wird benutzt, um die Phase des Eingangsignals zu verändern, d.h. das Signal abhängig von der Frequenz mehr oder weniger zu verzögern.

Ein Allpass erster Ordnung ist in Abb. 2.28 dargestellt.

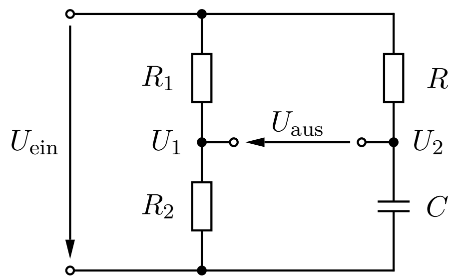


Abb. 2.28 Allpass erster Ordnung.

Es gilt

$$U_{\text{aus}} = U_2 - U_1 = U_{\text{ein}} \left(\frac{Z}{Z+R} - \frac{R_1}{R_1+R_1} \right) \quad .$$

Die Übertragungsfunktion für $R_1 = R_2$ ist

$$\begin{aligned} \underline{Z} &= \frac{dU_{\text{aus}}}{dU_{\text{ein}}} = \frac{Z}{Z+R} - \frac{R_1}{R_1+R_1} = \frac{1}{1+sRC} - \frac{1}{2} \\ &= \frac{1}{2}\left(\frac{1-sCR}{1+sCR}\right) \\ &= \frac{1}{2}\left(\frac{(1-sRC)^2}{1-s^2R^2C^2}\right) \\ &= \frac{1}{2}\left(\frac{1-2j\omega RC-\omega^2R^2C^2}{1+\omega^2R^2C^2}\right) \quad . \end{aligned} \tag{2.24}$$

Damit gilt

$$|\underline{Z}| = \frac{1}{2} \quad \text{und} \quad \tan\varphi = \frac{-2\omega RC}{1-\omega^2R^2C^2} \quad .$$

Hier ist die Amplitude der Übertragungsfunktion konstant. Die Phase variiert von 0° bei $\omega = 0$ bis zu $-180°$ bei $\omega \to \infty$. Man kann sich das Verhalten des Allpasses veranschaulichen, in dem man die Extremfälle $\omega = 0$ und $\omega = \infty$ betrachtet, die jeweils die Ausgangsspannung $U_{\text{aus}} = U_{\text{ein}}\,(1 - \frac{1}{2}) = U_{\text{ein}}/2$ und $U_{\text{aus}} = U_{\text{ein}}\,(0 - \frac{1}{2}) = -U_{\text{ein}}/2$ erzeugen.

Auch der Grenzfall

$$\omega = \frac{1}{RC}$$

ist schnell überprüft und ergibt

$$U_{\text{aus}} = U_{\text{ein}} \left(\frac{1}{1+j} - \frac{1}{2} \right) = U_{\text{ein}} \left(\frac{1-j}{2} - \frac{1}{2} \right) = \frac{-j}{2}\, U_{\text{ein}}$$

und damit

$$|\underline{Z}| = \frac{1}{2} \quad .$$

Diese drei Spezialfälle entsprechen der Phase von 0°, $-90°$ und $-180°$. Das Bodediagramm ist in Abbildung 2.29 dargestellt.

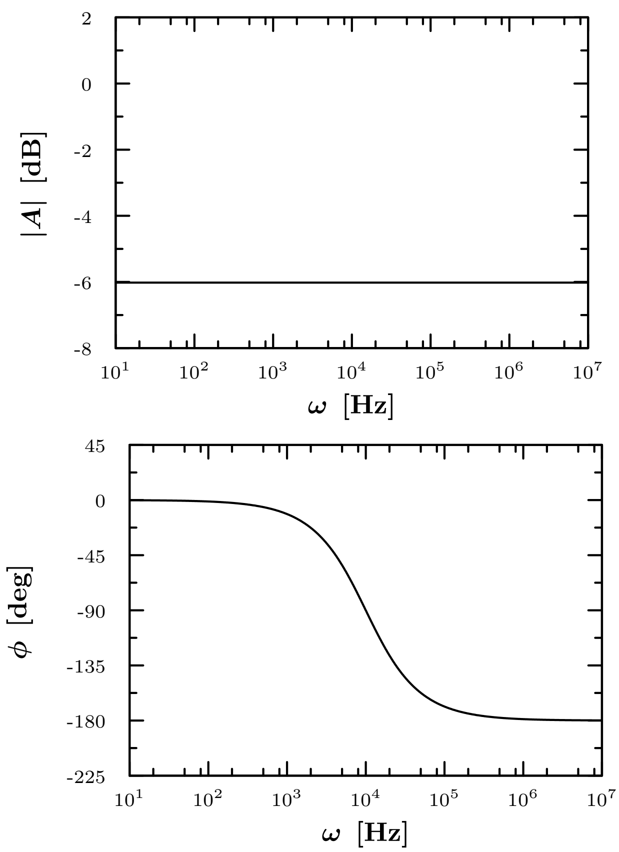


Abb. 2.29 Bodediagramm eines Allpasses erster Ordnung mit Grenzfrequenz $\omega = 10$ kHz.

Ein Allpassfilter zweiter Ordnung, eine symmetrische T-Topologie ist in Abb. 2.30 dargestellt. Auch hier lässt sich die Übertragungsfunktion elementar berechnen.

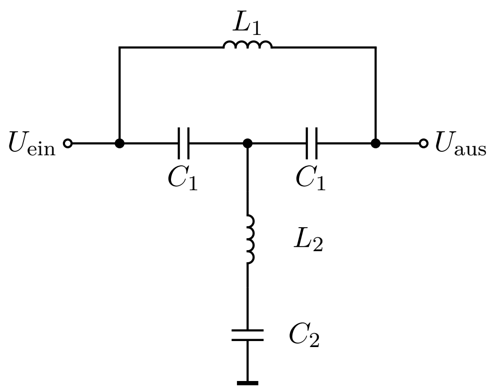


Abb. 2.30 Beispiel eines symmetrischen Allpassfilters zweiter Ordnung.

Wir stellen zuerst alle unabhängigen Gleichungen auf.

$$U_{ein} = U_{aus} + Z_1 I_1$$

$$U_{ein} = Z_2 I_2 + (Z_4 + Z_5) I_4$$

$$I_2 = I_3 + I_4$$
$$U_{ein} = U_{aus} + Z_2 I_2 + Z_3 I_3$$
$$U_{aus} = I_4(Z_4 + Z_5) - Z_3 I_3$$

Es gibt fünf Unbekannte, nämlich die Größen U_a, I_2, I_3, I_4, I_1. Die Übertragungsfunktion ergibt sich zu

$$\underline{Z} = \frac{s^2 - \frac{\omega_0}{Q}\, s + \omega_0^2}{s^2 + \frac{\omega_0}{Q}\, s + \omega_0^2} \quad .$$

Das Bodediagramm ist in Abb. 2.31 dargestellt. Hier ist die Phasenverschiebung größer als beim Allpass erster Ordnung (Abb. 2.28).

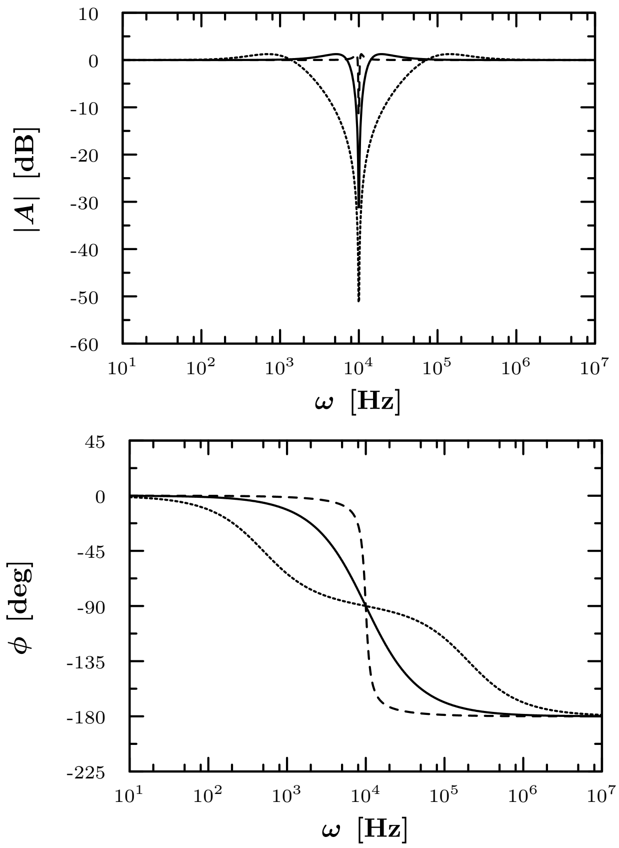


Abb. 2.31 Bodediagramm eines Allpasses zweiter Ordnung.

2.13 Komplexe Filter*

In Abb. 2.32 ist eine etwas unübersichtliche Bandsperre dargestellt (siehe Texas Instruments, „Op Amps for Everyone“). Es ist jetzt leicht, auch derartige Schaltungen zu analysieren.

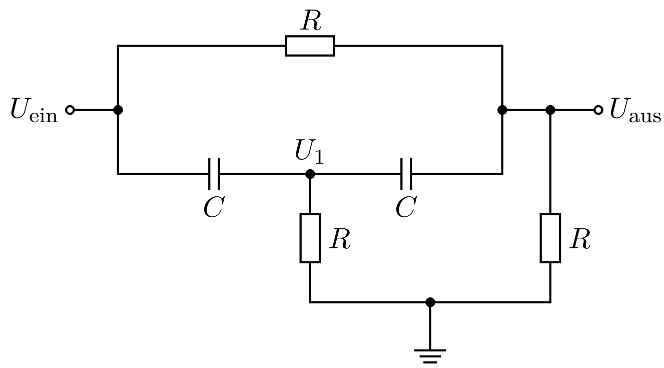


Abb. 2.32 Bandsperre.

Die komplexe Übertragungsfunktion ergibt sich zu (siehe Appendix 19.15)

$$\underline{Z}(j\omega) = \frac{\underline{U}_{aus}}{\underline{U}_{ein}} = \frac{(1+j\omega\,\tau)\,(1+j\omega\,\tau)}{2\;(1+j\omega\,\tau/a)\;(1+j\omega\,\tau/b)} \quad \text{mit} \quad \tau = RC \quad . \tag{2.25}$$

Die Übertragungsfunktion ist ein Polynom mit einer doppelten Nullstelle (zero) im Zähler und zwei Polen (poles) im Nenner. (Die Nullstelle und Pole liegen in der linken Hälfte der komplexen Ebene mit $Re(\omega) < 0$.) Die Konstanten a und b ergeben sich zu

$$a = \frac{5+\sqrt{17}}{2} \approx 4,561 \quad \text{und} \quad a \cdot b = 2 \quad , \quad \text{also} \quad b \approx 0,438 \quad .$$

Das Bodediagramm der Übertragungsfunktion ist in Abb. 2.33 dargestellt. Der Amplitudenverlauf ist deutlich konturenreicher als ein Tief- oder Hochpass und hat ein Minimum bei $\omega = 1/RC$. Ausgehend von kleinen Frequenzen fällt die Übertragungsfunktion zunächst wie ein einfacher Tiefpass ab. Dies wird durch den Pol bei $\omega = 0.438/RC$ bewirkt. Bei Erreichen der Frequenz $\omega = 1/RC$, führt die doppelte Nullstelle der Übertragungsfunktion zu einem Anstieg von 40 db/Dekade. Wird dann die Frequenz $\omega = 4,561/RC$ erreicht, so flacht die Übertragungsfunktion ab, da jetzt der zweite Pol relevant wird.

Das Phasenverhalten ist noch unübersichtlicher und soll hier nicht betrachtet werden.

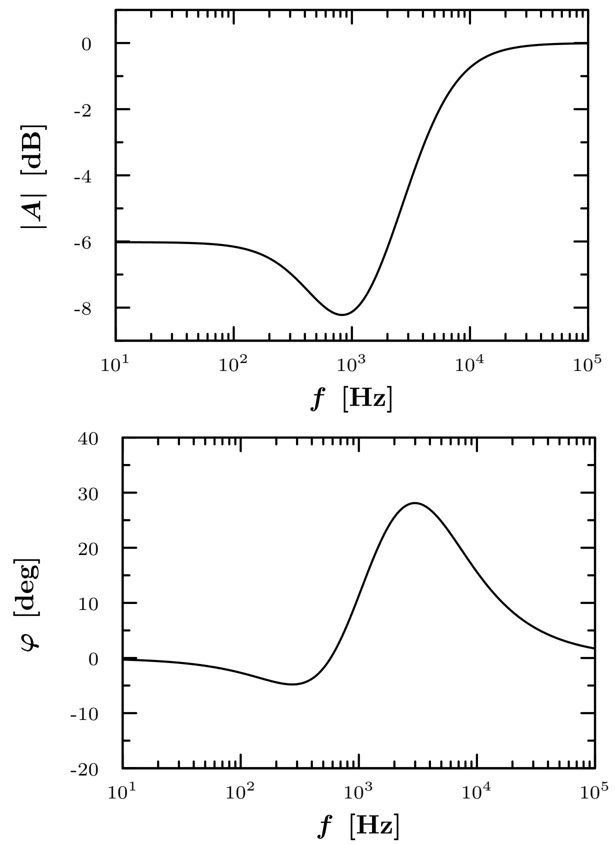


Abb. 2.33 Bodediagramm der Funktion aus Gleichung 2.25.

2.14 Schwingkreise

Aus der Kombination von Kondensator, Spule und Widerstand kann ein Schwingkreis aufgebaut werden. Dies ist ein weiteres Arrangement von Spule und Kondensator. Ein Schwingkreis besteht aus einer Parallel- oder Reihenschaltung von Spule und Kondensator (Abb. 2.34 und 2.35), die durch eine Strom- oder Spannungsquelle angeregt werden. Diese Schaltung entspricht bis auf die Position des Widerstandes der Bandsperre und dem Bandfilter aus Kapitel 2.13.

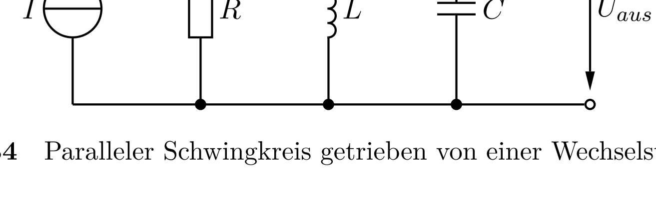


Abb. 2.34 Paralleler Schwingkreis getrieben von einer Wechselstromquelle.

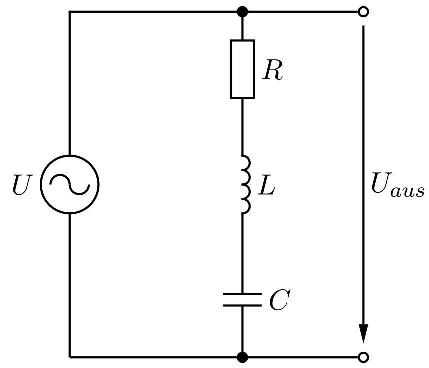


Abb. 2.35 Reihenschwingkreis getrieben von einer Wechselspannungsquelle.

Ein Schwingkreis muss nicht mit einer kontinuierlichen Wechselspannung betrieben werden, sondern kann auch durch einen einmaligen Spannungs- oder Stromstoß angeregt werden. Alternativ kann man nach Anlegen der Wechselspannung bzw. nach Anregen des Schwingkreises die Spannungsquelle kurzschließen bzw. die Stromquelle entfernen. Die Kopplung der Energiespeicher Kondensator und Spule hält dann auch ohne äußere Energiezufuhr für eine Zeit harmonische Schwingungen von Strom und Spannung aufrecht, daher die Bezeichnung Schwingkreis.

Wir behandeln den Schwingkreis gegen Ende der Vorlesung zur Illustration der Lösung von Einschwingvorgängen mittels des Laplace-Formalismus. Hier interessiert uns nur das mathematisch einfachste Verhalten der Schaltung bei angelegten Wechselspannungen.

Wir berechnen zuerst den komplexen Widerstand, die Impedanz, der Kombination von Widerstand, Kondensator und Spule. Aus Kenntnis des Gesamtwiderstands lassen sich in einer konkreten Schaltung die Ausgangsspannung oder andere relevante Größen leicht ableiten.

2.15 Parallelschwingkreis

Im Parallelschwingkreis gilt

$$\underline{Z} = \frac{1}{\frac{1}{\underline{Z}_R} + \frac{1}{\underline{Z}_C} + \frac{1}{\underline{Z}_L}} = \frac{1}{\frac{1}{R} + \mathrm{j}\,\omega\,C - \mathrm{j}\,\frac{1}{\omega\,L}} = \frac{\frac{1}{R} - \mathrm{j}\left(\omega\,C - \frac{1}{\omega\,L}\right)}{\frac{1}{R^2} + (\omega\,C - \frac{1}{\omega\,L})^2}$$

mit dem Betrag

$$Z = \frac{1}{\sqrt{\frac{1}{R^2} + \left(\omega\,C - \frac{1}{\omega\,L}\right)^2}} \quad \text{und der Phase} \quad \tan\varphi = \left(\frac{1}{\omega\,L} - \omega\,C\right) R \quad . \tag{2.26}$$

Der Betrag Z wird für die Resonanzfrequenz $\boldsymbol{\omega_0 = \frac{1}{\sqrt{LC}}}$ maximal. Für $\omega = \omega_0$ wird $\underline{Z}$ reell. $Z = R$ und $\varphi = 0$.

Der Betrag Z und die Phase sind in Abbildung 2.36 gegen die Frequenz $\frac{\omega}{\omega_0}$ aufgetragen. Es ist plausibel, dass Z für mittlere Frequenzen maximal ist. Denn für sehr kleine Frequenzen schließt die Spule die Parallelschaltung kurz ($Z \approx \omega L \approx 0$) und für sehr große Frequenzen der Kondensator ($Z \approx \frac{1}{\omega C} \approx 0$). Für mittlere Frequenzen haben beide einen von 0 verschiedenen Widerstand.

Aus $\tan\varphi = (\frac{1}{\omega L} - \omega C)\,R$ folgt, $\varphi = \pi/2$ bzw. 90° für $\omega = 0$, $\varphi = -90°$ für $\omega = \infty$, und $\varphi = 0$ für $\omega = \omega_0 = \frac{1}{\sqrt{L C}}$.

Auch der Phasenverlauf ist plausibel und entspricht für kleine Frequenzen einer Spule für große Frequenzen einem Kondensator und bei der Resonanzfrequenz einem Widerstand mit $\varphi = 0$.

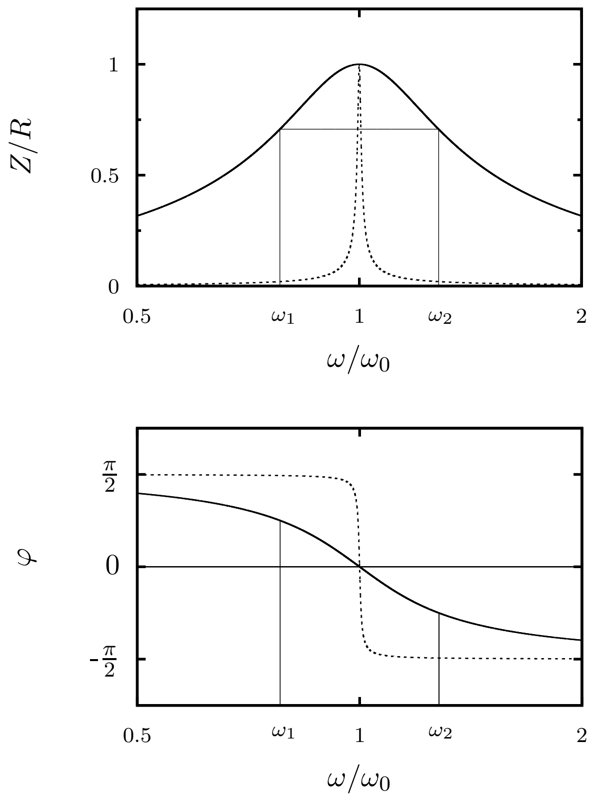


Abb. 2.36 Impedanz und Phase für den Parallelschwingkreis für zwei unterschiedliche Bandbreiten (siehe unten) $\Delta\omega = 0,3\,\omega_0$ (durchgezogen) und $\Delta\omega = 0,1\,\omega_0$ (gestrichelt).

2.16 Reihenschwingkreis

Im Reihenschwingkreis gilt

$$\underline{Z} = \underline{Z}_R + \underline{Z}_C + \underline{Z}_L = R + \mathrm{j}\,\omega L - \mathrm{j}\,\frac{1}{\omega C}$$

mit dem Betrag

$$Z = \sqrt{R^2 + \left(\omega L - \frac{1}{\omega C}\right)^2} \quad \text{und der Phase} \quad \tan\varphi = \frac{\omega L - \frac{1}{\omega C}}{R} \quad . \tag{2.27}$$

Hier wird der Betrag Z für die Resonanzfrequenz $\omega_0 = \frac{1}{\sqrt{LC}}$ minimal. Auch das ist plausibel. Da für kleine und große Frequenzen entweder der Kondensator oder die Spule einen großen Widerstand bieten, muss bei mittleren Frequenzen ein Minimum liegen. Auch hier gilt für $\omega = \omega_0$ wieder: $\underline{Z}$ wird reell, $Z = R$ und $\varphi = 0$.

Der Phasenverlauf entspricht hier bei kleinen Frequenzen einem Kondensator und bei großen Frequenzen einer Spule. Das Bodediagramm des Reihenschwingkreis ist in Abbildung 2.37 dargestellt.

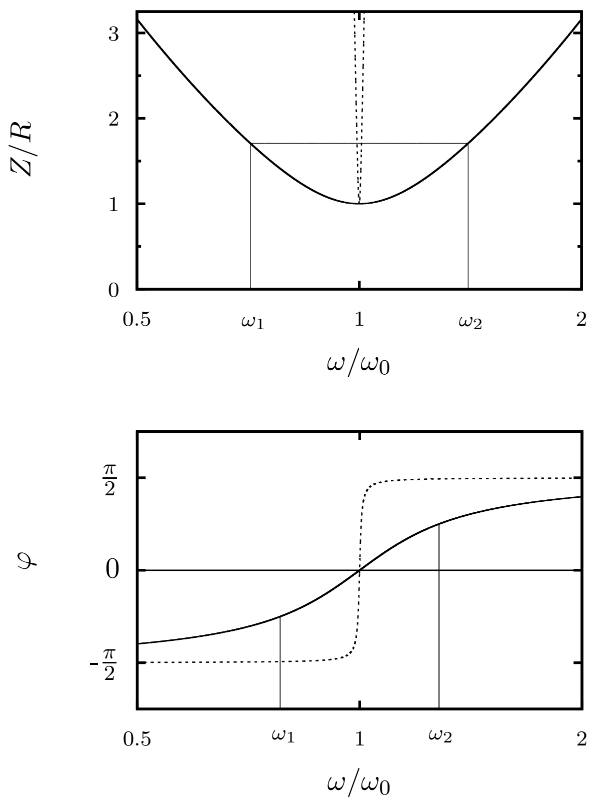


Abb. 2.37 Impedanz und Phase für den Reihenschwingkreis für zwei unterschiedliche Bandbreiten (siehe unten) $\Delta\omega = 0,3\,\omega_0$ (durchgezogen) und $\Delta\omega = 0,1\,\omega_0$ (gestrichelt).

2.17 Güte und Bandbreite

Die Güte Q eines Schwingkreises ist definiert als das Verhältnis von Resonanzfrequenz ω_0 und Bandbreite $\Delta\,\omega$

$$Q = \frac{\omega_0}{\Delta\omega} = \frac{\omega_0}{\omega_2 - \omega_1} \quad . \tag{2.28}$$

Dabei sind ω_2 und ω_1 die Grenzfrequenzen, bei denen der Wert der Spannung oder des Stroms auf $\frac{1}{\sqrt{2}}$ des Spitzenwertes, der bei der Resonanzfrequenz ω_0 erreicht wird, abgefallen ist. Die Bandbreite $\Delta\,\omega$ ist das Frequenzintervall, in dem eine „nennenswerte" Ausgangsspannung auftritt. Dieselben Frequenzen erhält man aus der äquivalenten Definition von $\omega_{1,2}$ als denjenigen Frequenzen, bei denen die Phase $\pm\,45°$ ist.

Anmerkung: Die Güte Q kann äquivalent aber allgemeiner definiert werden als die Zahl der Schwingungen mit einer nennenswerten Amplitude bzw. über den relativen Energieverlust pro Periode des freischwingenden Schwingkreises.

Wir bestimmen zuerst ω_2 und ω_1 für den Parallelschwingkreis. Aus $\underline{U} = \underline{Z}\,\underline{I}$ ergibt sich die Ausgangsspannung, die der Erregerstrom bewirkt. Wir suchen also die Werte ω_2 und ω_1, für die Z auf $\frac{1}{\sqrt{2}}$ des Maximalwertes gesunken ist. Mit

$$Z = \frac{1}{\sqrt{\frac{1}{R^2} + \left(\omega\,C - \frac{1}{\omega\,L}\right)^2}}$$

gilt:

- Z ist maximal, falls $\omega\,C - \frac{1}{\omega\,L} = 0 \quad \Rightarrow \quad Z = R$.
- Z ist auf $\frac{1}{\sqrt{2}}$ gesunken, falls $\left(\omega\,C - \frac{1}{\omega\,L}\right)^2 = \frac{1}{R^2}$.

Dies führt zu der quadratischen Gleichung

$$\pm\left(\omega^2 - \frac{1}{L\,C}\right) - \frac{\omega}{R\,C} = 0$$

mit den Lösungen

$$\omega_{1,2} = \frac{1}{2\,R\,C} \pm \sqrt{\frac{1}{(2\,R\,C)^2} + \frac{1}{L\,C}} \quad \text{und} \quad \omega_{3,4} = -\frac{1}{(2\,R\,C)} \pm \sqrt{\frac{1}{(2\,R\,C)^2} + \frac{1}{L\,C}} \quad .$$

Nur ω_1 und ω_3 sind positive Frequenzen. Dadurch ergibt sich die Bandbreite als $\omega_1 - \omega_3 = \boldsymbol{\Delta\omega = \frac{1}{R\,C}}$. Entsprechende Gleichungen erhält man durch den alternativen Ansatz $\tan(\pm 45°) = \pm 1 = \left(\frac{1}{\omega\,L} - \omega\,C\right)\,R$

Die Güte des Parallelschwingkreises ist also $Q_P = \frac{R\,C}{\sqrt{L\,C}} = R\sqrt{\frac{C}{L}}$. Es ist intuitiv einleuchtend, dass sich die Güte proportional zu R verhält. Schließlich bewirkt ein sehr kleiner Widerstand eine Entladung oder sogar einen Kurzschluss von C und L. Dagegen entspricht ein unendlich hoher Widerstand dem idealen Schwingkreis.

Entsprechend leitet man die Bandbreite $\omega_2 - \omega_1 = \Delta\omega = \frac{R}{L}$ und die Güte des Reihenschwingkreises $Q_R = \frac{1}{\sqrt{L\,C}}\,\frac{L}{R} = \frac{1}{R}\sqrt{\frac{L}{C}}$ ab (siehe Appendix).

In Tabelle 2.2 sind einige der bisherigen Ergebnisse zu Schwingkreisen zusammengefasst. Für einen Beispielwerte von C = 100 nF, L = 10 μH und R = 1 kΩ ergibt sich für einen Parallelschwingkreis ω_0 = 1 MHz, $\Delta\omega$ = 10 kHz und Q = 100.

	ω_0	$\Delta\omega$	Q
Reihenschwingkreis	$\frac{1}{\sqrt{L\,C}}$	$\frac{R}{L}$	$\frac{1}{R}\sqrt{\frac{L}{C}}$
Parallelschwingkreis	$\frac{1}{\sqrt{L\,C}}$	$\frac{1}{R\,C}$	$R\sqrt{\frac{C}{L}}$

Tab. 2.2 Übersicht: Schwingkreise

Durch den Einsatz von Schwingkreisen in Schaltungen lassen sich Signalkomponenten der Resonanzfrequenz ω_0 verstärken oder unterdrücken.

Frage: Wie erklärt man sich anschaulich das Auftreten der Größen L und C in den Formeln für die Bandbreite des Reihenschwingkreises und des Parallelschwingkreises?

2.18 Relle Widerstände, Kondensatoren und Spulen

Es gibt die unterschiedlichsten Widerstände, Kondensatoren und Spulen, entsprechend der Vielfalt der möglichen Anwendungen. Bei der Auswahl der Komponente kommt es nicht nur auf den Wert des Widerstandes, der Kapazität oder Induktivität an, sondern auf einiges mehr. Wichtig sind auch

- die maximal verträglichen Leistungen, Ströme und Spannungen des Bauteils
- das Hochfrequenzverhalten
- die Toleranz, das heißt die zulässige Abweichung z. B. des Widerstands vom Sollwert
- die Temperaturstabilität und der erlaubte Temperaturbereich
- die Bauform und Größe
- bei größeren Stückzahlen auch der Preis.

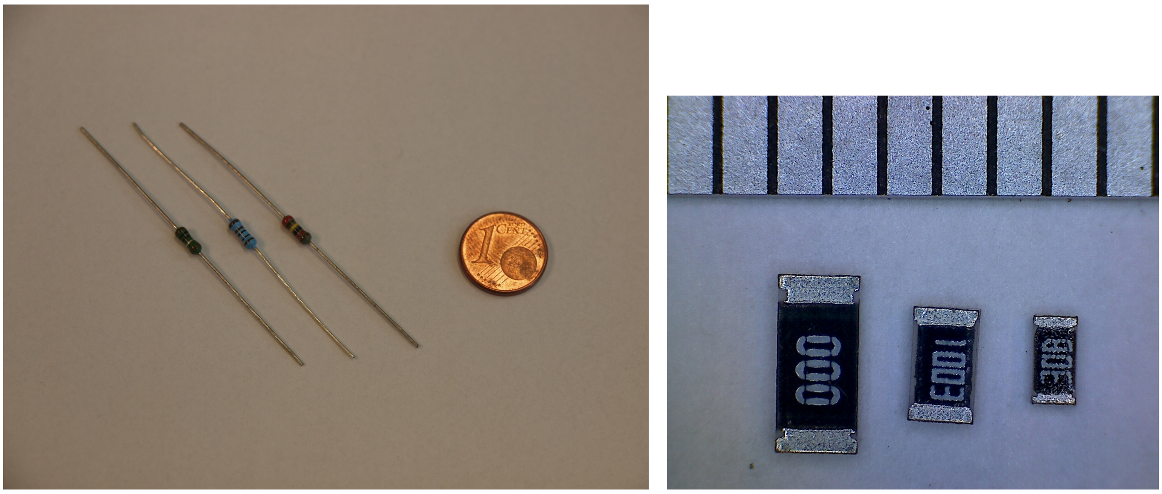
Die aufgezählten Eigenschaften sind stark abhängig von der Bauart der Komponente.

2.19 Reelle Widerstände

Zwei wichtige Typen sind in Abbildung 2.38 dargestellt. Der zylindrische Widerstand mit Anschlussdrähten *(through-hole)* ist vielleicht vertrauter. Man kann die Drahtenden durch Löcher der Platine stecken und leicht festlöten. Die miniaturisierten SMD-Widerstände werden dagegen nur auf der Oberfläche der Platine befestigt und eignen sich besonders für kompakte Aufbauten und maschinelle Bestückung. SMD steht für *Surface Mounted Device*. SMD-Bauteile sind leichter zu handhaben, als man auf den ersten Blick meinen könnte. Für die Detektorinstrumentierung und industrielle Produkte, z. B. Verbraucherelektronik, werden fast ausschließlich SMDs benutzt.

Wie kann man den Widerstand eines Bauteils einstellen? Entsprechend der Formel $\boldsymbol{R = \rho\,\frac{l}{A}}$ variiert man den spezifischen Widerstand, die Länge oder den Querschnitt des Leiters. Alle diese Möglichkeiten werden oder wurden in der Praxis verwendet.

So wurden in **Kohlewiderständen** Kohlepulver als Leiter mit Keramikteilchen als Nichtleiter vermischt und verklebt. Durch Variation der Anteile von Kohle und Keramik lässt sich der spezifische Widerstand einstellen. Diese Widerstände sind nicht mehr gebräuchlich.



(a) (b)

Abb. 2.38 (a) Drahtwiderstand. (b) SMD-Widerstand mit Lineal.

Man kann auch einfach einen dünnen **Metalldraht**, z. B. eine Nickel-Chrom-Legierung, über einen Trägerstift aus Plastik, Hartglas oder Keramik wickeln. Die Länge und der Querschnitt des Drahtes bestimmen dann den Widerstand. Diese Widerstände sind für mittlere und höhere Leistungen ($> 1\,\mathrm{W}$) geeignet und verbreitet. Es ist wichtig, dass man beim Wickeln nicht gleichzeitig eine Spule erzeugt! Das kann man durch Bifilarwickelungen (Wickelung einer Drahtschleife, statt eines Einzeldrahtes) vermeiden, trotzdem sind diese Widerstände aufgrund ihrer induktiven Komponente für Hochfrequenzanwendungen nicht geeignet.

Schichtwiderstände werden hergestellt, indem man auf einen zylindrischen Träger eine dünnen Kohle, Metall oder Metalloxidschicht aufbringt. Man kann die Schichtdicke variieren ($\frac{1}{1000}\,\mu$m bis $10\,\mu$m), z. B. bei Kohleschichtwiderständen, oder aus der Vollschicht Windungen strukturieren (Metallschichtwiderstände) und damit den Widerstand einstellen.

Zur Herstellung von **Dickschichtwiderständen** werden auf ein flaches Keramiksubstrat leitende Schichten aufgedruckt und bei hohen Temperaturen ($> 850\,°\mathrm{C}$) verfestigt (gebacken). Für **Dünnschichtwiderstände** wird das Metall unter Vakuum aufgedampft. In beiden Technologien können die Schichtdicke und Fläche und damit der Widerstand kontrolliert werden.

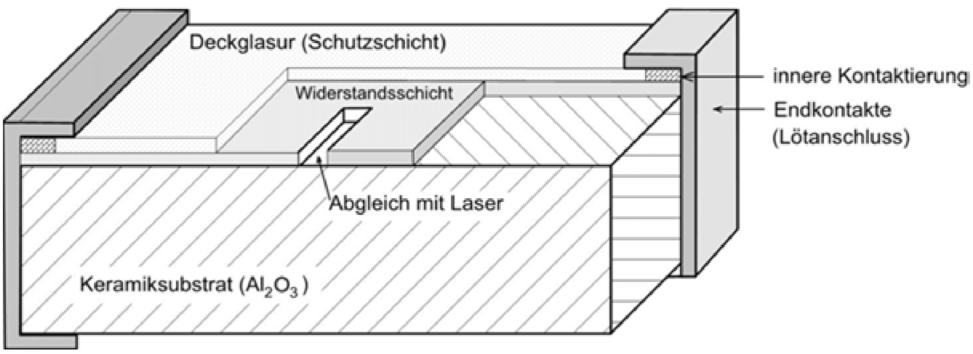


Abb. 2.39 Aufbau eines SMD-Widerstands in Dickfilmtechnologie (Quelle: Leonhard Stiny).

Der reelle Widerstand hat nicht nur einen ohmschen Widerstand, sondern auch kapazitive und induktive Komponenten. Daher sieht das Ersatzschaltbild des reellen Widerstands wie in Abbildung 2.40 dargestellt aus.

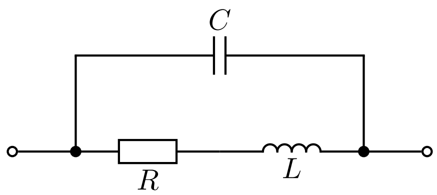


Abb. 2.40 Ersatzschaltbild eines reellen Widerstands.

Die Spule in Reihe zu dem ohmschen Widerstand ist für gewickelte Widerstände und Widerstände mit Anschlussdrähten besonders wichtig und kann leicht Induktivitäten von $1\,\mu$H bis $10\,$nH annehmen. Der Kondensator parallel zu Spule und Widerstand wird zum Beispiel durch die Endflächen eines Bauteiles gebildet.

Was bedeutet dies qualitativ für das Freuenzverhalten eines Widerstands?

Bei sehr hohen Freuenzen schließt der Kondensator die Reihenschaltung von Widerstand und Spule kurz. Die Impedanz nimmt bei hohen Frequenzen also immer ab. Hohe Frequenzen liegen im Bereich von Gigaherz bis zu hunderten von GHz. Bei mittleren Frequenzen beeinflussen sowohl der Kondensator wie die Spule die Impedanz. Die Schaltung ähnelt einem Parallelschwinkreis. Es kann zu einer Resonanzerhöhung der Impedanz gegenüber dem nominallen Wert des Widerstands kommen. Das Verhalten bei mittleren bis kleinen Frequenzen ist im folgenden exemplarisch ausgeführt.

Für den komplexen Leitwert des realen Widerstands ergibt sich

$$\underline{G} = \frac{1}{R + \mathrm{j}\,\omega\,L} + \mathrm{j}\,\omega\,C \quad .$$

Mit den Definitionen $\tau_C = R\,C$ und $\tau_L = \frac{L}{R}$ folgt

$$\underline{G}\,R = \frac{1}{1 + \mathrm{j}\,\omega\,\tau_L} + \mathrm{j}\,\omega\,\tau_C$$

und damit

$$\frac{Z}{R} = \frac{1}{|\underline{G}\,R|} = \left[\frac{1}{1 + \mathrm{j}\,\omega\,\tau_L} + \mathrm{j}\,\omega\,\tau_C\right]^{-1} = \frac{1 + \mathrm{j}\,\omega\,\tau_L}{1 + \mathrm{j}\,\omega\,\tau_C - \omega^2\,\tau_C\,\tau_L}$$

Multiplikation des Zählers und Nenners mit $[1 - \mathrm{j}\,\omega\,\tau_C - \omega_2\,\tau_C\,\tau_L]$ ergibt

$$\frac{Z}{R} = \frac{(1 + \mathrm{j}\,\omega\,\tau_L)\,(1 - \mathrm{j}\,\omega\,\tau_C - \omega^2\,\tau_C\,\tau_L)}{(1 - \omega^2\,\tau_C\,\tau_L)^2 + \omega^2\,\tau_C^2} \quad .$$

Der Ausdruck ist etwas unübersichtlich, deshalb betrachten wir im folgenden nur kleine und mittlere Frequenzen und vernachlässigen Terme der Ordnung ω^3 und ω^4 (siehe M. Reisch).

$$\frac{Z}{R} \approx \frac{1 + \mathrm{j}\,\omega\,(\tau_L - \tau_C)}{1 + \omega^2\,\tau_C\,(\tau_C - 2\,\tau_L)} \quad .$$

Wichtig an diesem Ausdruck ist zweierlei:

1. Für $\tau_C = \tau_L$ verschwindet der Imaginärteil. Dies kann man für nicht zu große Widerstände im Bereich von $10\,\Omega - 1\,\mathrm{K}\Omega$ durch geeignete Dimensionierung der Bauform erreichen.

2. Das Frequenzverhalten des Realteils von Z, der bei kleinem Imaginärteil den Betrag Z bestimmt, hängt von der Differenz $\tau_C - 2\,\tau_L$ ab. Für $\tau_C > 2\,\tau_L$ nimmt $\mathrm{Re}(\underline{Z})$ mit der Frequenz ab und der kapazitive Anteil dominiert. Dies ist besonders bei großem R der Fall. $\mathrm{Re}(\underline{Z})$ kann auch mit der Frequenz zunehmen, falls ab $\tau_C < 2\,\tau_L$. Dies ist bei kleinem R wahrscheinlicher.

Das Frequenzverhalten von typischen Dünn- und Dickschichtwiderständen ist in Abbildung 2.41 gezeigt. Bei Frequenzen unterhalb von 10 MHz muss man sich bei nicht zu großen Widerständen keine großen Gedanken um die Frequenzabhängigkeit machen.

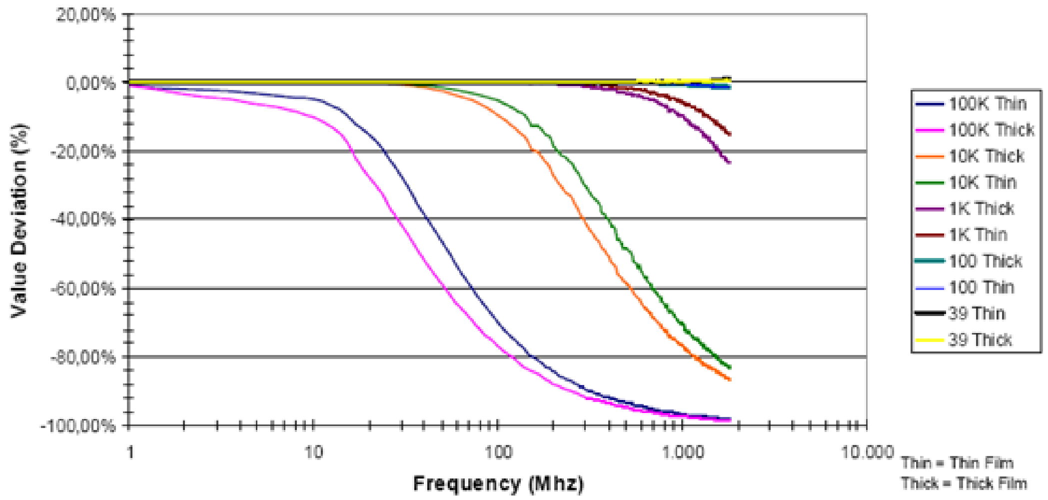


Abb. 2.41 Frequenzverhalten von Widerständen (Quelle: Vishay).

2.20 Reelle Kondensatoren

Auch für den Kondensator gibt es eine Vielzahl von Typen mit sehr unterschiedlichen Eigenschaften. Eine Auswahl ist in Abb. 2.42 gezeigt.

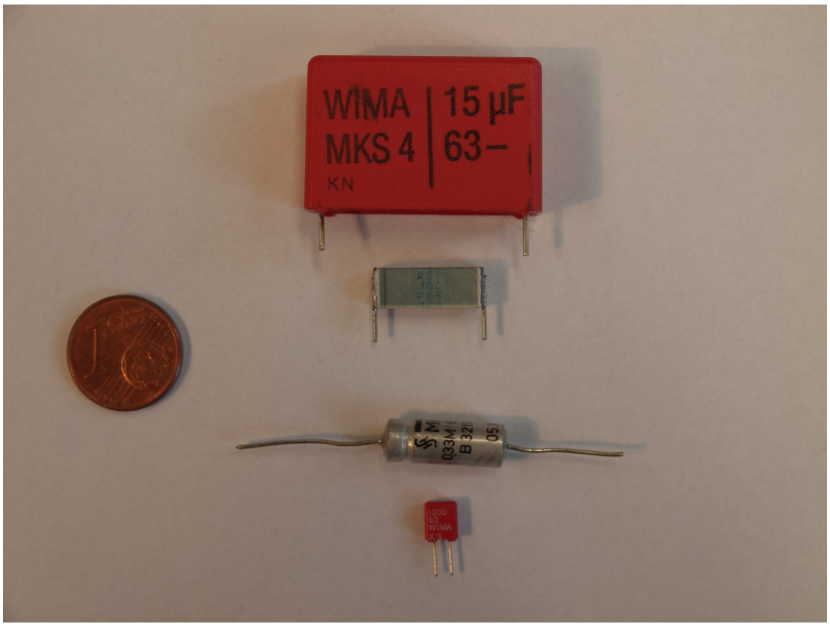


Abb. 2.42 Beispiele für bedrahtete Kondensatoren.

Der Wert der Kapazität eines idealen Plattenkondensators ist durch $C = \epsilon_0\,\epsilon_r\,\frac{A}{d}$ gegeben. In praktischen Einheiten gilt

$$C = 8,9\,[\mathrm{fF}]\;\frac{\epsilon_r\,A\,[\text{in mm}^2]}{d\,[\text{in mm}]} \qquad .$$

Um eine große Kapazität zu erreichen, kann also die Fläche A der Platten, ihren Abstand d oder die relative Dielektrizitätskonstante ϵ_r optimiert werden.

Die Dielektrizitätskonstanten verschiedener Materialien sind in Tabelle 2.3 zusammengestellt.

Material	Dielektrizitätskonstante
Luft	1
Teflon	$2,1$
Epoxy	$3,6$
Al_2O_3	10
Ta_2O_5	27
Nb_2O_5	42
Kapton	$3,3$
$BaTi_4O_9$	38
Mineralöl	$2,2$
$PbZr_xTi_{1-x}$ (PZT) [8]	$300 \ldots 4000$
Si	$11,8$
SiO_2	$4,5$
Diamant	$16,5$
Porzelan	7
Wasser (18 °C)	81

Tab. 2.3 Dielektrizitätskonstanten.

Neben dem einfachen Plattenkondensator gibt es noch andere Bauformen, z. B. den Zylinderkondensator (Abb. 2.43) mit der Kapazität

$$C = 2\,\pi\,\epsilon_0\,\epsilon_r\,\frac{l}{\ln\left(\frac{r_2}{r_1}\right)} \quad . \tag{2.29}$$

Hier sind R_1 und R_2 der Innen- und der Außendurchmesser der konzentrischen Zylinderschalen, l ist die Zylinderlänge. Durch das Auftreten des Logarithmus ist die Abhängigkeit der Kapazität von den Radien r_1 und r_2 klein.
Mit der Gleichung 2.29 kann z. B. die Kapazität einer abgeschirmten Leitung [9] berechnet werden.

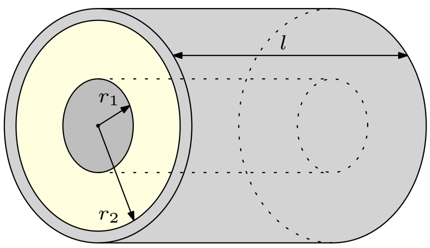


Abb. 2.43 Zylinderkondensator.

Man unterscheidet folgende Kondensatortypen:

[8] Blei-Zirkonat-Titanat
[9] Dabei ist die Signalleitung am Innenleiter angeschlossen. Der Außenleiter (Mantel) liegt an Masse.

- Papierkondensatoren
- Keramikkondensatoren
- Kunststofffolienkondensatoren
- Elektrolytkondensatoren: Aluminium, Tantal, jeweils nass oder trocken.

Um hohe Kapazitätswerte bei kleiner Baugröße zu erreichen, wird oft durch Schichten oder Wicklungen die Kondensatorfläche A vergrößert (Abb. 2.44. Es wechseln sich Leiter, Isolator, der Leiter des zweiten Pols, Isolator usw. ab. Da jede Schicht in der Mitte des Kondensators zu beiden Seiten die andere Elektrode „sieht“ gilt für die Kapazität eines Folienkondensators

$$C \approx 2\,\epsilon_0\,\epsilon_r\,\frac{A}{d} \quad . \tag{2.30}$$

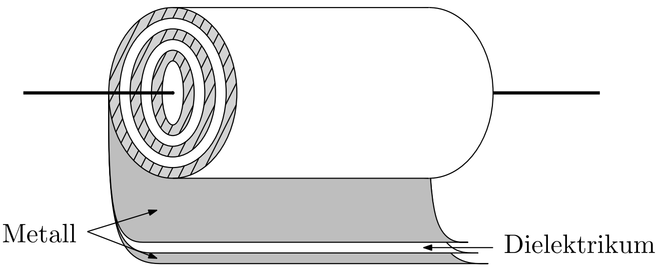


Abb. 2.44 Folienkondensator.

Die Zuleitungen werden seitlich an den Wicklungen und nicht etwa an den Enden der Folie angeschlossen, um die Leitungslänge möglichst klein zu halten. Dies reduziert auch die störende Leitungsinduktivität.

Die ersten massengefertigten Kondensatoren waren gewickelte **Papierkondensatoren** aus zwei Schichten ölgetränktem Papier und Metallfolien. Das Öl ist ein sehr guter Isolator.

Kunststofffolienkondensatoren decken heutzutage einen weiten Bereich an Kapazitätswerten ab. Je nach Konstruktion reicht die Spannungsfestigkeit bis in den kV-Bereich. Bei zu hohen Spannungen kann es zu Durchschlägen kommen, die jedoch nicht zu einem dauerhaften Kurzschluss führen („selbstheilend“).

Bei einem Schichtkondensator erhöht sich mit jeder Lage die Kapazität

$$C_{Schicht} \approx \epsilon_0\,\epsilon_r\,(n-1)\frac{A}{d} \quad . \tag{2.31}$$

Keramikkondensatoren sind als Einschicht- oder Vielschichtkondensatoren aufgebaut. Speziell als Vielschichtkondensatoren[10] zur SMD-Bestückung finden sie weite Verbreitung. Die Keramik bildet das Dielektrikum und ist mit anderen Stoffen gemischt, um bestimmte Eigenschaften, wie Frequenzunabhängigkeit oder geringe Temperaturabhängigkeit, zu erhalten. Zur Herstellung wird Keramikpulver mit den entsprechenden Zusatzstoffen gemischt, in Form gebracht und gesintert.

[10]Engl. Multi-layer ceramic capacitor (MLCC).

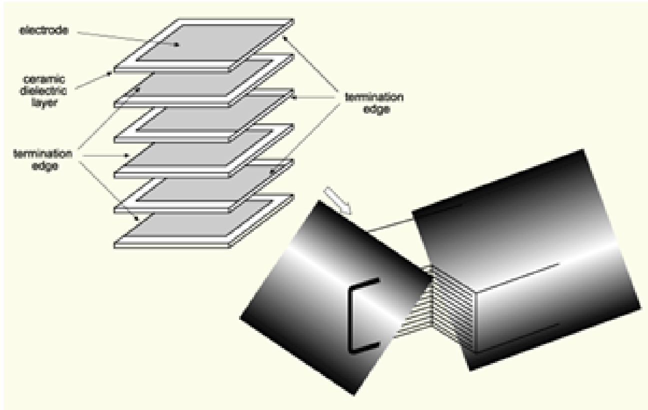


Abb. 2.45 Schematischer Aufbau eines Vielschichtkondensators.

Elektrolytkondensatoren

Elektrolytkondensatoren (Elkos) sind in der Praxis auf Grund ihrer großen Kapazitäten bei relativ kleiner Baugröße sehr beliebt. Anders als bei den bisher geannnten Kondensatoren muss man bei Anschließen von Elkos ihre Polung beachten. Bereits eine falsche Anwendung führt zur Explosion und Zerstörung der Elkos.

Die Anode eines Elkos ist rau und hat deshalb eine große Oberfläche. Zur Erzeugung des Dielektrikums wird die Anodenoberfläche oxidiert. Die Oxidschicht liegt in der Größenordnung von nur 1 - 10 nm. Bei „nassen“ Alu- und Tantal-Elkos wird nun als Kathode eine leitende Flüssigkeit benutzt, die alle Bereiche der Grenzfläche erreichen kann. Diese wird durch eine feste Kathodenkontaktierung umschlossen. Hat der Gleichanteil der Spannung bei Benutzung die richtige Polung, bleibt das Dielektrikum erhalten bzw. wird weiter aufgebaut.

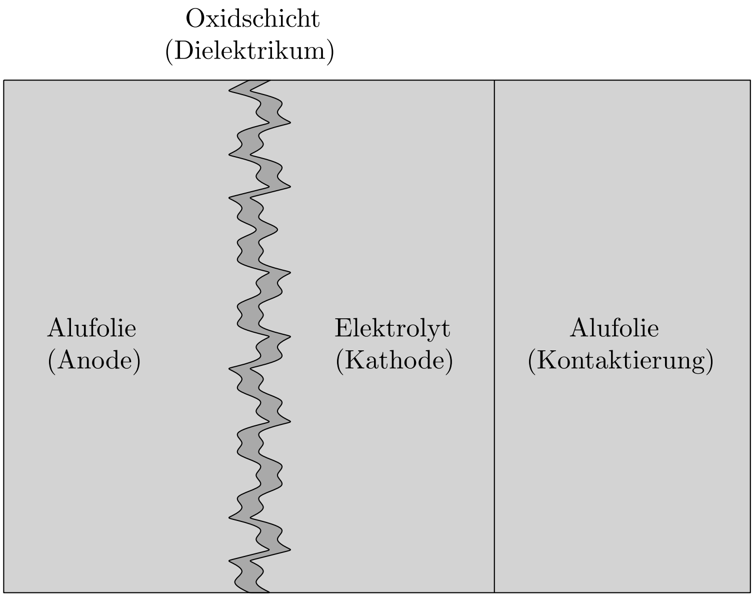


Abb. 2.46 Schematischer Aufbau eines Elektrolytkondensators.

Aluminium-Elektrolytkondensatoren besitzen eine Aluminiumoxidschicht (Al_2O_3) als Dielektrikum. Die Anodenfolie wird aufgewickelt, um noch mehr Oberfläche auf kleinem Platz unterzubringen. Die Lebensdauer ist dadurch begrenzt, dass der (oft) flüssige Elektrolyt im Laufe der Zeit nach außen diffundiert, so dass der Kondensator austrocknet.

Bei **Tantal-Elektrolytkondensatoren** besteht die Anode aus rauem, gesintertem, porösem Tantal, Dielektrikum ist Tantalpentoxid (Ta_2O_5). Neben flüssigen Elektrolyten gelingt auch die Herstellung mit festem Elektrolyt aus Braunstein oder Polymer. Dies hat den Vorteil, dass keine Verdunstung stattfindet und somit eine lange Lebensdauer möglich ist. Dafür ist der Preis des Tantals hoch.

Ersatzschaltbild eines Kondensators

Das Ersatzschaltbild eines Kondensators ist in Abbildung 2.47 dargestellt.

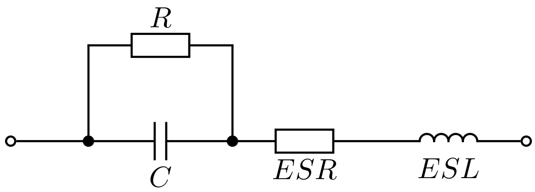


Abb. 2.47 Ersatzschaltbild eines reellen Kondensators.

Der Kondensator C entlädt sich mit der Zeit über den Widerstand R. Je größer R, desto größer ist die Entladungszeit.

ESR steht für *Equivalent Series Resistance*, ESL für *Equivalent Series Inductance*. Sowohl ESR wie ESL sollten möglichst klein sein. Der ESR wird durch den Leitungswiderstand in den Elektroden (oder auch im Elektrolyt) und Anschlüssen verursacht. Die Induktivität ESL wird im Wesentlichen durch die Zuleitungen bewirkt. Typische Werte für ESR sind $0,01 - 0,1\,\Omega$ für Folienkondensatoren und $0,05 - 1\,\Omega$ für Elkos.

Das **Frequenzverhalten** reeller Kondensatoren weicht von dem idealen Verlauf $-\mathrm{j}\frac{1}{\omega C}$ deutlich ab (Abb. 2.48), da

$$\underline{Z} = \frac{1}{\mathrm{j}\,\omega\,C} + \mathrm{j}\,\omega\,L + ESR$$

.

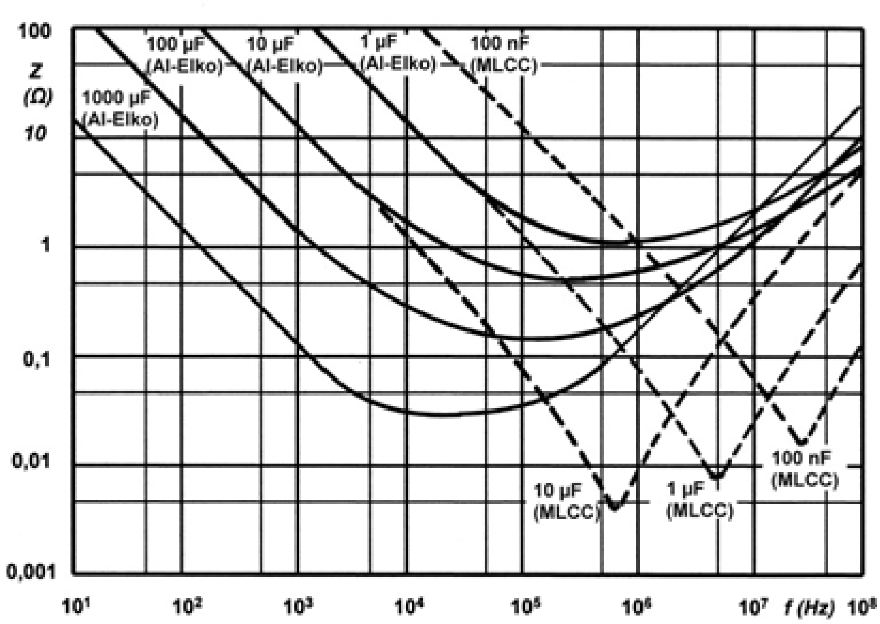


Abb. 2.48 Frequenzverhalten von Kondensatoren in doppeltlogarithmischer Auftragung (Quelle: Wikipedia).

Man sieht deutlich den Einfluss der induktiven (ESL) und resistiven (ESR) Komponenten. Für kleine ω dominiert der kapazitive Term $\frac{1}{\omega C}$, für hohe Frequenzen der induktive Term ωL. Das Frequenzverhalten eines Kondensators ist oft wichtiger als der exakte Wert seiner Kapazität.

Güte und Verlustfaktor

Die Güte Q eines Kondensators ist als das Verhältnis der Impedanz $\frac{1}{\omega C}$ zu ESR definiert, also als $\frac{1}{\omega C \,\mathrm{ESR}}$. Man bezeichnet den Tangens des Verhältnisses von ESR zur Impedanz $\frac{1}{\omega C}$ des idealen Kondensators als Verlustfaktor [11] (Abb. 2.49). Der Verlustfaktor ist der Kehrwert der Güte. Ist der ESR klein, sind auch die Verluste klein.

$$\text{Verlustfaktor} = \tan\delta = \mathrm{ESR}\,\omega\, C \quad . \tag{2.32}$$

Für kleine Werte von ω bzw. kleine ESR gilt somit näherungsweise $\tan\delta \approx \delta$.

[11] Engl. *dissipation factor*.

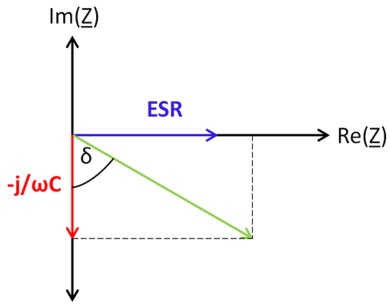


Abb. 2.49 Illustration des Verlustfaktors eines reellen Kondensators.

2.21 Reelle Spulen

Typische Spulentypen sind:

- Zylinderspule (Solenoid)
- Ringspule (Toroid)
- Drahtring
- Scheibe.

Die Zylinderspule besteht aus einem schraubenförmig gewickelten Leiter. Das Magnetfeld und damit die Induktivität der Spule werden von der Spulengeometrie und von der magnetischen Permeabilität des Mediums, das die Spule einschließt, bestimmt. Die Induktivität einer (unendlich) langen Zylinderspule ist gegeben durch

$$L = \mu_0 \, \mu_r \, N^2 \, \frac{A}{l} \quad . \tag{2.33}$$

Dabei ist μ_0 die magnetische Permeabilität des Vakuums bzw. die magnetische Feldkonstante mit $\mu_0 = 4\,\pi\,10^{-7}\,[\frac{\mathrm{H}}{\mathrm{m}}] = 12{,}566\ldots\,10^{-7}\,[\frac{\mathrm{H}}{\mathrm{m}}]$ [12], μ_r die relative magnetische Permeabilität bzw. die magnetische Feldkonstante des Spulenkerns bzw. des Mediums, das die Spule einschließt, l die Länge und A die Querschnittsfläche der Zylinderspule.

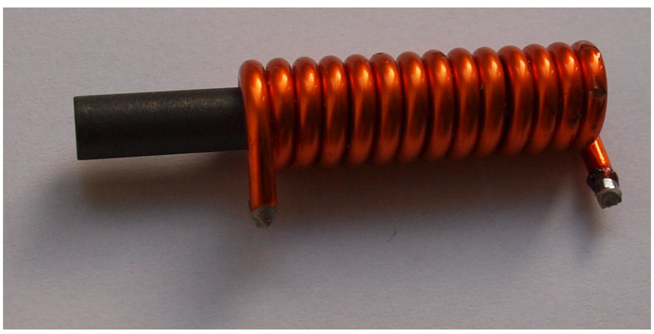


Abb. 2.50 Zylinderspule mit teilweise herausgezogenem Eisenkern.

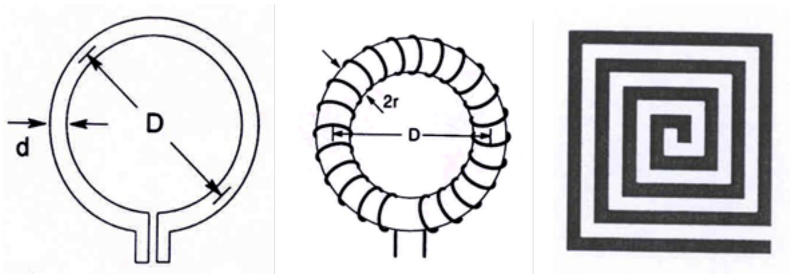
[12] In SI-Einheiten: $[\mathrm{H/m}] = 1\frac{\mathrm{kg\,m}}{\mathrm{A^2 s^2}}$

Die Induktivität L einer endlich langen Zylinderspule ergibt sich zu

$$L \approx \mu_0\,\mu_r\,N^2\,\frac{A}{l}\,\frac{1}{1+0,45\,\frac{D}{l}} \quad , \tag{2.34}$$

wobei D der Drahtdurchmesser ist.

Anmerkung: Das induzierte Magnetfeld zeigt in Richtung der Zylinderachse.



(a) (b) (c)

Abb. 2.51 (a) Ringspule. (b) Toroidpule. (c) Scheibenspule (Quelle: Michael Reisch, Elektronische Bauelemente, 2. Auflage).

Ein **Toroid (Ringkernspule)** besteht aus einem Kern mit hoher magnetischer Permeabilität, um den schraubenförmig ein Leiter gewickelt ist. Wird der Leiter von einem Gleichstrom durchflossen, bildet sich insbesondere in dem Kern ein starkes, kreisförmig geschlossenes Magnetfeld aus. Dies setzt voraus, dass der Kern keine Unterbrechungen hat. Deshalb ist eine Toroidspule schwer zu wickeln bzw. herzustellen. Die Induktivität einer Toroidspule beträgt

$$L = \mu_0\,\mu_r\,N^2\,\frac{r^2}{D}\left[1+\frac{r^2}{D^2}\right] \quad . \tag{2.35}$$

Eine sehr einfache Form einer Spule ist ein einzelner **Drahtring**, der z. B. auch auf Leiterplatten gedruckt werden kann

$$L[nH] = 2\,\pi\,\mu_0\,\mu_r\left(D-\frac{d}{2}\right)[cm]\left[\ln\left[\frac{D}{d}\sqrt{1-\frac{d}{D}}\right]+0,047\right] \quad . \tag{2.36}$$

Zwei sehr unterschiedliche Toroidspulen sind in Abb. 2.52 dargestellt.

(a) (b)

Abb. 2.52 Einfacher Ringkerntrafo zum Einsatz auf Leiterplatten (a) und Spule für ITER (b) auf dem Gelände des KIT (Campus Nord).

Scheibenspulen eignen sich besonders als gedruckte Spulen auf Leiterplatten. Für den Anschluss in der Mitte ist ein Via für die Rückleitung auf einer zweiten Leiterplattenlage nötig. Die Induktivität einer quadratischen Scheibenspule mit Kantenlänge a ist näherungsweise

$$L = \mu_0 \, \mu_r \, N^2 \, 0,85 \, a \quad . \tag{2.37}$$

Die Permeabilität einiger typischen Materialien sind in Tabelle 2.4 inklusive der Frequenzabhängigkeit aufgeführt:

Material	**relative Permeabilität** μ_r
Vakuum	1
Luft	$1 + 0,4\,10^{-6}$
Kobalt	$80 \ldots 200$
Eisen	$300 \ldots 10\,000$
Ferrite	$4 \ldots 15\,000$
μMetall (NiFe)	$50\,000 \ldots 140\,000$
Supraleiter	0

Tab. 2.4 Permeabilität inklusive Frequenzabhängigkeit (vgl. Wikipedia).

Die magnetische Permeabilität hängt vom äußeren Magnetfeld ab. Wird die Spule von einem Wechselstrom durchflossen, ist auch die Frequenzabhängigkeit von μ zu beachten.

Spulenkerne

Spulenkerne dienen zum Erhöhen der Induktivität. So wird z. B. ein Eisenkern durch das Magnetfeld magnetisiert. Dies bedeutet bei Wechselströmen, dass der Kern laufend ummagnetisiert wird. Insbesondere bei stärkeren Feldern kann bei dem Kern eine Hysterese in der Magnetisierung beobachtet werden, die Verlusten entspricht. Bei noch stärkeren Feldern geht die Magnetisierung des Kerns in die Sättigung und der magnetische Fluss wird durch den Kern nicht mehr vergrößert. Verluste treten bei Wechselströmen auch auf, indem sich durch das veränderliche Magnetfeld im Kern Wirbelströme ausbilden. Zur Reduzierung der Wirbelströme kann der Kern z. B. aus mehreren dünnen, voneinander isolierten Platten zusammengesetzt sein.

Güte und Verlustfaktor

Die Güte Q einer Spule ist als das Verhältnis $\frac{\omega L}{\mathrm{ESL}}$ definiert, also als Blindwiderstand der Spule zu Wirkwiderstand bzw. Verlustwiderstand derselbigen. Der Verlustfaktor ist analog zu dem des Kondensators als

$$\tan\delta = \frac{R}{\omega L} \tag{2.38}$$

definiert.

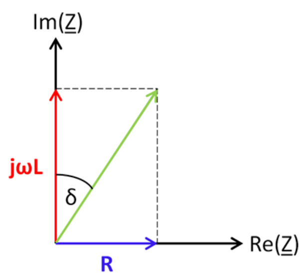


Abb. 2.53 Größen zur Berechnung der Spulengüte.

Induktivitäten

Bei sehr hohen Frequenzen können sich auch kleinste Induktivitäten auf das Schaltverhalten auswirken, z. B. solche von Drahtverbindungen (*wirebonds*) integrierter Schaltkreise. Eine Drahtverbindung hat je nach Dicke eine Induktivität/Länge von $0{,}5 \ldots 1\,\frac{\mathrm{nH}}{\mathrm{mm}}$ [3]. Geringere Induktivitäten treten bei der Flip-chip Montage auf. Hier liegen die (unerwünschten) Induktivitäten nur noch im Bereich von pH bis 100 fH [4]. Dies ist in der Hochfrequenztechnik und für das Rauschverhalten von ladungsempfindlichen Vorverstärkern von Bedeutung, bei denen das Rauschen proportional zur Eingangskapazität zunimmt.

Ersatzschaltbild einer Spule

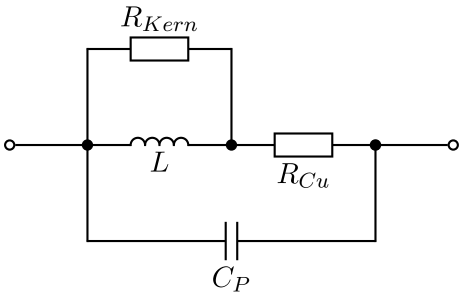


Abb. 2.54 Ersatzschaltbild einer reellen Spule.

Jede Spule hat auch einen ohmschen Widerstand, der mit der Zahl der Windungen N bzw. der Länge l des gewickelten Leiters zunimmt.
Dieser Widerstand ist in Abbildung 2.54 mit R_{Cu} bezeichnet. Er kann durch Anlegen von Gleichspannung an die Spule leicht bestimmt werden. Bei Wechselspannung treten zusätzliche Effekte auf, die in R_{Kern} und C_P zusammengefasst sind. R_{Kern} fasst Verluste wie das Ummagnetisieren des Kerns und Wirbelströme im Kern zusammen. C_P wird zum Beispiel bei einer Zylinderspule durch kapazitive Kopplung zwischen Windungen verursacht und macht sich bei höheren Frequenzen bemerkbar.

3 Diode

3.1 Diode

Die Diode, deren Schaltsymbol in Abbildung 3.1 dargestellt ist, ist ein Zweipol.

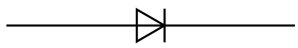


Abb. 3.1 Schaltsymbol einer Diode.

Im Gegensatz zu Widerständen, Kondensatoren und Spulen ist die Diode ein **nichtlineares** Bauelement. Der Strom durch die Diode und die an der Diode anliegende Spannung sind nicht proportional zueinander. Die Diode ähnelt in ihrem Verhalten einem Ventil. Abhängig von der Polarität der angelegten Spannung ist die Diode durchlässig oder nahezu undurchlässig für Strom.

Die Strom-Spannung-Kennlinie der Diode wird durch die Dioden- oder Shockley-Gleichung

$$I = I_S \left[\mathrm{e}^{\frac{q_e\, U}{k\, T}} - 1 \right] = I_S \left[\mathrm{e}^{\frac{U}{U_T}} - 1 \right] \tag{3.1}$$

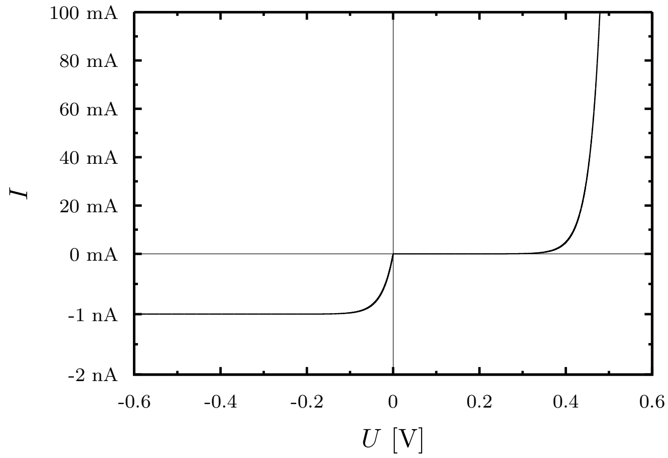
beschrieben. Die Spannung $\frac{k\,T}{q_e} = U_T$ beträgt bei Raumtemperatur ($T = 300\,\mathrm{K}$) rund 26 mV, I_S ist der Sperrstrom, U die externe Spannung, q_e ist die Elementarladung und k die Boltzmannkonstante. Die Shockley-Gleichung drückt die generische Strom-Spannungskennlinie einer Diode aus, wird aber der Vielfalt der möglichen Dioden nicht gerecht. Deshalb führt man einen Parameter n ein und modifiziert Gl. 3.1 zu

$$I = I_S \left[\mathrm{e}^{\frac{U}{n U_T}} - 1 \right] \tag{3.2}$$

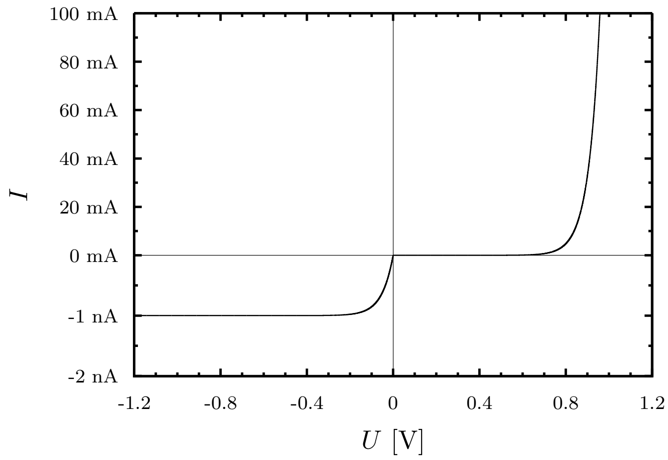
Der Parameter n nimmt Werte zwischen 1 und 2 an.

Die Shockley-Gleichung ist in Abbildung 3.2 für zwei Werte von n graphisch dargestellt. Der Sperrstrom ist um viele Größenordnungen kleiner als der Durchlassstrom. Damit der Sperrstrom in einem Strom-Spannungsdiagramm überhaupt sichtbar wird, werden die negative und positive Ordinate meistens unterschiedlich dimensioniert, z. B. [mA] auf der positiven Achse und [nA] auf der negativen Achse.

Vorsicht: Später diskutieren wir noch das Verhalten bei sehr großen negativen Spannungen, das die Shockley-Gleichung nicht beschreibt.



(a)



(b)

Abb. 3.2 (a) Diodenkennlinie entsprechend der Shockley-Gleichung für n=1. Durchlassstrom in [mA], Sperrstrom in [nA]. (b) Diodenkennlinie für n=2.

Anmerkung für die Vorlesung: Vergleiche Abbildung 3.2 mit dem Datenblatt einer Diode, z. B. $TS4448RZ150\,\mathrm{mW}$ *High Speed SMD Switching.*

Was bedeutet dieser Kennlinienverlauf?

Es gibt zwei Lesarten:

1. Betreibt man die Diode mit einem nennenswerten festen Strom in Durchlassrichtung, z.B. bei 1/10 des zulässigen Maximalstroms, so fällt fast unabhängig vom Betrag des Stroms an der Diode die sogenannte Schwellenspannung U_S ab. Für Silizium-pn-Dioden ist $U_S \approx 0,7\,\mathrm{V}$.

2. Legt man eine feste positive Spannung größer der Schwellenspannung an, so fließt ein großer Strom durch die Diode. In einer realistischen Schaltung wird der Strom wahrscheinlich nicht durch die Diode, sondern eher durch andere Bauelemente begrenzt werden, z. B. durch die externe Spannungsquelle, Lastwiderstände, Vorwiderstände etc.

Die wichtigsten Funktionen von Dioden sind in der Praxis folgende:

- Gleichrichtung, also der Umwandlung von Wechselspannung in Gleichspannung
- Spannungsvervielfachung, z. B. als Hochspannungsgenerator oder zur Verdopplung von Gleichspannungen
- Spannungsreglung (Zenerdioden)
- Spannungsbegrenzung (Schutzdioden)

Außerdem dienen Dioden als Lichtquelle (Leuchtdiode, Laserdiode) und Optokoppler, als Lichtdetektor (Photodiode, PIN-Diode, Avalanche-Diode, Solarzelle), steuerbarer Kondensator für Schwingkreise, und „last, but not least“ als Detektor für geladene Teilchen.

Dioden sind also sehr vielseitige, aber auch komplexe Bausteine.

3.2 Halbleitertechnologie für Physiker

Man kann elektronische Schaltungen mit Dioden und Transistoren (Kapitel 5) entwerfen, ohne viel über Halbleitereigenschaften oder die Physik von Halbleiterbausteinen zu wissen. Die Kenntnis der Kennlinien einer Diode oder eines Transistors und einiger anderer Eigenschaften aus dem Datenblatt des Herstellers führt schon recht weit.

Aber es ist natürlich hilfreich und interessant, die Vorgänge in einem Halbleiter zu kennen, ein grundlegendes Verständnis für die Funktionsweise von Diode und Transistor zu entwickeln und mit elementaren Fertigungstechniken vertraut zu sein. Trotz der immer noch exponentiell ansteigenden Transistordichte gibt es physikalische Grenzen der Leistungsfähigkeit moderner Bausteine. Es ist insbesondere für Physiker wichtig, ein Gefühl für das Mögliche zu entwickeln.

Anspruchsvolle Elektronikprojekte sind oft auf viele Jahre angelegt. Die Entwicklung der Elektronik für die Teilchendetektoren am Large Hadron Collider LHC in Genf hat rund zehn Jahre gedauert. Auch industrielle Projekte, wie die Entwicklung eines neuen Prozessors, dauern viele Jahre. Eine der großen Schwierigkeiten der Elektronikentwicklung bei LHC bestand darin, die richtige Transistortechnologie auszuwählen und zu verstehen, welche Verbesserungen eine modernere Transistorgeneration ermöglicht. Nun ist das nicht ganz einfach, und die Experimente und Gruppen des LHC haben sehr unterschiedliche Ansätze gewählt. Nicht alle sind gleich gut!

3.3 Halbleiter im thermischen Gleichgewicht

Halbleiter leiten Strom deutlich schlechter als die besten metallischen Leiter und deutlich besser als gute Isolatoren. Der spezifische Widerstand einiger wichtiger Materialien ist in Tabelle 3.1 zusammengestellt. Typische Halbleitermaterialien sind z.B. Silizium, Germanium und Kohlenstoff (Diamant), alles Elemente der IV. Hauptgruppe. Relevant sind auch Verbindungshalbleiter wie GaAs, InSb, SiC, GaN, HgCdTe, CdZnTe, etc.

Das wichtigste Halbleitermaterial der Elektronik ist aber bei weitem Silizium. Silizium ist ein sehr häufiges und daher billiges Material. Es läßt sich außerdem sehr gut verarbeiten. Die Zahl der Siliziumatome in einem Volumen von einem Kubikzentimeter ist $5 \cdot 10^{22}$. Die Reinheit von Si-Wafern in der Chipproduktion beträgt typischerweise 10^{-10} (oder 99,9999999%, „9 Neunen“). Die Reinheit von detektortauglichem Silizium ist noch größer.

Der spezifische Widerstand variiert nicht nur sehr stark zwischen unterschiedlichen Halbleitern, sondern ist für alle Halbleiter stark temperaturabhängig. Durch thermische Anregung werden Elektronen aus dem Valenzband in das Leitungsband angehoben und können sich dort frei bewegen. Die dazu benötigte thermische Energie hängt von der Bandlücke ab. Durch das gezielte Einbringen von Fremdatomen vor allem der III. oder der V. Hauptgruppe, dem sogenannten **Dotieren**, lässt sich die Leitfähigkeit von Halbleiterkristallen gezielt erhöhen. Es gibt dann zusätzliche Elektronen, die zum Aufbau des Kristallgitters nicht benötigt werden, oder freie Gitterplätze, die sogenannten Löcher, die als „Sprungbrett“ für den Elektronentransport dienen. Im Bändermodel spricht man von mit Elektronen gefüllten Energiezuständen in der Nähe des Leitungsbands bzw. von Lochzuständen in der Nähe des Valenzbands. Typisch sind Abstände von einigen Dutzend meV.

Häufige Dotierungsmaterialen sind Bor (B, Z=5), Aluminium (Al, Z=13), Gallium (Ga, Z=31) oder Indium (In, Z=49) aus der dritten Hauptgruppe und Phosphor (P, Z=15), Arsen (As, Z=33) oder Antimon (Sb, Z=51) aus der 5. Hauptgruppe. Typische Dotierungsdichten variieren von $10^{12} - 10^{18}$ Fremdatomen/cm^3. Die Dichte der freien Ladungsträger in intrinsischen (reinem) Silizium aufgrund der thermischen Eigenleitung beträgt dagegen bei Raumtemperatur nur rund 10^{10} Elektronen/cm^3.

Einen Halbleiter mit einem Überschuss an Elektronen aufgrund der Dotierung nennt man n-dotiert, einen mit einem Überschuss von Löchern p-dotiert. Abhängig vom Grad der Dotierung unterscheidet man n$^-$-dotierte, n-dotierte, n$^+$- und n^{++}-dotierte Halbleiter. Dabei steht n$^-$ für schwach dotierte Halbleiter mit einer Dotierungsdichte von $10^{13} - 10^{15}$ Fremdatomen/cm^3, n steht für $10^{15} - 10^{18}$, n$^+$ für $10^{18} - 10^{20}$ und n^{++} für noch extremere Werte. Selbst in einem schwach dotierten Halbleiter ist die Eigenleitung nahezu irrelevant im Vergleich zu der Leitung aufgrund der Donatoren oder Akzeptoren, so dass die Dotierung den spezifischen Widerstand des Halbleiters deutlich verändert. Eine schwache Dotierung entspricht einem spezifischen Widerstand von einigen kΩ. So dotiertes Silizium wird zur Herstellung von Siliziumdioden zur Teilchendetektion verwendet. Hier kommt rund ein Fremdatom auf 10^7 Siliziumatome.

Die Zahl der Elektronen n und Löcher p ist durch die Zahl der (möglichen) Zustände im Leitungsband N_C und Valenzband N_V und der Besetzungswahrscheinlichkeit der Zustände bestimmt. Die Besetzungswahrscheinlichkeit $F(E)$ ist temperatur- und energieabhängig und durch die Fermiverteilung gegeben

$$F(E) = \frac{1}{1 + \exp(\frac{E-E_F}{kT})} \quad , \tag{3.3}$$

E_F ist die Fermienergie, bei der die Besetzungswahrscheinlichkeit $F(E_F) = 50\%$. Für intrinsische Halbleiter gilt in guter Näherung $E_F = (E_C + E_V)/2$. Die Fermieenergie liegt in der Mitte der Bandlücke. Falls $E - E_F \gg kT$, kann man $F(E)$ durch die Boltzmannverteilung

$$F(E) \approx \exp(-\frac{E - E_F}{kT}) \tag{3.4}$$

annähern. Für Silizium mit einer Bandlücke E_G von 1.12 eV ist diese Bedingung bei Raumtemperatur ($U_T = 26$ mV) sehr gut erfüllt.

Mit Gleichung 3.4 gilt

$$n = N_C\, e^{-\frac{E_C - E_F}{kT}} \quad \text{und} \quad p = N_V\, e^{-\frac{E_F - E_V}{kT}} \tag{3.5}$$

In einem intrinsischen, also reinen bzw. nicht dotierten Halbleiter, entspricht die Zahl der Löcher genau der Zahl der freien Elektronen $n = p$. Daher ist es hilfreich, die sogenannte intrinsische Ladungsträgerdichte n_i zu definieren. Es gilt

$$n_i = \sqrt{N_C\, N_V}\, e^{-\frac{E_G}{kT}} \tag{3.6}$$

wobei E_G die Breite der Bandlücke ist: $E_G = E_C - E_V$.

Damit gilt auch

$$n = n_i\, e^{\frac{E_F - E}{kT}} \quad , \quad p = n_i\, e^{\frac{E - E_F}{kT}} \tag{3.7}$$

und

$$np = {n_i}^2 \tag{3.8}$$

Die Gleichung 3.8 gilt auch für dotierte Halbleiter und wird als Massenwirkungesetz[1] bezeichnet. In n-dotierten Silizium ist damit die Dichte der Löcher p gegeben durch $p = n_i^2/N_D$, N_D ist die Dichte der Donatoratome. Analog ist in p-dotierten Silizium damit die Dichte der Elektronen im Leitungsband n durch $n = n_i^2/N_A$ gegeben; N_A ist die Dichte der Akzeptoratome.

Material	$\rho\,[\Omega\,\mathbf{cm}]$
Ag	$1,6\,10^{-6}$
Cu	$1,7\,10^{-6}$
Au	$2,2\,10^{-6}$
Al	$2,9\,10^{-6}$
Ge	$2\,10^{-4}\ldots 20$
Si	$2\,10^{-4}\ldots 10^{4}$
GaAs	$2\,10^{-4}\ldots 2\,10^{8}$
Kohlenstoff	$3,5\,10^{-3}$
Epoxidharz	$10^{10}\ldots 10^{17}$
Kapton	$10^{13}\ldots 10^{18}$
SiO_2	$10^{14}\ldots 10^{16}$
Teflon	10^{18}
Porzellan	10^{14}
Glas	$10^{12}\ldots 10^{17}$

Tab. 3.1 Spezifischer Widerstand einiger wichtigen Leiter, Halbleiter und Isolatoren.

[1] Engl. *mass action law*

Die Bandlücke E_G und intrinsischen Ladungsträgerdichten für einige wichtige Halbleiter sind in Tabelle zusammengestellt.

	Si	Ge	GaAs	Diamant
E_G [eV]	$1,12$	$0,67$	$1,42$	$5,5$
n_i [cm^{-3}]	$1,1\,10^{10}$	$2,4\,10^{13}$	$1,8\,10^{6}$	$< 10^3$

Tab. 3.2 Bandbreite E_G und intrinsische Ladungsträgerdichte von Silizium, Germanium, Galliumarsenid und Diamant bei Raumtemperatur.

3.4 Funktionsweise einer Diode*

Beim Zusammenfügen zweier unterschiedlich dotierter, elektrisch neutraler Hableiter passiert folgendes (siehe Abb. 3.3 - 3.5):

1. Die jeweiligen Majoritätsladungsträger diffundieren auf die andere Seite.

   Elektronen diffundieren aus dem n-dotierten Halbleiter in den p-dotierten Halbleiter. Löcher diffundieren aus dem p-dotierten Halbleiter in den n-dotierten Halbleiter.

2. Durch die Dotierung ändert sich die elektrische Ladung des Halbleiters nicht. Die zugefügten Donator- und Akzeptoratome sind ja jeweils neutral. Durch die Diffusion entsteht aber ein räumliches Ladungsungleichgewicht, und es baut sich eine elektrische Spannung zwischen den unterschiedlich dotierten Halbleitern auf.

   Jetzt sind weniger Elektronen und mehr Löcher auf der n-dotierten Seite als vor dem Diffusionsvorgang. Entsprechend lädt sich dieser Bereich positiv auf, der p-dotierte Bereich lädt sich negativ auf. Das so verursachte elektrische Feld $\mathscr{E}$ wirkt der Diffusionsbewegung entgegen, da es die Elektronen in die der Diffusion entgegengesetzte Richtung driften lässt. Dadurch stellt sich eine Gleichgewicht und eine stabile Spannung, die Diffusionsspannung U_D, zwischen den beiden Bereichen ein. Die Diffusionsbewegung wird kompensiert.

3. Die auf die p-dotierte Seite diffundierten Elektronen füllen die Löcher. Man sagt, sie rekombinieren mit den Löchern. Dies führt zu einer Verarmung von freien Ladungsträgern. Auch auf der n-dotierten Seite führt die Diffusionsbewegung der Elektronen zu einem Mangel an freien Ladungsträgern. Man bezeichnet den Übergangsbereich deshalb als Verarmungszone oder auch als Raumladungszone (siehe Abb. 3.4 bzw. 3.5).

   Interessant wird das Verhalten bei Anlegen einer äußeren Spannung U, zusätzlich zu der Diffusionsspannung. Wichtig sind der Betrag und die Polarität des Feldes. Wirkt das Feld der Diffusion entgegen, so sperrt die Diode. Gleichzeitig verbreitert sich die Verarmungszone.

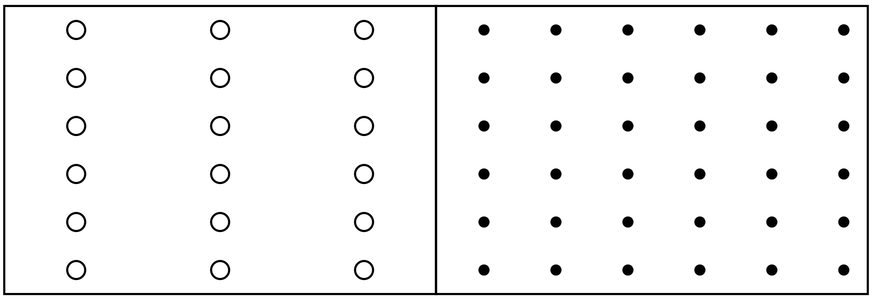


Abb. 3.3 Illustration eines pn-Übergangs „vor“ der Diffusion der freien Ladungsträger. Die offenen Kreise stellen die Löcher der p-dotierten Seite dar, die kleinen fetten Kreise die Überschusselektronen der n-dotierten Seite.

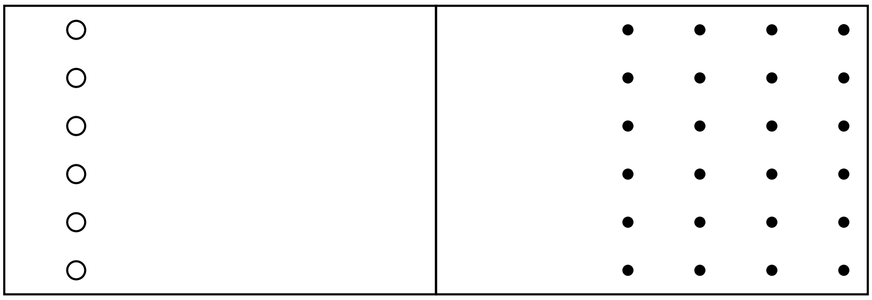


Abb. 3.4 Illustration eines pn-Übergangs nach der Diffusion der freien Ladungsträger. Die offenen Kreise stellen die Löcher der p-dotierten Seite dar, die kleinen fetten Kreise die Überschusselektronen der n-dotierten Seite. Der mittlere Bereich ohne Symbole ist die Verarmungszone.

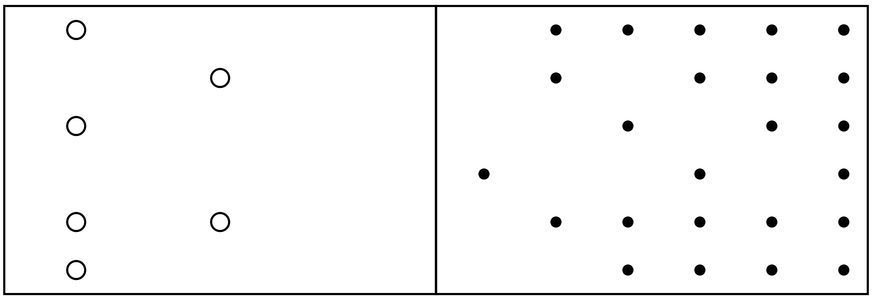


Abb. 3.5 Illustration eines pn-Übergangs nach der Diffusion der freien Ladungsträger mit einer unscharfen Grenze der Verarmungszone.

Wir wollen diese qualitativen Überlegungen quantifizieren und die Diffusionsspannung, das elektrische Feld und die Breite der Raumladungszone ausrechnen. Die Stromdichte J_{Diff} aufgrund der Diffusion von Elektronen beträgt

$$J_{Diff} = q_e \, D_n \, \frac{dn}{dx} \quad . \tag{3.9}$$

Dabei ist D_n die Diffusionskonstante, für die nach Einstein gilt:

$$D_n = \frac{kT}{q_e}\,\mu_n \quad , \tag{3.10}$$

μ_n ist die Elektronmobilität. Der Diffusionsstrom nimmt also linear mit der Temperatur, der Mobilität und der räumlichen Ladungsträgerasymmetrie dn/dx zu.

Die Stromdichte J_{Drift} aufgrund der Driftbewegung beträgt

$$J_{Drift} = q_e\,\mu_n\,n\,\mathscr{E} \quad . \tag{3.11}$$

Eine analoge Beziehung gilt für den Lochstrom. Insgesamt gilt für Elektronen und analog für Löcher:

$$J_{Diff} + J_{Drift} = q_e\,D_n\,\frac{dn}{dx} + q_e\,\mu_n\,n\,\mathscr{E} = 0 \quad . \tag{3.12}$$

Der Halbleiter ist im thermischen Gleichgewicht. Aus dieser Beziehung lässt sich die Diffusionsspannung ausrechnen. Es gilt

$$q_e\,\mu_n\,n\,\mathscr{E} = -q_e\,D_n\,\frac{dn}{dx} = -\mu_n\,kT\,\frac{dn}{dx} \quad \text{und damit} \quad \mathscr{E} = -\frac{d\phi}{dx} = -\frac{kT}{n}\,\frac{dn}{dx} \quad . \tag{3.13}$$

Durch Integration über dx und dn ergibt sich

$$U_D = \phi(x_n) - \phi(-x_p) = kT \int_{-x_p}^{x_n} \frac{dn}{n} = kT\,\ln\frac{n(x_n)}{n(-x_p)} \tag{3.14}$$

Am und jenseits des Randes x_n gilt $n(x_n) \approx N_D$ und entsprechend $n(-x_p) \approx n_i^2/N_A$. Damit folgt

$$U_D = kT\,\ln\left(\frac{N_D\,N_A}{n_i^2}\right) \quad . \tag{3.15}$$

Diese wichtige Beziehung gilt auch für $n - n^+$-Übergange, etc. Dies ist z.B. für CMOS-Pixelkameras relevant. U_D beträgt in pn-Dioden aus Silizium etwa $0,6\,\mathrm{V}$.

Aus Gl. 3.12 lässt sich auch herleiten, dass der Wert der Fermienergie im thermischen Gleichgewicht ortsunabhängig ist, also $\frac{dE_F}{dx} = 0$. Dies folgt mit Gl. 3.7 und $E = \int \mathscr{E} dx$ aus

$$\begin{aligned} J_{Diff} + J_{Drift} = q_e\,D_n\,\frac{dn}{dx} + q_e\,\mu_n\,n\,\mathscr{E} = q_e\,D_n\,\frac{n}{kT}\left(\frac{dE_F}{dx} - \frac{dE}{dx}\right) + q_e\,\mu_n\,n\,\mathscr{E} = \\ \mu_n\;n\left(\frac{dE_F}{dx} - q_e\,\mathscr{E}\right) + q_e\,\mu_n\,n\,\mathscr{E} = \mu_n\,n\,\frac{dE_F}{dx} = 0 \quad . \end{aligned} \tag{3.16}$$

Unter Annahme der in Abb. 3.6 illustrierten Näherungen ergibt sich die Feld- und Potentialverteilung (siehe auch Abb. 3.6). Wir nehmen nun an, dass die Verarmungszone scharfe Ränder (bei x_p, 0, und x_n) besitzt und dass die räumliche Ladungsverteilung homogen ist. Die Breite der Verarmungszone $d = x_p + x_n$, bzw. die Werte x_p und x_n müssen im folgenden bestimmt werden. Die Raumladung im Bereich $[x_p, 0]$ beträgt $-qN_A$, wobei N_A die Zahl der Akzeptoratome ist. Die Raumladung im Bereich $[0,\, x_n]$ beträgt qN_D, wobei N_D die Zahl der Donatoratome ist. Das elektrische Feld $\mathscr{E}$ bzw. Potential ϕ ergibt sich aus der (eindimensionalen) Poissongleichung:

$$-\frac{d^2\phi}{dx^2} = \frac{d\mathscr{E}}{dx} = \frac{\rho}{\epsilon_0\,\epsilon_r} = -\frac{q(N_D - N_A)}{\epsilon_0\,\epsilon_r} \tag{3.17}$$

Damit ergeben sich durch Integration der Ladungsdichte in den Bereichen $[x_p, x]$ und $[x, x_n]$ zwei Lösungen:

$$\mathscr{E}_-(x) = \int_{-x_p}^{x} \frac{\rho}{\epsilon_0\,\epsilon_r} dx' = \frac{-q_e}{\epsilon_0\,\epsilon_r}\left[N_A\,x'\right]_{-x_p}^{x} = -\frac{N_A(x + x_p)}{\epsilon_0\,\epsilon_r} \quad \text{und} \quad \mathscr{E}_+(x) = +\frac{N_D(x - x_n)}{\epsilon_0\,\epsilon_r} \tag{3.18}$$

Außerhalb der genannten Bereiche, also für $x < -x_p$ und $x > x_n$ gilt $E(x) = 0$. Sowohl das elektrische Feld $\mathscr{E}$ wie das Potential ϕ ist am Punkt $x = 0$ stetig. Durch Gleichsetzen der Terme in Gl. 3.18 folgt:

$$N_A\,x_p = N_D\,x_n \quad . \tag{3.19}$$

Mit der weiteren Beziehung $d^2\phi/dx^2 = -d\mathscr{E}/dx$ (siehe oben) folgt

$$\phi(x_p) = -\frac{q_e\,N_A\,x_p^2}{2\,\epsilon_0\,\epsilon_r} \quad \text{und} \quad \phi(x_n) = \frac{q_e\,N_D\,x_n^2}{2\,\epsilon_0\,\epsilon_r} \tag{3.20}$$

Die Potentialdifferenz $\phi(x_n) - \phi(-x_p)$ entspricht der Diffusionsspannung U_D. Liegt zusätzlich eine äußere Spannung U an, so gilt

$$U_D - U = \frac{q_e\,N_D\,x_n^2}{2\,\epsilon_0\,\epsilon_r} + \frac{q_e\,N_A\,x_p^2}{2\,\epsilon_0\,\epsilon_r} \tag{3.21}$$

Aus 3.19 und 3.21 folgt

$$x_n = \sqrt{\frac{2\,\epsilon_r\,\epsilon_0\,N_A}{q_e\,N_D\,(N_A + N_D)}\,(U_D - U)} \quad \text{und} \quad x_p = \sqrt{\frac{2\,\epsilon_r\,\epsilon_0\,N_D}{q_e\,N_A\,(N_A + N_D)}\,(U_D - U)} \tag{3.22}$$

Die Breite $d = x_N + x_p$ der Raumladungszone hängt damit gemäß

$$d = \sqrt{\frac{2\,\epsilon_r\,\epsilon_0}{q_e}\left(\frac{1}{N_A} + \frac{1}{N_D}\right)(U_D - U)} \tag{3.23}$$

von den Dotierungen und der angelegten Spannung U ab. Für eine negative Spannung U verbreitert sich wie oben gesagt die Raumladungszone.

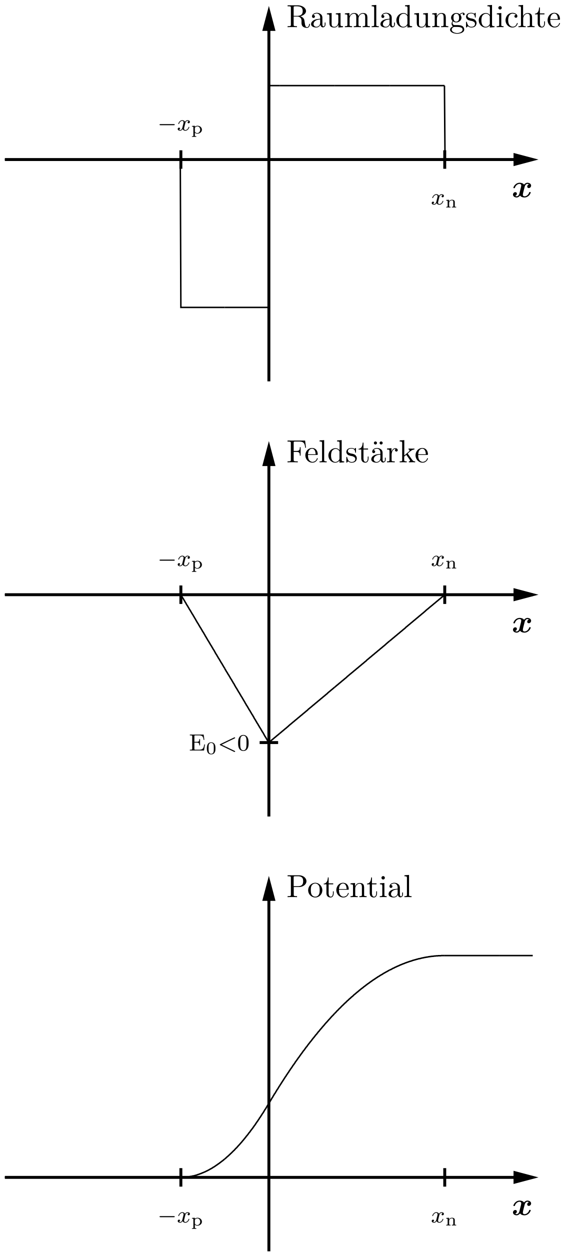


Abb. 3.6 Raumladungsdichte, Feldstärke und Potential eines pn-Übergangs.

3.5 Gleichrichterschaltungen

Eine Gleichrichterschaltung[2] wandelt Wechselspannung in Gleichspannung um. Die einfachste Gleichrichterschaltung, die Einwegschaltung, besteht aus nur einer Diode. Die Schaltung ist in Abbildung 3.7 dargestellt.

Die Diode lässt nur einen nennenswerten Stromfluss in Durchlassrichtung zu und zwar für

[2]Engl. *rectifier, AC-DC converter*

Spannungen U oberhalb der Schwellenspannung $U_S \approx 0,7\,\mathrm{V}$. Entsprechend treten am Ausgang und an der Last R_L keine negativen Spannungen auf.

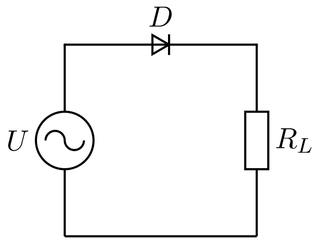


Abb. 3.7 Gleichrichterschaltung mit einer Diode.

Eine sinusförmige Eingangsspannung $U = U_0 \sin \omega\, t$ und die entsprechende Ausgangsspannung sind in Abbildung 3.8 dargestellt.

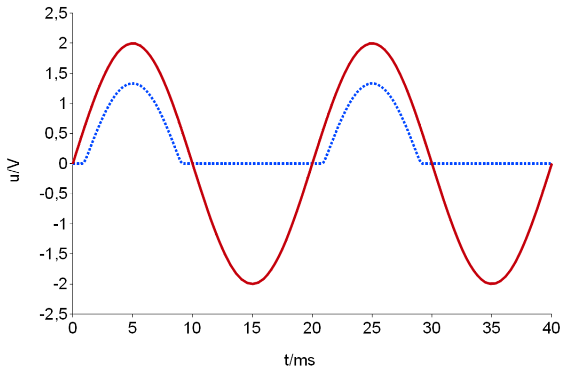


Abb. 3.8 Gleichrichterschaltung: Eingangsspannung (rot) und Ausgangsspannung (blau, gestrichelt).

Diese elementare Schaltung hat eine Reihe von Unzulänglichkeiten. Zum einen ist die Ausgangsspannung nicht konstant, sondern in Durchlassrichtung noch sinusförmig. Die Gleichrichtung ist auch recht ineffizient, da die Energie der gesamten negativen Halbwelle und eines kleinen Teiles der positiven Halbwelle nicht genutzt wird [3].

Anmerkung: Abbildung 3.8 gilt nur für nicht zu hohe Frequenzen. Beim sehr schnellen Wechsel der Stromrichtung kommt die Diode nicht mit und lässt eine kurze negative Spannung durch, bevor sie sperrt. Außerdem stellen die Diodensperrschicht, das Gehäuse und die Zuleitungen eine parallel geschaltete (nichtlineare) Kapazität dar.

Mit einem **Kondensator** C parallel zum Lastwiderstand R_L (Abb. 3.9) kann man die Ausgangsspannung glätten.

[3] Der Nutzungsgrad der Schaltung, das heißt die Leistung am Ausgang der Schaltung normiert auf die Eingangsleistung, beträgt ungefähr

$$\frac{1}{2}\left[\frac{U_0 - U_S}{U_0}\right]\frac{\pi - 2\,\alpha}{\pi} \quad \text{mit} \quad \alpha = \arcsin\frac{U_S}{U_0} \quad .$$

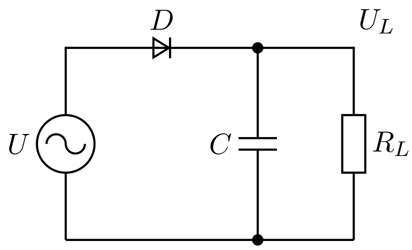


Abb. 3.9 Gleichrichterschaltung mit Diode, Kondensator und Lastwiderstand.

Wie funktioniert diese Schaltung?

Der Kondensator wird durch die Diode in einem oder mehreren Wechselspannungszyklen auf die Spitzenspannung $U_0 - U_S$ aufgeladen. Gleichzeitig wird der Lastwiderstand R_L mit Spannung und Strom versorgt. Sobald die Eingangsspannung unter den Wert $U_C + U_S$ sinkt, sperrt die Diode. Die Last ist jetzt von der Spannungsquelle isoliert. Jetzt liefert der Kondensator den Laststrom. Dabei entlädt sich der Kondensator mit der Zeitkonstanten $\tau = R_L\, C$, und die Spannung nimmt exponentiell mit $U_C = U_0\, \mathrm{e}^{-\frac{t}{R_L\, C}} \approx U_0 \left[1 - \frac{t}{R_L\, C}\right]$ ab (Abb. 3.10). Im nächsten Zyklus wird der Kondensator dann wieder auf die volle Spannung aufgeladen usw.

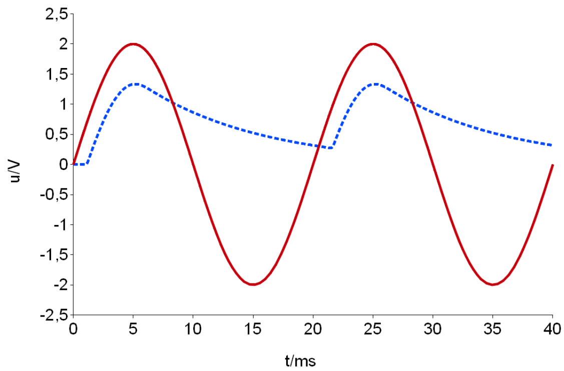


Abb. 3.10 Eingangs- und Ausgangsspannung der Gleichrichterschaltung aus Abb. 3.9.

Wenn man die Spannung möglichst konstant halten will, ist es wichtig, dass die Entladungszeitkonstante τ deutlich größer als die Periodendauer T der Eingangswechselspannung ist, also $\tau \gg T = \frac{1}{f}$.

Zahlenbeispiel: Die Frequenz f sei 50 Hz, die Periodendauer T also 20 ms. Der Lastwiderstand R_L sei 100 Ω. Hier ist ein Kondensator mit einer Kapazität von 1 mF ausreichend. Mit $\tau = 100\,\Omega \cdot 1\,\mathrm{mF} = 100\,\mathrm{ms}$ gilt $\tau \gg T = 20\,\mathrm{ms}$.

Die maximale Schwankung der Ausgangsspannung U_C wird auch Brummspannung U_{BS} genannt. Die Brummspannung sollte möglichst klein sein. Es gilt

$$U_{BS} = \frac{\Delta Q}{C} = \frac{1}{C} \int I_L\, dt \approx I_L\, \frac{T}{C} = \frac{I_L}{C\, f} \quad . \tag{3.24}$$

Eine große Kapazität ist also besser als eine kleine. Kleine Lastströme und Periodendauern auch, aber diese Größen sind durch die Umstände vorgegeben.

Die korrekte Wahl der Diode ist auch wichtig. Die Diode muss dauerhaft eine maximale Sperrspannung von rund $2\,U_0 - U_S - \frac{U_{BS}}{2}$ aushalten können.

Das Aufladen des Kondensators ist schnell im Vergleich zum Entladen, da hier die relevante Zeitkonstante $R_i\,C$ durch den kleinen Innenwiderstand R_i der Spannungsquelle gegeben ist. Dabei treten kurzfristig hohe Ströme auf. Bei der Wahl einer Diode muss also auch die Abwärme berücksichtigt werden.

Grätzschaltung

Der Nachteil der Schaltung 3.7 ist, dass die Leistung der negativen Halbwelle nicht genutzt wird. Die Lastleistung stammt nur aus dem Kondensator, die Spannungsquelle ist durch die Diode in Sperrrichtung von der Last getrennt. Dies vermeidet man mit der Grätzschaltung (Abb. 3.11). Hier kommen vier Dioden zum Einsatz. Die Grätzschaltung ist ein sog. „Vollweggleichrichter". Die obige Schaltung dagegen „Halbweggleichrichter".

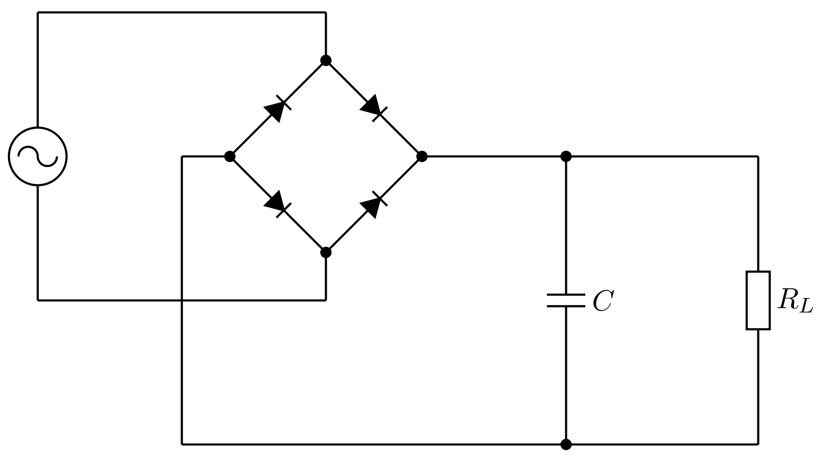


Abb. 3.11 Gleichrichtung mittels einer Grätzschaltung.

Der Spannungsverlauf einer Grätzschaltung ist in Abb. 3.12 dargestellt. Hier ist die Brummspannung geringer, da das Aufladen doppelt so häufig erfolgt.

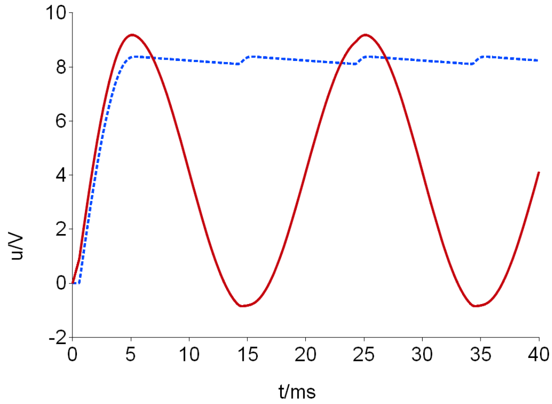


Abb. 3.12 Eingangs- und Ausgangsspannung der Gleichrichterschaltung aus Abb. 3.9.

Halbbrücke und Vollbrücke*

Die Grätzschaltung ist eine spezielle Realisierung der allgemeinen Brückenschaltung. Die bekannte Wheatstonesche Brücke zur Präzisionsbestimmung von Widerständen ist in Abb. 3.13 zu sehen.

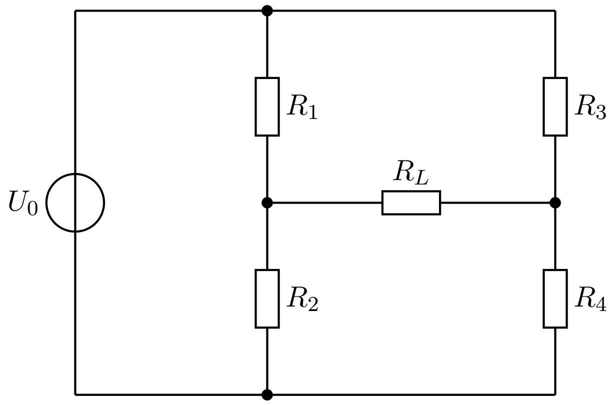


Abb. 3.13 Vollbrücke aus Widerständen, Spannungsquelle und Last R_L.

Diese Brücke besteht aus einer Parallelschaltung von zwei Spannungsteilern (R_1/R_2 und R_3/R_4). Hier hängt nicht nur der Betrag der Lastspannung von den Werten der Widerstände ab, sondern auch das Vorzeichen der Spannung und damit die Stromrichtung. Die sieht man leicht in Abb. 3.14, in der die Widerstände durch Schalter ersetzt sind. Dieses Arrangement wird als Vollbrücke oder suggestiv auch als H-Brücke bezeichnet.

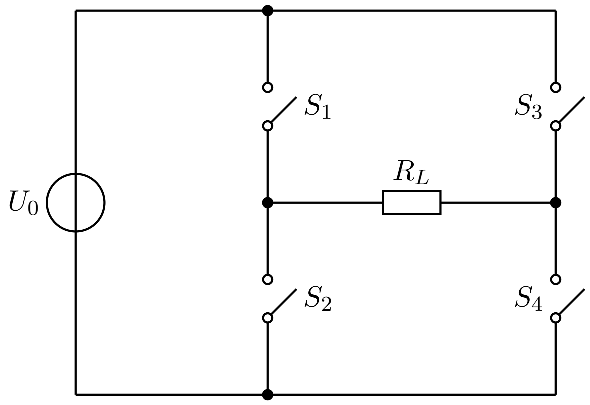


Abb. 3.14 Vollbrücke aus Widerständen und Spannungsquelle.

Sind die Schalter S_1 und S_4 geschlossen und S_2 und S_3 offen, liegt U_0 am Widerstand R_L an. Sind dagegen S_2 und S_3 geschlossen und S_1 und S_4 offen, liegt $-U_0$ an R_L an, und die Stromrichtung ändert sich. Die Schalter sind in der Regel als Dioden wie in der Grätzschaltung oder als Transistoren realisiert.

Als Halbbrücke bezeichnet man das Arrangement in Abb. 3.15.

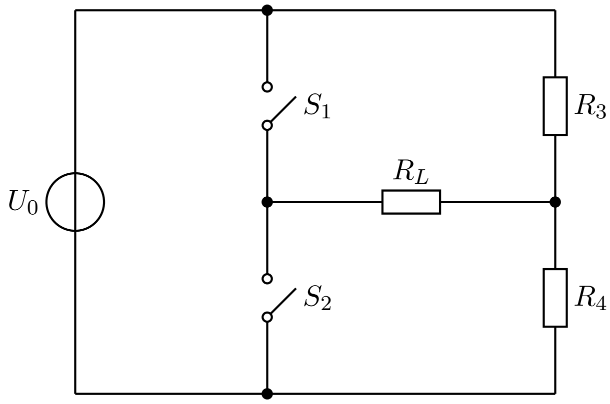


Abb. 3.15 Halbbrücke.

Hier werden entweder die Spannung $U_0 - U_0/2 = U_0/2$ an die Last R_L angelegt. (S_1 ist geschlossen und S_2 offen) oder die Spannung $0 - U_0/2 = -U_0/2$ (S_2 geschlossen und S_1 offen). Halbbrücken werden u. a. in Spannungswandlern (DC-DC und AC-DC) eingesetzt.

3.6 Schutzdiode

Dioden werden oft eingesetzt, um Überspannungen und damit verbundene Schäden an Bauteilen zu verhindern. Die Bauteile können Transistoren, integrierte Schaltkreise (ICs), Schalter, Spulen, Elektromotoren, supraleitende Magnete und vieles mehr sein. Ein einfaches Beispiel ist in Abbildung 3.16 gezeigt.

Beim Öffnen des Schalters S würde ohne die Diode D durch Selbstinduktion (siehe Spule) eine hohe Spannung $U = L\,dI/dt$ auftreten. Der positive Pol der Spannung liegt auf der Schalterseite.

Die Selbstinduktion der Spule versucht, den Stromfluss aufrecht zu erhalten. Die Diode arbeitet als sogenannte Freilaufdiode, begrenzt die Selbstinduktionsspannung auf etwa $0,7\,\mathrm{V}$ und stellt während des Abklingens der Selbstinduktion einen Pfad für den durch die Spule erzwungenen Strom zur Verfügung.

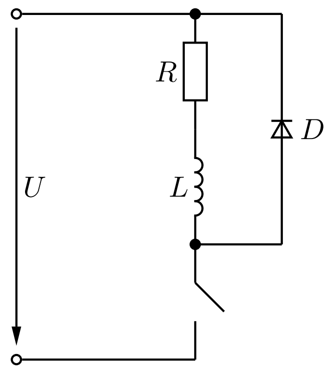


Abb. 3.16 Schutzdiode.

Die magnetische Energie wird sowohl im Vorwiderstand R, im ohmschen Widerstand der Spule als auch im statischen Innenwiderstand der Diode in Wärme umgewandelt.

Auch ein sogenannter Snubber, ein Widerstand und ein Kondensator in Reihenschaltung, anstelle der Diode ist als Schutz beliebt.

Bei Wechselspannungsbetrieb sind Spannungsspitzen beider Polarität zu unterdrücken. Dies geschieht wie in Abbildung 3.17 gezeigt durch zwei antiserielle Zenerdioden (Abs. 3.12).

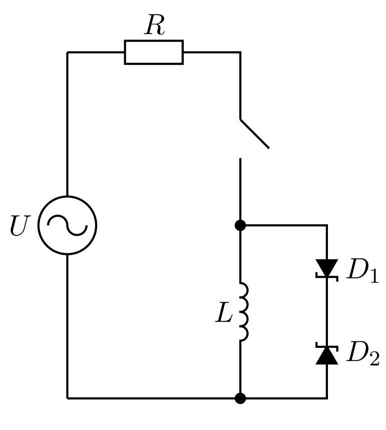


Abb. 3.17 Antiserielle Schutzdioden.

Diese Schaltung ist durchaus raffiniert, da die Zenerdiode ja sowohl bei positiven als auch bei negativen Spannungen leiten kann. Sie leitet aber kaum bei kleinen negativen oder bei kleinen positiven Spannungen. Im Normalbetrieb bei geschlossenem Schalter S fließt ein konstanter

Strom durch die Spule. Die Spulenspannung $U_L \approx 0$ V, so dass im Wesentlichen kein Strom durch die Dioden fließt.

Einen einfachen Schutz der hochohmigen Eingänge von integrierten Schaltkreisen gegen Spannungsspitzen (z. B. durch elektrostatische Überschläge bei Berühren mit der Hand) bietet die Schaltung in Abbildung 3.18. Spannungen größer als $U_{IC} = +0,7\,\mathrm{V}$ und kleiner als $-0,7\,\mathrm{V}$ werden vom Eingang des ICs weggelenkt.

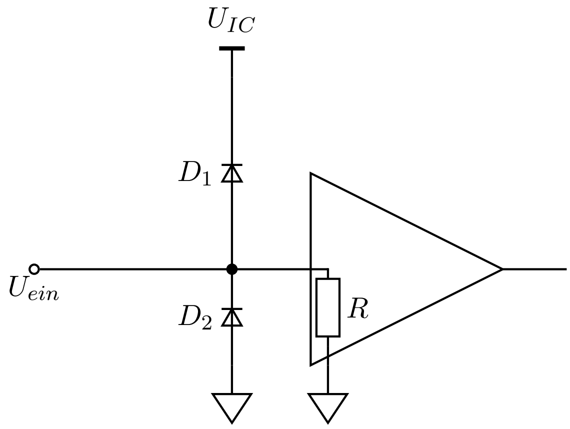


Abb. 3.18 Schutzdioden am Eingang von ICs.

3.7 Spannungsverdoppler

Die einfachste Schaltung zur Verdopplung einer Wechselspannung besteht aus einem Kondensator und einer Diode. Dies ist die Villard-Schaltung.

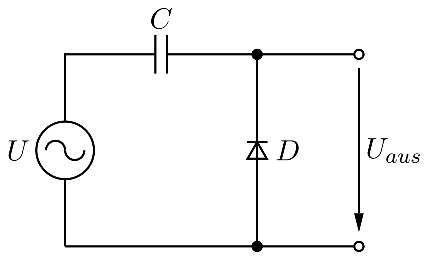


Abb. 3.19 Spannungsverdopplung mit einer Villard-Schaltung.

Der Kondensator wird in den negativen Halbschwingungen der Wechselspannung $U_{ein} = U_0 \sin(\omega\, t)$ auf $U_0 - U_S$ aufgeladen, allerdings nicht unbedingt in nur einer Halbschwingung auf seine Maximalspannung $U_0 - U_S$. Der positive Pol des Kondensators liegt dabei an der Diodenseite. Während der positiven Halbschwingungen sperrt die Diode und am Ausgang liegt der aktuelle Wert der Eingangswechselspannung plus die Kondensatorspannung an, da die Spannungsquelle und der Kondensator in Reihe geschaltet sind. Nach einer kurzen Einschwingzeit schwingt die Ausgangsspannung zwischen 0 und $2\,U_0 - U_S$ (Abb. 3.20). Die Einschwingzeit wird durch den Innenwiderstand der Spannungsquelle und die Kapazität des Kondensators bestimmt.

Die Villard-Schaltung verdoppelt den maximalen Spannungswert am Ausgang, indem sie den Wert der Ausgangsspannung um $U_0 - U_S$ gegenüber der Eingangsspannung verschiebt, $U_{aus}(t) = U_{ein}(t) + U_0 - U_S$. Es handelt sich trotzdem noch um eine harmonische Schwingung.

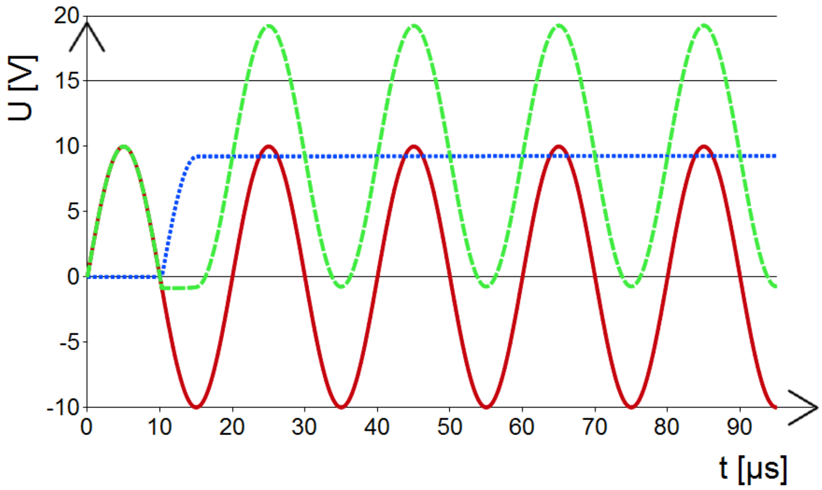


Abb. 3.20 Eingangsspannung *(rot, durchgezogen)* und Ausgangsspannung *(grün, gestrichelt)* der Villard-Schaltung (Abb. 3.19). Die blaue gepunktete Kurve entspricht der Spannung, die am Kondensator anliegt. In dieser Simulation beträgt der Innenwiderstand der Spannungsquelle 0 Ω.

Durch Hinzufügen einer weiteren Diode und eines Glättungskondensators, wie in Abbildung 3.9, kann man eine relativ konstante Ausgangsspannung erzeugen. Diese Schaltung ist die sogenannte Greinacher-Schaltung (Abb. 3.21).

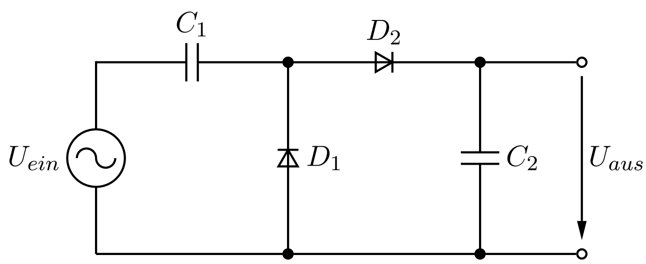


Abb. 3.21 Spannungsverdopplung mit Greinacher-Schaltung.

Die Brummspannung der belasteten Schaltung entspricht der der Gleichrichterschaltung $U_{BS} = I_{aus}/f\, C_2$. Die Leerlaufspannung der unbelasteten Schaltung ist $U_0 - 2\,U_S$.

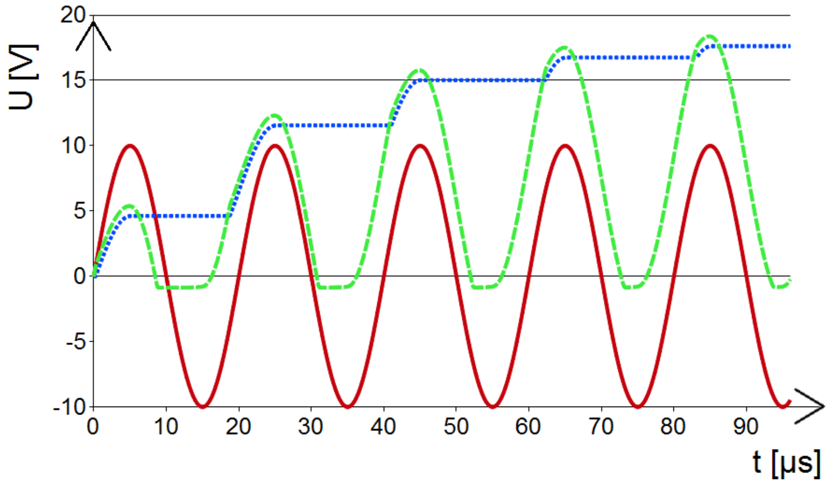


Abb. 3.22 Eingangsspannung *(rot, durchgezogen)* und Ausgangsspannung *(blau, gepunktet)* der Greinacher-Schaltung (Abb. 3.21). Die grüne, gestrichelte Kurve entspricht der Ausgangsspannung der dem Gleichrichter vorgeschalteten Villard-Schaltung.

Frage: Wie erklärt sich der Einschwingvorgang? Wie unterscheidet sich diese Kurve von Abb. 3.20? Warum reduziert sich der flache Bereich der grünen Kurve bei $U \approx 0$ mit der Zeit?

3.8 Kaskadenschaltung

Man kann mehrere Greinacher-Schaltungen hintereinander schalten und dadurch sehr hohe Gleichspannungen größer 100 kV erzeugen.

In Abbildung 3.23 ist eine zweistufige Kaskadenschaltung gezeigt. Jede Stufe besteht aus zwei Dioden und zwei Kondensatoren.

Die zweistufige Schaltung vervierfacht die Ausgangsspannung nahezu. (Wir vernachlässigen hier den Spannungsabfall an den Dioden.) Man kann auch noch mehr Stufen aneinander hängen und entsprechend größere Ausgangsspannungen erzeugen.

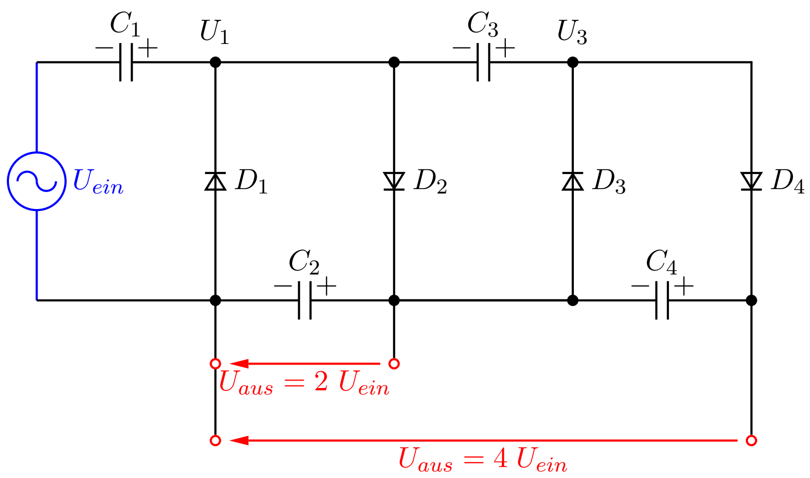


Abb. 3.23 Kaskadenschaltung.

Die Funktionsweise der Schaltung entspricht der Greinacher-Schaltung, nur sind die vielen Stufen verwirrend. Die Kondensatoren C_1 und C_3 nennt man auch Schubkondensatoren, C_2 und C_4 Glättungskondensatoren. Man versteht die Schaltung leichter, wenn man sich klar macht, dass sich einmal aufgeladene Kondensatoren aufgrund der geschickten Anordnung der Dioden nicht leicht entladen können. So ist der positive Pol des Kondensators C_1 durch die Diode D_1 von Masse isoliert. Wie bei der Greinacher-Schaltung braucht man mehrere Perioden, um alle Kondensatoren aufzuladen.

Wir gehen davon aus, dass zunächst alle Kondensatoren ungeladen sind. Also fällt an ihnen keine Spannung ab. Wir brauchen vier Schritte, um die maximale Ausgangsspannung zu erreichen. Diese Schritte sind in Tabelle 3.3 und Abbildung 3.24 dargestellt:

1. Zuerst wird der Kondensator C_1 durch die erste negative Halbwelle aufgeladen. Zur Vereinfachung gehen wir davon aus, dass eine Halbwelle ausreicht, um C_1 und später auch die anderen Kondensatoren voll aufzuladen. Die Maximalspannung von C_1 ist U_0, die Spitzenspannung der Eingangswechselspannung. (Wir vernachlässigen die Schwellenspannung der Diode). D_2 sperrt und damit ist C_2 isoliert und wird nicht aufgeladen, $U_{C_2} = 0$.

2. Die Spannung U_1 steigt während der positiven Halbwelle auf maximal $2\,U_0$ an. (Nämlich auf die positive Eingangsspannung U_0 plus der maximalen Kondensatorspannung U_0.) Zu diesem Zeitpunkt ist C_2 noch nicht geladen und liegt auf Masse, damit ist D_2 durchlässig und lädt jetzt C_2 auf maximal $2\,U_0$ auf. Da C_3 während dieser positiven Halbwelle noch nicht aufgeladen ist, sperrt D_3 und C_3 wird noch nicht aufgeladen, $U_{C_3} = 0$. Die Spannung U_{C_2} hat jetzt ihren Maximalwert $2\,U_0$ erreicht.

3. Während der nächsten negativen Halbwelle sinkt U_1 wieder auf $0\,\mathrm{V}$ ab. Damit wird D_3 durchlässig und C_3 auf $2\,U_0$ aufgeladen. Die Spannung $U_{C_3} = 2\,U_0$. Auch $U_{C_4} = 2\,U_0$, da C_4 noch leer ist. Damit sperrt D_4 und C_4 wird jetzt noch nicht aufgeladen, $U_{C_4} = 0$. (Abhängig von der Spannung an C_1 wird auch C_1 wieder nachgeladen.)

4. Erst mit der nächsten positiven Halbwelle steigt U_3 auf den Wert $4\,U_0$, D_4 wird durchlässig und C_4 wird auf $2\,U_0$ aufgeladen. C_4 kann sich nur durch den Lastwiderstand

entladen, der in Abbildung 3.23 nicht eingezeichnet ist. (Abhängig von der Spannung an C_2 wird auch C_2 wieder nachgeladen.)

Die Schaltung funktioniert nur bei kleinen Lastströmen. Erstens, da sonst die Leistung zu groß wird und zweitens, da die Ausgangsimpedanz mit zunehmender Stufenzahl steigt. (Es werden mehr Kondensatoren in Reihe geschaltet.)

Bis auf den Kondensator C_1 sind alle Kondensatoren auf $2\,U_0$ aufgeladen. Auch die Dioden müssen $2\,U_0$ Sperrspannung aushalten, aber eben nicht mehr!

Schritte	U_{ein}	U_{C_1}	U_{C_2}	U_{C_3}	U_{C_4}	U_{aus}
1	-100 V	100 V	0 V	0 V	0 V	0 V
2	100 V	100 V	200 V	0 V	0 V	200 V
3	-100 V	100 V	200 V	200 V	0 V	200 V
4	100 V	100 V	200 V	200 V	200 V	400 V

Tab. 3.3 Spannungen nach den in Abb. 3.24 illustrierten Schritten.

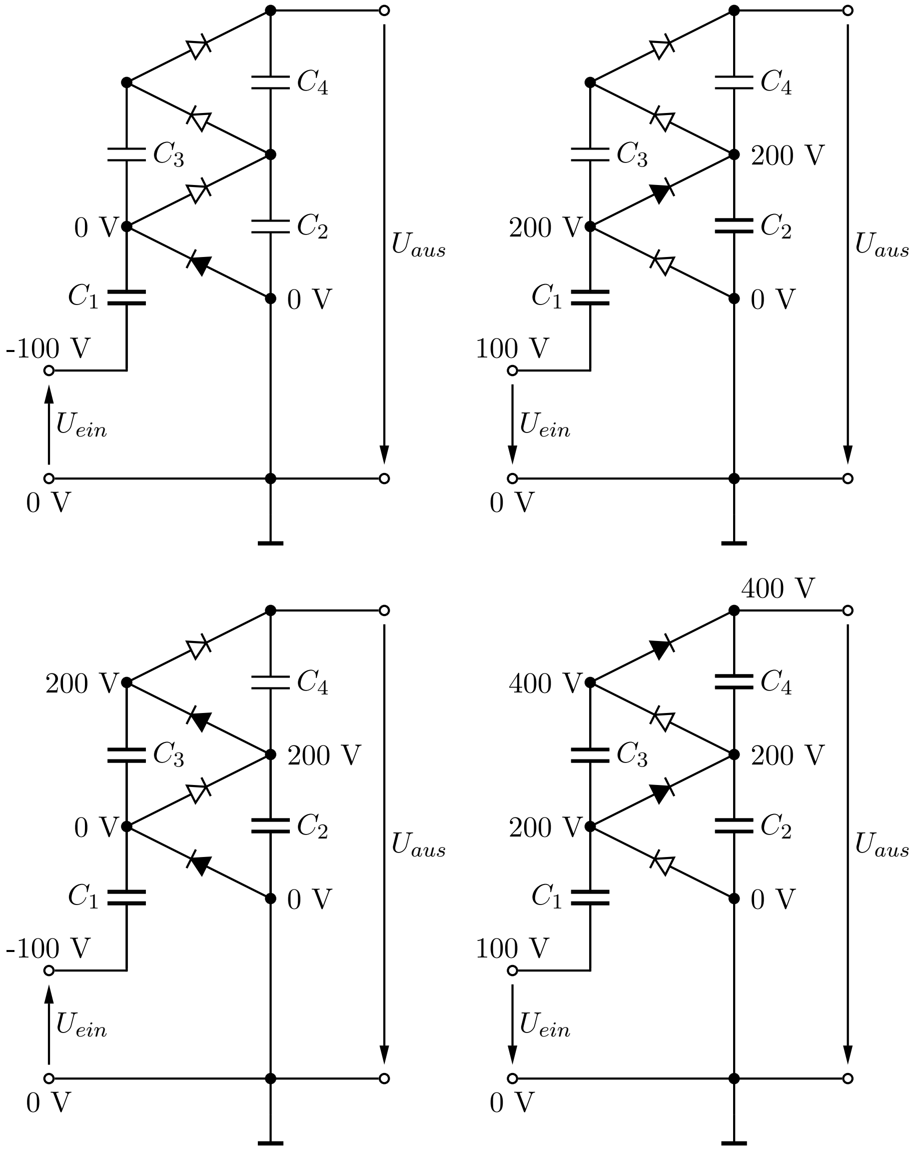


Abb. 3.24 Illustration der Funktionsweise der Kaskadenschaltung in vier Schritten (von links oben, rechts oben, links unten, etc.). Fett gemalte Kondensatoren sind *nach* dem jeweiligen Schritt aufgeladen. Fett gemalte Dioden leiten *während* des Schrittes.

3.9 Gleichspannungswandler[4]

Dasselbe Schaltungsprinzip kann man auch zur Verdopplung von Gleichspannungen anwenden (Abb. 3.25).

[4]Engl. *DC-DC converter*

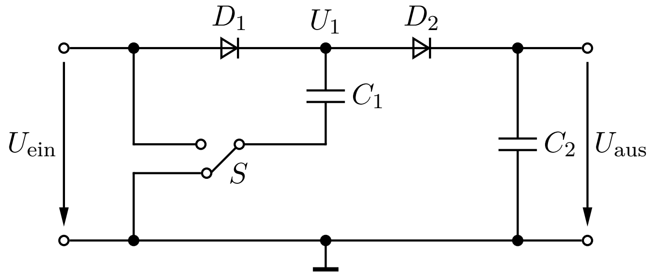


Abb. 3.25 Ladungspumpe zur Gleichspannungsverdopplung.

Der Schalter S wird periodisch hin und her geschaltet. Dies entspricht in gewisser Weise einer Eingangswechselspannung ohne Schalter wie in Abbildung 3.21.

In der gezeigten Schalterposition lädt sich der Kondensator C_1 auf die Spannung $U_{ein} - U_S$ auf. Wird der Schalter umgelegt, liegt C_1 und U_{ein} in Serie, die Diode D_1 sperrt und D_2 leitet. Dadurch wird C_2 auf $2\,U_{ein} - 2\,U_S$ aufgeladen. C_1 wird als Pumpkondensator und C_2 als Glättungskondensator bezeichnet. Diese und ähnliche Schaltungen nennt man Ladungspumpen. In der Praxis werden statt Dioden meistens Transistoren verwendet. Der Schalter S ist natürlich auch als Transistor realisiert. Ladungspumpen kommen ohne Spulen als Energiespeicher aus. Das hat beim Einsatz in hohen Magnetfeldern Vorteile. Ladungspumpen sind besonders effizient bei nur kleinen Strömen.

3.10 Abwärtswandler[5]

Gleichspannungswandler kann man auch mit Spulen aufbauen. In Abb. 3.26 ist ein Abwärtswandler, der auch „buck converter“ genannt wird, gezeigt.

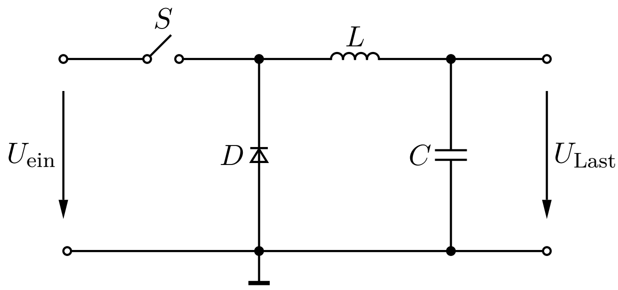


Abb. 3.26 Abwärtswandler mit Spule.

Der Schalter S wird periodisch mit Periode T hin und her geschaltet. Das Tastverhältnis des Schalters $D = T_{an}/T \quad \text{mit} \quad D = [0, 1]$ und $T = T_{an} + T_{aus}$ bestimmt die Ausgangsspannung zu $U_{aus}/U_{ein} = D$. Warum?

Das Magnetfeld der Spule wird bei geschlossenem Schalter S durch den Strom I_L aufgeladen; $U_L = U_{\text{ein}} - U_{\text{aus}}$. Bei offenem Schalter liefert die Spule den Laststrom, und das Magnetfeld baut sich ab; $U_L = -U_{\text{aus}}$ (bei Vernachlässigung der Diodenspannung). Es gilt immer

[5] Engl. *Step-down converter*

$$I_L = \frac{1}{L}\int U_L\, dt \quad , \quad \text{da} \quad U_L = L\frac{dI_L}{dt} \quad .$$

Fordert man, dass Auf- und Entladeströme gleich groß sind, ergibt sich

$$\int\limits_0^{T_\text{an}} \frac{U_L}{L}\,\text{dt} = \frac{U_\text{ein} - U_\text{aus}}{L}\,T_\text{an} = \int\limits_{T_\text{an}}^{T} \frac{U_L}{L}\,\text{dt} = \frac{U_\text{aus}}{L}\,(T - T_\text{an}) = \frac{U_\text{aus}}{L}\,T_\text{aus}$$

und damit die gesuchte Beziehung

$$(U_\text{ein} - U_\text{aus})\,T_\text{an} = U_\text{aus}\,(T - T_\text{an}) \;\Leftrightarrow\; U_\text{ein}\,D\,T = U_\text{aus}\,T \quad .$$

3.11 Temperaturabhängigkeit einer Diode

Durchlassrichtung

Die Strom-Spannungskennlinie einer Diode hängt von der Temperatur ab. Dies gilt es beim Schaltungsentwurf zu beachten.

Die Temperaturabhängigkeit der Diodeneigenschaften ist komplex und stammt vor allen aus zwei Quellen: der Abhängigkeit der intrinischen Ladungsträger von der Temperatur und der exponentiellen Abhängigkeit über U_T. Gleichung 3.25 stellt die Temperaturabhängigkeit der Spannung, die bei konstantem Strom an einer in Durchlassrichtung geschalteten Diode abfällt, dar.

$$\frac{dU}{dT} = \frac{U_S - U_G - 3\,U_T}{T} \tag{3.25}$$

Dabei ist $U_G \approx 1,12\,\text{V}$ die Bandlücke von Silizium und U_S die Schwellenspannung der Diode. Damit ergibt sich für eine Temperatur $T = 300\,\text{K}$ eine Temperaturabhängigkeit von

$$\frac{dU}{dT} = -\,1,87\,\frac{\text{mV}}{^\circ\text{C}} \quad . \tag{3.26}$$

Dieser Effekt ist wichtiger als man beim ersten Blick vermuten könnte. In Schaltkreisen können leicht Temperaturdifferenzen von 50 °C auftreten (an/aus), und Spannungsunterschiede von einigen Hundert mV sind ausreichend, um einen Feldeffekttransistor an- bzw. auszuschalten.

Sperrrichtung

Der Leckstrom I_S einer in Sperrrichtung geschalteten Diode ist stark temperaturabhängig

$$I_S \sim T^{\frac{3}{2}}\,\mathrm{e}^{-\frac{q_e\,U_G}{2\,k\,T}} \quad . \tag{3.27}$$

Jede Temperaturerhöhung von rund 7 °C verdoppelt den Leckstrom. Dies ist z. B. für Siliziumdioden, die als Teilchen- oder Röntgenstrahlungsdetektoren verwendet werden, sehr relevant. Ein zu großer Leckstrom führt zur Erhitzung und somit zum „thermal run-away“. Außerdem trägt der Leckstrom zum Rauschen des Detektors bei.

3.12 Zenerdiode

Bei großen Sperrspannungen weicht das Verhalten einer Diode von der Shockley-Gleichung ab und die Diode wird wieder durchlässig (Abb. 3.27).

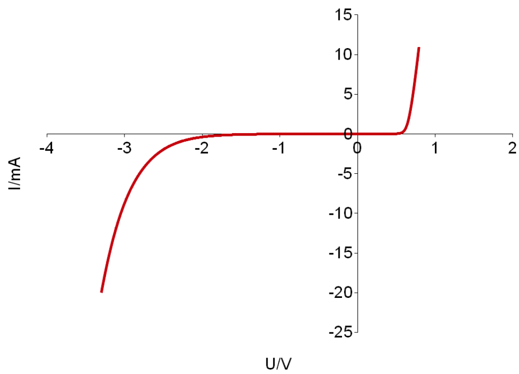


Abb. 3.27 Kennlinie einer Zenerdiode.

Zuerst setzt der sogenannte Zenereffekt ein ($U < -5$ V). In der Verarmungszone, die den Stromfluss blockiert, treten so hohe Feldstärken auf, dass Elektronen aus dem Valenzband in das Leitungsband tunneln können. Bei gegebener Sperrspannung wird das Feld durch die Breite der Verarmungszone bestimmt ($\mathscr{E} = dU/dx$). Bei hoher Dotierung ist die Verarmungszone dünn und damit die Feldstärke besonders hoch.

Anmerkung: Die dünne Verarmungszone führt dazu, dass die Sperrschichtkapazität der Zenerdiode groß ist.

Für noch größere Sperrspannungen ($U < -6$ V) setzt der Lawinendurchbruch[6] ein. Die thermisch in der Verarmungszone angeregten Elektronen werden ballistisch und durch das anliegende elektrische Feld so stark beschleunigt, dass sie in der Raumladungszone Elektronen aus dem Valenzband schlagen können, ohne selbst zu rekombinieren. Die befreiten Elektronen werden ebenfalls beschleunigt und so weiter. Dadurch steigt die Zahl der freien Ladungsträger lawinenartig an, und die Diode wird durchlässig[7]. Der Effekt ist reversibel und zerstört die Diode nicht, falls die auftretenden Ströme und Leistungen nicht zu groß werden.

Dementsprechend sind Zenerdioden sind für den Dauerbetrieb im Durchbruchsbereich ausgelegt. Der Wert der Zener- oder Durchbruchsspannung U_Z kann durch die Dotierung der Zenerdiode sehr genau eingestellt werden (Abb. 3.28)

Im Durchbruchsbereich der Zenerdiode verläuft die Strom-Spannungskennlienie sehr steil. Der differentielle Widerstand im Durchbruchsbereich $r_Z = \frac{dU_D}{dI_D}$ nimmt typischerweise Werte von 5 bis 50 Ω an.

U_Z hängt von der Temperatur ab

$$U_Z(T) = U_Z(T_0)\ [1 + C_T(T - T_0)] \quad .$$

[6] Engl. *avalanche effect*.
[7] Ähnliche Effekte treten auch bei der Gasverstärkung in Driftkammern und in Leuchtstoffröhren auf.

Der Temperaturkoeffizient C_T ist negativ im Zenerbereich ($\approx -6 \cdot 10^{-4}\,1/\mathrm{K}$), bei höheren Temperaturen tunneln die Elektronen leichter und der Absolutwert des Feldes kann kleiner sein.

Der Temperaturkoeffizient im Lawinenbereich ist positiv ($\approx 10^{-3}\,1/K$), die freie Weglänge der Elektronen im Halbleiter nimmt mit zunehmender Temperatur aufgrund der Gitterschwingungen ab, analog dem Verhalten eines ohmschen Widerstandes. D. h. man braucht größere Beschleunigungsfelder für denselben Strom. Im Übergangsbereich von Zener- und Lawineneffekt ist die Temperaturabhängigkeit sehr klein.

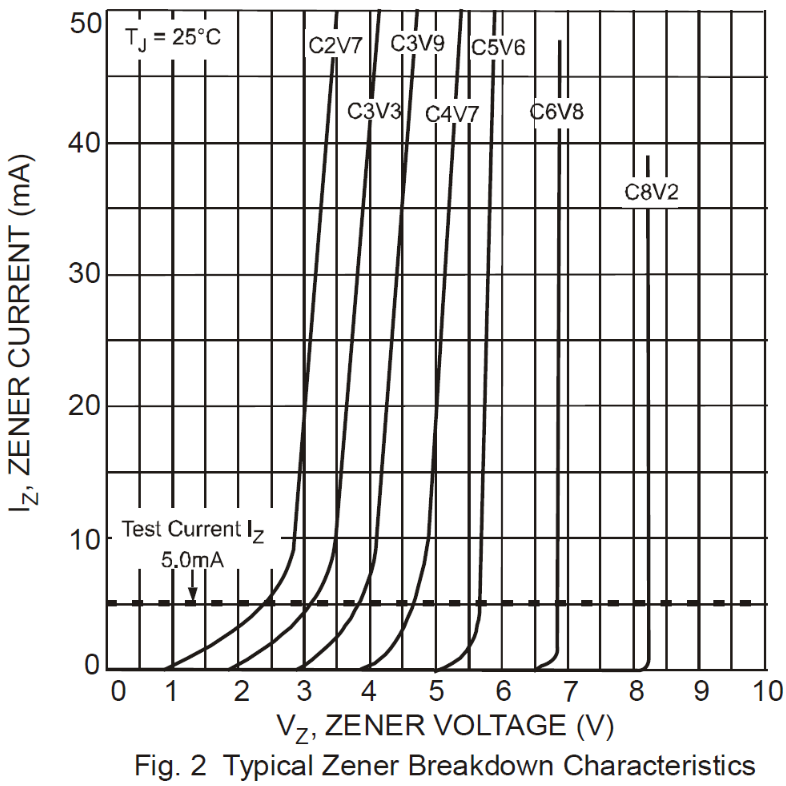


Abb. 3.28 Stromspannungskurven einer Zenerdiodenfamilie [5]. Im Gegensatz zu Abb. 3.27 sind Durchbruchstrom und -spannung hier als positive Werte dargestellt.

Zenerdioden werden gerne als Schutzdioden eingesetzt, wie im vorherigen Abschnitt (Abb. 3.16) gezeigt. Sie sind auch sehr verbreitet als Spannungsreferenz und zur Spannungsstabilisierung.

3.13 Zenerdioden zur Spannungsstabilisierung

Zenerdioden eignen sich als einfache Spannungsregler, d. h. zur Stabilisierung einer Ausgangsspannung bei schwankendem Laststrom oder bei nichtstabiler Eingangsspannung. Eine einfache Schaltung ist in Abbildung 3.29 gezeigt.

Bei ausreichend hohen Eingangsspannungen leitet die Zenerdiode und die Ausgangsspannung erreicht den Wert $U_{aus} \approx U_Z$. Da die Durchbruchskennlinie der Zenerdiode sehr steil verläuft, ist die Ausgangsspannung recht unabhängig von möglichen Variationen des Laststroms. Falls sich der Laststrom bei konstantem Gesamtstrom verringert, fließt einfach mehr Strom durch die Zenerdiode, ohne dass sich U_{aus} nennenswert verschiebt. Das ist ein Vorteil dieser Schaltung.

Unabhängig davon werden mögliche Schwankungen der Eingangsspannung am Ausgang um den Faktor

$$\frac{dU_{aus}}{dU_{ein}} = \frac{R_Z \parallel R_L}{(R_Z \parallel R_L) + R_V} \approx \frac{R_Z}{R_V} \tag{3.28}$$

unterdrückt. Dabei ist R_Z der differentielle Widerstand der Zenerdiode, und R_Z ist klein gegen R_V und R_L.

Anmerkung: Hier gehen wir davon aus, dass der Laststrom und damit der Lastwiderstand konstant sind.

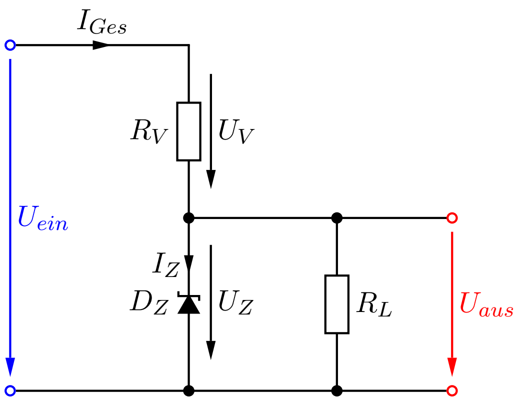


Abb. 3.29 Zenerdiode als Spannungsregler. Die Schaltung ist sowohl mit einer Eingangsspannungsquelle wie einer Eingangsstromquelle interessant [6].

Für kleine Eingangsspannungen U_{ein} sperrt die Zenerdiode und die Ausgangsspannung $U_{aus} = U_{ein}\,\frac{R_L}{R_V+R_L}$. Wie groß muss die Eingangsspannung sein, damit die Zenerdiode leitet? Es soll gelten $U_{aus} = U_Z$, damit folgt aus der obigen Gleichung $U_{ein} > U_Z\left(1 + \frac{R_V}{R_L}\right)$.

Die Schaltung 3.29 ist nur bei kleinen Lastströmen sinnvoll, da sonst die Verlustleistung durch R und die Zenerdiode zu groß wird.

3.14 Schottkydiode

Die Schottkydiode besteht aus einem Metall-Halbleiter-Übergang (Metall-n-Halbleiter) statt aus einem pn-Übergang.

Schottkydioden haben eine deutlich kleinere Schwellenspannung $(0,2 - 0,4\,V)$ als konventionelle pn-Dioden. Ihre Sperrkapazität ist klein. Nur die Majoritätsladungsträger, also Elektronen, tragen zur Leitung bei.

Die kleinen Sperrkapazitäten und die hohe Beweglichkeit der Elektronen (im Vergleich zu Löchern) machen Schottkydioden schneller als konventionelle Halbleiterdioden. Schaltfrequenzen $> 10\,\mathrm{GHz}$ werden ermöglicht. Die geringe Schwellenspannung ermöglicht die Gleichrichtung großer Ströme auch bei eher niedrigen Spannungen, wo konventionelle Dioden zu ineffizient wären.

Ein Nachteil von Schottkydioden sind ihre relativ hohen Leckströme.

Anmerkung für die Vorlesung: Zeige ein Datenblatt einer Schottkydiode z. B. 1N5711WS.

3.15 Metall-Halbleiter-Kontakt

Metall-Halbleiter-Kontakte verwendet man, um Dioden und Transistoren mit der Außenwelt zu verbinden. Hier wäre eine Gleichrichterwirkung äußerst unerwünscht.

Erfreulicherweise lässt sich die Dicke der Verarmungszone durch eine starke Dotierung so reduzieren, dass sich der Metall-Halbleiter-Kontakt wie ein kleiner „ohmscher“ Widerstand verhält.

M	n^{++}	n	p	p^{++}	M

Tab. 3.4 Schematischer Aufbau eines Metall-Halbleiter-Kontakts in einer pn-Diode.

4 Operationsverstärker

Es ist zu aufwendig und platzraubend, Verstärkerschaltungen aus sogenannten diskreten Einzeltransistoren aufzubauen. Viel einfacher ist es, käufliche integrierte Verstärkerbausteine zu benutzen. Ein Operationsverstärker (OPV), kurz Op-amp[1] genannt, ist dazu hervorragend geeignet.
OPV sind relativ billig und haben sehr gute Eigenschaften. Durch Verschaltung mit nur wenigen passiven Bausteilen, lassen sich OPV auf einfache Weise in die gewünschte Verstärkerschaltung verwandeln.

4.1 Grundlagen

Ein OPV ist ein hochempfindlicher, in der Regel mehrstufiger Differenzverstärker mit einem Ausgang. Das Schaltungssymbol eines OPVs wird in Abbildung 4.1 gezeigt.

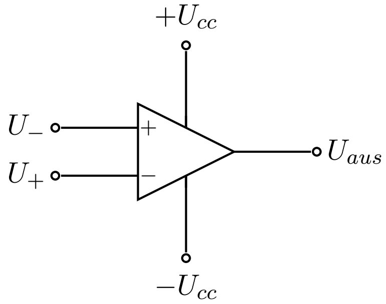


Abb. 4.1 Schaltsymbol eines Operationsverstärkers. „+“ symbolisiert den nichtinvertierenden, „-“ den invertierende Eingang.

Der OPV ist ein aktiver Baustein und hat deshalb eine Versorgungsspannung. Der OPV hat einen Ausgang und zwei Eingänge, einen invertierenden und einen nichtinvertierenden. Die Ausgangsspannung U_{aus} ist proportional zur Differenz der Eingangsspannungen

$$\boldsymbol{U_{aus} = A_D\,(U_+ - U_-)} \tag{4.1}$$

mit der Spannungsverstärkung A_D.

Anmerkung: Später werden wir zwischen der Leerlaufverstärkung[2] A_D und der Verstärkung eines OPVs bei Rückkopplung unterscheiden.

Warum hat der OPV zwei Eingänge?

Man kann einen Verstärker natürlich auch mit einem Eingang bauen, der die Masse der

[1] Engl. für *operational amplifier*
[2] Engl. *open loop gain*

Versorgungsspannung als Bezugspunkt hat. Schließlich hat der OPV auch nur einen Ausgang. Differenzsignale sind aber unempfindlicher gegenüber Störungen auf den Zuleitungen oder Schwankungen der Versorgungsspannung im vorgeschalteten Baustein. Außerdem ergeben sich mit zwei Eingängen viel mehr Möglichkeiten der Rückkopplung, wie wir bald sehen werden.

Der OPV ist zur Verstärkung von Gleich- oder Wechselspannungen geeignet. Wie immer sind zeitlich variierende Signale interessanter. Die Verstärkung A_D des OPVs sollte im Arbeitsbereich möglichst groß und konstant sein. Konstant meint unabhängig von der Amplitude des Eingangssignals (Linearität), frequenzunabhängig und temperaturunabhängig.

Da die Ausgangsspannung die Versorgungsspannung aber nicht überschreiten kann, nimmt A_D für große Eingangssignale ab (siehe Abb. 4.2), und der OPV ist vor allem für kleine Eingangssignale als Vorverstärker geeignet.

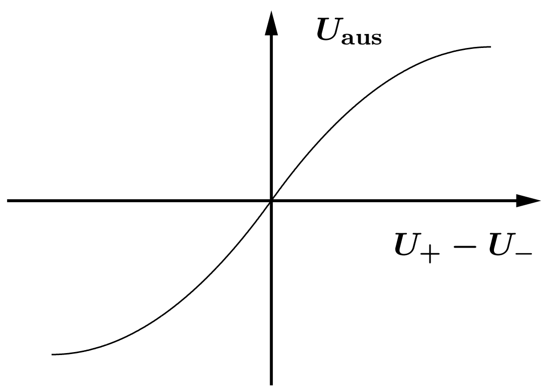


Abb. 4.2 Verstärkung eines generischen Operationsverstärkers als Funktion der Eingangsspannung.

Die Verstärkung ist bei kleinen Frequenzen frequenzunabhängig, fällt bei hohen Frequenzen f aber deutlich ab (Abb. 4.3). Der Wert von f bei dem A_D auf $-3\,\mathrm{dB} = \frac{1}{\sqrt{2}}$ abgefallen ist, wird als Bandbreite bezeichnet. Die Bandbreite eines OPVs bestimmt seinen Einsatzbereich. Für schnelle Signale muss man breitbandige OPV verwenden.

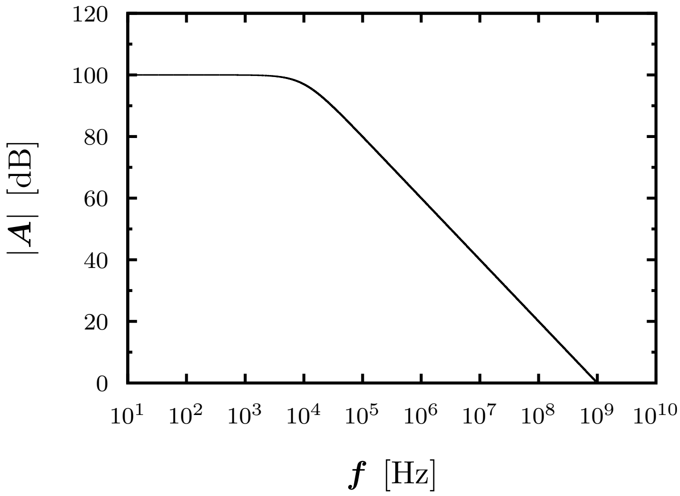


Abb. 4.3 Verstärkung eines generischen Operationsverstärkers als Funktion der Frequenz.

Der Eingangswiderstand des OPVs ist sehr hoch und beträgt mindestens MΩ bis GΩ. Das bedeutet, dass der Eingangsstrom, der in den OPV fließt, meistens vernachlässigbar klein ist. Auch der Widerstand zwischen den Eingängen ist sehr groß. Das ist wichtig, damit der OPV eine „kleine“ Last darstellt und die Eingangsspannungen durch das Anschließen des OPVs nur unwesentlich beeinflusst werden.

Die Ausgangsimpedanz des OPVs ist dagegen klein, damit die Ausgangsspannung des OPVs möglichst unabhängig vom Ausgangsstrom bzw. dem Eingangswiderstand der folgenden Verstärkerstufe ist.

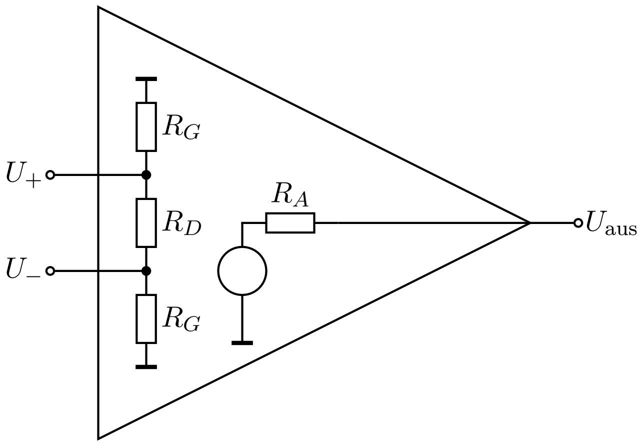


Abb. 4.4 Vereinfachtes Funktionsschaltbild eines Operationsverstärkers. Die Versorgungsspannung ist nicht eingezeichnet.

OPV werden hauptsächlich eingesetzt als:

- flexible Universalverstärker von Spannungen und Strömen
- ladungsempfindliche Vorverstärker
- Regler

- Impedanzwandler
- Komparatoren
- aktive Filter
- Sample and Hold-Bausteine
- Bestandteile von ADCs und DACs und vieles mehr.

4.2 Rückkopplung

Man kann die Eigenschaften eines OPVs wesentlich verbessern, indem man einen Bruchteil des Ausgangssignals auf das Eingangssignal zurückkoppelt (siehe z. B. Abb. 4.5 und 4.7). Man unterscheidet zwischen negativer und positiver Rückkopplung[3], abhängig davon, ob das Ausgangssignal auf den invertierenden oder den nichtinvertierenden Eingang gekoppelt wird.

4.3 Positive Rückkopplung

Die einfachste Form der positiven Rückkopplung ist in Abbildung 4.5(a) dargestellt. Hier ist der nichtinvertierende Eingang des OPVs direkt mit dem Ausgang verbunden. Der invertierende Eingang liegt auf Masse. Realistischer ist die Schaltung in Abbildung 4.5(b), in der noch die beiden Widerstände R_1 und R_2 eingefügt sind. Diese Schaltung wird als Schmitt-Trigger bezeichnet. In dieser Schaltung hat der Ausgang nur zwei stabile Zustände, die bei der positiven und negativen Versorgungsspannung des OPVs, $+U_{CC}$ und $-U_{CC}$, liegen. Wieso?

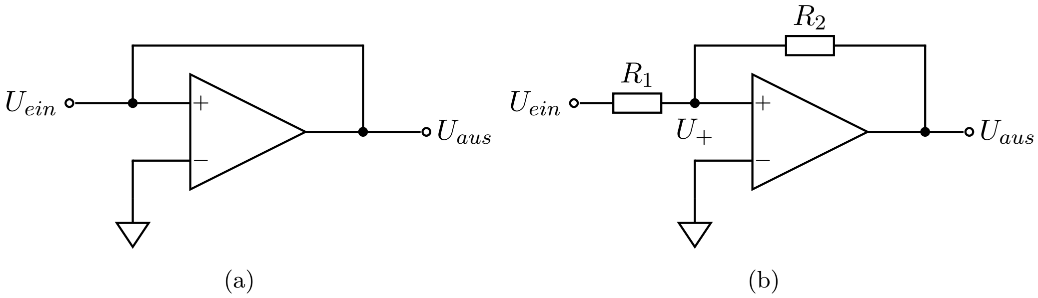


Abb. 4.5 Positive Rückkopplung am Beispiel eines (a) Komparators und (b)Schmitt-Triggers. Die Versorgungsspannung ist nicht eingezeichnet.

Wir betrachten drei kurz aufeinanderfolgende Zeitpunkte t_1, t_2 und t_3. Zum Zeitpunkt t_1 sei o.E.d.A. $U_{aus} \approx 0$. Ist $U_{ein}(t_1) > 0$, so stellt sich am Ausgang die Spannung

$$U_{aus}(t_2) = A_D\,(U_+ - 0\,\mathrm{V}) \gg U_{ein}(t_1)$$

ein.

Anmerkung: Der Wert U_+ wird durch den Spannungsteiler R_1 und R_2 bestimmt und liegt

[3]Engl. *feed-back*

zwischen U_{aus} und U_{ein}. Dies erhöht dank der positiven Rückkopplung die Spannung $U_+(t_3)$, was zu einer noch größeren Ausgangsspannung führt. Für positive Startwerte von U_{ein} nimmt U_{aus} also sehr schnell seinen Maximalwert an.

Anmerkung: Durch interne Spannungsabfälle liegt der Maximalwert meistens deutlich, z.B. ein bis zwei Volt, unter $+U_{CC}$ und der Minimalwert entsprechend über $-U_{CC}$. Dieses Detail ignorieren wir im Folgenden.

Den stabilen Zustand $U_{aus} \approx +U_{CC}$ kann man nur verlassen, wenn zu einem späteren Zeitpunkt ein negatives Signal auftritt. Wie groß muss dieses negative Signal sein?

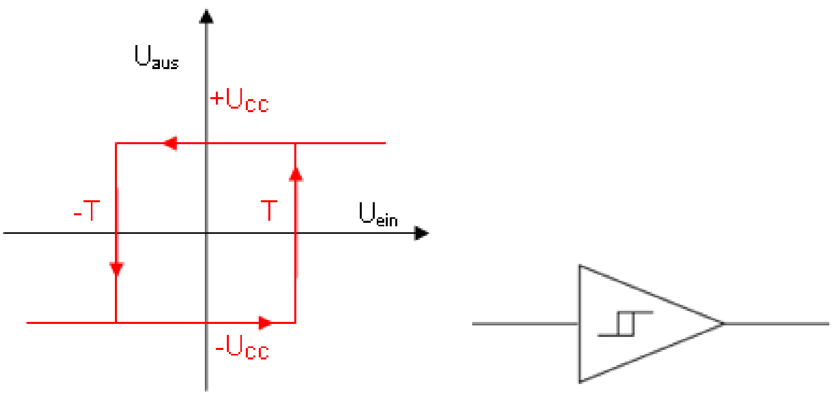
Aus der Knotenregel folgt

$$I_1 = I_2 \quad \text{und} \quad \frac{U_{ein} - U_+}{R_1} = \frac{U_+ - U_{aus}}{R_2} \qquad .$$

Anmerkung: Die Richtung des Stroms durch die Widerstände R_1 und R_2 wird hierbei willkürlich als von links nach rechts fließend definiert.

Für $U_+ < 0$ ändert sich der Ausgangszustand. Das entspricht $U_{ein} < -U_{aus}\,\frac{R_1}{R_2}$. Dieser Ansatz gilt für beide Ausgangszustände. Für negative U_{aus} muss U_{ein} also den Wert $U_{CC}\,\frac{R_1}{R_2}$ überschreiten, für positive U_{aus} den Wert $-U_{CC}\,\frac{R_1}{R_2}$ unterschreiten. Je kleiner R_1 und je größer R_2 ist, desto leichter lässt sich der Ausgangszustand ändern.

Der Schmitt-Trigger weist ein Hystereseverhalten auf. Dies ist in Abbildung 4.6 illustriert. Zum Einstellen des positiven Ausgangszustands $U_{aus} \approx U_{CC}$ benötigt man eine größere Triggerspannung T als zum Umschalten auf $U_{aus} \approx -U_{CC}$. (Es sein denn, R_1 ist sehr klein.) Das Schaltsymbol für einen Schmitt-Trigger ist in Abbildung 4.6 rechts gezeigt, es soll an die Hysteresekurve erinnern.



(a) (b)

Abb. 4.6 (a) Verlauf der Ausgangsspannung eines Schmitt-Triggers als Funktion der Eingangsspannung. (b) Schaltungssymbol des Schmitt-Triggers.

4.4 Negative Rückkopplung

Beispiele negativer Rückkopplung sind in den Abbildungen 4.7, 4.9 und 4.10 gezeigt. Bei negativer Rückkopplung wirkt die Ausgangsspannung dem Eingangssignal entgegen. Warum ist das sinnvoll?

Wir betrachten den invertierenden Verstärker im folgenden Kapitel als repräsentatives Beispiel für die Vorteile der negativen Rückkopplung.

4.5 Invertierender Verstärker

Beim invertierenden Verstärker (Abb. 4.7) liegt der nichtinvertierende Eingang auf Masse. Der Ausgang ist über den Spannungsteiler R_1, R_2 an den invertierenden Eingang des OPVs gekoppelt. Sei $U_{aus} \approx 0$ und $U_{ein}(t_1) > 0$. Dann ist o.E.d.A.

$$U_-(t_1) = U_{ein}(t_1)\,\frac{R_2}{R_1 + R_2} > 0 \quad \text{und} \quad U_+ - U_- < 0 \quad .$$

Durch die Verstärkung und die negative Rückkopplung wird der Wert von U_- verringert, bis sich der stabile Zustand $U_- = U_+$ einstellt, hier mit $U_+ = 0\,\mathrm{V}$.

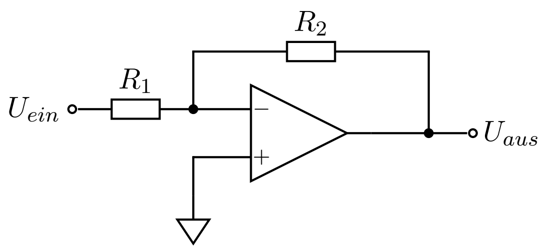


Abb. 4.7 Invertierender Verstärker.

In Abbildung 4.8 ist die Antwort des invertierenden Verstärkers auf eine Sprungfunktion der Spannung am Eingang U_{ein} für zwei repräsentative OPV gezeigt. Für beide OPV stellt sich U_- sehr schnell auf den Wert von U_+, hier $0\,V$, ein. Entsprechend schnell stabilisiert sich die Ausgangsspannung. U_- bildet eine sogenannte „virtuelle Masse“. Die Bezeichnung „virtuell“ soll ausdrücken, dass zwar das Potential auf $0\,\mathrm{V}$ liegt, aber kein Strom abgeleitet wird. Der Strom fließt durch den Rückkopplungswiderstand.

Anmerkung: Oft schaltet man parallel zu dem Widerstand R_2 einen Kondensator. Dies verbessert die Stabilität des Verstärkers und unterdrückt Oszillationen.

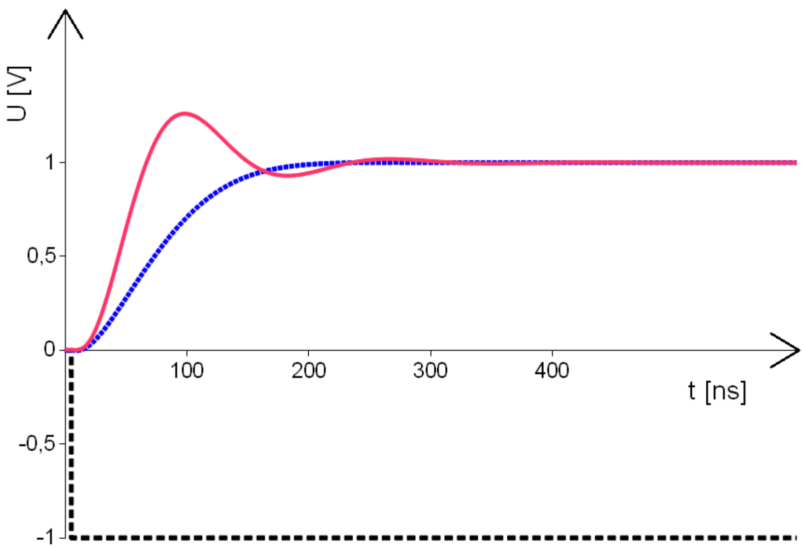


Abb. 4.8 Eingangssignal U_{ein} (gestrichelte schwarze Line) und Ausgangsspannung U_{aus} für einen schnellen (rote Kurve, THS4032CD) und einen langsamen OPV (blaue Kurve).

4.6 Goldene Regeln

OPV-Schaltungen mit negativer Rückkopplung lassen sich leicht mittels der sogenannten goldenen Regeln analysieren. In guter Näherung gilt:

1. Die Spannungen an beiden Eingängen des OPVs sind gleich: $U_+ = U_-$.
2. Die Eingangsimpedanz ist unendlich hoch. Das bedeutet, es fließt kein Strom in den OPV.
3. Die Ausgangsimpedanz des OPVs ist sehr klein. Das heißt, die Ausgangsspannung hängt minimal von dem Ausgangsstrom und damit den nachgestellten Schaltungselementen ab.

Angewandt auf den invertierenden Verstärker gilt dann

$$U_- = 0\,V \quad , \quad I_1 = I_2 \quad , \quad \frac{U_{ein}}{R_1} = -\frac{U_{aus}}{R_2}$$

und damit

$$\frac{U_{aus}}{U_{ein}} = -\frac{R_2}{R_1} \quad . \tag{4.2}$$

Anmerkung: Auch hier wird die Stromrichtung des Stroms durch die Widerstände R_1 und R_2 als von links nach rechts fließend definiert.
Der invertierende OPV invertiert das Eingangssignal und verstärkt es um den Faktor $\frac{R_2}{R_1}$.

Was bedeutet das?

Die **Leerlaufverstärkung** A_D taucht in der obigen Formel nicht auf. Dies ist durchaus ein

Vorteil, da die Leerlaufverstärkung, wie wir schon gesehen haben, nicht immer konstant ist. Die Widerstände dagegen sind passive Bausteine, die unabhängig vom OPV frei gewählt werden können. Da zudem das Verhältnis von baugleichen Widerständen auftritt, reduziert sich z. B. auch die Temperaturabhängigkeit der Verstärkung.

Bevor wir uns überlegen, wieso die Leerlaufverstärkung doch noch wichtig ist, wollen wir weitere Schaltungen mit negativer Rückkopplung kennenlernen.

4.7 Nichtinvertierender Verstärker

Hier liegt das Eingangssignal am nichtinvertierenden Eingang des OPV (Abb. 4.9). Die Spannung U_- wird durch den Spannungsteiler R_1 und R_2 eingestellt. Hier gilt

$$U_{ein} = U_+ = U_- = U_{aus} \frac{R_1}{R_1 + R_2}$$

und damit

$$\frac{U_{aus}}{U_{ein}} = 1 + \frac{R_2}{R_1} \quad . \tag{4.3}$$

Der Eingangswiderstand des idealen nichtinvertierenden Verstärkers ist sehr groß.

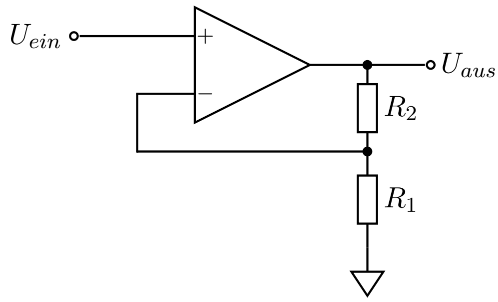


Abb. 4.9 Nichtinvertierender Verstärker.

4.8 Impedanzwandler

Ein Spezialfall des nichtinvertierenden OPVs ist der Impedanzwandler oder Spannungsfolger. Hier wird der Ausgang direkt mit dem invertierenden Eingang verbunden, $R_2 = 0$. Dann spielt der Wert von R_1 keine Rolle mehr (vorausgesetzt R_1 ist nicht selbst sehr klein), und daraus ergibt sich die Schaltung in Abbildung 4.10 ohne R_1.

Die Spannungsverstärkung des Impedanzwandlers

$$\frac{U_{aus}}{U_{ein}} \approx 1 \quad ,$$

entsprechend der obigen Formel

$$\frac{U_{aus}}{U_{ein}} = 1 + \frac{R_2}{R_1} \quad \text{mit} \quad R_2 = 0 \quad \text{und} \quad R_1 = \infty \quad .$$

Der Nutzen des Impedanzwandlers liegt nicht in der Spannungsverstärkung, sondern in seiner sehr hohen Eingangsimpedanz bei sehr kleiner Ausgangsimpedanz. Man kann also den Impedanzwandler an die Signalquelle anschließen, ohne sie nennenswert zu belasten. Gleichzeitig liefert der Impedanzwandler große Ausgangsströme.

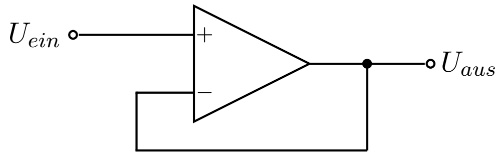


Abb. 4.10 Impedanzwandler.

Anmerkung: In der Praxis würde man oft einen Widerstand R, z. B. 10 kΩ, in die Rückkopplungsschleife schalten.

4.9 OPV als Regelkreis

Es ist bequem, den Effekt der negativen Rückkopplung eines OPVs abstrakt durch den Formalisimus eines Regelkreises zu beschreiben. Dazu betrachten wir Abbildung 4.11.

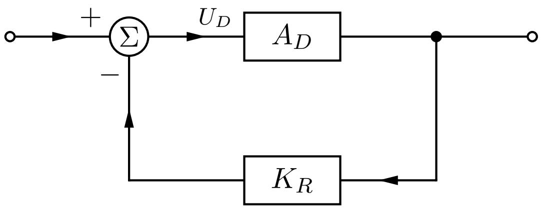


Abb. 4.11 Der OPV als Regelkreis.

Die um den Faktor K_F abgeschwächte Eingangsspannung U_{ein} und die um K_R abgeschwächte und negativ zurückgekoppelte Ausgangsspannung U_{aus} bilden die Differenzspannung U_D mit

$$U_D = K_F\, U_{ein} - K_R\, U_{aus} \quad . \tag{4.4}$$

Die Differenzspannung U_D wird um die Leerlaufverstärkung A_D verstärkt

$$U_{aus} = A_D\, U_D \quad .$$

Damit ergibt sich

$$U_{ein} = U_D \frac{1 + K_R\, A_D}{K_F} \quad \text{und} \quad A = \frac{U_{aus}}{U_{ein}} = \frac{K_F\, A_D}{1 + K_R\, A_D} \approx \frac{K_F}{K_R} \quad . \tag{4.5}$$

Hierbei ist A die Verstärkung mit Rückkopplung[4], A_D die Leerlaufverstärkung[5] und $K_R\, A_D$ die sogenannte Schleifenverstärkung[6]. Diese elementaren aber anfangs verwirrenden Begriffe muß man sich merken, um OPVs zu verstehen.

[4] Engl. closed-loop gain
[5] Engl. open-loop gain
[6] Engl. loop gain

Für den nichtinvertierenden Verstärker gilt

$$K_F = 1 \quad \text{und} \quad K_R = \frac{R_1}{R_1 + R_2} \quad ,$$

also folgt

$$A = 1 + \frac{R_2}{R_1} \quad \text{(vgl. Gl. 4.3)} \quad .$$

Für den invertierenden Verstärker gilt

$$K_F = -\frac{R_2}{R_1 + R_2} \quad \text{und} \quad K_R = +\frac{R_1}{R_1 + R_2} \quad .$$

Anmerkung: Wir wenden hier das Superpositionsprinzip an und setzen $U_{aus} = 0\,\mathrm{V}$ bei der Berechnung von K_F und setzen $U_{ein} = 0\,\mathrm{V}$ bei der Berechnung von K_R. Das negative Vorzeichen von K_F drückt aus, dass wir U_{ein} auf den invertierenden Eingang des OPVs legen.

Es folgt

$$A = -\frac{R_2}{R_1} \quad \text{(vgl. Gl. 4.2)} \quad .$$

4.10 Realer Operationsverstärker

Der reale Operationsverstärker weicht vom Ideal in einer Reihe von wichtigen Effekten ab. Viele dieser Abweichungen haben wir schon besprochen (Tab. 4.1). So sind die *Eingangswiderstände* und die *Leerlaufverstärkung* groß, aber endlich. Die *Ausgangsimpedanz* ist klein, aber nicht 0.

Eine weitere wichtige Abweichung ist die *Gleichtaktverstärkung*. Ein reeller OPV verstärkt nicht nur das Differenzsignal $(U_+ - U_-)$, sondern leider auch – allerdings mit geringer Verstärkung – das Summensignal $(U_+ + U_-)$. Das Ausgangssignal U_{aus} des reellen OPVs ist damit

$$U_{aus} = A_D\,(U_+ - U_-) + A_G\,(U_+ + U_-) \quad .$$

Das Verhältnis der Spannungsverstärkungen A_D/A_G nennt man Gleichtaktunterdrückung[7].

Die Ausgangsspannung des idealen OPVs für eine Eingangsspannung von $U_+ = U_- = 0\,\mathrm{V}$ ist $0\,\mathrm{V}$. Für den reellen OPV beobachtet man eine kleine Abweichung. Man definiert die Offsetspannung U_{OS} als die Spannung zwischen den beiden Eingängen, für die gilt $U_{aus} = 0\,\mathrm{V}$.
Die Offsetspannung kann man im Prinzip durch eine einstellbare Korrekturspannung kompensieren. Allerdings hängt die Offsetspannung auch von der Temperatur und der Versorgungsspannung ab und kann sich auch durch Alterungseffekte verschieben.

[7] Engl. *Common mode rejection ratio (CMRR)*

Größe	Symbol	idealer OPV	reeller OPV
Gleichtakteingangsimpedanz	R_G	∞	$10^8 \dots 10^{12}\,\Omega$
Differenzeingangsimpedanz	R_D	∞	$10^5 \dots 10^7\,\Omega$
Ausgangsimpedanz	R_{aus}	0	$0,1 \dots 100\,\Omega$
Differenzverstärkung	A_D	∞	$10^4 \dots 10^7$
Gleichtaktverstärkung	A_G	0	$1 \dots 10$
Eingangsruhestrom	I_{OG}	0	$1\,\text{pA} \dots 1\,\mu\text{A}$
Offsetspannung	U_{OS}	0	$10\,\mu\text{V} \dots 10\,\text{mV}$

Tab. 4.1 Parameter des reellen und idealen OPVs (Abb. 4.4). Die Werte der Tabelle verstehen sich ohne Rückkopplung.

4.11 Ersatzschaltung des Operationsverstärkers

Die Goldenen Regeln und auch das Verhalten reeller OPV versteht man am einfachsten anhand einer Ersatzschaltung. Der invertierende Verstärker und seine Ersatzschaltung sind in Abbildung 4.12 dargestellt.

Daraus kann man den Wert von U_- und U_{aus} in Abhängigkeit von den externen und internen Widerständen ausrechnen und wird in guter Näherung die goldenen Regeln bestätigt finden. Außerdem kann man die Eingangs- und Ausgangsimpedanz berechnen.

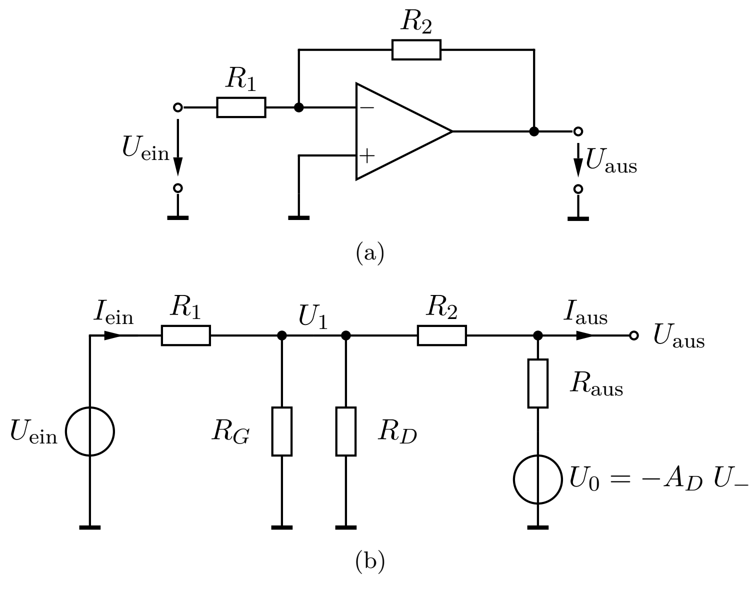


Abb. 4.12 (a) Invertierender Verstärker und (b) seine Ersatzschaltung. U_- entspricht der Spannung am negativen Eingang des OPVs.

Anmerkung: Die Berechnung der Impedanzen von Schaltungen mit externen bzw. internen, gesteuerten Spannungsquellen oder Stromquellen ist etwas komplizierter als für passive Netzwerke. Man stelle sich vor, dass man am Eingang der Schaltung eine (variable) Testspannung anschließt und den Strom, der sich einstellt, misst. Man kann genauso gut eine Stromquelle anschließen und die Spannung, die sich einstellt, bestimmen. Dasselbe Prozedere wendet man für die Bestimmung der Ausgangsimpedanz an. Alle externen unabhängigen Spannungsquellen und Stromquellen stören und werden durch Kurzschluss (Spannungsquellen) oder eine offene Leitung (Stromquelle) ersetzt. Gesteuerte, interne Quellen bleiben erhalten.

Die Eingangsimpedanz Z_{ein} definieren wir als $Z_{ein} = \frac{dU_{ein}}{dI_{ein}}$ bei unbelastetem offenen Ausgang $I_{aus} = 0$.

Die Ausgangsimpedanz Z_{aus} definieren wir als $Z_{aus} = \frac{dU_{aus}}{dI_{aus}}$ bei kurzgeschlossener Eingangsspannung $U_{ein} = 0$. Diese Definitionen sind plausibel, da der Wert der Impedanzen des OPVs nicht von den Eigenschaften der voran- oder nachgeschalteten Bausteinen abhängen soll.

Virtuelle Masse des invertierenden Verstärkers

Für den Eingangsstrom gilt

$$I_{ein} = \frac{U_{ein} - U_-}{R_1} \quad .$$

Hier taucht leider die noch unbekannte Größe U_- auf. Wir berechnen U_- nach dem Superpositionsprinzip (für $I_{aus} = 0$) und erhalten mit der Annahme $R_G \gg R_D$

$$U_- = U_{ein} \frac{(R_2 + R_{aus}) \parallel R_d}{R_1 + (R_2 + R_{aus}) \parallel R_D} - A_D \, U_- \frac{R_1 \parallel R_D}{R_2 + R_{aus} + R_1 \parallel R_D} \quad .$$

Mit der Abkürzung

$$R = R_2 + R_{aus}$$

und den bequemen Näherungen

$$R_1 \parallel R_D \approx R_1 \quad \text{und} \quad R \parallel R_D \approx R$$

folgt

$$U_- = U_{ein} \frac{R}{R_1 + R} - A_D \, U_- \frac{R_1}{R + R_1} \quad \text{und damit} \quad U_- = U_{ein} \frac{R}{R + R_1 + R_1 \, A_D} \quad .$$

Da A_D sehr groß ist, gilt $U_- \ll U_{ein}$.

Dies bestätigt das U_- eine virtuelle Masse bildet und entspricht den goldenen Regeln. Entsprechend gilt $Z_{ein} = R_1$. Dieser Wert ist also nicht so hochohmig wie ohne Rückkopplung. Das ist ein Nachteil des invertierenden Verstärkers.

Anmerkung: Versteht man den invertierenden Verstärker als einen Strom-Spannungswandler ist ein geringer Eingangswiderstand erwünscht (siehe Kapitel 4.12.)

Eingangsimpedanz des nichtinvertierenden Verstärkers

Die Eingangsimpedanz des nichtinvertierenden Verstärkers ist durch R_D dominiert. Ein Eingangsstrom I_{ein} verursacht die Differenzspannung

$$U_D = I_{ein} \, R_D = \frac{U_{aus}}{A_D} \quad . \tag{4.6}$$

Mit der Beziehung $U_{aus} \approx \frac{U_{ein}}{K_R}$ und $\frac{1}{K_R} = 1 + \frac{R_2}{R_1}$ ergibt sich

$$Z_{ein} = \frac{dU_{ein}}{dI_{ein}} = R_D\,K_R\,A_D = \frac{R_D\,A_D}{1 + \frac{R_2}{R_1}} \quad . \tag{4.7}$$

Die Rückkopplung vergrößert also den intrinsischen Eingangswiderstand R_D deutlich.

4.12 Eingangs- und Ausgangsimpedanz eines OPV mit Rückkopplung*

Mit einem OPV können durch die Art der Rückkopplung vier unterschiedliche Verstärkertypen aufgebaut werden (siehe Tabelle 4.2), von denen wir bisher nur den invertierenden und den nicht-invertierenden Verstärker betrachtet haben. Hier führen wir einen Bruchteil β bzw. K_R der Ausgangsspannung über das Rückkopplungsnetzwerk als Korrektur der *Eingangsspannung* (nicht-invertierender Verstärker) oder des *Eingangsstroms* (invertierender Verstärker) auf den Eingang zurück. Im folgenden sollen diese Verstärkertypen vorgestellt und plausibel gemacht werden. Die angegebenen Ergebnisse werden aber nicht hergeleitet.

Dieses Prinzip ist in Abb. 4.14 bis 4.17 abstrakter dargestellt. In der Abbildung ist im Interesse der Klarheit ein OPV mit einem differentiellen Ausgang gezeichnet. Die Ergebnisse dieses Betrachtung gelten aber ganz analog für einen OPV mit nur einem Ausgang (siehe Abb. 4.13).

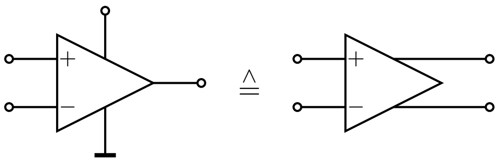


Abb. 4.13 Schematischer Vergleich von OPVs mit differentiellen und nicht-differentiellen Ausgang.

Grundsätzlich kann man den Ausgang des OPV parallel (shunt) oder seriell (series) an den Eingang des Rückkopplungsnetzwerkes anschließen. Im ersten Fall handelt es sich um eine Spannungsrückkopplung, im zweiten Fall um eine Stromrückkopplung. Auch für die Kopplung des Ausgangs des Rückkopplungsnetzwerkes an den Eingang des OPV gibt es diese beiden Möglichkeiten. Schließt man den Ausgang parallel (shunt) an den Eingang des OPV an, so handelt es sich um eine Stromrückkopplung. Schließt man den Ausgang seriell (series) an den Eingang des OPV an, so handelt es sich um eine Spannungsrückkopplung.

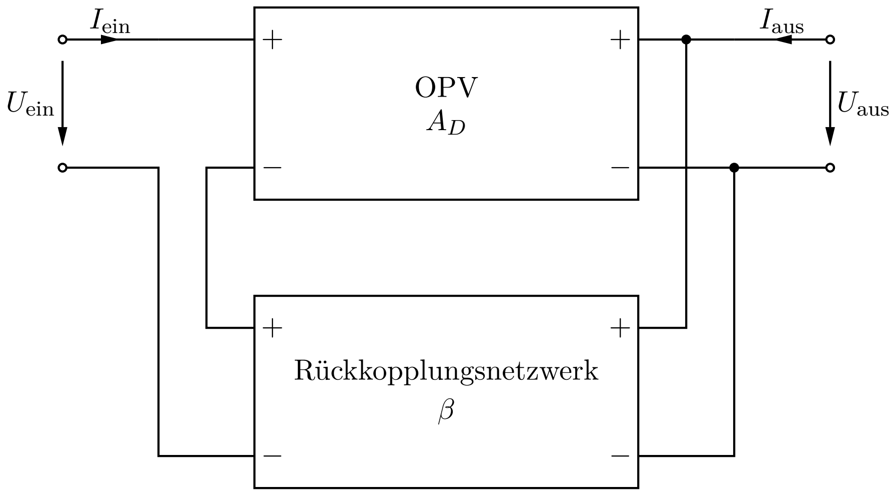


Abb. 4.14 Rückkopplung eines Spannungsverstärkers (nicht-invertierender OPV) mit Spannungs-Spannungs-Rückkopplung (series-shunt). Der Eingang des Rückkopplungsnetzwerkes ist mit dem Ausgang des OPV verbunden. Der Ausgang des Rückkopplungsnetzwerkes mit dem Eingang des OPV.

Durch die Rückkopplung wird die Eingangsimpedanz des Spannungsverstärkers massiv erhöht und die Ausgangsimpedanz massiv reduziert (siehe Gl. 4.7 und Tabelle 4.2). Dadurch wird der OPV verbessert, denn sein Eingang belastet die in der Regel schwache Signalspannungsquelle jetzt noch weniger. Gleichzeitig stellt der Ausgang des OPV eine bessere Spannungsquelle dar.

In Abb. 4.15 ist der invertierende Verstärker dargestellt. Wir betrachten den invertierenden Verstärker hier als einen Strom-zu-Spannungs-Wandler. Damit wird die Spannungsquelle und der Widerstand R_1 in Abb. 4.12 durch eine Stromquelle ersetzt. In dieser Sicht hat die Verstärkung A bzw. A_D des OPV die Einheit [V/I = Ω]

Hier wird durch die Rückkopplung die Eingangsimpedanz *und* die Ausgangsimpedanz reduziert. Die Reduktion der Eingangsimpedanz ist wichtig, damit der Signalstrom „in“ den Verstärker fließen kann.

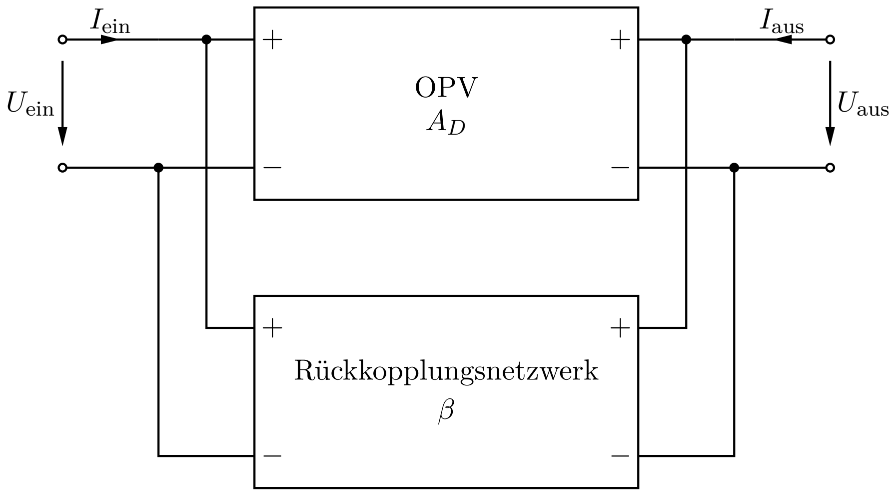


Abb. 4.15 Rückkopplung eines invertierenden Verstärkers mit Strom-Spannungs-Rückkopplung (shunt-shunt).

Analog kann man durch die Art der Rückkopplung noch einen Stromverstärker 4.16 bzw. einen Transkonduktanzverstärker 4.17 aufbauen. Der Stromverstärker sollte eine kleine Eingangsimpedanz besitzen, damit der Signalstrom „in" den Verstärker fließen kann. Gleichzeitig ist eine große Ausgangsimpedanz (Stromquelle!) erwünscht. Beide Eigenschaften werden durch die Rückkopplung verbessert. Der Transkonduktanzverstärker ist ein Spannungs-zu-Strom-Wandler. Hier wird durch die Rückkopplung die Eingangsimpedanz *und* die Ausgangsimpedanz erhöht.

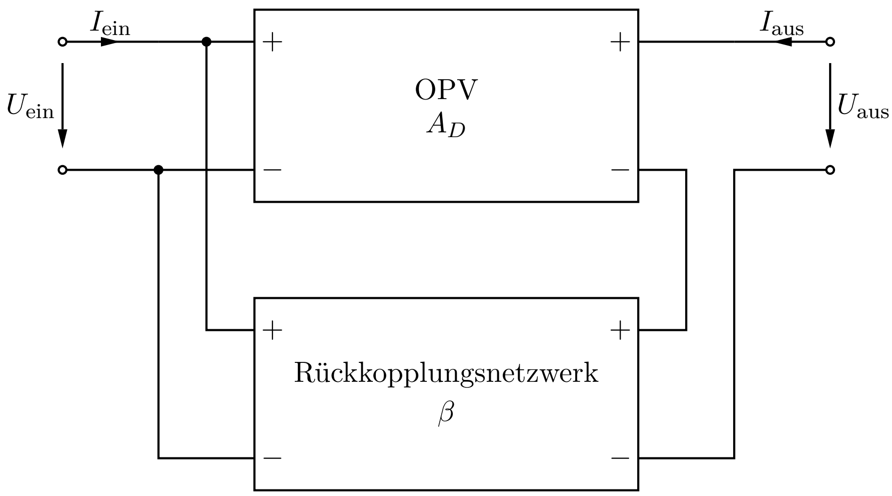


Abb. 4.16 Rückkopplung eines Stromverstärkers mit Strom-Strom-Rückkopplung (shunt-series).

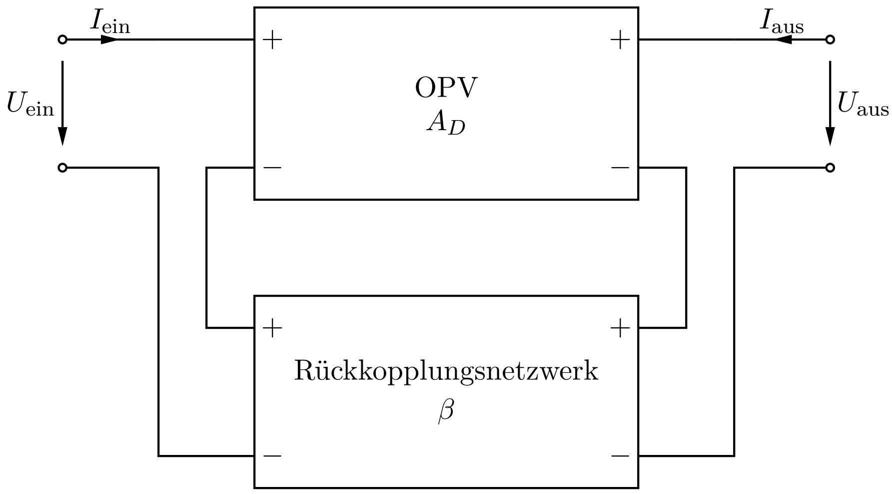


Abb. 4.17 Rückkopplung eines Transkonduktanzverstärkers mit Spannungs-Strom-Rückkopplung (series-series).

Verstärker	Typus	Verstärkung	R_{ein}	R_{aus}
Nicht-invertierender OPV (series-shunt)	Spannung zu Spannung	$\frac{A_D}{1+\beta\,A_D}$	$(1+\beta\,A_D)\;R_{\mathrm{ein}}$	$\frac{R_{\mathrm{aus}}}{1+\beta\,A_D}$
Invertierender OPV (shunt-shunt)	Strom zu Spannung	$\frac{A_D}{1+\beta\,A_D} \approx -R_2$	$\frac{R_{\mathrm{ein}}}{1+\beta\,A_D}$	$\frac{R_{\mathrm{aus}}}{1+\beta\,A_D}$
Stromverstärker (shunt-series)	Strom zu Strom	$\frac{A_D}{1+\beta\,A_D}$	$\frac{R_{\mathrm{ein}}}{1+\beta\,A_D}$	$(1+\beta\,A_D)\;R_{\mathrm{aus}}$
Transkonduktanz-verstärker (series-series)	Spannung zu Strom	$\frac{A_D}{1+\beta\,A_D}$	$(1+\beta\,A_D)\;R_{\mathrm{ein}}$	$(1+\beta\,A_D)\;R_{\mathrm{aus}}$

Tab. 4.2 Verstärkung, Eingangs- und Ausgangsimpedanzen von OPVs mit Rückkopplung. Die Leerlaufverstärkung ist A_D, Eingangs- und Ausgangsimpedanzen ohne Rückkopplung sind R_{ein} und R_{aus}, $\beta\,A_D$ ist die Schleifenverstärkung.

4.13 Frequenzverhalten des OPVs

Die Verstärkung A_D eines OPVs nimmt mit zunehmender Signalfrequenz ab (Abb. 4.3). Im einfachsten Fall entspricht dieser Abfall dem eines Tiefpasses (Gl. 2.13) und beträgt jenseits der Grenzfrequenz f_g 20 dB pro Frequenzdekade

$$\underline{A}_D = \frac{A_{D_0}}{1+\mathrm{j}\,\frac{f}{f_g}} \quad . \tag{4.8}$$

Der Frequenzgang $A(f)$ des gegengekoppelten (nichtinvertierenden) Verstärkers ergibt sich aus Gleichungen 4.5 und 4.8 und

$$\frac{U_{aus}}{U_{ein}} = \underline{A} = \frac{A_D}{1 + K_R\, A_D} \tag{4.9}$$

zu

$$\underline{A} = \frac{\frac{A_{D_0}}{1+\mathrm{j}\,\frac{f}{f_g}}}{1 + K_R\,\frac{A_{D_0}}{1+\mathrm{j}\,\frac{f}{f_g}}} = \frac{A_{D_0}}{1+\mathrm{j}\,\frac{f}{f_g} + K_R\, A_{D_0}} \approx \frac{A_{D_0}}{\mathrm{j}\,\frac{f}{f_g} + K_R\, A_{D_0}} \quad , \tag{4.10}$$

da typischerweise $K_R\, A_{D_0} >> 1$. Der Betrag von $\underline{A}$ ist

$$|\underline{A}| = \frac{A_{D_0}}{\sqrt{(1 + K_R\, A_{D_0})^2 + (\frac{f}{f_g})^2}}$$

Für kleine Frequenzen $f << f_g$ vernachlässigen wir den Imaginärterm f/f_g im Nenner und $|\underline{A}| \approx \frac{1}{K_R}$. Im Grenzfall sehr großer Frequenzen $f >> f_g$ dominiert der Imaginärterm den Betrag der Verstärkung und

$$|\underline{A}| \approx \frac{A_{D_0}}{\frac{f}{f_g}} \quad .$$

Dies gilt auch für Gleichung 4.8 und deshalb

$$|\underline{A}_D| \approx \frac{A_{D_0}}{\frac{f}{f_g}} \quad .$$

Die Frequenzabhängigkeit bei hohen Frequenzen ist also näherungsweise unabhängig von der Gegenkopplung.

Die Grenzfrequenz des gegengekoppelten Verstärkers ist $f_g\, K_R\, A_{D_0}$, da

$$\underline{A} \approx \frac{A_{D_0}}{\mathrm{j}\,\frac{f}{f_g} + K_R\, A_{D_0}} = \frac{\frac{1}{K_R}}{1 + \mathrm{j}\,\frac{f}{f_g\, K_R\, A_{D_0}}}$$

näherungsweise gilt. Die Grenzfrequenz mit Gegenkopplung ist zu höheren Frequenzen verschoben und zwar genau um den Faktor $A_{D_0}\, K_R$, d. h. Leerlaufverstärkung A_{D_0} durch Verstärkung mit Gegenkopplung $\frac{1}{K_R}$.

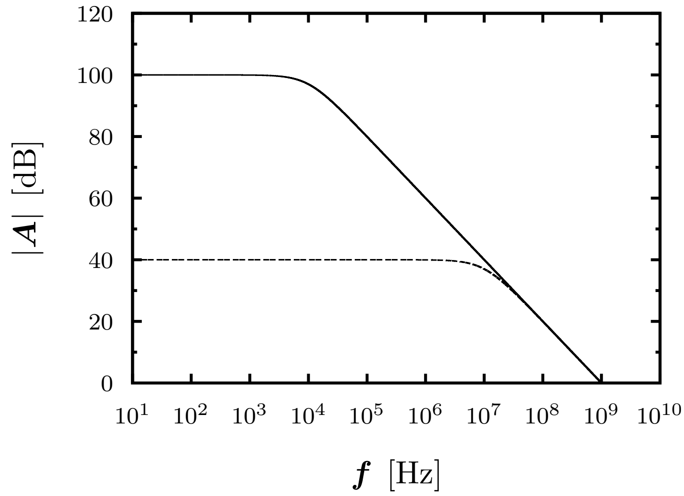


Abb. 4.18 Frequenzverhalten des OPVs ohne Rückkopplung (durchgezogene Linie) und mit Rückkopplung (gestrichelte Linie).

Man kann also durch Wahl einer reduzierten Verstärkung das Frequenzverhalten des Verstärkers verbessern und die Bandbreite erhöhen. Diese drückt man gelegentlich durch das sogenannte „Verstärkung-Bandbreite-Produkt“ aus. Das Verstärkung-Bandbreite-Produkt[8] ist eine charakteristische Größe eines OPVs und hängt nicht von der Gegenkopplung ab. Es gilt

$$\underbrace{A_{D_0} \cdot f_g}_{ohne\,Rückkopplung} \quad = \quad \underbrace{\frac{1}{K_R} \cdot f_g \, K_R \, A_{D_0}}_{mit\,Rückkopplung} \quad = \quad f_0 \quad . \tag{4.11}$$

Hier ist f_0 als die Frequenz definiert, bei der die Verstärkung A bzw. A_D den Wert 1 (= 0 db) annimmt.

4.14 Integrierer

Die Schaltung aus Abbildung 4.19(a), ist die aktive Implementierung des Tiefpasses oder Integrierglieds (Abb. 4.19(b)), das wir im Kapitel 1.14 besprochen haben.

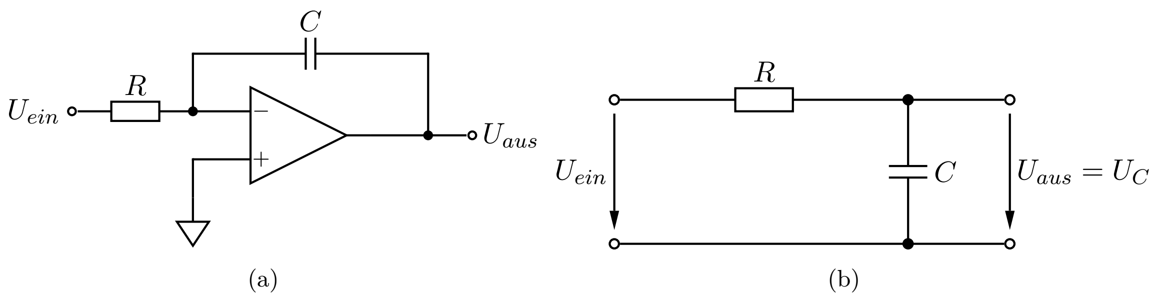


Abb. 4.19 (a) Integrierer. (b) Passiver Tiefpass.

[8]Engl. *gain bandwidth product* (GBP)

Im Vergleich zum invertierenden Verstärker ist hier statt des ohmschen Widerstands R_2 ein Kondensator C eingebaut. Für Eingangssignale einer festen Frequenz ω gilt die obige Ableitung mit $|Z| = \frac{1}{\omega C}$ statt R_2, also

$$\frac{U_{aus}}{U_{ein}} = -\frac{Z}{R} = -\frac{1}{\omega RC} \quad .$$

Hier ist $\frac{U_{aus}}{U_{ein}}$ das Amplitudenverhältnis von Eingangs- und Ausgangssignal. Die Phaseninformation ist nicht enthalten. Die Schaltung wirkt auch als Integrierer. Der Strom durch den Widerstand R lädt den Kondensator C auf. Damit z. B. bei einer positiven Spannung am Eingang der Schaltung U_- auf Masse liegen kann wie es der Goldenen Regel entspricht, muss sich U_{aus} absenken.

Analog zum invertierenden Verstärker gilt

$$\frac{U_{ein}}{R_1} = -C\,\frac{dU_{aus}}{dt}$$

und damit

$$\boldsymbol{U_{aus} = -\frac{1}{RC}\int U_{ein}\,dt} \quad . \tag{4.12}$$

Diese Beziehung gilt ohne die Einschränkung an die Integrationszeit $t \ll RC$, die beim passiven Integrierer auftritt. Allerdings kann auch hier die Ausgangsspannung nicht beliebig groß werden und ist durch die Versorgungsspannung limitiert.

Die Schaltung in Abbildung 4.19 spielt eine sehr große Rolle im Detektorbau und wird als ladungsempfindlicher Vorverstärker eingesetzt, z. B. bei der Auslese von Halbleiterdetektoren. Hier ist es wichtig den Kondensator zwischen aufeinanderfolgenden Signalen wieder zu entladen, z. B. durch einen parallel zum Kondensator geschalteten Widerstand oder aktiv über einen parallelgeschaltenen Transistor, der entsprechend gesteuert wird.

Der Name ladungsempfindlicher Vorverstärker erklärt sich leicht, wenn man die Eingangsspannungsquelle und den Widerstand R durch eine einfache Stromquelle ersetzt. Dann gilt

$$\boldsymbol{U_{aus} = -\frac{1}{C}\int I_{ein}\,dt} = -\frac{Q}{C} \quad , \tag{4.13}$$

wobei Q die im betrachteten Zeitintervall integrierte Ladung ist. Die Verstärkung dU_{aus}/dQ hat dann die Einheit [V/C] und beträgt $-1/C$. Typische Verstärkungen im Detektorbau liegen in der Größenordnung von einigen Dutzend mV/fC.

Frage: In welchem Bereich liegt dann der Wert des Kondensators C?

4.15 Differenzierer

Durch Vertauschen von Widerstand und Kondensator erhält man einen Differenzierer (Abb. 4.20). Hier gilt

$$\boldsymbol{U_{aus} = -RC\,\frac{dU_{ein}}{dt}} \quad . \tag{4.14}$$

Auch diese Schaltung wird in der Praxis gerne modifiziert, indem man einen Widerstand in Reihe zum Kondenstor C anbringt. Dies vergrößert die Phasenreserve und vermeidet mögliche Oszillationen bei schnellen Wechseln der Eingangsspannung.

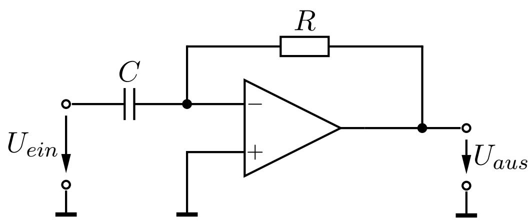


Abb. 4.20 Differenzierer.

4.16 Aktive Filter*

Aktive Filter sind nicht nur aus passiven Komponenten, sondern zusätzlich aus Operationsverstärkern als aktiven Bausteinen aufgebaut. Dies bietet im wesentlichen zwei Vorteile:

- Man kann mehrere Filterstufen in Reihe schalten (siehe Abbildung 4.21) ohne dass sich die Stufen gegenseitig belasten. Dadurch lassen sich leistungsfähige Filter realisieren.
- Man kann Spulen durch eine geeignete Kombinationen von Kondensatoren und einem OPV ersetzen und dadurch bei Anwendungen im Niederfrequenzbereich (< 1 MHz) viel Platz sparen.

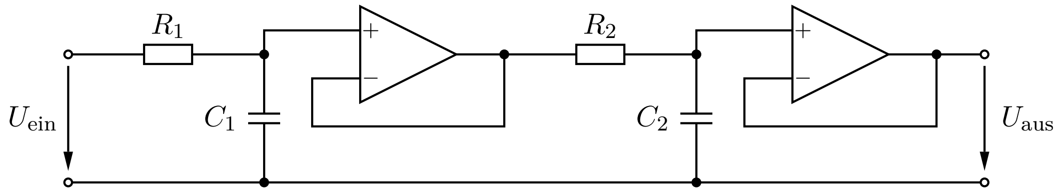


Abb. 4.21 Tiefpass zweiter Ordnung aus zwei passiven RC-Gliedern und zwei OPVs als Puffer.

Wir haben den aktiven Hochpass und Tiefpass erster Ordnung schon kennengelernt (Abb. 4.19 und 4.20). Es gibt zwei häufige Varianten aktiver Filter zweiter Ordnung, den Sallen-Key-Filter (Abb. 4.22) und den Multiple-Feed-back-Filter (Abb. 4.23). In beiden Varianten wird genau ein OPV verwendet.

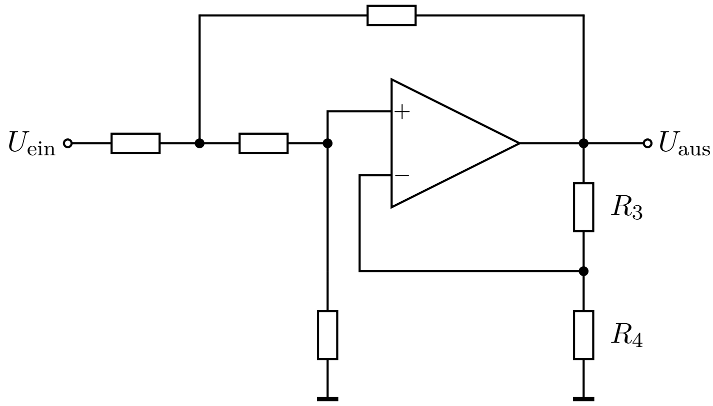


Abb. 4.22 Generischer Sallen-Key-Filter.

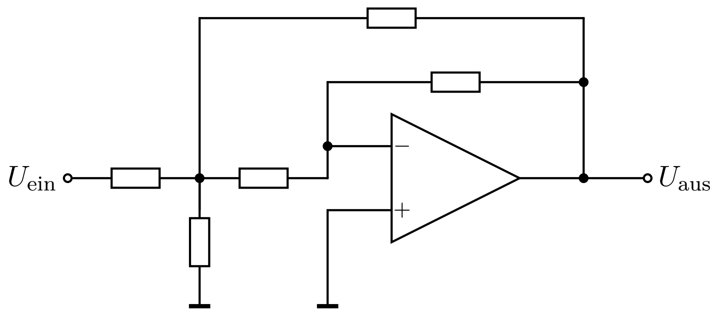


Abb. 4.23 Generischer Multi-Feedback-Filter.

Tiefpass als Sallen-Key-Filter

Allgemein gilt für jeden Tiefpass zweiter Ordnung

$$Z(s) = \frac{1}{as^2 + bs + 1} \quad .$$

Bei Realisierung mit Widerständen und Kondensatoren entsprechend Abb. 4.24 erhalten wir einen Tiefpass.

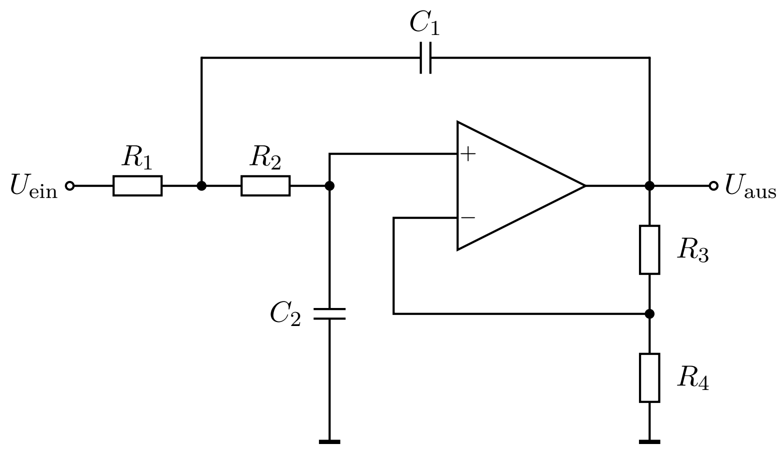


Abb. 4.24 Sallen-Key-Tiefpass.

Es gilt

$$a = C_1 C_2 R_1 R_2 = \frac{1}{\omega_C^2} \quad , \quad b = C_2 \cdot (R_1 + R_2) = \frac{1}{\omega_C\, Q} \quad ,$$

$$2\pi f_c = \omega_c = \frac{1}{\sqrt{R_1 R_2 C_1 C_2}} \quad \text{und} \quad Q = \frac{\sqrt{R_1 R_2 C_1 C_2}}{C_2 (R_1 + R_2)} \quad .$$

Das Verhalten des Sallen-Key-Tiefpasses ist plausibel. Bei kleinen Frequenzen $f \approx 0$ sperren die Kondensatoren und $Z(s) = 1$. Für hohe Frequenzen liegt U_+ über C_2 auf Masse, sodass nach den goldenen Regeln auch U_{aus} auf Masse gezogen wird. Wie sieht es bei mittleren Frequenzen aus? Die Schaltung ohne C_2 mit $R_2 = 0\,\Omega$ ist der uns bekannte Tiefpass erster Ordnung. Er wird mit R_2 und C_2 durch einen weiteren Tiefpass ergänzt. Es ist intuitiv plausibel, dass sich ein Tiefpass zweiter Ordnung ergibt.

Hochpass als Sallen-Key-Filter

Werden die Widerstände und Kondensatoren wie in Abbildung 4.25 plaziert, ergibt sich ein Hochpass.

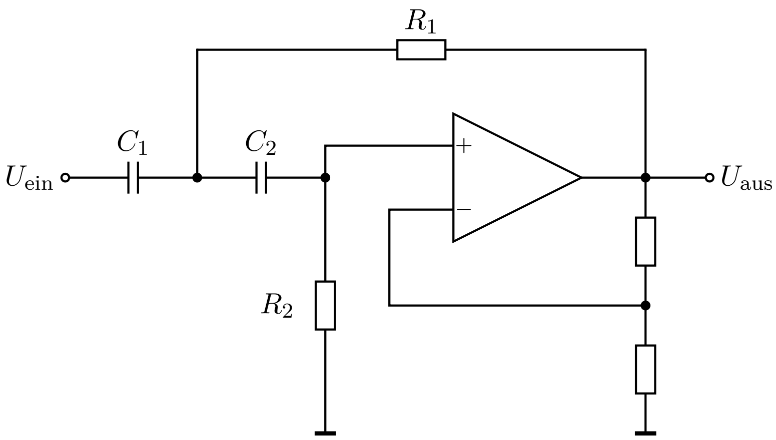


Abb. 4.25 Sallen-Key-Hochpass.

Allgemein gilt für jeden Hochpass zweiter Ordnung

$$Z(s) = \frac{s^2}{s^2 + as + b} \quad .$$

Bei Realisierung entsprechend Abb. 4.25 gilt

$$a = \frac{1}{R_2C_1} + \frac{1}{R_2C_2} \quad , \quad b = \frac{1}{R_1R_2C_1C_2}$$

$$2\pi f_c = \omega_c = \frac{1}{\sqrt{R_1R_2C_1C_2}} \quad \text{und} \quad Q = \frac{\sqrt{R_1R_2C_1C_2}}{R_2(C_1 + C_2)} \quad .$$

Die Grenzfrequenz hat also denselben Wert wie für den Sallen-Key-Tiefpass.

Man kann sich den Ausdruck gut merken, da er für $R_1 = R_2 = R$ und $C_1 = C_2 = C$ der Grenzfrequenz $\omega_c = \frac{1}{RC}$ des Tiefpasses und Hochpasses erster Ordnung entspricht.

Ein **Sallen-Key-Bandpass** ist in Abbildung 4.26 dargestellt.

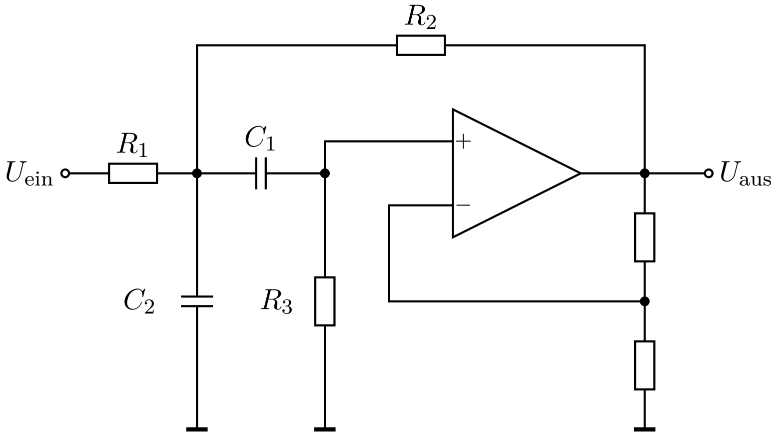


Abb. 4.26 Sallen-Key-Bandpass.

Filtertypen

Filter sind durch ihre Funktionalität (z.B. Tiefpass, Hochpass, etc.), die Filterordnung, die Filtertopologie (z.B. Sallen-Key oder Multi-Feedback) und den Filtertypus gekennzeichnet.

Hat man sich für eine Filtertopologie entschieden, müssen die Werte der Widerstände und Kondensatoren festgelegt werden. Insbesondere, wenn man Filter höherer Ordnung baut, ist es wichtig, dass die einzelnen Filter sich sinnvoll ergänzen. Es haben sich vier Filtertypen etabliert:

- der Butterworth-Filter
- der Tschebyscheff-Filter
- der Besser-Filter und der
- elliptische oder Cauer-Filter.

Diese Typen werden sowohl für Tiefpässe, Hochpässe, Bandpässe und Bandsperren eingesetzt. Wir diskutieren sie hier nur am Beispiel eines Tiefpasses.

In Abbildung 4.27 ist der Betrag der Übertragungsfunktion einiger Filter vierter Ordnung dargestellt.

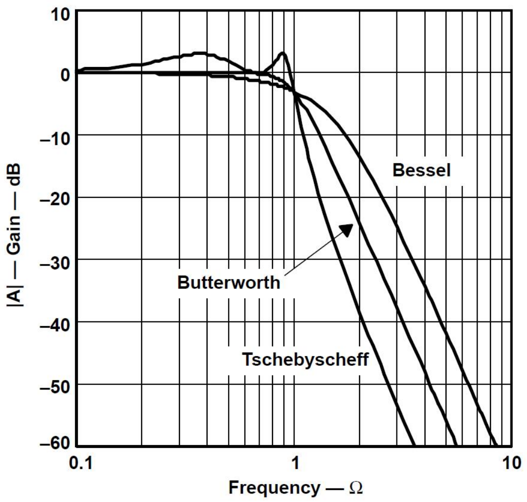


Abb. 4.27 Vergleich des Amplitudengangs verschiedener Tiefpassfilter 4. Ordnung [7].

Hier fällt folgendes auf: Die Übertragungsfunktion fällt sehr schnell mit der Frequenz ab, wie es sich für einen Tiefpass höherer Ordnung gehört (mindestens -60 db/Dekade). Die Abnahme ist unterschiedliche schnell. Der Durchlassbereich ($f < f_{-3\mathrm{db}}$) ist nicht für alle Filtertypen gleichmäßig flach. Auch die Phasenverläufe sind unterschiedlich (nicht gezeigt).

Der **Butterworthfilter** ist dadurch charakterisiert, dass er im Durchlassbereich von allem Typen am flachsten verläuft.

Der **Tschebyschefffilter** fällt schnell ab. Dafür variiert er aber im Durchlassbereich. (Dies muss nicht unbedingt ein Nachteil sein, falls das Eingangssignal nur wenig relevante Frequenzen in einem Bereich aufweist.)

Der **Besselfilter** fällt nicht so steil ab, wie die anderen Filter. Seine charakteristische Eigenschaft ist die lineare Abhängigkeit der Phasenverschiebung von der Frequenz. Dadurch ist die Phasengeschwindigkeit für $f < f_{-3\mathrm{db}}$ konstant und alle Frequenzanteile kleiner $f_{-3\mathrm{db}}$ um eine konstante Zeit verzögert. Die Verzögerung des Eingangssignals ist minimal.

Elliptische Filter fallen am steilsten ab. Dafür ist die Übertragungsfunktion weder im Durchlassbereich noch im Sperrbereich konstant, sondern wellig.

Die Übertragungsfunktionen dieser Filtertypen sind tabelliert (siehe Tabelle 4.3) und können auch mit geeigneten Programmen bestimmt und simuliert werden.

n	
1	$(1+s)$
2	$(1+1.414\,s+s^2)$
3	$(1+s)(1+s+s^2)$
4	$(1+0.765\,s+s^2)(1+1.848\,s+s^2)$
5	$(1+s)(1+0.618\,s+s^2)(1+1.618\,s+s^2)$
6	$(1+0.518\,s+s^2)(1+1.414\,s+s^2)(1+1.932\,s+s^2)$
7	$(1+s)(1+0.445\,s+s^2)(1+1.247\,s+s^2)(1+1.802\,s+s^2)$
8	$(1+0.390\,s+s^2)(1+1.111\,s+s^2)(1+1.663\,s+s^2)(1+1.962\,s+s^2)$
9	$(1+s)(1+0.347\,s+s^2)(1+s+s^2)(1+1.532\,s+s^2)(1+1.879\,s+s^2)$
10	$(1+0.313\,s+s^2)(1+0.908\,s+s^2)(1+1.414\,s+s^2)(1+1.782\,s+s^2)(1+1.975\,s+s^2)$

Tab. 4.3 Butterworthfilterpolynome für die Ordnungen n = 1 bis 10 für $\omega_c = 1$ rad.

Die Übertragungsfunktion eines Tiefpasses dritter Ordnung ergibt sich also zu

$$Z(s) = \frac{1}{(s+1)(s^2+s+1)} \quad .$$

Will man eine beliebige Grenzfrequenz erzielen, muss man in der obigen Filtertabelle s durch s/ω_c ersetzen.

Der praktische Nutzen von Tabelle 4.3 liegt darin, das durch das Hintereinanderschalten von aktiven Filtern 1. und 2. Ordnung komplexe Filter höherer Ordnung zusammensetzen kann. So kann man einen Filter sechster Ordnung durch drei aufeinander folgende Filter zweiter Ordnung aufbauen, und einen Filter siebter Ordnung durch einen zusätzlichen Filter erster Ordnung. Dem liegt mathematisch zugrunde, dass man jede komplexe Übertragungsfunktion als Verhältnis von Polynomen mit reellen Koeffizienten bzw. als Verhältnis der entsprechenden Linearfaktorzerlegungen schreiben kann:

$$\begin{aligned} Z(s) &= Z_0 \cdot \frac{s^n + \alpha_{n-1}s^{n-1} \dots + \alpha_1 s + \alpha_0}{s^n + \beta_{n-1}s^{n-1} \dots + \beta_1 s + \beta_0} && (4.15) \\ &= \frac{(s^2+a_1 s+b_1)(s^2+a_2 s+b_2)\dots}{(s^2+c_1 s+d_1)(s^2+c_2 s+d_2)\dots} && (4.16) \\ &= \frac{(s-z_0)(s-z_1)(s-z_2)\dots(s-z_n)}{(s-p_0)(s-p_1)(s-p_2)\dots(s-p_n)} \quad . && (4.17) \end{aligned}$$

Das beschriebene systematische Vorgehen ist vorteilhaft, aber natürlich könnte man alternativ auch Konfigurationen wie in Abb.4.28 betrachten.

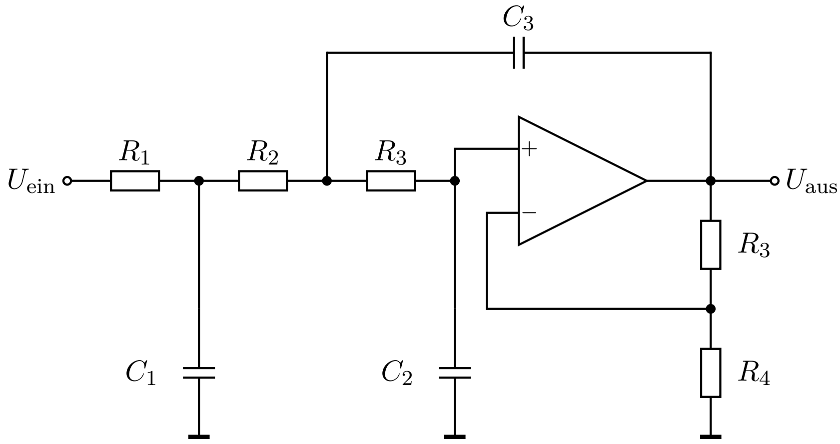


Abb. 4.28 Sallen-Key Filter 3. Ordnung mit einem OPV.

4.17 Differenzverstärker

Der Differenzverstärker (Abb. 4.29) wirkt komplizierter als der invertierende Verstärker, da der nichtinvertierende Eingang nicht auf Masse liegt.

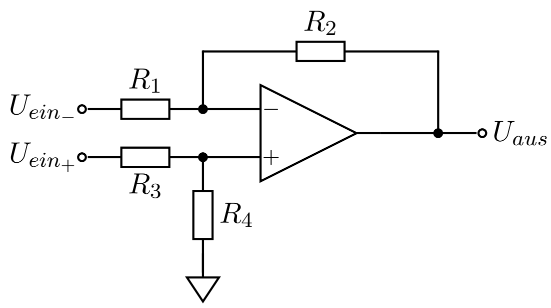


Abb. 4.29 Differenzverstärker.

Wir vernachlässigen wieder die Eingangsströme in den Verstärker, entsprechend der 2. goldenen Regel. Dadurch ergibt sich

$$U_+ = U_{ein_+} \frac{R_4}{R_3 + R_4} \quad \text{und} \quad U_- = U_{ein_-} \frac{R_2}{R_1 + R_2} + U_{aus} \frac{R_1}{R_1 + R_2}$$

Die zweite Beziehung erhält man durch Anwendung des Superpositionsprinzips auf den Schaltkreis mit den Widerständen R_1, R_2 und den beiden Spannungsquellen U_{ein_+} und U_{aus} (Abb. 4.30).

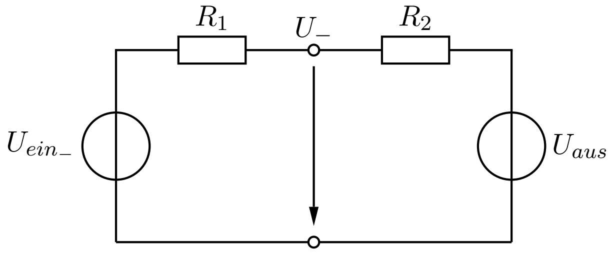


Abb. 4.30 Anwendung des Superpositionsprinzips.

Entsprechend der 1. goldenen Regel setzten wir $U_+ = U_-$. Es ergibt sich

$$U_{ein_+} \frac{R_4}{R_3 + R_4} = U_{ein_-} \frac{R_2}{R_1 + R_2} + U_{aus} \frac{R_1}{R_1 + R_2}$$

und

$$U_{ein_+} \frac{R_4 (R_1 + R_2)}{R_1 (R_3 + R_4)} - U_{ein_-} \frac{R_2}{R_1} = U_{aus} \quad .$$

Die Widerstandsverhältnisse der Eingänge U_{ein_+} und U_{ein_-} sind unübersichtlich und im Allgemeinen verschieden voneinander. Wählt man die Widerstände aber so aus, dass gilt

$$\frac{R_2}{R_1} = \frac{R_4}{R_3} = \alpha \quad ,$$

so gilt [9]

$$U_{aus} = (U_{ein_+} - U_{ein_-})\, \alpha \quad .$$

Am einfachsten setzt man $R_3 = R_1$ und $R_4 = R_2$ und verwendet Präzisionswiderständen bzw. Widerstände aus derselben Packung. Im Gegensatz zur Gleichung 4.1 ist die Verstärkung des Differenzverstärkers mit negativer Rückkopplung durch ein konstantes Widerstandsverhältnis gegeben.

4.18 Logarithmierer und Exponentierer*

Ein OPV mit einer Diode in der Rückkopplungsschleife (Abb. 4.31(a)) erzeugt eine zum Logarithmus der Eingangsspannung proportionale Ausgangsspannung. Wieso?

Der negative Eingang stellt eine virtuelle Masse da, und es gilt

$$U_{ein} = R\, I_{ein} \quad \text{und} \quad I_{ein} = I_D = I_S \left[\mathrm{e}^{-\frac{q_e U_{aus}}{k T}} - 1 \right] \quad .$$

Anmerkung: Wir definieren I_D als Vorwärtsstrom durch die Diode. Entsprechend definiert sich die Diodenspannung als $U_- - U_{aus} = -U_{aus}$.

Daraus folgt

$$\frac{U_{ein}}{R\, I_S} + 1 = \mathrm{e}^{-\frac{q_e U_{aus}}{k T}} \quad \text{und} \quad U_{aus} = -\frac{k\, T}{q_e} \ln \left(\frac{U_{ein}}{R\, I_S} + 1 \right) \approx -\frac{k\, T}{q_e} \ln \left(\frac{U_{ein}}{R\, I_S} \right) \quad .$$

[9] $\frac{R_4}{R_1} \frac{R_1+R_2}{R_3+R_4} = \frac{1+\alpha}{\frac{1}{\alpha}+1} = \alpha$

Durch Vertauschen von Diode und Widerstand (Abb. 4.31(b)) kann man die Exponentialfunktion der Eingangsspannung bilden. Jetzt gilt

$$I_{ein} = I_D = I_S \left[\mathrm{e}^{\frac{q_e U_{ein}}{k T}} - 1 \right] \quad \text{und} \quad I_R = I_{ein} = -\frac{U_{aus}}{R} \qquad .$$

Daraus folgt

$$U_{aus} \approx -R\, I_S\, \mathrm{e}^{\frac{q_e U_{ein}}{kT}} \qquad .$$

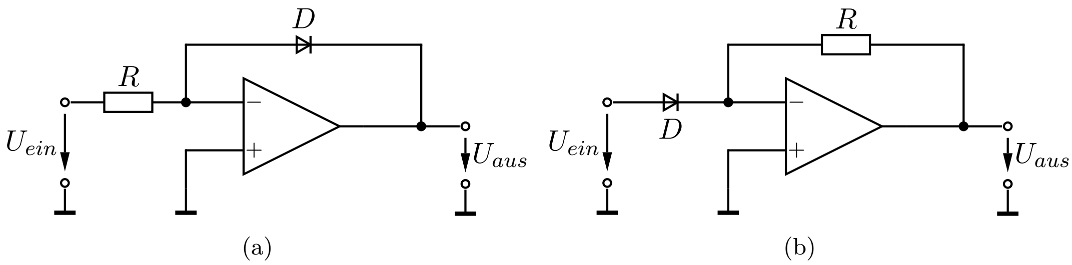


Abb. 4.31 (a) Logarithmierer. (b) Exponentierer.

Über den Umweg des Logarithmus, seiner Umkehrfunktion und der Addition kann man jetzt auch die analoge Multiplikation verwirklichen, da $\log a + \log b = \log (a\, b)$.
Die entsprechende Schaltung wird im Praktikum untersucht.

4.19 Driftkompensation

Der Eingangswiderstand des realen OPVs ist groß, aber endlich. Typische Eingangsströme variieren von einigen 1 pA für OPVs aus Feldeffekttransistoren bis zu einigen 10 nA für bipolare OVPs.

Größere Eingangsströme führen zu Abweichungen von den einfachen Formeln für Verstärkung, Eingangs- und Ausgangsimpedanzen, die aus den goldenen Regeln folgen. Viel wichtiger ist aber, dass ein endlicher Eingangsstrom bei geerdetem Eingang zu einer Differenzeingangsspannung am Eingang führt und damit eine von Null verschiedene Ausgangsspannung hervorruft.

Diesem Effekt kann man entgegenwirken, indem man an beide Eingänge gleichgroße externe Widerstände anbringt. Dies ist in Abbildung 4.32 für den invertierenden Verstärker illustriert. Der nichtinvertierende Eingang liegt nicht direkt an Masse, sondern ist durch den Widerstand R_3 getrennt.

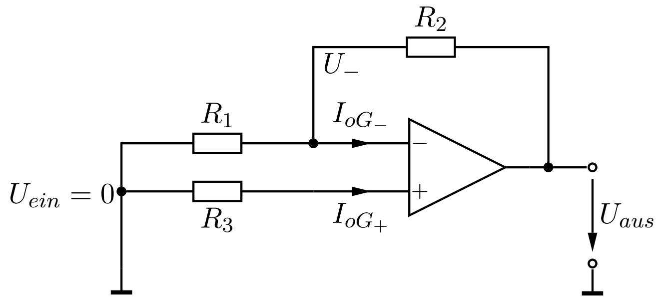


Abb. 4.32 Invertierender Verstärker mit Kompensation des Offsetstroms durch den Widerstand R_3.

Wir berechnen den Effekt, den ein an beiden Eingängen identischer Ruhestrom $I_{oG} = I_{oG\,-} = I_{oG\,+}$ bewirkt. Der Effekt eines Offsetstroms, also einer kleinen Abweichung der beiden Ströme ist dann oft vernachlässigbar.

Es gilt

$$I_{oG\,+} = -\frac{U_+}{R_3} \quad \text{und} \quad I_{oG\,-} = \frac{U_{0\,aus} - U_-}{R_2} - \frac{U_-}{R_1} \quad , \tag{4.18}$$

da wir den Ruhestrom, dass heißt den Strom für $U_{ein} = 0\,\mathrm{V}$ betrachten.

Auch dieser OPV versucht die beiden Eingänge auf dieselbe Spannung zu ziehen, also $U_+ = U_- = -I_{oG_+}\,R_3$. Einsetzen in 4.18 ergibt

$$I_{oG} = \frac{U_{0\,aus} + I_{oG}\,R_3}{R_2} + I_{oG}\,\frac{R_3}{R_1} \quad \text{und damit} \quad U_{0\,aus} = I_{oG}\,R_2 \left(1 - \frac{R_3}{R_1 \parallel R_2}\right) \quad .$$

Der Effekt eines symmetrischen Eingangsruhestroms wird also für $R_3 = R_1 \parallel R_2$ minimiert. Dafür zahlt man allerdings einen Preis. Der invertierende Eingang entspricht jetzt nur noch in Näherung einer (virtuellen) Masse. Der Widerstand führt auch zu Rauschen, deshalb ist es sinnvoll parallel zu R_3 einen Kondensator anzubringen und ihn für Wechselspannungen kurzzuschließen.

Entsprechend kann man auch den nichtinvertierenden Verstärker kompensieren. Hier muss man zwischen Signalquelle und dem nichtinvertierenden Eingang der OPVs einen Vorwiderstand R_G setzen mit $R_G = R_1 \parallel R_2$.

In beiden Fällen ist es auch hilfreich, den Rückkopplungswiderstand R_2 nicht zu groß zu wählen!

4.20 Oszillatoren

Oszillatoren haben viele Anwendungen. Insbesondere sind sie in der Nachrichtentechnik und im Audiobereich wichtig. Gute Oszillatoren zeichnen sich aus durch

- hohe Frequenzstabilität
- geringe Abweichungen von der idealen z.B. harmonischen Schwingung (Verzerrung)

Die folgenden Beispiele ermöglichen Schwingungsperioden von einigen kHz. Für hohe Frequenzen benutzt man Schwingquarzoszillatoren.

Wir wollen uns zuerst überlegen wie sich der Schmitttrigger (Abb. 4.33(a)) verhält, wenn wir den invertierenden Eingang nicht direkt auf Masse legen, sondern über einen Kondensator und einen Widerstand an den Ausgang koppeln (Abb. 4.33(b)).

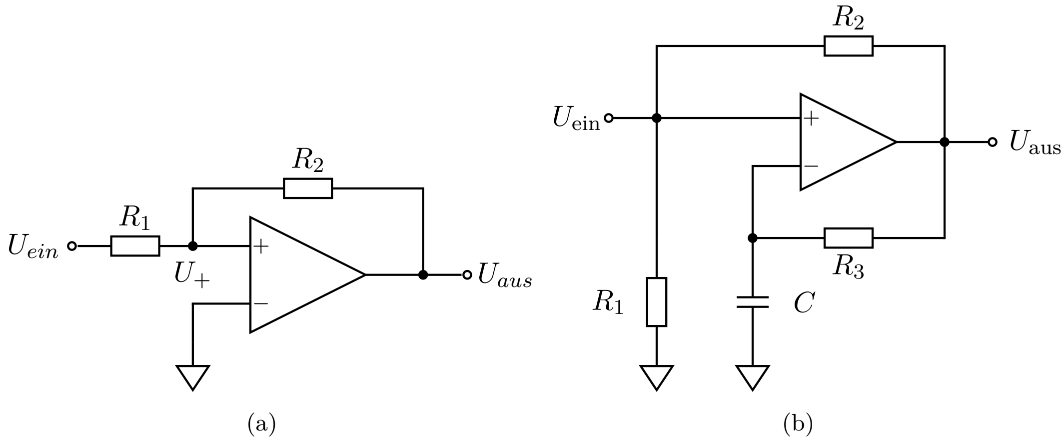


Abb. 4.33 (a) Schmitttrigger. (b) Rechteckpulsgenerator.

Wie funktioniert diese Schaltung? Für $t = 0$ sei $U_{aus} = 0$ V und $U_C = 0$ V, der Kondensator ist nicht geladen. Eine positive Spannung am Eingang bewirkt aufgrund der positiven Rückkopplung wie beim Schmitttrigger eine große Ausgangsspannung von $\approx +U_{CC}$. Dadurch wird der Kondensator aufgeladen und die Spannungsdifferenz am Eingang verringert. Sobald U_C den Wert $U_{CC}\,R_1/(R_1 + R_2)$ überschreitet, dreht sich das Vorzeichen von U_D um, und das Vorzeichen von U_{aus} ändert sich entsprechend. Durch die positive Rückkopplung springt der Ausgang schnell auf die Spannung $\approx -U_{CC}$. Jetzt wird der Kondensator mit einer Zeitkonstanten $\tau = RC$ entladen, bis $U_C < -U_{CC}\,R_1/(R_1 + R_2)$ ist, und das Vorzeichen von U_D wieder wechselt.

Die Schaltung erzeugt periodische Rechteckpulse. Die Periode lässt sich elementar ausrechnen und beträgt

$$T = 2\,R_3\,C\,\ln\left(\frac{2R_1 + R_2}{R_2}\right) \quad .$$

4.21 Phasenschieberoszillator*

Die Schaltung in Abbildung 4.34 eignet sich zur Erzeugung von harmonischen Schwingungen. Der linke OPV ist als Integrator geschaltet. Sein invertierender Eingang ist über den Widerstand R_1 an die Ausgangsspannung U_{aus} des rechten OPVs, eines Spannungsfolgers, geschaltet. Die Schaltung hat keinen externen Eingang. Zwischen den OPVs liegen zwei Tiefpässe. Diese Schaltung wird im Praktikum untersucht.

Die Schaltung arbeitet als Oszillator mit Periodendauer T, falls die Bedingung $U_{aus}(t) = U_{aus}(t + T)$ erfüllt ist. Das heißt nach einem vollen Durchlauf durch den Verstärker samt der

Rückkopplungsschleife ist die Phase und Amplitude des Ausgangssignals unverändert. Wann ist dies der Fall?

Der Spannungsfolger beeinflusst die Amplitude und die Phase nicht: $U_{aus} = U_3$. Der Integrator verstärkt das Signal U_{aus} und dreht die Phase um $+90°$ (die Invertierung entspricht $\pm 180°$ Phasenverschiebung, die Integration $-90°$). Die beiden Tiefpässe verschieben im Oszillationsfall die Phase um jeweils $-45°$, wie wir gleich sehen werden.

Wir berechnen zuerst den Effekt der Tiefpässe, die wir als doppelten Spannungsteiler mit Impedanzen $Z = \frac{1}{\mathrm{j}\,\omega\,C}$ und R ansehen. Das ergibt

$$U_3 = U_2 \frac{Z}{Z+R} = U_1 \frac{Z}{Z+R} \frac{Z \parallel (R+Z)}{R + Z \parallel (R+Z)} = \frac{U_1}{1 + \left(\frac{R}{Z}\right)^2 + 3\,\frac{R}{Z}} \quad . \tag{4.19}$$

Der Integrator impliziert für $U_{aus}(t) = U_{aus} \exp(j\,\omega\,t)$

$$U_1 = -\frac{1}{R_1\,C_1} \int U_{aus}(t)\,dt = -\frac{1}{\mathrm{j}\,\omega\,R_1\,C_1}\,U_{aus} \quad . \tag{4.20}$$

Fordert man wie oben $U_{aus}(t=0) = U_{aus}(T) = U_3(T)$ und ersetzt man U_1 durch Kombination von Gleichung 4.19 und 4.20:

$$U_3(T) = \frac{U_1(T)}{1 + \left(\frac{R}{Z}\right)^2 + 3\,\frac{R}{Z}} = -\frac{U_{aus}(T)}{1 + \left(\frac{R}{Z}\right)^2 + 3\,\frac{R}{Z}} \frac{1}{\mathrm{j}\,\omega\,R_1\,C_1} \quad ,$$

so folgt damit die Bedingung

$$3\,\omega^2\,RC\,R_1\,C_1 - \mathrm{j}\,R_1\,C_1\,\left[1 - (\omega\,R\,C)^2\right] = 1 \quad . \tag{4.21}$$

Der Imaginärteil der linken Seite muss identisch 0 sein. Also gilt

$$\omega = \frac{1}{R\,C} \quad . \tag{4.22}$$

Daraus folgt dann

$$R_1\,C_1 = \frac{R\,C}{3} \quad . \tag{4.23}$$

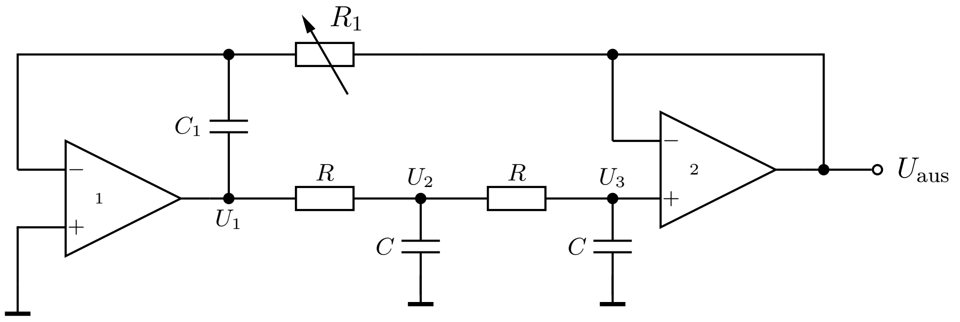


Abb. 4.34 Phasenschieberoszillator mit OPVs 1 und 2, zwei Tiefpässen und einem einstellbaren Widerstand R_1.

Fragen: Wozu braucht man den Spannungsfolger? Wie bringt man den Oszillator zum Schwingen und vermeidet den stationären Zustand $U_{aus} = 0$? Wie sieht das Einschwingverhalten des Oszillators aus und was bestimmt die Amplitude der Schwingung.

Es gibt viele Spielarten des Phasenschieberoszillator. Eine andere Variante mit nur einem OPV und drei Tiefpässen ist in Abb. 4.35 dargestellt.

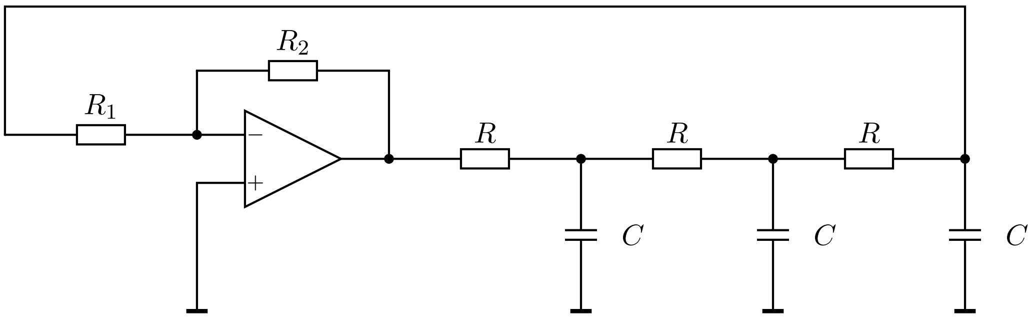


Abb. 4.35 Phasenschieberoszillator mit einem invertierenden OPV und drei Tiefpässen.

Die Resonanzfrequenz liegt hier bei $\omega \approx 1,73/RC$. Warum ?

4.22 Wien-Brücken-Oszillator*

Die Schaltung in Abbildung 4.36 ist ein weiterer beliebter, guter und einfach aufgebauter Oszillator für harmonische Schwingungen. Der Ausgang des OPV ist über einen Spannungsteiler R_{10}, R_{11}, R_{12} an den invertierenden Eingang gekoppelt. Dies bewirkt eine frequenzunabhängige Verstärkung. Der nichtinvertierende Eingang ist über einen Bandpass, bestehend aus den Widerständen R_1, R_2 und den Kondensatoren C_1, C_2 an den Ausgang gekoppelt.

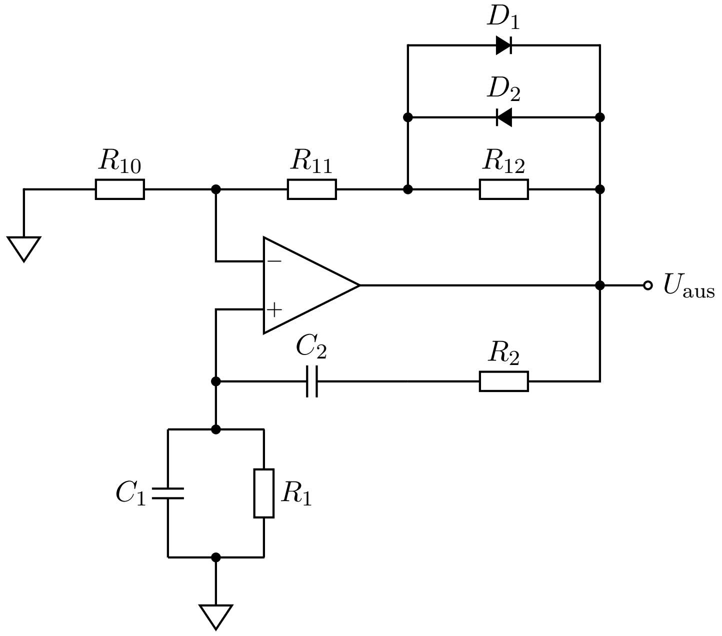


Abb. 4.36 Wien-Brücken-Oszillator.

Der Bandpass, der in Abb. 4.37 getrennt dargestellt ist, ist durchlässig für Frequenzen zwischen $\omega_1 = \frac{1}{R_1C_1}$ und $\omega_2 = \frac{1}{R_2C_2}$, falls $\omega_1 < \omega_2$. Es gilt mit der Vereinfachung $R_1 = R_2 = R$ und $C_1 = C_2 = C$

$$\begin{aligned}
U_{\text{ein}} &= U_{\text{aus}} \frac{Z||R}{Z+R+Z||R} = U_{\text{aus}} \frac{\frac{ZR}{Z+R}}{Z+R+\frac{ZR}{Z+R}} = U_{\text{aus}} \frac{ZR}{(Z+R)^2+ZR} \\
&= U_{\text{aus}} \frac{\frac{R}{j\omega C}}{(\frac{1}{j\omega C}+R)^2+\frac{R}{j\omega C}} = U_{\text{aus}} \frac{R}{\frac{(1+j\omega RC)^2}{j\omega C}+R} = U_{aus} \frac{j\omega RC}{1+3j\omega RC-\omega^2R^2C^2} \\
&= \frac{j\omega RC(1-\omega^2R^2C^2-3j\omega RC)}{(1-\omega^2R^2C^2)^2+9\,\omega^2R^2C^2} \quad .
\end{aligned}$$

Dieser Ausdruck strebt für $\omega \to 0$ und $\omega \to \infty$ gegen 0. Der Ausdruck hat für $\omega = \frac{1}{RC}$ ein Maximum bei $U_{ein} = U_{aus}/3$. Gleichzeitig ist hier der Imaginärteil und damit die Phasenverschiebung nach dem Durchlaufen der Rückkopplungsschleife null.

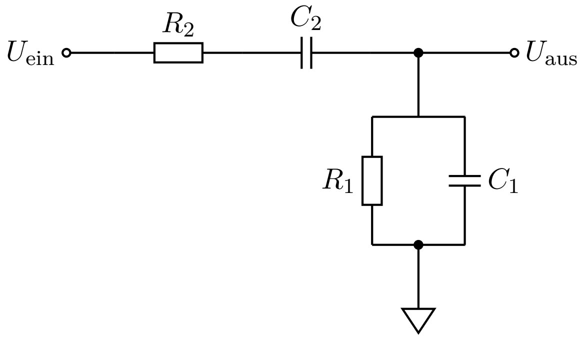


Abb. 4.37 Bandpass des Wien-Brücken-Oszillators. Hier bezeichnet U_{ein} die Spannung am Ausgang als Operationsverstärker und U_{aus} die Spannung am nichtinvertierenden Eingang des Operationsverstärkers.

Den Wien-Brücken-Oszillator versteht man am Besten, wenn man die Schaltung als nichtinvertierenden Verstärker interpretiert. Statt einer externen Signalquelle wird hier allerdings das Ausgangssignal über den Bandpass auf einen nichtinvertierenden Eingang des Operationsverstärkers gelegt. Falls sich die Abschwächung des Bandpasses und die Verstärkung aufgrund der Widerstände R_{10}, R_{11}, R_{12} am invertierenden Eingang kompensieren, entsteht eine stabile Oszillation. Das ist bei

$$\frac{R_{10}}{R_{11} + R_{12}} = 3$$

gegeben. Um Instabilitäten und eine entweder zu kleine oder zu große Verstärkung zu vermeiden, kann man diese Schaltung durch ein weiteres Rückkopplungselement stabilisieren. Dies können wie in Abb. 4.36 Dioden sein. R_{12} wird so gewählt, dass $R_{10}, R_{11} << R_{12}$. Die Dioden verhindern, dass die Verstärkung den Wert $A = 3$ überschreitet, indem sie bei richtiger Dimensionierung der Widerstände den Widerstand R_{12} kurzschließen bzw. effektiv verringern. Damit verringert sich auch die Verstärkung, etc.

4.23 Phasenreserve von Verstärkern*

Eine Verstärkerschaltung kann relativ leicht ins Schwingen kommen, falls das der Design der Schaltung nicht robust genug ist oder die Last bzw. die Signalquelle nicht den Annahmen entsprechen. Ausgehend von der Verstärkung (entsprechend Gl. 4.9)

$$A = \frac{A_D}{1 + \beta\, A_D}$$

können wir drei wesentliche Szenarien für OPVs mit negativer Rückkopplung unterscheiden

1. Die Schleifenverstärkung $|\beta\, A_D| >> 1$: Hier ist die Rückkopplung stark und damit die Verstärkung des OPV begrenzt. Es gilt

   $$A = \frac{A_D}{1 + \beta\, A_D} \approx \frac{1}{\beta}$$

   Der OPV ist stabil und schwingt nicht.

2. Die Schleifenverstärkung $|\beta\, A_D| << 1$: Hier ist die Rückkopplung schwach und die Verstärkung $A \approx A_D$ sehr groß. Dies ist für Verstärker nicht ideal, aber gut für Komparatoren bzw. Schmitttrigger.

3. $|\beta\, A_D| \approx 1$: Jetzt müssen wir das Vorzeichen bzw. die Phase von $\beta\, A_D$ beachten. Falls $\beta\, A_D = +1$, entspricht dies in etwa der Bedingung 2.

   Falls $\beta\, A_D = -1$ (Barkhausensches Oszillationskriterium) schwingt der OPV. Dies ist nur bei Oszillatoren erwünscht und muss bei Verstärkern zuverlässig vermieden werden.

Warum ist der Wert $\beta\, A_D = -1$ problematisch?

Hier wird die Verstärkung sehr groß, da der Nenner in Gleichung 4.23 divergiert. Wir können die Signalspannung des OPV vernachlässigen und es gilt

$$U_{aus}(T) = -\beta\, A_D\, U_-(t=0) = -\beta\, A_D\, U_{aus}(t=0) = +U_{aus}(t=0)$$

Dies entspricht der Oszillationsbedingung der vorangegangenen Abschnitte.

Man kann mithilfe des Bodediagramms anschaulich machen, wann diese Bedingung erfüllt ist (siehe Abb. 4.38).

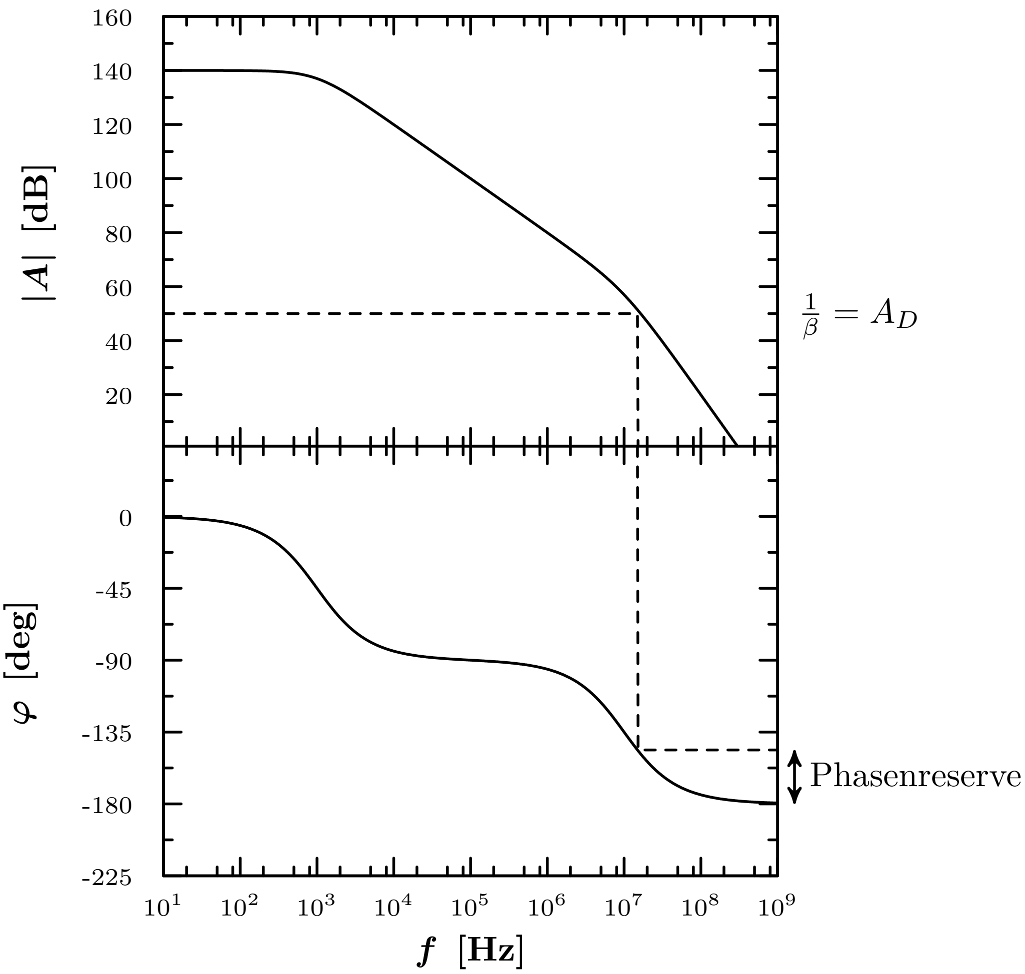


Abb. 4.38 Bodediagramm zur Illustration der Phasenreserve. Die durchgezogenen Kurve im oberen Bild zeigt die Leerlaufverstärkung A_D eines OPV mit einem dominanten Pol bei $\omega_1 \approx 1$ kHz und einem weiteren Pol bei $\omega_2 \approx 1$ MHz. Die gestrichelte Linie zeigt die Verstärkung $A = 1/\beta$ mit Rückkopplung an. Im unteren Bild ist die Phasenverschiebung und die Phasenreserve dargestellt.

Hier ist das Bodediagramm eines OPV gezeigt, dessen Lehrlaufverstärkung zunächst wie ein einfacher Tiefpass mit der Grenzfrequenz ω_1 abfällt, dann aber einen zweiten Pol bei der Frequenz ω_2 zeigt. Dadurch fällt zum einen die Verstärkung bei großen Frequenzen $\omega > \omega_2$ schneller ab (-40 db/Dekade). Zum anderen beträgt die Phasenverschiebung bei hohen Frequenzen 180°.

Der Schnittpunkt der Graphen *Leerlaufverstärkung* A_D und *Verstärkung mit Rückkopplung* $A \approx \frac{1}{\beta}$ entspricht der Bedingung $|\beta\, A_D| = 1$. Falls der Wert der Phase am Schnittpunkt der Graphen bei $\phi = 180°$ liegt, impliziert dies $\beta\, A_D = -1$.

Anmerkung: Die beiden Pole entstehen zum einen durch den dominanten Pol in dem frequenzkompensierten OPV und durch (mindestens) einen weiteren höheren Pol innerhalb des OPV oder einen weiteren Pol in der Beschaltung des OPV.

Um die Stabilität eines OPV in einer gegebenen Schaltung zu beurteilen, finden wir zuerst die Frequenz, bei der sich A und A_D schneiden. Die Differenz zwischen der Phase bei dieser Frequenz und dem kritischen Wert $\phi = 180°$ bezeichnet man als *Phasenreserve*. Die Phasenreserve sollte 45° nicht unterschreiten.

4.24 Negative Widerstände*

Ein ohmscher Widerstand genügt dem ohmschen Gesetz. Strom und Spannung sind proportional zueinander und der Strom fließt von einem positiven zu einem negativen Potential. Bei einem negativen Widerstand ist die Stromrichtung umgekehrt!

Wie ist das möglich? Strom und Spannung genügen den Maxwellgleichungen (siehe Appendix) und den Gesetzen der Elektrodynamik. Einen negativen Widerstand kann man nur durch einen Trick, nämlich ein speziell geschaltetes aktives Bauteil erzeugen, einen sogenannten „Negative Impedance Converter“ (NIC). Die Beschaltung des OPV ist interessant, da hier gleichzeitig negative und positive Rückkopplung vorliegt (Abb. 4.39).

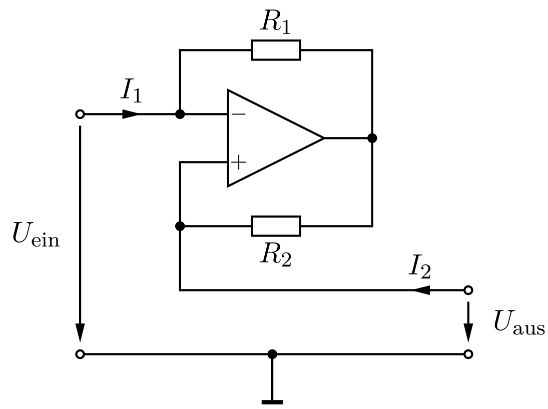


Abb. 4.39 Schaltkreis eines NICs.

Schließt man an den Ausgang des NIC eine Impedanz Z an, so ist die Eingangsimpedanz $-Z$. Um dies zu überprüfen, legen wir eine Testspannung U_{ein} an und berechnen den resultierenden Eingangsstrom I_1.

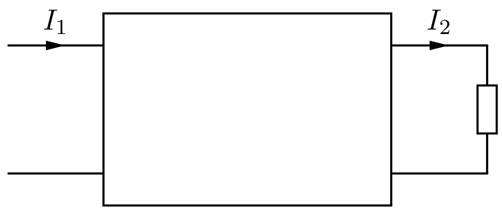


Abb. 4.40 Ersatzschaltbild eines NIC mit Last Z.

Wir gehen davon aus, dass der OPV nicht in Sättigung ist und dass die negative Rückkopplung dominiert. Es ist plausibel, dass diese Bedingung erfüllt ist, falls $R_2 > R_1$ ist. (Genauer gesagt muss gelten: $R_i/R_1 > Z/R_2$, wobei R_i der Innenwiderstand der Spannungsquelle U_{ein} ist.) Es gilt

$$U_{\text{ein}} = U_+ = U_{aus}\,\frac{Z}{Z+R_2} \quad .$$

Mit den goldenen Regeln gilt

$$U_+ = U_S = U_- \quad \text{und} \quad R_1 I_1 = R_2 I_2 \quad .$$

Gleichzeitig gilt

$$U_{aus} = U_{ein} - R_1\,I_1 = U_{ein}\,\frac{Z+R_2}{Z} \quad , \quad -\,R_1\,I_1 = U_{ein}\,\frac{R_2}{Z} \tag{4.24}$$

und damit

$$Z_{ein} = \frac{dU_{ein}}{dI_{ein}} = -Z\,\frac{R_1}{R_2} \quad . \tag{4.25}$$

Für $R_1 = R_2$ sieht man am Eingang des NICs also die negative Ausgangsimpedanz. Umgekehrt würde man am Ausgang die negative Eingangsimpedanz sehen.

Man kann mit negativen Widerständen so rechnen wie mit positiven. Die Formeln zur Berechnung von Parallel- und Reihenschaltung bleiben gültig.

Wozu sind NICs gut? Man kann mit einem negativen Widerstand den Effekt eines positiven Widerstands kompensieren. Zum Beispiel könnte man den Innenwiderstand einer Spannungsquelle verringern oder den Innenwiderstand einer Stromquelle vergrößern (Abb. 4.41).

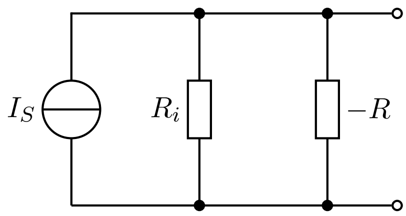


Abb. 4.41 Negativer Widerstand angeschlossen an Stromquelle I_S.

Wählt man z. B. den Widerstand des NICs so, dass $R_{NIC} = -\,R$ sich R_i annähert und dabei gilt $|R_{NIC}| > R_i$ gilt, so ergibt sich

$$lim_{R\to R_i}\,R_i \,\|\, -R = lim_{R\to R_i}\,\frac{R_i\,(-\,R)}{R_i - R} = \infty \quad .$$

4.25 Gyrator*

Die vermutlich wichtigste Anwendung von NICs ist der Gyrator. Mit dem Gyrator kann man z. B. Kapazitäten in Induktivitäten umwandeln und umgekehrt. Das ist in Hochfrequenzschaltungen mitunter hilfreich, wenn kein Platz für konventionelle Spulen zur Verfügung steht.

Ein idealer Gyrator ist in Abbildung 4.42 dargestellt. Es ist ein Vierpol mit zwei Eingängen und zwei Ausgängen.

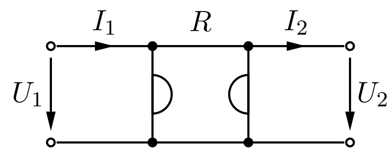


Abb. 4.42 Schaltsymbol eines Gyrators.

Das interessante am Gyrator ist die Beziehung zwischen den Strömen und Spannungen am Eingang und Ausgang. Es gilt

$$I_1 = 0 \cdot U_1 + \frac{1}{R} U_2 \quad \text{und} \quad I_2 = \frac{1}{R} U_1 + 0 \cdot U_2 \quad . \tag{4.26}$$

Damit hängt der Eingangstrom I_1 also nur von der Ausgangsspannung und der Ausgangsstrom nur von der Eingangsspannung ab. Gibt es einen Vierpol mit diesen Eigenschaften? Ja. Eine entsprechende Schaltung aus zwei NICs ist in Abbildung 4.43 dargestellt. Hier sind die Widerstände der zwei NICs gleich groß, $R_1 = R_2 = R$.

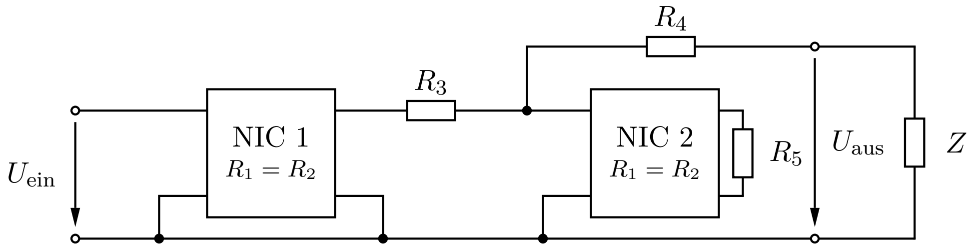


Abb. 4.43 Gyratorschaltung aus zwei NICs.

Diese Schaltung hat wie gewünscht zwei Eingangs- und Ausgangspole. Um das Verhalten des Gyrators zu verstehen, schließen wir am Ausgang eine Impedanz Z an. Wie groß ist die Eingangsimpedanz Z_{ein}?

Wir gehen von recht nach links vor. Der NIC N_2 hat die Eingangsimpedanz $-R_5$. Diese Eingangsimpedanz liegt parallel zu $R_4 + Z$, und $[-R_5 \| (R_4 + Z)]$ liegt in Serie zu R_3 am Ausgang von N_1. Damit ergibt sich die Eingangsimpedanz des NIC N_1 zu

$$-\{\,(R_3 + [-R_5 \| (R_4 + Z)])\} \quad .$$

Insgesamt gilt also mit $R_3 = R_4 = R$

$$Z_{ein} = -\{\,(R_3 + [-R_5 \| (R_4 + Z)])\} = \frac{R^2}{Z} \quad . \tag{4.27}$$

Überprüfen Sie die elementare Rechnung in 4.29.

Was bedeutet das? Sei Z eine Kapazität mit $Z = \frac{1}{\mathrm{j}\,\omega\, C}$, dann ist $Z_{ein} = \mathrm{j}\, R^2\, \omega\, C = j\, \omega\, L$ mit $L = R^2\, C$ und hat damit effektiv das Frequenzverhalten einer Induktivität L. Durch den Vorfaktor R^2 kann diese Induktivität recht groß werden! Mit $C = 10\,\mu\text{F}$ und $R = 10\,\text{k}\,\Omega$ wird $L = 1\,\text{kH}$.

Schaltet man an den Eingang von N_1 einen Kondensator C_1, so hat man einen Schwingkreis gebaut. Aus $\tau = \sqrt{L\,C_1} = R\,\sqrt{C_1\,C}$ sieht man, dass sich sehr große Schwingungsperioden von Sekunden oder größer erzielen lassen.

Nebenrechnung zu Gleichung 4.27

$$\begin{aligned} Z_{\text{ein}} &= -\{R_3 + [-R_5||(R_4 + Z)\} \\ &= -\{R + [-R||(R + Z)]\} \\ &= -\{R + \frac{[-R(R + Z)]}{-R + R + Z}\} \\ &= -\{\frac{RZ}{Z} - \frac{R^2 + RZ}{Z}\} \\ &= \frac{R^2}{Z} \end{aligned} \tag{4.28}$$

Alternative Rechnung mittels Gl. 4.26

Aus $U_2 = Z\,I_2$ folgt $I_1 = \frac{Z}{R} \cdot I_2 = \frac{Z}{R} \cdot \frac{U_1}{R}$ und damit $Z_{ein} = \frac{R^2}{Z}$.

5 Transistor: Grundlagen und 1T-Schaltungen

5.1 Einleitung

Die Entdeckung des Transistors gehört zu den großen Errungenschaften des 20. Jahrhunderts[1] und hatte weitreichende Konsequenzen. Der Transistor ist zentraler Baustein fast aller analogen und digitalen Schaltungen und ICs. Ganze Industriebereiche und Landstriche leben von der Produktion und Weiterentwicklung von immer kleineren, leistungsfähigeren und billigeren Transistorstrukturen. Auch die Informationsgesellschaft mit e-mail, Internet und vielleicht dem Internet der Dinge sind ohne den hohen Stand der Halbleiter- und Transistortechnologien undenkbar. Der Transistor ist deutlich komplexer als Widerstand, Kondensator, Spule und Diode, und das Verständnis von elementaren Transistorschaltungen ist ein wichtiges Ziel der Vorlesung.

Es gibt drei grundlegende Transistortypen:

- bipolarer Transistor (BJT) [2]
- Junction Field Effect Transistor (JFET)
- Metal Oxid Semiconductor Field Effect Transistor (MOSFET).

Der Transistor hat drei Pole. Sie werden abhängig vom Transistortyp als Basis, Kollektor und Emitter (bipolarer npn-Transistor) oder als „Gate“, „Source“ und „Drain“ bezeichnet (JFET oder MOSFET). Die gängigen Symbole für einen npn- und einen NMOS-Transistor sind in Abbildung 5.1 dargestellt.

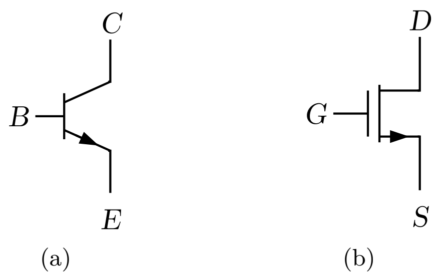


Abb. 5.1 (a) Symbol eines npn-Transistors mit Basis B, Kollektor C und Emitter E. (b) Schaltbild eines NMOS-Transistors mit Gate G, Drain D und Source S.

Der Transistor ist wie auch die Diode ein nichtlinearer Baustein. Im Gegensatz zu Widerstand, Kondensator, Spule und Diode ist er ein aktiver Baustein. Das heißt, mit Transistoren kann

[1] Nobelpreis in Physik für Shockley, Bardeen und Brattain im Jahr 1956
[2] Engl. *Bipolar Junction Transistor*

man Ströme und Spannungen verstärken.

Trotz großer Unterschiede fungieren alle Transistortypen als:

- Schalter
- Stromverstärker
- Spannungsverstärkung
- regulierbarer Widerstand
- Grundelement digitaler Logik

Wir betrachten zuerst ausführlich den bipolaren Transistor und seine Grundschaltungen.

5.2 Funktionsweise des Bipolartransistors

Der Transistor besteht aus drei unterschiedlich dotierten Bereichen (Abb. 5.2) mit einem pn- und einem np-Übergang. Im Normalbetrieb ist die Basisspannung eines npn-Transistors größer als die Emitterspannung, also $U_B > U_E$.

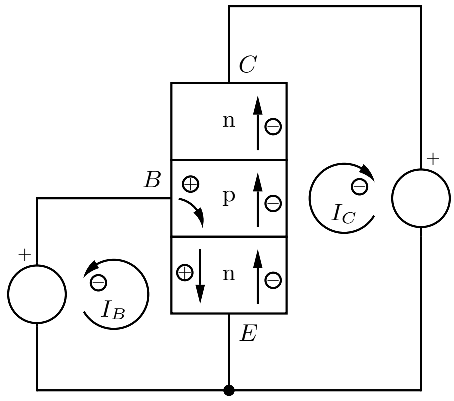


Abb. 5.2 Funktionsweise eines npn-Transistors (Quelle: H. Spieler).

Der Basis-Emitter-Übergang entspricht dann einer Diode in Durchlassrichtung und es fließt ein Basisstrom I_B in den Emitter. Wie bei einer Diode setzt sich der Strom durch den pn-Übergang aus einem Lochstrom aus der p-dotierten Basis in den Emitter und einem Elektronenstrom aus dem Emitter in die Basis zusammen. Sobald die Elektronen die Basis erreichen, haben sie allerdings im Gegensatz zu einer Diode nun auch die Möglichkeit, in den Kollektor zu fließen. Dies setzt eine positive Kollektorspannung $U_C > U_B$ voraus.

Das Verhältnis von Loch- und Elektronenstrom und damit von Basisstrom I_B und Kollektorstrom I_C ist durch die Dotierung von Basis und Emitter bestimmt. Der Emitter ist sehr viel stärker dotiert als die Basis. Für die Dichte der Akkzeptoratome $N_{A_{Basis}}$ und die Dichte der Donatoratome des Emitters $N_{D_{Emitter}}$ gilt also $N_{A_{Basis}} \ll N_{D_{Emitter}}$ und damit

$$\frac{I_C}{I_B} \sim \frac{N_{D_{Emitter}}}{N_{A_{Basis}}} \gg 1 \quad .$$

Der Kollektor-Basis-Übergang entspricht einer in Sperrrichtung geschalteten Diode. Je kleiner die Dicke der Basis, desto leichter können die Emitterelektronen die Basis durchqueren. Das Erhöhen der Spannung U_{CB} vergrößert die Sperrschicht und verringert die effektive Basisdicke.

Die grundlegenden Schaltungen mit einem Transistor sind in Abbildungen 5.9 - 5.10 dargestellt.

5.3 Kennlinien des Bipolartransistors

Die Transistorenkennlinien beschreiben das Verhalten des Transistors und sind oft alles, was man zum Schaltungsentwurf braucht. Da der Transistor drei Pole hat, wird die Darstellung der Kennlinien etwas komplizierter als bei der Diode.

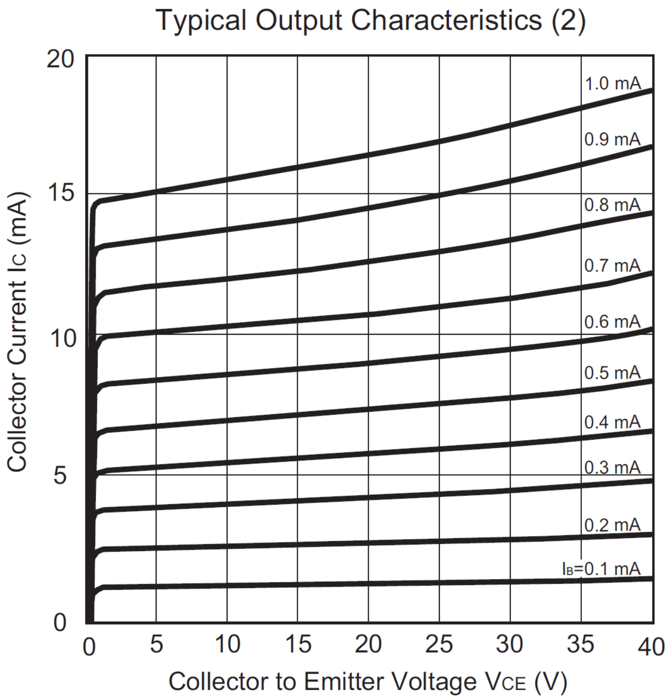


Abb. 5.3 Kollektorstrom I_C als Funktion der Kollektor-Emitter-Spannung U_{CE} für verschiedene Basisströme (2N3904)[8].

Wir diskutieren zuerst den Kollektorstrom I_C als Funktion der Kollektor-Emitter-Spannung U_{CE} für eine Schar von Basisströmen I_B (Abb. 5.3). Der genaue Verlauf der Kurven hängt von dem Transistortyp, der Transistorgeometrie und der Dotierung ab. Allgemein gilt aber, dass für sehr kleine Basisströme ein sehr großer Kollektorstrom erzielt werden kann. Die Abhängigkeit des Stromes von der Spannung U_{CE} zwischen Kollektor und Emitter ist eher klein, zumindest für nicht zu kleine Werte von U_{CE}. Der flache Bereich der I_C-U_{CE}-Kennlinie ist der bevorzugte Arbeitsbereich von Transistoren. Er wird auch der „aktive Bereich“ genannt.

Der Bereich sehr kleiner Spannungen $U_{CE} < 0.2\,\mathrm{V}$ ist der sogenannte Sättigungsbereich. Hier ist die Kollektorspannung nicht groß genug, um die Elektronen von der Basis in den Kollektor zu ziehen. Stattdessen rekombinieren die meisten Elektronen mit den Löchern der Basis und der Kollektorstrom ist (für einen festen Basisstrom) sehr viel kleiner als bei großen U_{CE}.

Wie für alle Bausteine gibt es Maximalwerte für Ströme, Spannungen und Leistung. In Abbildung 5.3 entspricht eine konstante Leistung einer Hyperbel. Zu große Werte dauerhaft oberhalb der entsprechenden Kurve zerstören den Transistor. Dasselbe gilt natürlich auch für zu hohe Eingangs- und Versorgungsspannungen[3], die allerdings durch Schutzdioden abgeleitet werden sollten.

In Abbildung 5.4 ist der Kollektorstrom gegen die Spannung zwischen Basis und Emitter U_{BE} aufgetragen. Wir beobachten, dass der Strom I_C durch den Transistor exponentiell von der Spannung zwischen Basis und Emitter U_{BE} abhängt

$$I_C = \beta \, I_S \left[e^{\frac{U_{BE}}{U_T}} - 1 \right] \quad . \tag{5.1}$$

Dies entspricht der Diodengleichung.

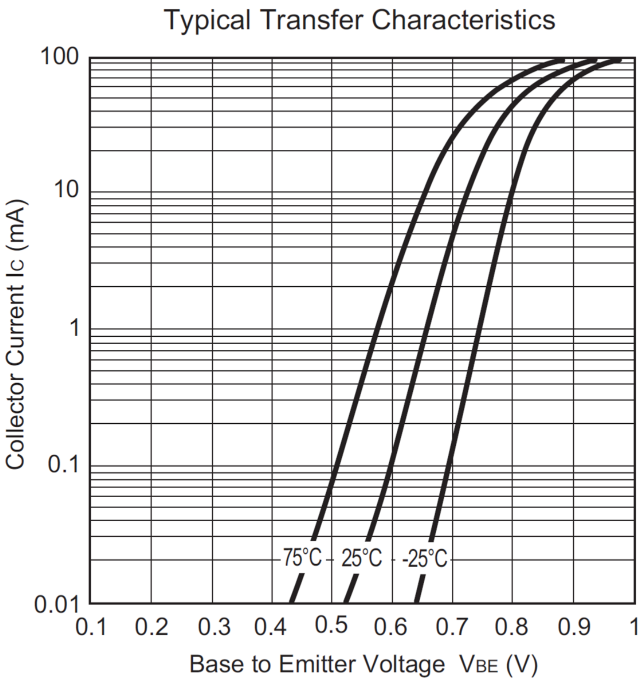


Abb. 5.4 Kollektorstrom I_C als Funktion der Basis-Emitter-Spannung U_{BE} für einen npn-Transistor für $U_{CE} = 1$ V und verschiedene Temperaturen (2N3904)[8].

Die Stromverstärkung β des Transistors mit

$$\beta = \left. \frac{dI_C}{dI_B} \right|_{U_{CE}=konst.} \tag{5.2}$$

ist in Abbildung 5.5 als Funktion des Kollektorstroms für verschiedene Temperaturen aufgetragen. Die Kurve verläuft über mehrere Größenordnungen des Kollektorstroms recht flach.

[3] Es hängt von der konkreten Transistorschaltung ab, an welchem Pol die Eingangsspannung bzw. Versorgungsspannung liegt (siehe Kapitel 5.4).

Die Stromverstärkung β variiert zwischen 100 und 200. Dies sind typische Werte, die aber stark von Fabrikationsparametern und vom Transistortyp abhängen. Es gibt auch Transistoren mit einer Stromverstärkung von 1000 oder höher. Für das Verständnis der meisten Transistorschaltungen ist die Näherung $I_C = \beta\, I_B$ ausreichend.

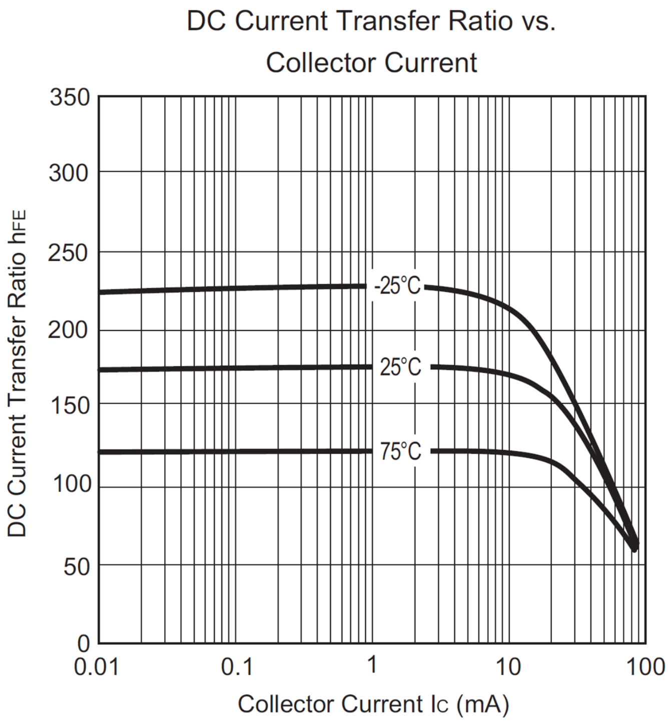


Abb. 5.5 Gleichstromverstärkung β (hier als h_{FE} bezeichnet) als Funktion des Kollektorstroms I_C für einen npn-Transistor bei $U_{CE} = 1$ V (2N3904)[8].

5.4 Grundlegende Transistorschaltungen

Den Transistor kann man als eine stromgesteuerte oder als eine spannungsgesteuerte Stromquelle auffassen. Das heißt der Strom I_C durch den Transistor hängt z.B. über Gl. 5.2 von I_B ab. Diese Stromquelle kann auf unterschiedliche Weise genutzt werden (siehe Abb. 5.6 und 5.7).

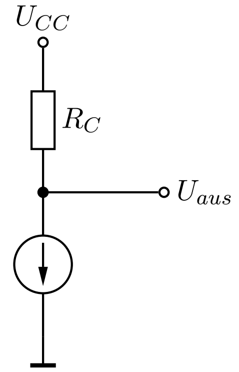


Abb. 5.6 Transistor als Stromsenke mit Widerstand als Spannungsverstärker.

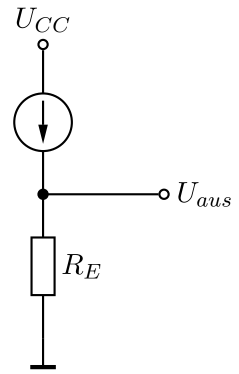


Abb. 5.7 Transistor als Stromquelle mit Widerstand.

In beiden Bildern ist der Widerstand in Reihe mit der Stromquelle geschaltet. In Bild 5.6 ist der Transistor eine Stromsenke, in Bild 5.7 eine Stromquelle. Die Stromsenke in Verbindung mit dem Widerstand und dem gezeigten Ausgang in Abb. 5.6 wirkt als invertierender Spannungsverstärker der Eingangsspannung U_{BE} (nicht eingezeichnet). Hier ist die sinnvolle Last eher hochohmig, da sonst der Transistor einfach kurzgeschlossen würde und die Ausgangsspannung nur durch U_{CC}, R_C und R_L bestimmt würde. In Abb. 5.7 könnte man auch eine niederohmige Last anschließen. Der Strom fließt ja in jedem Fall durch den Transistor.

Wir betrachten im größeren Detail zuerst Schaltungen mit nur einem Transistor, einem Schutzvorwiderstand R_V an der Basis und einem weiteren Kollektor- oder Emitterwiderstand. Zwei Anschlüsse des Transistors, meistens Kollektor und Emitter, werden an eine Gleichspannungsquelle angeschlossen. Dies ermöglicht es, ein vergleichsweise kleines variables Signal, das am dritten Pol des Transistors anliegt, zu verstärken.

Bei der Emitterschaltung (Abb. 5.8) und der Kollektorschaltung (Abb. 5.9) liegt das Eingangssignal an der Basis an. Das Ausgangssignal wird am Kollektor bzw. am Emitter abgegriffen. Die Grundschaltungen unterscheiden sich auch in der Position des Widerstands, der für die Emitterschaltung am Kollektor und für die Kollektorschaltung am Emitter liegt.

Anmerkung: Dies gilt nur für die Grundversion der Schaltungen. In Schaltungen mit Rückkopplung tauchen weitere Widerstände auf.

Die Basisschaltung (Abb. 5.10) ergibt sich aus der Emitterschaltung durch das Vertauschen der Anschlüsse für das Eingangssignal und der Masse.

Die Schaltungen sind nach dem Pol benannt, der in der Grundform an der Versorgungsspannung oder Erde liegt und somit weder als Eingang noch Ausgang der Schaltung dient.

Äquivalente Schaltungen gibt es auch für unipolare Transistoren, z. B. für MOSFETs (Kap. 9).

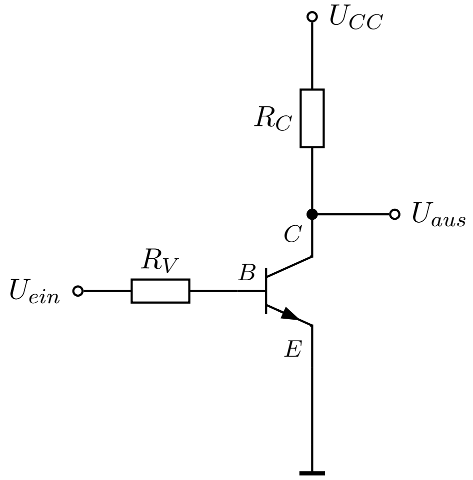


Abb. 5.8 Emitterschaltung (common emitter amplifier) mit Eingangsspannung U_{ein}, Ausgangsspannung U_{aus} und Versorgungsspannung U_{CC}.

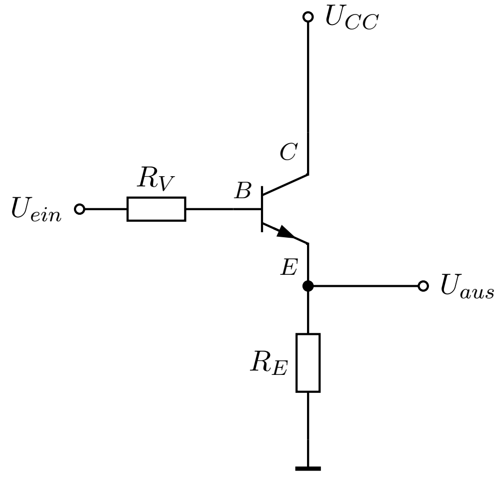


Abb. 5.9 Kollektorschaltung (common collector amplifier).

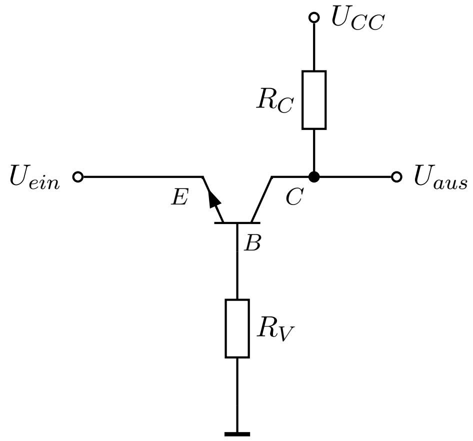


Abb. 5.10 Basisschaltung (common base amplifier).

5.5 Kenngrößen von Transistorschaltungen

Bei der Analyse der obigen Transistorverstärkerschaltungen interessieren wir uns häufig für die Abhängigkeit des Basisstroms I_B und des Kollektorstroms I_C von den Spannungen U_{BE} und U_{CE}, die in den Abbildungen 5.3 und 5.4 diskutiert wurde. Da wir im Allgemeinen kleine Eingangssignale verstärken, ist es naheliegend die Änderungen von Basisstrom und Kollektorstrom durch das totale Differential auszudrücken:

$$dI_C\,(U_{BE}, U_{CE}) = \frac{\partial I_C}{\partial U_{BE}}\,dU_{BE} + \frac{\partial I_C}{\partial U_{CE}}\,dU_{CE} = S\,dU_{BE} + \frac{1}{R_{CE}}\,dU_{CE} \tag{5.3}$$

und

$$dI_B\,(U_{BE}, U_{CE}) = \frac{\partial I_B}{\partial U_{BE}}\,dU_{BE} + \frac{\partial I_B}{\partial U_{CE}}\,dU_{CE} \approx \frac{1}{R_{BE}}\,dU_{BE} \quad , \tag{5.4}$$

wobei der Term $\frac{\partial I_B}{\partial U_{CE}}$ vernachlässigt wurde.

Hier haben wir den differentiellen Basis-Emitter-Widerstand

$$R_{BE} = \frac{\partial U_{BE}}{\partial I_B} \quad , \tag{5.5}$$

den differentiellen Kollektor-Emitter-Widerstand

$$R_{CE} = \frac{\partial U_{CE}}{\partial I_C} \tag{5.6}$$

und die Steilheit

$$S = \frac{\partial I_C}{\partial U_{BE}} \tag{5.7}$$

definiert[4].

[4]Die differentiellen Widerstände R_{CE} und R_{BE} werden gerne auch als r_c und r_b bezeichnet.

Anmerkung: In der englischsprachigen Literatur wird das Symbol g_m für die Steilheit[5] benutzt.

In Gleichung 5.3 wird der Kollektorstrom als Funktion der Spannungen ausgedrückt.In diesem Bild entspricht der Transistor einer spannungsgesteuerten Stromquelle und die Steilheit S ist die bestimmende Größe.

Die einfachste quantitative Beschreibung der Kennlinien eines Transistors liefert das Ebers-Moll-Modell, in dem der Transistor als zwei entgegengesetzte Dioden aufgefasst wird. Es hat nur drei Parameter. Hier machen wir es uns noch einfacher und benutzen nur die Diodengleichung des Basis-Emitter-Übergangs und die Größe β:

$$I_B = I_S \left[e^{\frac{U_{BE}}{U_T}} - 1 \right] \approx I_S \, e^{\frac{U_{BE}}{U_T}} \tag{5.8}$$

mit

$$U_T = \frac{k\,T}{q_e} \quad \text{und} \quad I_C = \beta\, I_B \tag{5.9}$$

zu

$$S = \frac{\partial I_C}{\partial U_{BE}} = I_C \, \frac{q_e}{k\,T} = \frac{I_C}{U_T} \quad . \tag{5.10}$$

Bei Raumtemperatur gilt $\frac{k\,T}{q_e} = 25\,\text{mV}$ und damit

$$S[\frac{1}{\Omega}] = 40\, I_C[\text{A}] \quad . \tag{5.11}$$

Diese Abhängigkeit ist sehr praktisch und entspricht Abbildung 5.4. Je größer U_{BE} bzw. I_C sind, desto steiler ist die Kurve und desto größer die Steilheit. Diese Abhängigkeit gilt universell. Für einen gegebenen Kollektorstrom ist die Steilheit unabhängig von den Details des Transistors.

Der bipolare Transistor ist unabhängig von den obigen Überlegungen in guter Näherung auch eine stromgesteuerte Stromquelle. Es handelt sich um eine Stromquelle, da der Strom I_C relativ unabhängig von U_{CE} ist. Aufgrund der Abhängigkeit des Kollektorstroms vom Basisstrom (Gl. 5.9) können wir die Stromquelle als stromgesteuert bezeichnen. Willkommene Eigenschaften sind außerdem die großen Stromverstärkung und ihre relativ geringe Abhängigkeit vom Kollektorstrom (Abb. 5.5). In dieser Sichtweise ist es natürlicher, statt der Steilheit die Stromverstärkung zu benutzen. Die Steilheit und die Stromverstärkung hängen wie folgt voneinander ab:

$$S = \frac{\partial I_C}{\partial U_{BE}} = \beta \frac{\partial I_B}{\partial U_{BE}} = \frac{\beta}{R_{BE}} \quad \text{und} \quad R_{BE} = \frac{k\,T}{q_e\, I_B} = \frac{U_T}{I_B} \quad . \tag{5.12}$$

Die einzige Größe der Gleichungen 5.3 und 5.4, die wir noch nicht näher betrachtet haben, ist R_{CE}. Eine Vergrößerung der U_{CE} verändert die effektive Breite der Basis und erhöht damit den Kollektorstrom. Für ein genaueres Verständnis betrachten wir Abbildung 5.11.

[5]Engl. *transconductance*

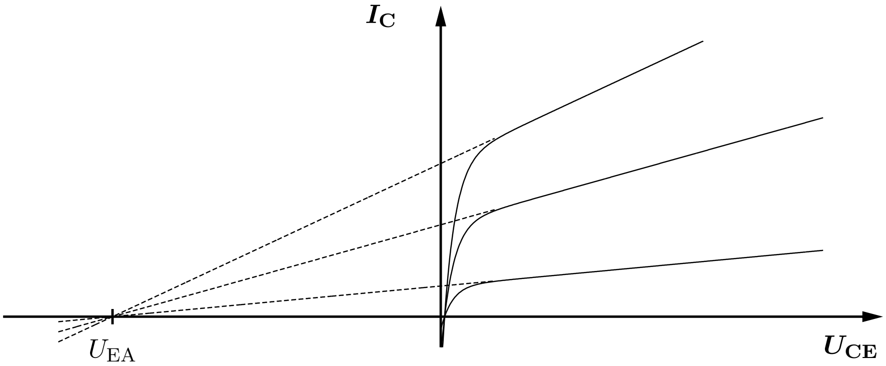


Abb. 5.11 Kollektorstrom I_C als Funktion der Kollektor-Emitter-Spannung U_{CE} für verschiedene Basisströme.

Die Steigung der I_C-U_{CE}-Kurve in Abbildung 5.3 entspricht $1/R_{CE}$. Extrapoliert man die Kurven aus dem Bereich positiver Werte von U_{CE} zu negativen U_{CE}, so treffen sich alle Kurven in einem Punkt. Dies entspricht der Early-Spannung U_{EA}. Damit gilt

$$R_{CE} = \frac{U_{EA}}{I_{C0}} \quad , \tag{5.13}$$

wobei I_{C0} der Schnittpunkt der extrapolierten Kennlinie mit der vertikalen Achse $U_{CE} = 0$ ist.

Die Early-Spannung variiert grob von 15 V bis 150 V. Abhängig vom Kollektorstrom kann R_{CE} recht groß sein, d. h. einige 100 kΩ, und $R_{BE} \ll R_{CE}$ eher klein.

Hier ist es wichtig, zwischen dem Widerstand $R = U_{CE}/I_C$, der auch die Leistung bestimmt, und dem differentiellen Widerstand $R_{CE} = dU_{CE}/dI_C$ zu unterscheiden.

In Lehrbüchern wird gerne die Darstellung aus Abbildung 5.12 gewählt, in der diverse Kennlinien zusammengefasst werden.

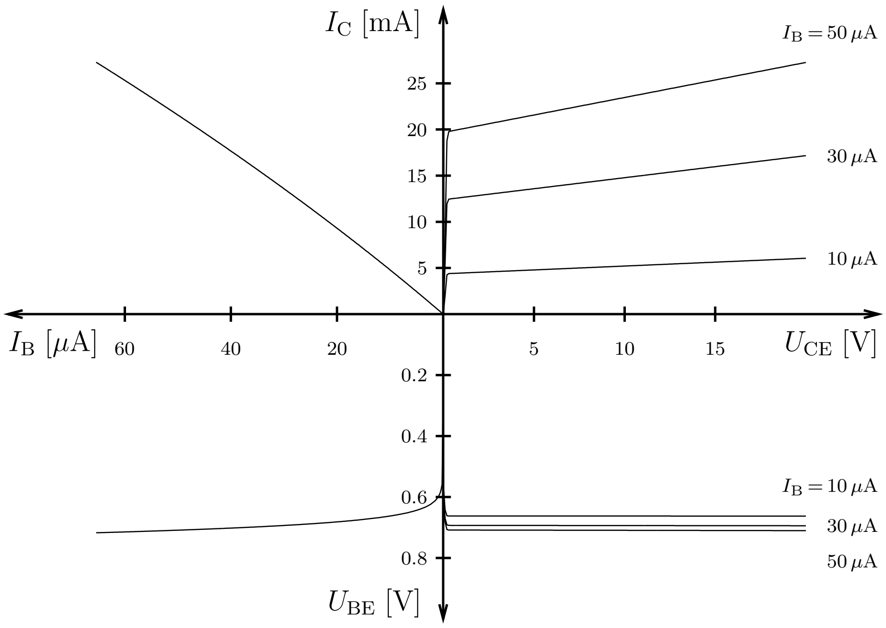


Abb. 5.12 Übersicht der Transistorkennlinien eines Bipolartransistors für exemplarische Basisströme I_B von 10, 30 und 50 μA.

5.6 Emitterschaltung

Die Emitterschaltung (Abb. 5.8 und 5.13) besteht aus einem Transistor und einem Kollektorwiderstand. Der Eingang der Schaltung liegt an der Basis, der Ausgang am Kollektor. (Daher der Name Emitterschaltung!)

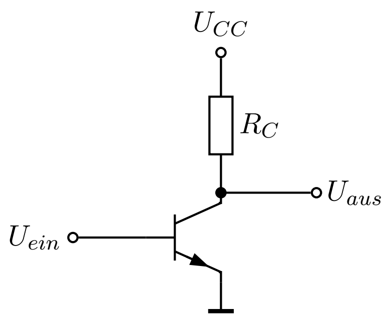


Abb. 5.13 Emitterschaltung.

Die einfache Emitterschaltung versteht man am schnellsten im Spannungsbild. Da $U_{aus} = U_{CC} - I_C\,R_C$ gilt, gilt für die Spannungsverstärkung

$$V_U = \frac{dU_{aus}}{dU_{ein}} = -\,R_C\,\frac{dI_C}{dU_{ein}} = -\,S\,R_C = -\,\beta\,\frac{R_C}{R_{BE}} \quad . \tag{5.14}$$

In diesem Ausdruck ist der Einfluss des Lastwiderstands R_L und des Widerstands R_{CE} noch nicht berücksichtigt. Am Kleinsignalersatzschaltbild (Abb. 5.14) sieht man, dass man dazu den Parallelwiderstand von R_C, R_{CE} und R_L (nicht eingezeichnet) bilden muss.

Das Kleinsignalersatzschaltbild ist äußerst hilfreich zum Verständnis von elektronischen Schaltungen. Im Ersatzschaltbild werden Quellen von konstanten Strömen und Spannungen durch ihren Innenwiderstand ersetzt. Für die äußere Spannungsquelle U_{CC} ist der Innenwiderstand sehr klein, so dass man diese Quelle durch einen Kurzschluss ersetzen kann. Nichtlineare Bauteile ersetzt man durch ihren differentiellen Widerstand im Arbeitspunkt, da man nur Strom- und Spannungsänderungen betrachtet. So wird die Basis-Emitter-Diode durch den Widerstand R_{BE} ersetzt. Zwischen Kollektor und Emitter liegt im Ersatzschaltbild eine gesteuerte Stromquelle mit parallelgeschaltetem Innenwiderstand R_{CE}.

Anmerkung: Man kann hier eine spannungsgesteuerte Stromquelle mit dem Strom $S\,dU_{BE}$ oder eine stromgesteuerte Stromquelle mit Strom $\beta\,dI_B$ benutzen.

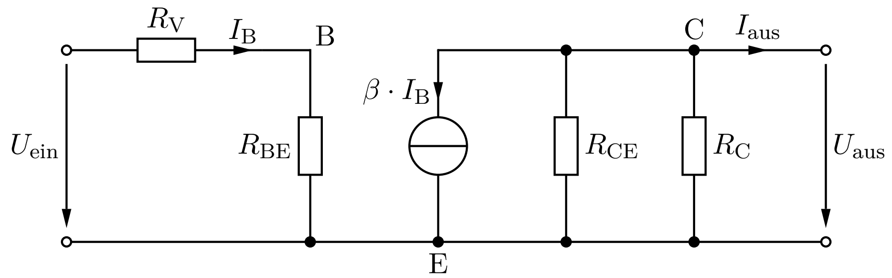


Abb. 5.14 Kleinsignalersatzschaltbild der Emitterschaltung. R_V ist ein Schutzvorwiderstand der Basis. Die Spannungsquelle U_{CC} ist kurzgeschlossen. Die Buchstaben B, C, E bezeichnen Basis, Emitter und Kollektor. Hier und in den folgenden Kleinsignalersatzschaltbildern bezeichnen die Größen U_ein, I_B, I_aus, U_aus Strom- und Spannungsänderungen.

Für typische Werte von $\beta = 100$, $I_B = 10\,\mu\mathrm{A}$ und $R_{CE} = 100\,\mathrm{k\Omega}$ ergibt sich folgendes Bild.

Der **Eingangswiderstand** dieser Schaltung ist

$$R_{ein} = R_V + R_{BE} \quad \text{mit} \quad R_{BE} = \frac{U_T}{I_B} = \frac{25}{10}\,\frac{\mathrm{mV}}{\mu\mathrm{A}} = 2,5\,\mathrm{k\Omega} \quad .$$

Mit

$$S = \frac{\beta}{R_{BE}} \quad \text{folgt} \quad S = \frac{100}{2,5}\,\frac{1}{\mathrm{k\Omega}} = 40\,\frac{\mathrm{mA}}{\mathrm{V}} = 40\,\mathrm{mS} \quad (\text{vgl. Gl. 5.10}).$$

Wählt man

$$R_C = 5\,\mathrm{k\Omega} \quad ,$$

so ergibt sich nach Gleichung 5.14 die Spannungsverstärkung

$$V_U = \frac{dU_{aus}}{dU_{ein}} = -\,200 \quad .$$

Dasselbe erhält man aus dem Ersatzschaltbild, da

$$dU_{aus} = -\beta\,dI_B\,(R_C \parallel R_{CE}) \approx -\beta\,dI_B\;R_C \quad \text{und} \tag{5.15}$$

$$dU_{ein} = dI_B \, (R_V + R_{BE}) \approx dI_B \, R_{BE} \quad . \tag{5.16}$$

Also gilt

$$V_U = \frac{dU_{aus}}{dU_{ein}} \approx -\frac{\beta \, R_C}{R_{BE}} = -S \, R_C \quad . \tag{5.17}$$

Hier wird der Lastwiderstand R_L als groß angenommen. Stimmt dies nicht, müsste man R_C durch $R_C \parallel R_L$ bzw. noch genauer durch $R_C \parallel R_{CE} \parallel R_L$ ersetzen.

Der **Ausgangswiderstand** der Emitterschaltung ist

$$R_C \parallel R_{CE} \approx R_C = 5\,\mathrm{k}\Omega \quad .$$

Ein Nachteil der einfachen Emitterschaltung ist die Abhängigkeit der Verstärkung von dem Transistorparametern β bzw. S. Außerdem ist weder der Eingangswiderstand besonders hoch, noch der Ausgangswiderstand besonders klein. Dies kann man durch eine Schaltung mit Rückkopplung vermeiden, wie wir später sehen werden.

5.7 Kollektorschaltung

Die Kollektorschaltung (Abb. 5.9 und 5.15) besteht aus einem Transistor und einem Emitterwiderstand. Der Eingang der Schaltung liegt an der Basis, der Ausgang am Emitter. (Daher der Name Kollektorschaltung!)

Die Funktionsweise der Kollektorschaltung lässt sich mit der Beziehung $I_C = \beta \, I_B$ und der Annahme $U_{BE} \approx 0,7\,\mathrm{V}$ in guter Näherung verstehen.

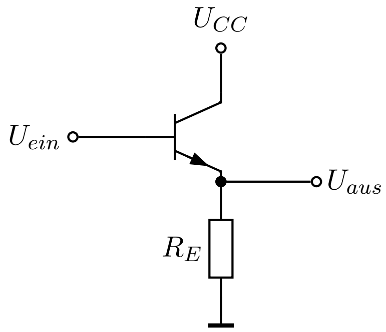


Abb. 5.15 Kollektorschaltung.

Aus

$$U_{ein} - U_{aus} = U_B - U_E = U_{BE} \approx 0,7\,\mathrm{V}$$

folgt

$$dU_{ein} = dU_{aus} \quad ,$$

das heißt die Spannungsverstärkung

$$V_U = \frac{dU_{ein}}{dU_{aus}} \approx 1 \quad . \tag{5.18}$$

Deshalb wird diese Schaltung auch als Emitterfolger bezeichnet. Die Ausgangsspannung „folgt“ der Eingangsspannung.

Diese Eigenschaft ist noch nicht sehr nützlich. Aber dafür verstärkt die Schaltung den Eingangsstrom. Aus

$$I_C = \beta\, I_B \quad \text{folgt} \quad \boldsymbol{I_E = (\beta + 1)\, I_B} \quad . \tag{5.19}$$

Das heißt, die Stromverstärkung ist groß. Welcher Anteil des Emitterstroms durch den Lastwiderstand fließt (hier nicht eingezeichnet) und welcher durch den Emitterwiderstand R_E fließt, hängt vom Verhältnis R_L/R_E ab. Diese Schaltung eignet sich zum Treiben von Signalen über lange Leitungen. In mehrstufigen Verstärkern mag ein Vorverstärker eine hohe Spannungsverstärkung bewirken. Zusammen mit einer nachgeschalteten Kollektorschaltung zur Stromverstärkung wird dann eine sehr große Leistungsverstärkung erzielt.

Zur Bestimmung der Eingangsimpedanz und der Ausgangsimpedanz des Emitterfolgers betrachten wir das Großsignalersatzschaltbild Abb. 5.16 und das daraus abgeleitete Kleinsignalersatzschaltbild Abb. 5.17.

Im Großsignalersatzschaltbild sind die externen und internen Spannungen noch nicht kurzgeschlossen.

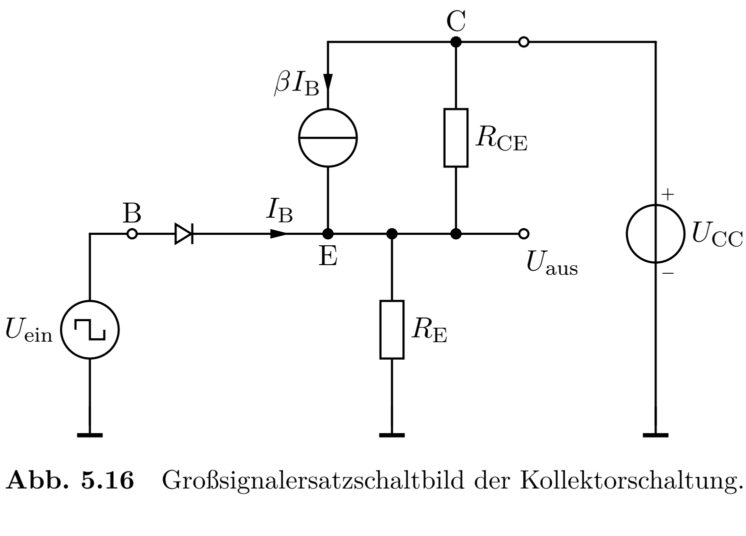


Abb. 5.16 Großsignalersatzschaltbild der Kollektorschaltung.

Im Kleinsignalersatzschaltbild sind die Diode mit Spannung U_{BE} und die Versorgungsspannung dagegen durch Kurzschlüsse ersetzt.

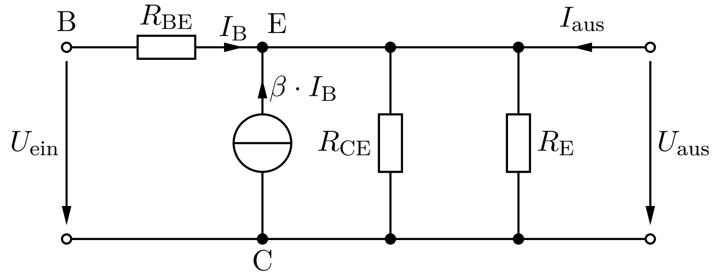


Abb. 5.17 Kleinsignalersatzschaltbild der Kollektorschaltung.

Zur Bestimmung der **Eingangsimpedanz**

$$Z_{ein} = \frac{dU_{ein}}{dI_{ein}}$$

lassen wir einen kleinen Teststrom dI_{ein} in die Basis fließen und berechnen (bzw. messen) die Änderung der Eingangsspannung $\mathrm{d}U_{\mathrm{ein}}$. Es gilt

$$\mathrm{d}U_{\mathrm{ein}} = R_{\mathrm{BE}} \cdot \mathrm{d}I_{\mathrm{ein}} + (\beta+1)\ \mathrm{d}I_{\mathrm{ein}} \cdot R_{\mathrm{E}} \| R_{\mathrm{CE}}\ , \tag{5.20}$$

da der Teststrom einen Emitterstrom $(\beta + 1)\ \mathrm{d}I_{\mathrm{ein}}$ bewirkt. Bei Vernachlässigung von R_{CE} gegenüber von R_{E} ergibt sich

$$Z_{\mathrm{ein}} = \frac{\mathrm{d}U_{\mathrm{ein}}}{\mathrm{d}I_{\mathrm{ein}}} = R_{\mathrm{BE}} + (\beta+1)\ R_{\mathrm{E}} \approx (\beta+1)\ R_{\mathrm{E}}\,. \tag{5.21}$$

Die Eingangsimpedanz ist also groß, da der Wert von R_E, der typischerweise schon im Bereich von $100\,\Omega$ bis $1\,\mathrm{k}\Omega$ liegt, mit dem Faktor $(\beta + 1)$ multipliziert wird.

Anmerkung: Hier wird wieder vorausgesetzt, dass R_L sehr groß ist. Sonst müsste man mit $R_L \parallel R_E$ bzw. $R_L \,||\, R_E \,||\, R_{CE}$ rechnen.

Wieso taucht der Wert von R_E, dem Widerstand auf der Ausgangsseite des Transistors, in der Formel für den Eingangswiderstand auf?

Deshalb, weil R_E und U_{aus} den Emitterstrom festlegen und weil die beiden Größen I_E und U_{aus} mit den Eingangsgrößen I_{ein} und U_{ein} in einer festen Beziehung stehen.

Die **Ausgangsimpedanz** berechnen wir, indem wir bei konstantem U_{ein} die Ausgangsspannung um $\mathrm{d}U_{\mathrm{aus}}$ variieren und den resultierenden Strom dI_{aus} in den Ausgang berechnen. Die Änderung $\mathrm{d}U_{\mathrm{aus}}$ verursacht einen Strom

$$\frac{\mathrm{d}U_{\mathrm{aus}}}{R_{\mathrm{E}} \| R_{\mathrm{CE}}} \tag{5.22}$$

durch R_E und R_{CE}. Dieser Strom setzt sich aus zwei Anteilen zusammen: aus dem gesuchten dI_{aus} und aus der Änderung des Emitterstroms $(\beta + 1)\, dI_{\mathrm{B}}$, den $dU_{\mathrm{aus}} = dU_{\mathrm{EB}} = -dU_{\mathrm{BE}}$ bewirkt. Es gilt

$$dI_\mathrm{B} = \frac{dU_\mathrm{BE}}{R_\mathrm{BE}} = -\frac{dU_\mathrm{aus}}{R_\mathrm{BE}} \quad , \tag{5.23}$$

wobei R_S der Innenwiderstand der Eingangsspannungsquelle ist, der in Abb. 5.16 nicht eingezeichnet ist. Aufgrund der Knotenregel im Punkt E gilt

$$\mathrm{d}I_\mathrm{aus} + \mathrm{d}I_\mathrm{B}\,(\beta+1) = \frac{dU_\mathrm{aus}}{R_\mathrm{CE} \| R_\mathrm{E}} \quad . \tag{5.24}$$

Mit Gl. 5.23 folgt

$$\mathrm{d}I_\mathrm{aus} \quad = \quad \frac{\beta+1}{R_\mathrm{BE}} dU_\mathrm{aus} + \frac{dU_\mathrm{aus}}{R_\mathrm{CE} \| R_\mathrm{E}} \tag{5.25}$$

und damit

$$\mathrm{Z}_\mathrm{aus} \quad = \quad \frac{\mathrm{d}U_\mathrm{aus}}{\mathrm{d}I_\mathrm{aus}} \quad = \quad \frac{R_\mathrm{BE}}{\beta+1} \| R_\mathrm{CE} \| R_\mathrm{E} \quad \approx \quad \frac{R_\mathrm{BE}}{\beta+1} \, . \tag{5.26}$$

Berücksichtigt man, dass die Eingangsspannung aus einer Signalquelle mit endlichen Widerständen R_S kommt, so gilt

$$Z_\mathrm{aus} \approx \frac{R_\mathrm{BE} + R_S}{\beta+1} \tag{5.27}$$

Hier sieht man aber schon, dass der Emitterfolger gut als Impedanzwandler geeignet ist. Er stellt für eine Signalquelle eine hochohmige Eingangsstufe dar und verfälscht dadurch nicht das Messsignal. Gleichzeitig leitet er das Signal niederohmig weiter.

5.8 Basisschaltung

Die Basisschaltung (Abb. 5.10 und 5.18) besteht aus einem Transistor und einem Kollektorwiderstand. Der Eingang der Schaltung liegt am Emitter, der Ausgang am Kollektor. (Daher der Name Basisschaltung!)

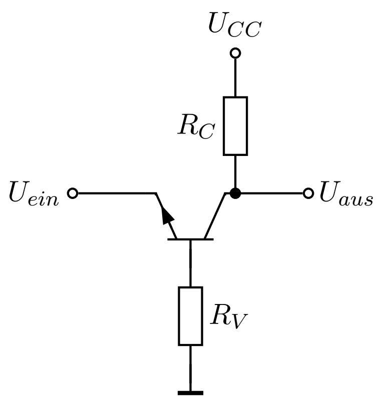


Abb. 5.18 Basisschaltung mit Schutzvorwiderstand R_V.

Die Funktionsweise der Basisschaltung lässt sich durch den Vergleich mit der Emitterschaltung verstehen. Da die Anschlüsse von Basis und Emitter vertauscht sind, hat die Spannungsverstärkung denselben Betrag aber das umgekehrte Vorzeichen wie der Emitterverstärker und beträgt $V_U = S\,R_C$. Da durch den Eingang der Schaltung fast derselbe Strom fließt wie durch den Kollektor ($I_E \approx I_C$), ist die Stromverstärkung $\frac{dI_c}{dI_{ein}} \approx 1$. Der Ausgangswiderstand $Z_{aus} \approx R_C$ wie bei der Emitterschaltung.

Die Basisschaltung wird interessant durch ihren kleinen Eingangswiderstand

$$Z_{ein} \approx \frac{dU_{BE}}{dI_E} \approx \frac{dU_{BE}}{dI_C} = \frac{1}{S} = \frac{k\,T}{q_e\,I_C} \quad . \tag{5.28}$$

Für Raumtemperatur und einen Kollektorstrom von 1 mA wie in unseren obigen Beispielen ist $Z_{ein} \approx 25\,\Omega$. Der Eingangswiderstand ist damit um den Faktor $\frac{1}{\beta}$ kleiner als für den Emitterverstärker. Damit die Bedingung $U_{BE} \approx 0,7\,\mathrm{V}$ erfüllt ist, muss man das Gleichspannungspotential des Eingangs auf eine negative Vorspannung $\approx -0,7\,\mathrm{V}$ legen. Die Basisschaltung ist besonders für Hochfrequenzanwendungen geeignet. Es tritt kein Millereffekt auf (siehe unten).

5.9 AC- oder DC-Kopplung?

Eine prinzipielle Frage, die der Designer beantworten muss, ist, ob er das Signal direkt oder über einen Kondensator an den Eingang des Verstärkers koppelt (siehe Abb. 5.19). Diese Frage gilt für jede Verstärkerstufe und stellt sich damit auch für den Ausgang des Verstärkers.

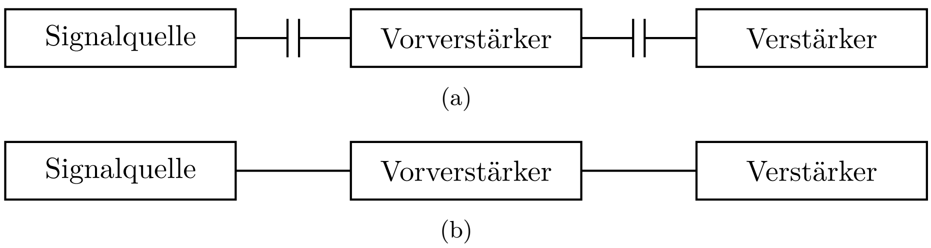


Abb. 5.19 (a) AC-Kopplung. (b) DC-Kopplung.

Der Koppelkondensator blockiert den Gleichspannungsanteil des Signals. Durch ein geeignetes Netzwerk wird der Gleichspannungs- bzw. Gleichstromwert am Verstärkereingang unabhängig vom Signal eingestellt, so dass der Arbeitspunkt des Verstärkers optimal liegt. Ein einfaches Beispiel ist die stromgegengekoppelte Emitterschaltung (Abb. 5.20).

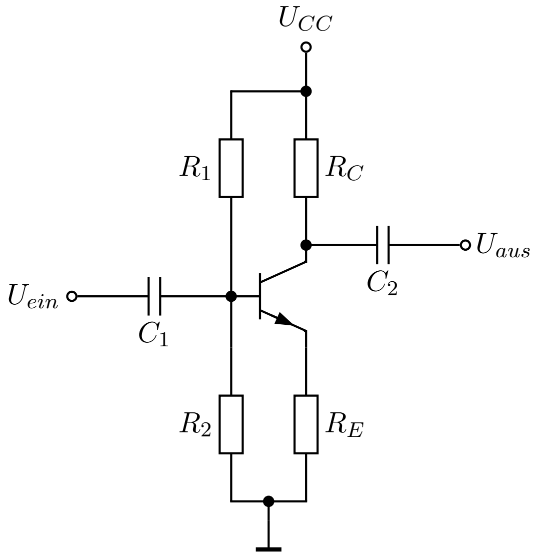


Abb. 5.20 Stromgegengekoppelter Verstärker mit AC-Kopplung des Eingangs- und Ausgangssignals über die Koppelkondensatoren C_1 und C_2.

Der Spannungsteiler R_1 und R_2 zieht den Eingang des Transistors auf die gewünschte Vorspannung.

Besonders vorteilhaft bei einer AC-Kopplung ist die Unabhängigkeit vom Gleichspannungsanteil des Eingangssignal, der oft zu hoch ist, d.h. deutlich größer als $0,7$ V. Zudem werden mögliche Verschiebungen (Drift) des Gleichstromanteils, z. B. mit der Zeit oder Temperatur, unterdrückt.

Anmerkung: Bei Siliziumhalbleiterdetektoren kann z. B. der Leckstrom durch Strahlenschäden um viele Größenordnungen steigen, gleichzeitig ist er stark temperaturabhängig. Durch Koppelkondensatoren isoliert man den Verstärker von diesen Effekten.

Ein Nachteil kann sein, dass der Koppelkondensator mit der Eingangsimpedanz des Verstärkers einen Hochpass bildet, der niederfrequente Signale abschwächt.

5.10 Einstellung des Arbeitspunktes

Wir haben die Transistorschaltungen bisher betrachtet, ohne uns um den Arbeitspunkt des Transistors zu kümmern oder die Details des Eingangssignals zu berücksichtigen. Bei jeder konkreten Schaltung muss aber sichergestellt sein, dass der Transistor überhaupt leiten kann und dass seine Ausgangsströme und -spannungen vernünftige Werte annehmen. Eine stabile und an die mögliche Variation des Eingangssignals angepasste Positionierung des Arbeitspunktes ist wichtig.

In Abbildung 5.21 wird eine beliebte und effektive Methode gezeigt. Das Potential des Basiseingangs wird über den Spannungsteiler R_1 und R_2 mithilfe der Versorgungsspannung festgelegt.

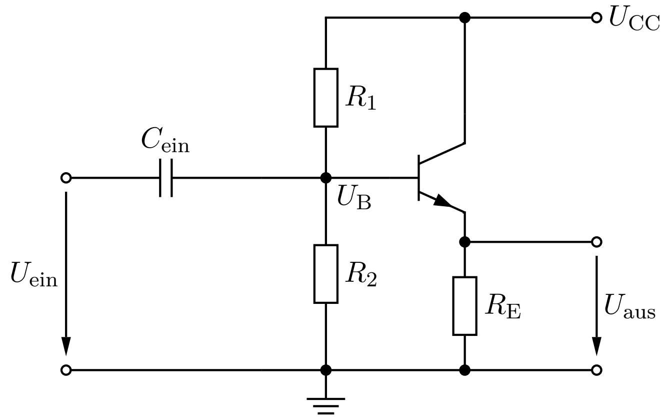


Abb. 5.21 Arbeitspunkteinstellung der Kollektorschaltung.

Bei Einstellungen des Arbeitspunktes kann man folgendermaßen vorgehen.

1. Wähle I_C und U_{CE}.

2. Berechne R_E als
$$R_E = \frac{U_{CC} - U_{CE}}{I_C} \quad .$$

3. Wähle R_1 und R_2 so, dass für die Ruhespannung U_B ohne Eingangssignal gilt
$$U_B = U_{CC} - U_{CE} + 0,7\,\mathrm{V} \quad .$$

4. Wähle den Koppelkondensator C_{ein} groß genug, so dass alle relevanten Eingangsfrequenzen durchkommen.

Der optimale Strom I_C wird stark von der Last abhängen. Falls es sich um einen hochohmigen Verstärker handelt, können der Ausgangsstrom und I_C klein sein. Falls der Transistor ein Signal über ein langes Kabel treiben soll, wird I_C eher groß sein. Noch größer ist I_C für Leistungstransistoren.

Ähnliche Überlegungen gelten für U_{CE}. Oft wird man U_{CE} im Arbeitspunkt symmetrisch zur Versorgungsspannung wählen, also
$$U_{CE} \approx \frac{U_{CC}}{2} \quad .$$

Die Wahl von U_{CE} wird auch von der Polarität des Eingangssignals abhängen.

Die Widerstände R_1, R_2 sollte man nicht zu klein wählen, da sonst der sogenannte Querstrom durch die Widerstände R_1, R_2 und damit die Leistung und Abwärme zu groß werden. Außerdem reduzieren kleine R_1 und R_2 die Eingangsimpedanz der Schaltung. Der Querstrom sollte aber groß genug sein, so dass eine Variation des Basisstroms den Wert U_B nicht wesentlich beeinflusst.

5.11 Emitterschaltung mit Stromgegenkopplung

Man kann die Eigenschaften der einfachen Emitterschaltung durch negative Rückkopplung deutlich verbessern. Die allgemeinen Vorteile der negativen Rückkopplung haben wir schon am Beispiel des Operationsverstärkers diskutiert. Sie gelten genauso für elementare 1-Transistor-schaltungen. Wir unterscheiden die Emitterschaltung mit Stromgegenkopplung und die Emitterschaltung mit Spannungsgegenkopplung. Die Emitterschaltung mit Stromgegenkopplung erhält man aus der einfachen Emitterschaltung durch das Einfügen eines weiteren Widerstands am Emitter (Abb. 5.22).

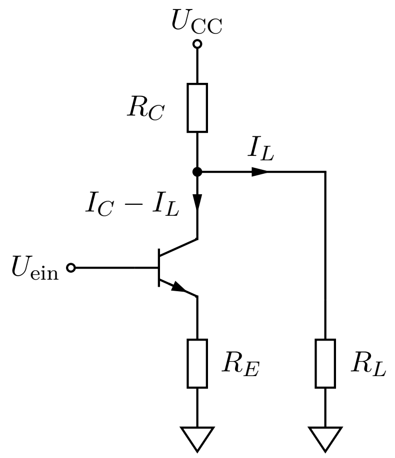


Abb. 5.22 Emitterschaltung mit Stromgegenkopplung. (Hier weicht die Definition von I_C von den bisherigen und folgenden Darstellungen ab.)

Diese Schaltung ist mit einigen Näherungen leicht zu verstehen: Es gilt unabhängig von dem eingefügten Emitterwiderstand

$$U_{aus} = U_C = U_{CC} - I_C\,R_C \quad .$$

U_{CC} ist konstant, also gilt

$$dU_{aus} = -\,dI_C\,R_C \quad .$$

Außerdem gilt

$$dI_E = \frac{dU_E}{R_E} \quad , \quad dU_E = dU_{ein} \quad \text{und} \quad I_E = I_C - I_L + I_B \approx I_C \quad .$$

An dieser Stelle haben wir vorausgesetzt, dass R_L sehr groß ist und damit I_L klein. Aus

$$dU_{aus} = -\,dI_C\,R_C \approx -\,dI_E\,R_C = -\,dU_E\,\frac{R_C}{R_E} = -\,dU_{ein}\,\frac{R_C}{R_E}$$

ergibt sich die Spannungsverstärkung

$$V_U = \frac{dU_{aus}}{dU_{ein}} = -\,\frac{R_C}{R_E} \quad . \tag{5.29}$$

Dieses Ergebnis müssen wir mit der Verstärkung

$$V_U = -\,S\,R_C = -\,\beta\,\frac{R_C}{R_{BE}}$$

der einfachen Emitterschaltung vergleichen. Durch die Gegenkopplung ist die starke Abhängigkeit der Verstärkung vom Arbeitspunkt durch die entsprechende Abhängigkeit von R_{BE} aufgehoben. Jetzt hängt die Verstärkung nur noch von dem Verhältnis der externen Widerstände R_C/R_E ab. Auch die Abhängigkeit von β und die damit verbundene Temperaturabhängigkeit und Exemplarstreuung ist verschwunden. Dafür ist die Verstärkung in der Regel deutlich kleiner, z. B. ergibt sich mit $I_C = 1\,\mathrm{mA}$ die Steilheit $S[\mathrm{S}] = 40 \cdot I_C[\mathrm{A}]$ zu $40\,\mathrm{mS}$; dies ist nach Gl. 5.29 mit $1/R_E$ zu vergleichen. Für $R_E = 1\,\mathrm{k\Omega}$ entspricht dies einem $1/R_E = 1/\mathrm{k\Omega}$ von nur $1\,\mathrm{mS}$.

Der Reduktionsfaktor ist $S\,R_E$ (genau genommen $1 + S\,R_E$).

Ganz anschaulich erklärt funktioniert die Schaltung so:
Die Variation der Signalspannung dU_{ein} entspricht dU_E und stellt über R_E den Strom dI_E ein. Dieser Strom wiederum bestimmt den Spannungsabfall an R_C und damit die Ausgangsspannung

$$dU_{aus} \sim -\,dI_C\,R_C \quad .$$

Wieso handelt es sich hier um eine Gegenkopplung? Durch eine positive Variation der Eingangsspannung U_{ein} erhöht sich *ohne* einen Emitterwiderstand die Spannung U_{BE} und damit der Strom I_C und I_E. *Mit* einem Emitterwiderstand erhöht sich der Strom auch. Da aber gleichzeitig die Emitterspannung $U_E = I_E\,R_E$ ansteigt, ist der Anstieg der Basis-Emitter-Spannung dU_{BE} und damit dI_C und dU_{aus} mit Stromgegenkopplung geringer.

Wie bei der Kollektorschaltung ist der Eingang hochohmig. Die Ausgangsimpedanz ist durch den Widerstand R_C bestimmt und ist daher nicht so niederohmig wie für die Kollektorschaltung.

Das Kleinsignalersatzschaltbild der Emitterschaltung mit Stromgegenkopplung ist in Abbildung 5.23 gezeigt.

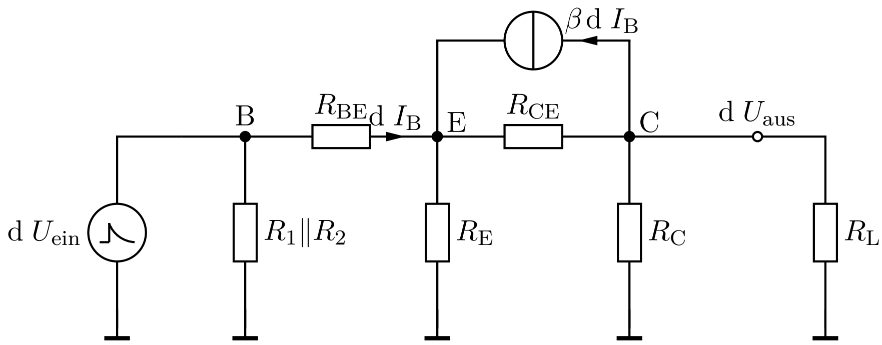


Abb. 5.23 Kleinsignalersatzschaltbild der Emitterschaltung mit Stromgegenkopplung mit Spannungsteiler R_1 und R_2 zur Arbeitspunkteinstellung.

Dem Ersatzschaltbild entnimmt man leicht die genaueren Ausdrücke für die interessanten Größen Spannungsverstärkung, Stromverstärkung, Eingangsimpedanz und Ausgangsimpedanz.

In Übereinstimmung mit Gleichung 5.29 gilt bei Vernachlässigung des kleinen Stroms durch R_{CE}

$$V_U = \frac{dU_{aus}}{dU_{ein}} \approx -\frac{dI_C}{dI_B}\,\frac{R_C \parallel R_L}{R_{BE} + R_E\,(\beta+1)} = -\frac{dI_C}{dI_E}\,\frac{R_C \parallel R_L}{\frac{R_{BE}}{\beta+1} + R_E} \approx -\frac{R_C}{R_E} \quad .$$

Aus

$$dU_{ein} = dI_B\,R_{BE} + dI_B\,(\beta+1)\,R_E \tag{5.30}$$

folgt für die **Eingangsimpedanz** [6]

$$Z_{ein} = (R_{BE} + (\beta+1)R_E) \parallel R_1 \parallel R_2 \quad . \tag{5.31}$$

Dies entspricht der Eingangsimpedanz des einfachen Emitterfolgers ohne Stromgegenkopplung.

Die **Ausgangsimpedanz** wird normalerweise durch R_C und R_L dominiert und $Z_\mathrm{aus} \approx R_C \| R_L$.

Für eine genauere Abschätzung muss man auch den Effekt eines Teststroms $\mathrm{d}I_x$ durch den Transistor, also durch die gesteuerte ideale Stromquelle und den Parallelwiderstand R_{CE} betrachten. Nach Erreichen des Emitters fließt $\mathrm{d}I_x$ über R_BE und R_E zu Masse. (Der Widerstand $R_1 \| R_2$ ist kurzgeschlossen, wenn wir den Innenwiderstand der Spannungsquelle zu Null setzen.) Dadurch verändert sich die Emitterspannung um $\mathrm{d}U_\mathrm{E} = \mathrm{d}I_x\,R_\mathrm{BE}\|R_\mathrm{E}$. Jetzt können wir den Anteil $\mathrm{d}I_{CE}$ von $\mathrm{d}I_x$, der durch R_CE fließt, ausrechnen:

$$\mathrm{d}I_{CE} = \mathrm{d}I_x - S\,\mathrm{d}U_\mathrm{BE} \;=\; \mathrm{d}I_x + S\,\mathrm{d}U_\mathrm{E} \qquad \text{bzw.} \qquad \mathrm{d}I_x + \frac{\beta\,\mathrm{d}U_\mathrm{E}}{R_\mathrm{BE}} . \tag{5.32}$$

Die Variation der Ausgangsspannung $\mathrm{d}U_x$ ergibt sich damit zu

$$\begin{aligned} \mathrm{d}U_x &= \mathrm{d}U_\mathrm{E} + \mathrm{d}U_\mathrm{CE} = \mathrm{d}I_x\,R_\mathrm{BE}\|R_\mathrm{E} + [\mathrm{d}I_x + S\,\mathrm{d}I_x\,R_\mathrm{BE}\|R_\mathrm{E}]\,R_\mathrm{CE} \\ &= \mathrm{d}I_x\,[R_\mathrm{BE}\|R_\mathrm{E} + R_\mathrm{CE}\,(1 + S\,R_\mathrm{BE}\|R_\mathrm{E})] \quad . \end{aligned} \tag{5.33}$$

Der Transistor stellt damit für $\mathrm{d}I_x$ die Impedanz

$$\begin{aligned} Z &= \frac{\mathrm{d}U_x}{\mathrm{d}I_x} = R_\mathrm{BE}\|R_\mathrm{E} + R_\mathrm{CE}\,(1 + S\,R_\mathrm{BE}\|R_\mathrm{E}) \\ &\approx R_\mathrm{CE}\,(1 + S\,R_\mathrm{BE}\|R_\mathrm{E}) \end{aligned} \tag{5.34}$$

dar. Jetzt muss man sich wieder an R_L und R_C erinnern und damit ergibt sich das Endergebnis

$$Z_\mathrm{aus} = R_\mathrm{C}\|R_\mathrm{L}\|R_\mathrm{CE}\,(1 + S\,R_\mathrm{BE}\|R_\mathrm{E}) \quad . \tag{5.35}$$

Wie groß ist die **Stromverstärkung** der Emitterschaltung mit Stromgegenkopplung? Wir interessieren uns hier für die Variation des Laststroms I_L mit dem Basisstrom, also für die Größe

$$V_I = \frac{dI_L}{dI_B} = \frac{(\beta+1)\,dI_L}{dI_E} \quad .$$

[6] Der Spannungsteiler aus R_1 und R_2 ist in Abbildung 5.22 im Gegensatz zu Abbildung 5.23 zur Vereinfachung weggelassen worden.

Diese berechnet man leicht unter Vernachlässigung des kleinen Stroms durch R_{CE} aus dem Kleinsignalersatzschaltbild.

Aus

$$dU_{aus} = -dI_C\, R_C||\, R_L = dI_L\, R_L \tag{5.36}$$

folgt

$$\frac{dI_L}{dI_C} = -\frac{R_C R_L}{R_L\,(R_C + R_L)} = -\frac{R_C}{R_C + R_L} \quad \text{und} \quad V_I = \frac{dI_L}{dI_B} = -\frac{\beta\, R_C}{R_C + R_L} \quad . \tag{5.37}$$

Die Ausgangstromänderung nimmt also mit zunehmendem Lastwiderstand R_L ab und beträgt maximal $-\beta\, dI_B$ wie man es intuitiv erwarten würde.

5.12 Emitterschaltung mit Spannungsgegenkopplung

Negative Rückkopplung wird auch bei der Emitterschaltung mit Spannungsgegenkopplung (Abb. 5.24 bzw. Abb. 5.25) angewandt.

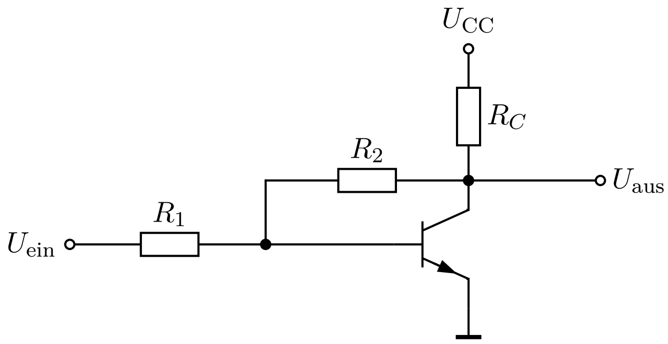


Abb. 5.24 Emitterschaltung mit Spannungsgegenkopplung.

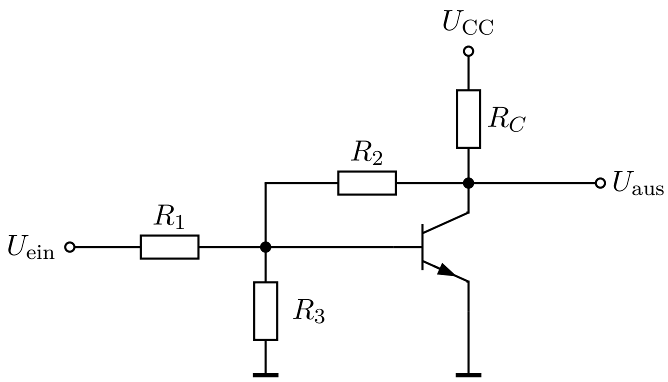


Abb. 5.25 Emitterschaltung mit Spannungsgegenkopplung.

Hier wird ein Teil der Ausgangsspannung über den Spannungsteiler R_1, R_2 bzw. R_2, R_3 auf den Eingang zurückgegeben. Schon äußerlich erinnert die Schaltung sehr an den invertierenden Operationsverstärker. Wie beim OPV fließt der Eingangsstrom fast vollständig zum Kollektor und nur ein sehr kleiner Teil des Eingangsstroms zur Basis. Hier ist in guter Näherung die Verstärkung

$$V_U = \frac{U_{aus}}{U_{ein}} = -\frac{R_2}{R_1} \quad .$$

Auch diese Schaltung ist komplexer als sie auf den ersten Blick aussieht. Der Widerstand R_2 dient zum einem der Rückkopplung, zum anderen der Einstellung des Arbeitspunkts.

5.13 Übersicht der Grundschaltungen mit einem Transistor

Eine schematische Übersicht der Schaltungen mit einem Transistor vermittelt Abb. 5.26. Durch eine Kombination dieser 1T-Schaltungen zu Schaltungen mit zwei oder mehreren Transistoren kann man die Eigenschaften der Gesamtschaltung noch deutlich verbessern.

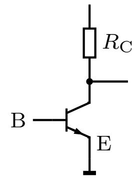
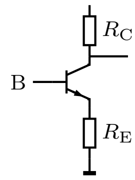
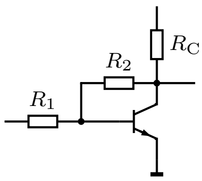
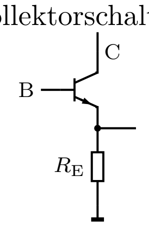
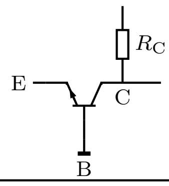
Grundschaltung	V_U	V_I	Z_{ein}	Z_{aus} für $\mathrm{d}U_{\text{ein}} = 0$
Emitterschaltung 	$-S\,R_{\text{C}} \parallel R_{\text{CE}}$	β	R_{BE}	$R_{\text{C}} \parallel R_{\text{CE}}$
Stromgegenkopplung 	$-\frac{R_{\text{C}}}{R_{\text{E}}}$	β	$R_{\text{BE}} + \beta\,R_{\text{E}}$	$R_{\text{C}} \parallel R_{\text{CE}}$
Spannungsgegengekoppelung 	$-\frac{R_2}{R_1}$	$\frac{R_2}{R_{\text{C}}}$	R_1	$R_{\text{C}} \parallel \frac{1}{S}\left(1 + \frac{R_2}{R_1}\right)$
Kollektorschaltung 	1	β	$\beta\,R_{\text{E}}$	$R_{\text{E}} \parallel \left(\frac{1}{S} + \frac{R_{\text{Signal}}}{\beta}\right)$
Basisschaltung 	$S\,R_{\text{C}}$	1	$\frac{1}{S}$	R_{C}

Abb. 5.26 Übersicht der Transistorgrundschaltungen mit den ungefähren Werten der Spannungs- und Stromverstärkung V_U und V_I, des Eingangs- und Ausgangswiderstandes Z_{ein} und Z_{aus}. Die Größe V_I steht für $dI_{Transistor}/dI_B$, nicht dI_{Last}/dI_B.

5.14 Stromquelle

Stromquellen sind wichtige Schaltungselemente und werden vielfältig in integrierten Schaltkreisen (ICs) und Vorverstärkern eingesetzt, so auch für Differenzverstärker.

Die ideale Stromquelle liefert einen Laststrom unabhängig von dem Spannungsabfall an der

Last. Nach Norton ist eine Spannungsquelle mit einem großen Innenwiderstand R_i einer Stromquelle äquivalent (Abb. 5.28).

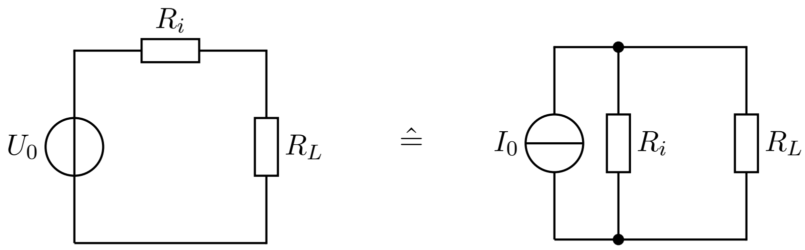


Abb. 5.27 Spannungsquelle und äquivalente Stromquelle nach Norton.

Auf diese Weise eine Stromquelle zu bauen, ist allerdings nicht praktikabel, da R_i den Wert des Laststroms stark begrenzt. So wäre für ein R_i von 10 MΩ (als typischer Wert für eine Stromquelle) und für einen immer noch kleinen Strom von 1 mA eine Spannung U_0 von 10 kV nötig. Hier ist es viel geschickter statt des passiven Widerstands R_i einen Transistor zu benutzen und den Arbeitspunkt so einzustellen, dass der differentielle Innenwiderstand dU_{aus}/dI_{aus} groß ist. Die Grundschaltung ist in Abbildung 5.28 gezeigt.

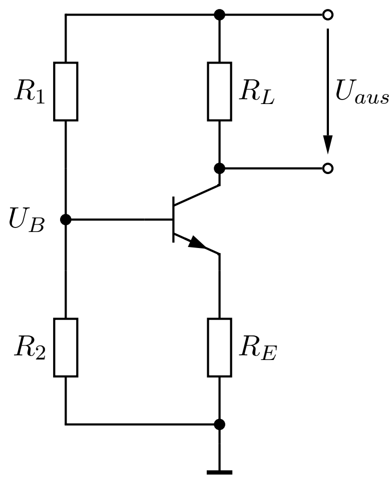


Abb. 5.28 Stromquelle realisiert mit einem Transistor.

Der Strom durch die Last $I_L = I_C$ ist in erster Näherung unabhängig von R_L und lässt sich durch die Wahl von U_B und R_E einstellen. Es gilt

$$I_L = I_C \approx I_E = \frac{U_E}{R_E} \approx \frac{U_B - 0{,}7\,\mathrm{V}}{R_E} \quad . \tag{5.38}$$

Wieso kann der Laststrom I_L unabhängig von R_L sein? Deshalb, weil sich U_{CE} und damit U_C so einstellt, dass gilt

$$I_L = \frac{U_{CC} - U_C}{R_L} \quad \text{und} \quad I_L = I_C = \beta\, I_B \quad .$$

Die Güte der Stromquelle hängt von der Stabilität des Stromes I_L bei Schwankungen von U_{aus}, d.h. der Variation von R_L ab. Bei fester Versorgungsspannung U_{CC} führt eine Variation dU_{aus} zu einer entgegengesetzten Änderung der Kollektor-Emitter-Spannung dU_{CE}. Abhängig von dem differentiellen Widerstand R_{CE} ist die Laststromvariation dI_L entsprechend klein:

Mit

$$dU_{BE} = -\, dU_E \quad (\text{da}\, dU_B = 0) = -\, R_E\, dI_E \approx -\, R_E\, dI_C \quad , \quad dU_{CE} \approx -dU_{aus}$$

und

$$dI_L = dI_C = S\, dU_{BE} + \frac{1}{R_{CE}}\, dU_{CE} \approx -S\, R_E\, dI_C - \frac{1}{R_{CE}}\, dU_{aus}$$

folgt

$$dI_L = -dU_{aus}\, [1 + S\, R_E] \quad .$$

Wichtig ist nicht nur ein hoher Innenwiderstand der Stromquelle, sondern auch Temperaturstabilität. Es gilt mit Gl. 5.38

$$\frac{dI_{aus}}{dT} = -\, \frac{1}{R_E}\, \frac{dU_{BE}}{dT} \approx \frac{2}{R_E}\, \frac{\text{mV}}{\text{K}} \quad .$$

Ein kleiner Emitterwiderstand R_E verringert also die Temperaturabhängigkeit. Durch Einsetzen einer Diode vor R_2 oder einer Zenerdiode anstelle von R_2 kann man die Temperaturabhängigkeit weiter reduzieren.

Für makroskopische Stromquellen ist der Wert von U_{CC} ein wichtiger Designparameter, da er die Lastspannung begrenzt bzw. den benötigten Spannungsabfall U_{CE} mitbestimmt. Ist U_{CC} zu groß, muss auch $U_{CE} = U_{CC} - U_{aus} - U_E$ groß sein und damit die Leistung des Transistors. Deshalb wird es sinnvoll sein, U_{CC} über eine variable Spannungsquelle an den Maximalwert U_{aus} der Last anzupassen (Abb. 5.29).

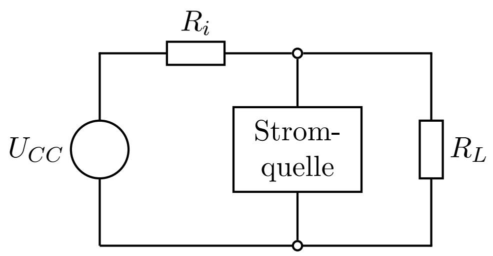


Abb. 5.29 Stromquelle realisiert mit vorgeschalteter regelbarer Spannungsquelle.

5.15 Millereffekt

Der Millereffekt hat mit der Kapazität C_{BC} zwischen Basis und Kollektor eines Transistors zu tun. Diese Kapazität liegt für den Emitterverstärker zwischen dem Eingang und dem Ausgang des Verstärkers. Dies ist schematisch in Abbildung 5.30 dargestellt. Durch Z tritt eine Rückkopplung auf.

Anmerkung: In unserem Fall ist Z eine Kapazität.

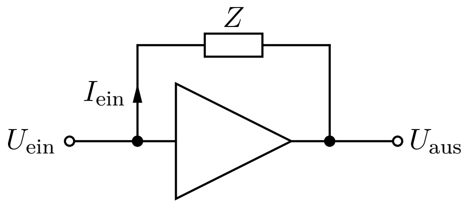


Abb. 5.30 Illustration des Millereffekts.

Welche Auswirkungen hat Z auf die Ausgangsimpedanz und die Eingangsimpedanz des Transistors? Dies ist nicht ganz offensichtlich, denn Z liegt nicht auf Masse, sondern verbindet Eingang und Ausgang!

Diese Fragestellung wurde von John M. Miller 1919 beantwortet [9]. Die Rückkopplung über Z entspricht einer Impedanz vom Eingang zu Masse von $\frac{\boldsymbol{Z}}{\boldsymbol{A+1}}$ zusammen mit einer weiteren Impedanz vom Ausgang zu Masse von $Z\,(\frac{1}{A}+1)$. Wir behandeln im folgenden nur die Eingangsimpedanz zu Masse.

Ihre Wirkung und Größe kann man leicht ausrechnen. Wie in Kapitel 6 bei der Diskussion des OPVs gehen wir davon aus, dass kein (oder nur ein kleiner) Strom in den Eingang des Verstärkers fließt. Dies ist nicht nur für OPVs, sondern auch für FETs in sehr guter Näherung und für Bipolartransistoren in guter Näherung erfüllt, da der Basisstrom klein ist.

Der Strom I_{ein} durch Z ist gegeben durch

$$I_{ein} = \frac{U_{ein} - U_{aus}}{Z} = \frac{U_{ein} + A\,U_{ein}}{Z} = U_{ein}\,\frac{A+1}{Z} \quad ,$$

wobei A die Spannungsverstärkung des invertierenden Verstärkers ist. Damit gilt

$$Z_{ein} = \frac{dU_{ein}}{dI_{ein}} = \frac{Z}{A+1} \quad .$$

Mit $Z = \frac{1}{\mathrm{j}\,\omega\,C}$ gilt

$$Z_{ein} = \frac{1}{\mathrm{j}\,\omega\,C\,(A+1)} \quad . \tag{5.39}$$

Dies bedeutet, dass sich die Kapazität zwischen Basis und Kollektor wie eine deutlich größere effektive Kapazität zwischen Eingang und Masse auswirkt (Abb. 5.31).

Dies ist der Millereffekt. Der Millereffekt ist für schnelle Signale recht störend, da eine Kapazität C_{BC} von einigen pF leicht um ein bis zwei Größenordnungen vergrößert erscheint, als Tiefpass wirkt und das Eingangssignal verlangsamt. Die Zeitkonstante τ ist durch den Innenwiderstand der Signalquelle R_S und die Millerkapazität C_M gegeben zu

$$\tau = R_S\,C_M \quad .$$

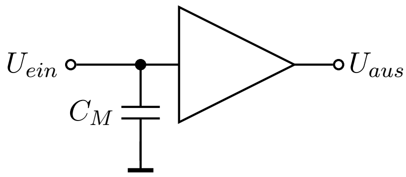


Abb. 5.31 Illustration des Millereffekts. $C_M = \mathrm{j}\,\omega\,C\,(A+1)$.

5.16 Frequenzverhalten von Verstärkerschaltungen

Wir haben bislang eine sehr wesentliche Charakteristik von Verstärkern ignoriert, nämlich ihr Frequenzverhalten. Um das Frequenzverhalten zu verstehen, müssen wir die vorhandenen externen und internen Kapazitäten in die Schaltbilder einfügen und bei der Analyse berücksichtigen. Folgenden Kapazitäten treten auf:

- Koppelkapazitäten C_{ein} und C_{aus} für Eingangssignal und Ausgangssignal
- Lastkapazität, falls vorhanden
- internen Transistorkapazitäten, nämlich C_{BE}, C_{BC} und C_{CE} (siehe Abb. 5.35)

Zur Illustration betrachten wir den Emitterverstärker mit Stromgegenkopplung (Abb. 5.32). Im Gegensatz zur Abbildung 5.20 ist hier noch ein Kondensator C_E parallel zum Emitterwiderstand R_E geschaltet. Dies ist eine Standardkonfiguration.

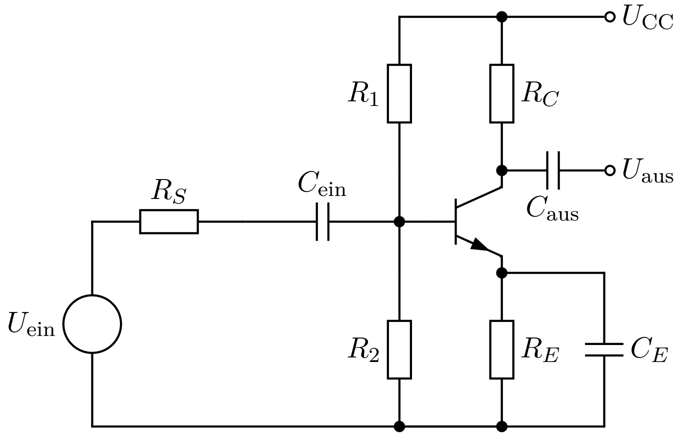


Abb. 5.32 Emitterverstärker mit Stromgegenkopplung.

Die Kleinsignalersatzschaltung in Abbildung 5.32 ist in Abbildung 5.33 dargestellt.

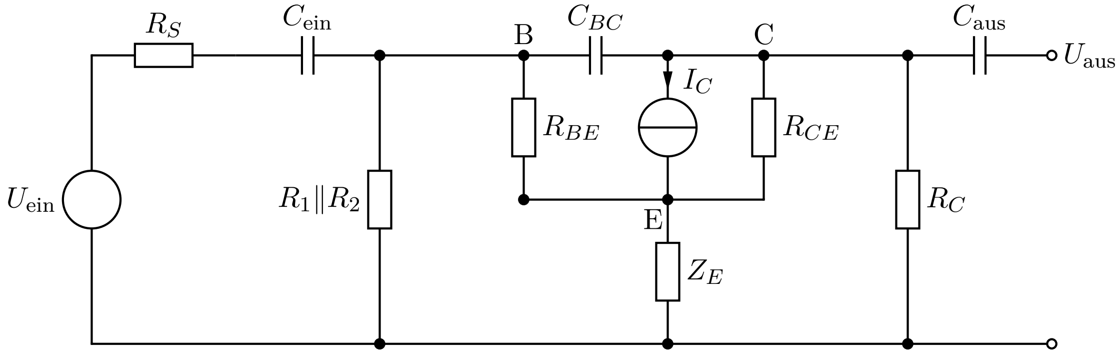


Abb. 5.33 Kleinsignalersatzschaltung zu Abb. 5.32. Z_E steht für $R_E \parallel \frac{1}{\mathrm{j}\,\omega\, C_E}$.

Das Frequenzverhalten dieser Schaltung sieht grob so aus wie in Abbildung 5.34. Wir identifizieren vier Grenzfrequenzen, bei denen sich die Verstärkung ändert.

Die Eingangskoppelkapazität bildet zusammen mit dem Eingangswiderstand R_{ein} des Transistors und den Widerständen R_1 und R_2 zur Einstellung der Basisruhespannung einen Hochpass. Das heißt, wir haben eine Grenzfrequenz

$$f_1 \approx \frac{1}{2\,\pi\, C_{ein}\,(R_S + R_1 \parallel R_2 \parallel [R_{BE} + (\beta+1)\, Z_E])} \approx \frac{1}{2\,\pi\, C_{ein}\,(R_S + R_{\text{ein}})}$$

identifiziert, unterhalb der nicht das volle Eingangssignal die Basis erreicht. Durch die Wahl eines genügend großen C_{ein} kann man f_1 so legen, dass alle relevanten Eingangsfrequenzen durchkommen; zu kleine Frequenzen werden unterdrückt. Dasselbe gilt auch für C_{aus}.

Anmerkung: Der Wert von R_S ändert nicht nur die Eingangsgrenzfrequenz, sondern auch den Signalpegel des an der Basis gesehenen Eingangssignals U_B um einen frequenzunabhängigen Faktor $R_{ein}/(R_{ein} + R_S)$, da

$$U_B = U_{\text{ein}}\,\frac{R_{\text{ein}}}{R_{\text{ein}} + R_S + Z_{\text{ein}}} = U_{\text{ein}}\,\frac{R_{\text{ein}}}{R_{\text{ein}} + R_S + \frac{1}{sC}} = U_{\text{ein}}\,\frac{\frac{R_{\text{ein}}}{R_S + R_{\text{ein}}}}{1 + \frac{1}{sC(R_{\text{ein}} + R_S)}}$$

statt des allgemeinen Ausdrucks $U_{\text{ein}}\,\frac{1}{1+1/sCR}$ für einen einfachen Hochpass aus R und C.

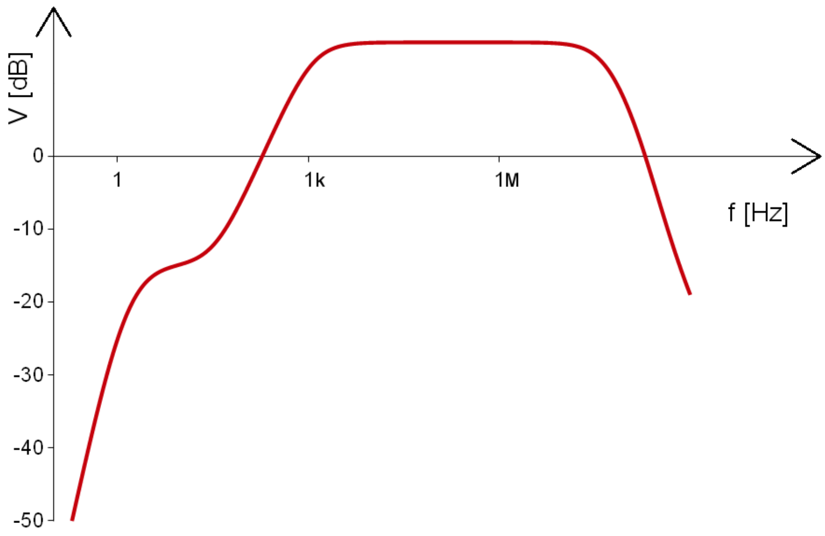


Abb. 5.34 Schematischer Verlauf der Spannungsverstärkung in dB des Emitterverstärkers mit Stromgegenkopplung aus Abb. 5.32 als Funktion der Frequenz.

Gleichzeitig wirkt C_E in Verbindung mit R_{BE} als Tiefpass mit

$$f_2 = \frac{1}{2\,\pi\,C_E\,R_{BE}} \quad .$$

Es gibt einen weiteren Hochpass in dieser Schaltung, der durch R_E und C_E gebildet wird. Die Gegenkopplung durch R_E legt die Verstärkung für kleine Frequenzen fest und minimiert Verschiebungen aufgrund von (langsamen) Temperaturschwankungen. Die Verstärkung beträgt $-R_C/R_E$. Für große Frequenzen wird C_E durchlässig, die Gegenkopplung nimmt ab und die Verstärkung nimmt zu. Dieser Übergang setzt mit der Grenzfrequenz

$$f_3 = \frac{1}{2\pi R_E C_E}$$

ein.

Mit einigen bequemen Vereinfachungen wie $\beta+1 \approx \beta$, $R_{CE} = \infty$ und $R_{BE} << \beta R_E$ lässt sich die Verstärkung in diesem Frequenzbereich leicht berechnen. Wir setzen $R_S = 0$, $Z_{\text{ein}} = 0$,

$Z_{\text{aus}} = 0$ und $Z_{\text{BC}} = \infty$. Dann gilt mit $Z_E = R_E/(sC_E R_E + 1)$

$$\begin{aligned} V_U &= \frac{dU_{\text{aus}}}{dU_{\text{ein}}} = \frac{-(\beta+1)I_B R_C}{I_B \cdot R_{BE} + (\beta+1)I_B \cdot Z_E} \\ &= -\frac{(\beta+1)R_C}{R_E} \frac{sC_E R_E + 1}{(sC_E R_E + 1)\frac{R_{BE}}{R_E} + (\beta+1)} \\ &= -\frac{R_C}{R_E} \frac{sC_E R_E + 1}{(sC_E R_E + 1)\frac{R_{BE}}{(\beta+1)R_E} + 1} \\ &\approx -\frac{R_C}{R_E} \frac{(sR_E C_E + 1)}{(sC_E \frac{R_{BE}+1}{\beta+1})} \\ &= -\frac{R_C}{R_E} \frac{j\omega/\omega_2 + 1}{j\omega/\omega_3 + 1} \end{aligned} \tag{5.40}$$

Man sieht, dass neben dem schon betonten Hochpassverhalten bei f_1 bei hohen Frequenzen $f_3 = \frac{\omega_3}{2\pi} \approx \frac{\beta}{2\pi\, R_{BE} C_E} = \frac{S}{2\pi\, C_E}$ eine Abflachung der Verstärkung einsetzt.

Die genaue Lage der Grenzfrequenzen in Abb. 5.34 hängt von den gewählten Werten C_{ein}, C_E, R_E und dem Arbeitspunkt ab und ist jetzt nicht von Interesse. Wichtig ist aber der Abfall der Verstärkung für Frequenzen größer als f_4. Dieser Abfall ist nicht durch die externen, sondern die internen Kapazitäten des Transistors verursacht, die durch die pn- und np-Übergänge zwischen Kollektor und Basis und Emitter gebildet werden (Abb. 5.35).

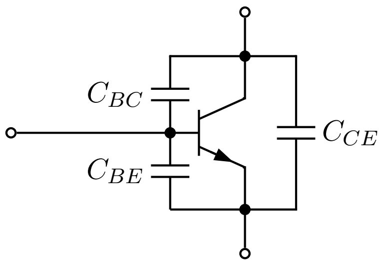


Abb. 5.35 Illustration der internen Transistorkapazitäten.

Die Werte dieser Kapazitäten liegen im Bereich von pF und sind daher klein. Wichtig ist aber die Kapazität zwischen Kollektor und Basis, die durch den Millereffekt (Abschnitt 5.15) vergrößert wird. Dies führt bei hohen Frequenzen zu einem Tiefpassverhalten, das typisch für jeden Verstärker ist und das wir schon am Beispiel des OPVs kennengelernt haben.

Wie kann man den Millereffekt umgehen? Man muss R_S oder C_M verkleinern. Wie? R_S kann man durch einen vorgeschalteten Emitterfolger begrenzen. Den Wert der Millerkapazität C_M kann man durch eine reduzierte Verstärkung begrenzen. Dies erreicht man durch negative Rückkopplung. Noch mehr Möglichkeiten ergeben sich bei Schaltungen mit zwei oder mehr Transistoren wie z.B. der Kaskodenschaltung (nächstes Kapitel).

6 Laplacetransformation*

Die Laplacetransformation (LT) ist ein sehr effektives Mittel, um Einschwingvorgänge von Hochpässen, Tiefpässen, Filtern und Schwingkreisen zu berechnen. Voraussetzung für die Anwendung des Laplaceformalismus ist eine Anregungsfunktion $f(t)$ mit $f(t) = 0$ für $t < 0$. Die Laplacetransformation von $f(t)$ ist

$$\mathscr{L}\{f(t)\} = F(s) = \int_0^\infty f(t)\,\mathrm{e}^{-s\,t}\,dt \tag{6.1}$$

wobei die komplexe Variable s durch $s = x + \mathrm{j}\,\omega$ definiert wird. Der Unterschied zur Fouriertransformation ist der Konvergenz erzeugende reelle Faktor e^{-x}. Die Umkehrfunktion der Laplacetransformation ist

$$f(t) = \mathscr{L}^{-1}\{F(s)\} = \frac{1}{\mathrm{j}\,2\,\pi}\int_{x-\mathrm{j}\,\infty}^{x+\mathrm{j}\,\infty} F(s)\,\mathrm{e}^{s\,t}\,ds \tag{6.2}$$

Der Vorteil des Laplaceformalismuses ist es, dass man im Laplaceraum statt mit Ableitungen höheren Grades mit Polynomen rechnen kann. Gleichzeitig kann die Übertragungsfunktion angewandt auf die Anregungsfunktion leicht mittels Tabellen in den Zeitraum transformiert werden. Eine typische Anregungsfunktion ist z. B. die Sprungfunktion

$$u_s = \begin{cases} 0 & \text{für} \quad t < 0 \\ 1 & \text{für} \quad t \geq 0 \end{cases} \tag{6.3}$$

Allgemeingültige Eigenschaften der Laplacetransformation sind:

1. **Linearität:**

$$\mathscr{L}\{a\,f_a(t) + b\,f_b(t)\} = a\,\mathscr{L}\{f_a(t)\} + b\,\mathscr{L}\{f_b(t)\} \tag{6.4}$$

2. **Verschiebungseigenschaft:**

$$\mathscr{L}\{f(t - t_0)\} = \mathrm{e}^{-s\,t_0}\mathscr{L}\{f(t)\} \tag{6.5}$$

   für $t_0 \geq 0$ und $f(t < t_0) = 0$. Dies ist relevant für Anregungsfunktionen, die nicht zum Zeitpunkt $t = 0$, sondern zu $t = t_0$ einsetzen.

3. **Ableitungseigenschaft:**

$$\mathscr{L}\left\{\frac{df(t)}{dt}\right\} = s\,\mathscr{L}\{f(t)\} - f(0) \tag{6.6}$$

   Zur Herleitung dieser und weitere Eigenschaften siehe Abschnitt 6.4.

   Durch Übergang in den Laplaceraum wird die Ableitung durch einen Vorfaktor s und einen konstanten Term $f(0)$ ersetzt. Durch wiederholte Anwendung des Ableitungssatzes kann man auch höhere Ableitungen vereinfachen.

$$\mathscr{L}\left\{\frac{d^n f(t)}{dt}\right\} = s^n\,\mathscr{L}\left\{f(t)\right\} - s^{n-1}\,f(0) - s^{n-2}\,\frac{df(0)}{dt}\cdots - \frac{d^{(n-1)}f(0)}{dt} \tag{6.7}$$

Allgemein gilt: Für die Anwendung des Laplaceformalismus muss das Integral

$$\int_0^\infty |f(t)\,\mathrm{e}^{-s\,t}|\,dt \tag{6.8}$$

konvergieren, $f(t)$ dürfte also z. B. keine exponentiell ansteigende Funktion sein

$$|f(t)| \leq M\,\mathrm{e}^{a\,t} \quad \Rightarrow \quad F(s) \text{ existiert für alle } s > a \tag{6.9}$$

4. **Skalierungseigenschaft:**

$$\mathscr{L}\left\{f(a\,t)\right\} = \frac{1}{a}\,F\left(\frac{s}{a}\right) \tag{6.10}$$

5. **Integraleigenschaft:**

$$\mathscr{L}\left\{\int_0^t f(t')\,dt'\right\} = \frac{1}{s}\,F(s) \tag{6.11}$$

6.1 Transformationstabelle

Zur Anwendung des Laplaceformalismus ist es nicht nötig, die auftretenden Integrale auszurechnen. Es ist ratsam, eine entsprechende Transformationstabelle (Tab. 6.1) anzuwenden, da sich so komplexe Ausdrüke sehr schnell und effizient lösen lassen.

Zeitraum	Laplace s-Raum
Stufenfunktion: $f(t) = \begin{cases} 0 & \text{für } t < 0 \\ 1 & \text{für } t \geq 0 \end{cases}$	$\frac{1}{s}$
$e^{a\,t}$	$\frac{1}{s-a}$
e^{-at}	$\frac{1}{s+a}$
$\sin(a\,t)$	$\frac{a}{s^2+a^2}$
$\cos(a\,t)$	$\frac{s}{s^2+a^2}$
Linearer Anstieg: $f(t) = \begin{cases} 0 & \text{für } t < 0 \\ t & \text{für } t \geq 0 \end{cases}$	$\frac{1}{s^2}$
$f(t) = \begin{cases} 0 & \text{für } t < 0 \\ \frac{t^n}{n!} & \text{für } t \geq 0 \end{cases}$	$\frac{1}{s^2}$
$\delta(t)$	1
$1 - e^{-\,a\,t}$	$\frac{a}{s\,(s+a)}$
$t\,e^{-\,a\,t}$	$\frac{1}{(s+a)^2}$
$e^{-\,a\,t}\,\sin(b\,t)$	$\frac{b}{(s+a)^2+b^2}$
$e^{-\,a\,t}\cos(b\,t)$	$\frac{s+a}{(s+a)^2+b^2}$

Tab. 6.1 Laplacetransformationstabelle.

6.2 Anwendungen der Laplacetransformation

Beispiel 1: Antwort eines Hochpasses auf eine Stufenspannung am Eingang

Die Antwortfunktion U_a im Laplaceraum ist gegeben durch die Multiplikation der Funktion der Eingangsspannung U_e und der Übertragungsfunktion $H(s)$ jeweils im Laplaceraum $U_a(s) = U_e(s) \cdot H(s)$.
Die Übertragungsfunktion des Hochpasses ist $\frac{R}{R+Z_C}$ im Zeitraum (Abs. 2.4). Im Laplaceraum transformieren sich die Impedanzen von Widerstand, Kondensator und Spule wie folgt: $R \to R$, $\mathrm{j}\,\omega\,L \to s\,L$, $\frac{1}{\mathrm{j}\,\omega\,C} \to \frac{1}{s\,C}$. Dies entspricht der Transformation der Differentialgleichungen:

$$i(t) = C\,\frac{du(t)}{dt} \to \mathscr{L}\,\{i(t)\} = C\,[s\,\mathscr{L}\,\{u(t)\} + u(0)] \tag{6.12}$$

$$u(t) = L\,\frac{di(t)}{dt} \to \mathscr{L}\,\{u(t)\} = L\,[s\,\mathscr{L}\,\{i(t)\} + i(0)] \tag{6.13}$$

und $Z = \frac{U}{I}$. Hier werden die Anfangsbedingungen $u(0) = 0$ und $i(0) = 0$ angenommen.

Die Übertragungsfunktion des Hochpasses ist damit

$$\frac{R}{R + \frac{1}{s\,C}} = \frac{1}{1 + \frac{1}{s\,R\,C}} = \frac{s}{s+a} \quad \text{mit} \quad a = \frac{1}{R\,C} \quad . \tag{6.14}$$

Dies ist uns schon vertraut, da wir Teile des Laplaceformalismus bei der Diskussion diverser Filter schon benutzt haben. Die Stufenfunktion U_e mit der Amplitude U_0 im Laplaceraum ist $\frac{U_0}{s}$. Also

$$U_a(s) = U_e(s)\,H(s) = \frac{U_0}{s+a} \tag{6.15}$$

Die Rücktransformation ergibt sich am einfachsten aus der Transformationstabelle.

$$u_a(t) = \mathscr{L}^{-1}\left\{U_a(s)\right\} = U_0\,\mathrm{e}^{-a\,t} = U_0\,\mathrm{e}^{-\frac{t}{RC}} \tag{6.16}$$

Dies ergibt also den wohlbekannten exponentiellen Abfall mit der Zeitkonstante $\tau = RC$.

Beispiel 2: Antwort eines Hochpasses auf eine exponentiell abfallende Funktion

Es sei

$$\mathscr{L}\left\{\mathrm{e}^{-a\,t}\right\} = \frac{1}{s+a} \quad .$$

Also folgt

$$U_a(s) = \frac{U_0}{s+a}\cdot\frac{s}{s+a} = U_0\,\frac{s}{(s+a)^2} = U_0\,\frac{s+a-a}{(s+a)^2} \tag{6.17}$$

$$\begin{aligned}\mathscr{L}^{-1}\left\{U_a(s)\right\} = \mathscr{L}^{-1}\left\{\frac{1}{s+a} - \frac{a}{(s+a)^2}\right\} &= U_0\left(\mathrm{e}^{-a\,t} - a\,t\,\mathrm{e}^{-a\,t}\right)\\ &= U_0\left(1-\frac{t}{RC}\right)\mathrm{e}^{-\frac{t}{RC}}\end{aligned} \tag{6.18}$$

Der Trick ist häufig, den Ausdruck im Laplace- oder Zeitraum so zu zerlegen, dass man auf eine bekannte Tabellengröße zurückgreifen kann.

Beispiel 3: CR-RC-Pulswandler

Ein CR-RC-Filter ist wichtig für die Detektortechnik. Die Wirkungsweise ist in Abbildung 6.1 illustriert. Ein Stufenimpuls wird erst vorverstärkt, dann differenziert, über einen Impedanzwandler gepuffert und dann integriert. Warum?

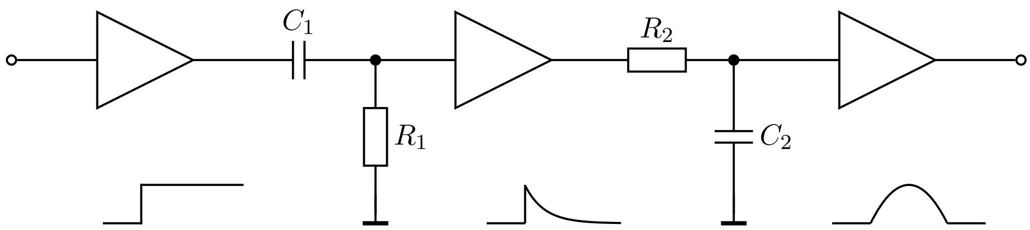


Abb. 6.1 CR-RC-Filter mit Vorverstärker, Hochpass C_1, R_1, Impedanzwandler, gefolgt von einem Tiefpass R_2, C_2 und einem weiteren Impedanzwandler oder Verstärker.

Die erste Stufe des Filters, der Hochpass, bewirkt, dass das Signal mit einer definierten Zeitkonstanten abklingt. Dies entspricht unserem Beispiel 1. Dadurch wird der Verstärker schneller aufnahmebereit für das nächste Signal und aufeinanderfolgende Signale überlagern sich

nicht (reduzierter *pile-up*). Die zweite Stufe führt zu einer langsamen Anstiegszeit des Signals und eliminiert die Stufe des Eingangsimpulses. Dadurch verringert sich die Bandbreite und dementsprechend auch das Rauschen.
Wir wissen aus Beispiel 1, dass die Stufenfunktion durch den Hochpass in einen exponentiell abfallenden Puls $\mathrm{e}^{-a\,t}$ mit $a = \frac{1}{C_1\,R_1} = \frac{1}{\tau_D}$ umgeformt wird. Der Puls $\mathrm{e}^{-a\,t}$ hat im Laplaceraum die Form $\frac{1}{s+a}$. Die Übertragungsfunktion des Tiefpasses $\frac{Z_{C_2}}{Z_{C_2}+R_2}$ ist im Laplaceraum

$$\frac{\frac{1}{s\,C_2}}{R_2 + \frac{1}{s\,C_2}} = \frac{1}{s\,R_2\,C_2 + 1} = \frac{b}{s+b} \tag{6.19}$$

mit $b = \frac{1}{R_2\,C_2} = \frac{1}{\tau_I}$. Damit ergibt sich $U_a(s) = \frac{b}{(s+a)(s+b)}$. Zur Transformation ist es hilfreich, diesen Ausdruck mit Hilfe der Partialbruchzerlegung zu vereinfachen und auf die Form $\frac{P_1(s)}{s+b} + \frac{P_2(s)}{s+a}$ zu bringen, wobei $P_1(s)$ und $P_2(s)$ Polynome von s sind. Es ergibt sich

$$U_a(s) = \frac{b}{(s+a)(s+b)} = \frac{b}{a-b}\left[\frac{1}{s+b} - \frac{1}{s+a}\right] \quad . \tag{6.20}$$

Jetzt ist die Rücktransformation einfach, da die benötigten Tabellenfunktionen in Tabelle 6.1 enthalten sind, und ergibt

$$u_a(t) = \left(\frac{b}{a-b}\right)\left[\mathrm{e}^{-b\,t} - \mathrm{e}^{-a\,t}\right] \quad .$$

Nach Einsetzen von $a = \frac{1}{\tau_D}$ und $b = \frac{1}{\tau_I}$ ergibt sich das Endergebnis

$$u_a(t) = \frac{\tau_D}{\tau_D - \tau_I}\left[\mathrm{e}^{-\frac{t}{\tau_D}} - \mathrm{e}^{-\frac{t}{\tau_I}}\right] \quad . \tag{6.21}$$

Für $\tau_I = \tau_D = \tau$ gilt

$$u_a(t) \approx \frac{t}{\tau}\,\mathrm{e}^{-\frac{t}{\tau}} \quad .$$

Anmerkung: Dies kann durch die Taylorentwicklung der Näherung und des Ausdrucks aus Gl. 6.21 überprüft werden.

Beispiel 4: Reihenschwingkreis angeregt durch eine Stufenfunktion

Wir betrachten Abbildung 6.2.

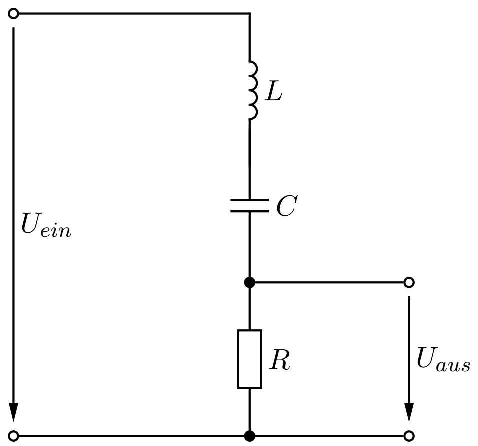


Abb. 6.2 Reihenschwingkreis mit Spule L, Kondensator C und Widerstand R.

Uns interessiert der Spannungsabfall an R, also betrachten wir die Übertragungsfunktion

$$\frac{R}{R + s\,L + \frac{1}{s\,C}} = \frac{R}{L}\,\frac{s}{s^2 + s\,\frac{R}{L} + \frac{1}{L\,C}} = P(s) \quad . \tag{6.22}$$

Hier, wie auch im Allgemeinen, ist es wichtig, die Pole und Nullstellen des Polynoms $P(s)$ zu kennen. Die Nullstellen liegen bei $s = 0$ und $|s| = \infty$. Die Pole ergeben sich aus der quadratischen Gleichung

$$s^2 + s\,\frac{R}{L} + \frac{1}{L\,C} = s^2 + 2\,a\,s + \omega_0^2$$

zu

$$s_{1/2} = -\,a \pm \sqrt{a^2 - \omega_0^2} \tag{6.23}$$

mit $a = \frac{R}{2\,L}$ und $\omega_0^2 = \frac{1}{L\,C}$.

Abhängig von dem Wert von a und ω_0 ergeben sich unterschiedliche Lösungen für $a^2 \geq \omega_0^2$, $a^2 < \omega_0^2$ und $a^2 = \omega_0^2$. Jetzt muss die Anregungsfunktion festgelegt und in den Laplaceraum transformiert werden. Für eine Stufenanregung gilt

$$u_a(t) = \frac{R}{L}\,\mathscr{L}^{-1}\left\{\frac{1}{s}\cdot\frac{s}{s^2 + 2as + \omega_0^2}\right\} \tag{6.24}$$

Diese Funktion hat keine Nullstelle bei $s = 0$ mehr. Die Lage der Pole ist unverändert.

Die Rücktransformation hängt vom Wert der Pole ab. Für $\omega_0^2 \geq a^2$ gilt

$$u_a(t) = \frac{R}{L\,\sqrt{\omega_0^2 - a^2}}\,\mathrm{e}^{-at}\,\sin\left(\sqrt{(\omega_0^2 - a^2)\,t}\right) = \frac{R}{L\,\omega_D}\mathrm{e}^{-\frac{R}{2\,L}\,t}\,\sin(\omega_D\,t) \tag{6.25}$$

mit $\omega_D^2 = \left(\frac{1}{LC}\right)^2 - \left(\frac{R}{2L}\right)^2$.

6.3 Pole und Nullstellen der Laplacetransformation

Die Übertragungsfunktion $H(s)$ beschreibt vollständig das Verhalten des Systems.

Für eine Impulsanregung $\delta(t)$ gilt, jeder Pol p_i von $H(s)$ führt zu einer Systemantwort mit Zeitverhalten der Form $\mathrm{e}^{p_i\,t}$.

Ist p_i reell und negativ, so ergibt sich ein exponentieller Abfall.

Ist p_i reell und positiv, so ergibt sich ein exponentieller Anstieg.

Hat p_i einen nicht verschwindenden Imaginärteil $p_i = -\,a + \mathrm{j}\,\omega$ mit $a > 0$, so verhält sich das System bei einer Impulsanregung wie $\mathrm{e}^{-\,a\,t}\cdot\sin(\omega\,t + \phi)$. Der Wert der Phase ergibt sich aus den Randbedingungen. Zu jedem nicht reellen Pol existiert auch der komplex konjugierte Pol. Imaginäre Pole treten nur bei quadratischen Termen im Nenner von $H(s)$ auf z. B. $\frac{1}{(s+a)^2+b^2}$.

Entsprechend hat ein rein imaginärer Pol ($p_i = \pm\mathrm{j}\,\omega$), die Systemantwort $\sin(\omega\,t + \phi)$.

Ist $p_i = 0$, so führt dies zu einer zeitlich konstanten Systemantwort.

Entsprechend führt eine Nullstelle von $H(s)$ zu einer verschwindenden Systemantwort.

6.4 Herleitung der Eigenschaften der Laplacetransformation

Herleitung der Ableitungseigenschaft

$$\mathscr{L}\left\{\frac{df(t)}{dt}\right\} = \int_0^\infty \frac{d}{dt}\{f(t)\,\mathrm{e}^{-st}\}dt$$

Es gilt

$$\begin{aligned}\int_0^\infty \{\frac{d}{dt}f(t)\}\,\mathrm{e}^{-st}dt &= \int_0^\infty \left\{\frac{d}{dt}f(t)\,\mathrm{e}^{-st}dt\right\} - s\int_0^\infty f(t)\,\mathrm{e}^{-st}dt\\ &= \mathscr{L}\left\{\frac{d}{dt}f(t)\right\} - s\mathscr{L}\{f(t)\}\end{aligned}$$

und gleichzeitig

$$\int_0^\infty \frac{d}{dt}\left\{f(t)\,\mathrm{e}^{-s\,t}\right\}\,dt = 0 - f(0) \quad .$$

Also gilt auch

$$\mathscr{L}\left\{\frac{d}{dt}f(t)\right\} = s\,\mathscr{L}\{f(t)\} - f(0) \quad .$$

Herleitung der Integraleigenschaft

$$\mathscr{L}\left\{\int_0^t f(\tau)d\tau\right\} = \int_0^\infty \left[\int_0^t f(\tau)d\tau\right]\mathrm{e}^{-st}dt$$

mit

$$x = \frac{-\mathrm{e}^{-st}}{s} \Rightarrow \frac{dx}{dt} = \mathrm{e}^{-st}$$

und $y = \int_0^t F(\tau)\,d\tau \Rightarrow \frac{dy}{dt} = F(\tau)$. Da F(0)=0, ergibt sich durch partielle Integration

$$\mathscr{L}\left\{\int_0^t f(\tau)d\tau\right\} = \left[\frac{-\mathrm{e}^{-st}}{s}\int_0^t f(\tau)d\tau\right]_0^\infty - \left(-\frac{1}{s}\int_0^\infty f(t)\mathrm{e}^{-st}dt\right) = \frac{1}{s}F(s) \quad ,$$

da der Ausdruck in den eckigen Klammern Null ergibt.

Stufenfunktion u(t)

$$\mathscr{L}\{u(t)\} = \int_0^\infty u(t)\mathrm{e}^{-s\,t}dt = \int_0^\infty \mathrm{e}^{-s\,t}\,dt = \left[-\frac{1}{s}\,\mathrm{e}^{-s\,t}\right]_0^\infty = \frac{1}{s}$$

$\mathbf{F(t) = e^{a\,t}}$

$$\mathscr{L}\left\{e^{a\,t}\right\} = \int_0^\infty e^{(a-s)t}\,dt = \frac{1}{a-s}\left[e^{(a-s)\,t}\right]_0^\infty = \frac{1}{s-a}$$

entsprechend für $\mathbf{f(t) = e^{-\,a\,t}}$

$\mathbf{f(t) = \sin(\omega\,t)}$

$$f(t) = \sin(\omega\,t) = \frac{e^{j\,\omega\,t} - e^{-j\,\omega\,t}}{2\,j}$$

$$\begin{aligned}\mathscr{L}\,\{\sin(\omega\,t)\} =& \frac{1}{2\,j}\left[\frac{1}{s-j\,\omega} - \frac{1}{s+j\,\omega}\right] = \frac{1}{2\,j}\left[\frac{s+j\,\omega-s+j\,\omega}{(s-j\,\omega)(s+j\,\omega)}\right]\\ =& \frac{2\,j\,\omega}{2\,j(s^2+\omega^2)} = \frac{\omega}{s^2+\omega^2}\end{aligned}$$

$\mathbf{f(t) = \cos(\omega t)}$

$$f(t) = \cos(\omega t) = \frac{e^{j\omega t} + e^{-j\omega t}}{2}$$

$$\mathscr{L}\,\{f(t)\} = \frac{1}{2}\left[\frac{1}{s-j\omega} + \frac{1}{s+j\omega}\right] = \frac{1}{2}\frac{2s}{s^2+\omega^2}$$

Linearer Anstieg: $\mathbf{f(t) = t}$ für $\mathbf{t \geq 0}$ und $\mathbf{0}$ für $\mathbf{t < 0}$

$$\mathscr{L}\,\{F(t)\} = \int_0^\infty te^{-st}dt = \int_0^\infty \frac{d}{dt}(te^{-st})dt = \left[te^{-st}\right]_0^\infty = 0$$

$$= \int_0^\infty e^{-st}dt - s\int_0^\infty te^{-st}dt$$

$$\Leftrightarrow -s\int_0^\infty te^{-st}dt = -\frac{1}{s}\left[e^{-st}\right]_0^\infty = -\frac{1}{s}$$

$$\Leftrightarrow \mathscr{L}\,\{f(a\,t)\} = \int_0^\infty te^{-st}dt = \frac{1}{s^2} \tag{6.26}$$

Skalierungseigenschaft

$$\mathscr{L}\,\{f(a\,t)\} = \frac{1}{a}\,F\left(\frac{s}{a}\right)$$

Da

$$\mathscr{L}\,\{f(at)\} = \int_0^\infty f(at)e^{-st}dt \overset{t'=at}{=} \frac{1}{a}\int_0^\infty f(t')\,e^{-\frac{s}{a}t'}dt' = \frac{1}{a}F(\frac{s}{a})$$

7 Fouriertransformation*

Die Fourierzerlegung von Signalen in ihre Frequenzkomponenten gehört zu den wichtigsten mathematischen Werkzeugen der Physik und Elektrotechnik. Dies liegt zum einen daran, dass harmonische Schwingungen und Signale in Natur und Technik oft vorkommen. Ein weiterer Grund ist, dass die meisten elementaren elektronischen Schaltungen und Bausteine zwar die Amplitude und Phase einer harmonischen Schwingung verändern, nicht aber die Frequenz.

Es gibt unterschiedliche Formen der Fourierzerlegung abhängig von den Eigenschaften des Eingangssignals.

7.1 Fourierreihe

Periodische kontinuierliche Signale $f(t)$ lassen sich als Fourierreihe darstellen, also

$$f(t) = \sum_{n=-\infty}^{\infty} c_n \, \mathrm{e}^{\mathrm{j}2\pi \frac{nt}{T}} = \sum_{n=-\infty}^{\infty} c_n \, \mathrm{e}^{\mathrm{j}\, n\, \omega\, t} \quad . \tag{7.1}$$

Hierbei ist T die Periodendauer, entsprechend folgt für die Kreisfrequenz $\omega = \frac{2\pi}{T}$, c_n sind die komplexen Koeffizienten, $n\omega = \frac{2\pi n}{T}$ ist die Kreisfrequenz der n-ten Frequenzkomponente.

Beliebt ist auch die reelle Zerlegung von $f(t)$ mit

$$\begin{aligned} f(t) &= a_0 + \sum_{n=1}^{\infty} a_n \cos\left(\frac{2\pi n t}{T}\right) + b_n \sin\left(\frac{2\pi n t}{T}\right) \\ &= a_0 + \sum_{n=1}^{\infty} a_n \cos(n\,\omega t) + b_n \sin(n\,\omega t) \quad . \end{aligned} \tag{7.2}$$

Die Werte der Koeffizienten c_n bzw. a_n und b_n werden mit

$$c_n = \frac{1}{T} \int_0^T f(t)\, \mathrm{e}^{-\mathrm{j}2\pi \frac{nt}{T}} \, dt \leq \frac{1}{T} \int_0^T f(t)\, \mathrm{e}^{-\mathrm{j}\, n\, \omega t} \, dt \tag{7.3}$$

$$a_0 = \frac{1}{T} \int_0^T f(t) \, dt \tag{7.4}$$

$$a_n = \frac{2}{T} \int_0^T f(t) \cos(n\,\omega t) \, dt, \quad n = 1, 2 \ldots \tag{7.5}$$

$$b_n = \frac{2}{T} \int_0^T f(t) \sin(n\,\omega t) \, dt, \quad n = 1, 2 \ldots \tag{7.6}$$

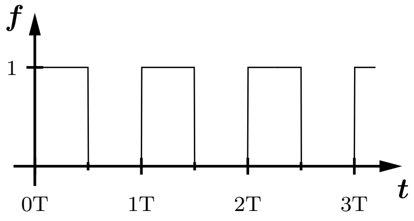
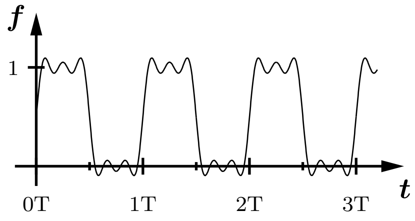
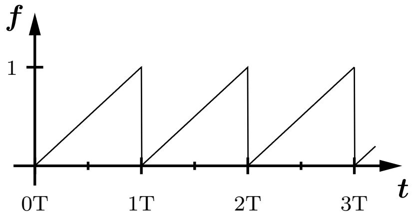
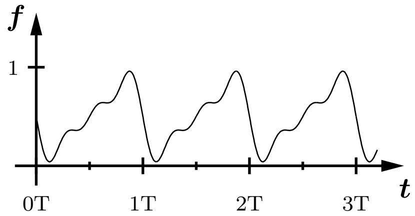
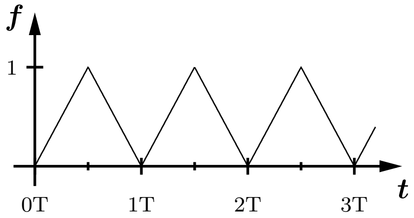
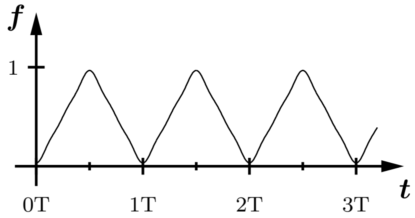
bestimmt. Die komplexe Zerlegung ist eine mathematisch kompakte Darstellung. Im Reellen muss man die leicht unterschiedlichen Normierungen $\frac{1}{T}$ für a_0 und $\frac{2}{T}$ für a_n, b_n mit $n \geq 1$ beachten. Wichtig ist, dass bei der Zerlegung von $f(t)$ nur ganzzahlige Vielfache der Kreisfrequenz $\omega = \frac{2\pi}{T}$ der Ausgangsfunktion $f(t)$ auftauchen. Die Koeffizienten a_n, b_n und c_n sind über die Beziehungen

$$\begin{aligned} a_0 &= c_0 \\ a_n &= 2\,\mathrm{Re}(c_n) \quad \text{und} \quad b_n = -2\,\mathrm{Im}(c_n) \quad \text{für } n \geq 1 \end{aligned}$$

miteinander verknüpft.

Die Kenntnis der Fourierkoeffizienten c_n bzw. a_n und b_n ist äquivalent zu der Kenntnis der Ausgangsfunktion $f(t)$.

Die Koeffizienten und Graphen einiger für die Elektronik relevanten periodischen Funktionen sind in Tabelle 7.1 dargestellt.

Signalformen	**Fourierreihenentwicklung**
Rechteckpulse  $f(t) = \begin{cases} 1 & \text{für} \quad t = [0, T/2] \\ 0 & \text{für} \quad t =]T/2, T] \end{cases}$	 $\frac{1}{2} + \frac{2}{\pi}\left[\sin\omega t + \frac{\sin 3\omega t}{3} + \frac{\sin 5\omega t}{5} + \cdots\right]$
Sägezahnpulse  $f(t) = \frac{t}{T} \quad \text{für} \quad t = [0, T]$	 $\frac{1}{2} - \frac{1}{\pi}\left[\sin\omega t + \frac{\sin 2\omega t}{2} + \frac{\sin 3\omega t}{3} + \cdots\right]$
Dreieckspulse  $f(t) = \begin{cases} \frac{2t}{T} & \text{für} \quad t = [0, T/2] \\ -2 - \frac{2t}{T} & \text{für} \quad t =]T/2, T] \end{cases}$	 $\frac{1}{2} - \frac{4}{\pi^2}\left[\cos\omega t + \frac{\cos 3\omega t}{9} + \frac{\cos 5\omega t}{25} + \cdots\right]$

Tab. 7.1 Fourierreihenentwicklungstabelle mit exemplarischen periodischen Funktionen. Die Graphen im rechten Teil der Tabelle entsprechen der Reihenentwicklung einschließlich des Terms $n = 3$.

Das Konvergenzverhalten der Reihe ist in Abb. 7.1 illustriert.

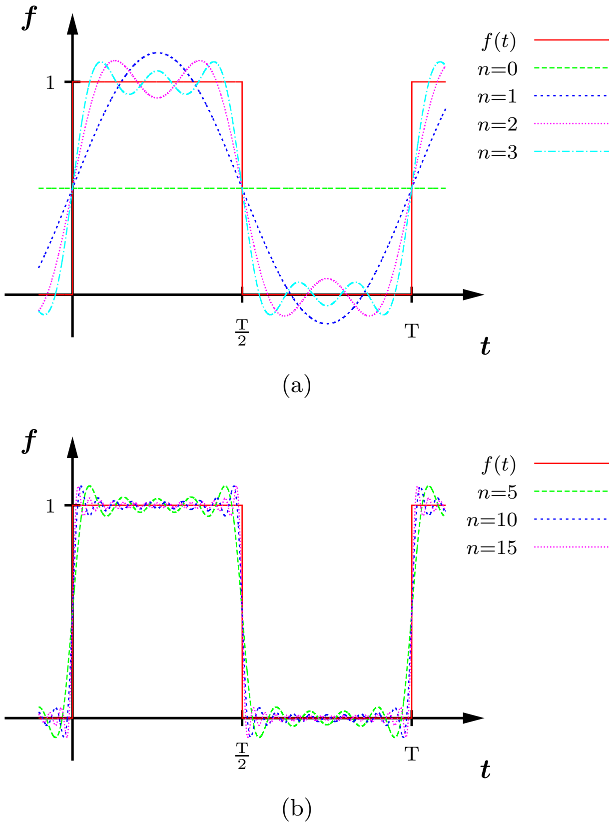


Abb. 7.1 (a) Fourierreihenentwicklung einer Rechteckfunktion für $n = 0, 1, ..., 3$ (b) Fourierreihenentwicklung einer Rechteckfunktion für $n = 5, 10, 15$

Für Funktionen wie die Rechteckfunktion konvergiert die Fourierreihe schlecht. Es treten zum einen beliebig hohe Frequenzen auf. An den Kanten treten für beliebig hohe Ordnungen endliche Abweichungen der Fourierreihe von der Ausgangsfunktion auf.

7.2 Fouriertransformation

Die Fouriertransformation ist die Erweiterung des Fourierformalismus auf *kontinuierliche nichtperiodische* Signale. Dies ist nur möglich, falls das Integral über $f(t)$ konvergiert:

$$\int_{-\infty}^{\infty} |f(t)|\, dt \quad < \infty \quad .$$

Dann lässt sich $f(t)$ ausdrücken mit

$$f(t) = \int\limits_{-\infty}^{\infty} F(\omega)\, \mathrm{e}^{\mathrm{j}\omega t}\, d\omega \quad . \tag{7.7}$$

Im Vergleich zu der Fourierreihe (Gl. 7.1) wird die Summe durch ein Integral über die kontinuierliche Frequenz ω ersetzt, die diskreten Koeffizienten c_n entsprechen einer kontinuierlichen Funktion $F(\omega)$, und der Vorfaktor ändert sich von $\frac{1}{T}$ zu $\frac{1}{2\pi}$.

$F(\omega)$ wird definiert über die Beziehung

$$F(\omega) = \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} f(t)\,\mathrm{e}^{-\mathrm{j}\omega t}\,dt \quad . \tag{7.8}$$

Auch hier gibt es neben der komplexen Fouriertransformation eine reelle Darstellung mit Termen $\cos(\omega t)$ und $\sin(\omega t)$.

Die Abbildung $F[f(t)](\omega)$ genügt den folgenden mathematischen Eigenschaften:

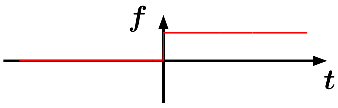
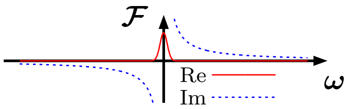
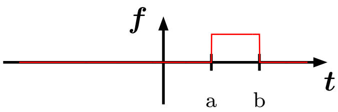
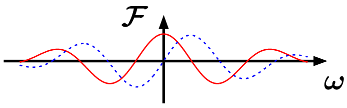
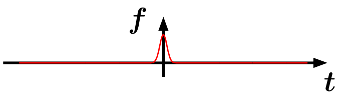
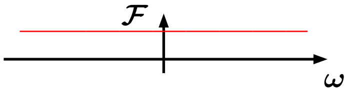
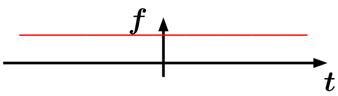
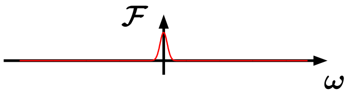
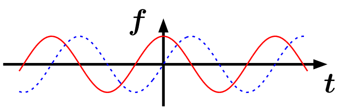
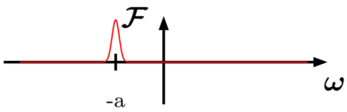
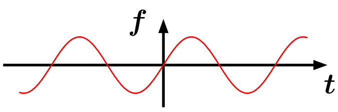
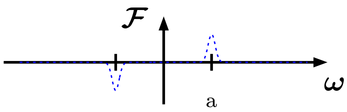
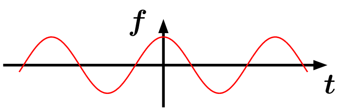
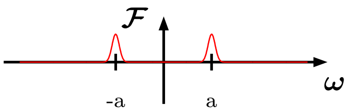
- Linearität: $F[a\,f + b\,g] = a\,F[f] + b\,F[g]$
- Ableitungseigenschaft: $F[\frac{d\,f}{dt}] = j\,\omega\,F[f]$
- Verschiebungseigenschaft: $F[f(t-t_0)] = \mathrm{e}^{-\mathrm{j}\,\omega\,t_0}\,F[f(t)]$
- Skalierungseigenschaft: $F[f(a\,t)] = \frac{1}{|a|}\,F[f(t)](\omega/a)$
- Faltungseigenschaft: Sei $f * g$ die Faltung $(f * g)(t) = \int\limits_{-\infty}^{\infty} f(t')\,g(t-t')\,dt'$, so gilt $F[f * g] = F[f] \cdot F[g] \quad .$

Die Fouriertransformationen $F(\omega)$ einiger wichtiger Funktionen $f(t)$ sind in Tabelle 7.2 dargestellt. Hier tritt die Diracsche Deltafunktion auf $\delta(t)$ auf, die über

$$\delta(\mathrm{t}) = \begin{cases} 0 & \mathrm{t} \neq 0 \\ \infty & \mathrm{t} = 0 \end{cases} \quad \text{und} \quad \int\limits_{-\infty}^{\infty} \delta(\mathrm{t})\,\mathrm{dt} = 1$$

definiert ist. Gleichzeitig gilt

$$\int\limits_{-\infty}^{\infty} \delta(\mathrm{t})\,\mathrm{f}(\mathrm{t})\,\mathrm{dt} = \mathrm{f}(0) \quad .$$

Zeitraum $f(t)$	**Frequenzraum** $F(\omega)$
Stufenfunktion  $\Theta(t)$	 $\frac{j}{\sqrt{2\pi}\omega} + \sqrt{\frac{\pi}{2}}\delta(\omega)$
Rechteckimpuls  $\Theta(t-a)\cdot\Theta(b-t)$	 $\frac{-j}{\sqrt{2\pi}}\frac{e^{jb\omega}-e^{ja\omega}}{\omega}$
Delta – Distribution  $\delta(t)$	 $\frac{1}{\sqrt{2\pi}}$
Konstante  1	 $\sqrt{2\pi}\delta(\omega)$
Exponentielle Schwingung  e^{jat}	 $\sqrt{2\pi}\delta(\omega+a)$
Sinus  $\sin(at)$	 $j\sqrt{\frac{\pi}{2}}\,[\delta(\omega-a)-\delta(\omega+a)]$
Cosinus  $\cos(at)$	 $\sqrt{\frac{\pi}{2}}\,[\delta(\omega-a)+\delta(\omega+a)]$

Tab. 7.2 Fouriertransformationstabelle. Der Realteil der Funktionen $f(t)$ und $F(\omega)$ ist als durchgezogene rote Linie und der Imaginärteil als gestrichelte blaue Linie dargestellt.

7.3 Diskrete Fouriertransformation

Die diskrete Fouriertransformation (DFT) ist die Erweiterung des Fourierformalismus auf diskrete Folgen f_k, wie sie z. B. bei der Digitalisierung von analogen Signalen auftreten. Die Bezeichnung f_k bzw. $f[k]$ ist eine Kurzform für $f(t_k)$ für diskrete Zeitpunkte t_k mit $k = -\infty, \infty$. Bei der praktischen Verarbeitung digitaler Signale ist die Zahl der Messwerte endlich, z. B. 100 oder 1000.

Statt des Integrals

$$F(\omega) = \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} f(t)\, e^{-j\omega t} dt$$

können wir hier wieder zu einer Summe übergehen, die für N Messpunkte wie folgt definiert ist:

$$F[k] = \frac{1}{N} \sum_{n=0}^{N-1} f[n]\, e^{-j2\pi kn/N} \quad . \tag{7.9}$$

Der erste Index der Summe ist $n = 0$, der höchste $N - 1$. Es treten die Kreisfrequenzen $\omega = \frac{2\pi k}{N}$, $k = 0, 1, ..., N - 1$ auf. Entsprechend gilt für die inverse DFT

$$f[n] = \sum_{k=0}^{N-1} F[k]\, e^{j2\pi kn/N} \quad . \tag{7.10}$$

Die DFT wird für endlich lange Signalsequenzen mit z.B. 1024 Messpunkten angewandt, entspricht aber der Fouriertransformation und nicht der Fourierreihe. Ein nützliche Relation ist

$$F[N-k] = \frac{1}{N} \sum_{n=0}^{N-1} f[n]\, e^{-j2\pi(N-k)\cdot\frac{n}{N}} = \overline{F}[k] = \frac{1}{N} \sum_{n=0}^{N-1} f[n]\, e^{+j2\pi k\cdot\frac{n}{N}} \quad ,$$

da

$$e^{\pm j2\pi\frac{N}{N}\cdot n} = e^{\pm 2\pi\cdot n} = 1$$

und $f[n]$ reel ist.

8 Elementare 2-Transistorschaltungen

Bisher haben wir nur Verstärker aus einem Transistor betrachtet. Noch deutlich bessere Verstärkereigenschaften erreicht man mit 2-Transistorschaltungen und später mit mehrstufigen Verstärkern. Attraktive Verstärkereigenschaften können z. B. sein:

- große Verstärkung
- hohe Eingangsimpedanz
- geringe Ausgangsimpedanz
- große Bandbreite
- hohe Linearität des Verstärkers
- geringes Rauschen
- geringer Ruhestrom und damit reduzierter Verbrauchsleistung
- reduzierte Temperaturabhängigkeit

Elementare Verstärkerschaltungen mit nur zwei Transistoren sind der Differenzverstärker, der Darlingtonverstärker und die Kaskode. Eine weitere wichtige Schaltung, die in ihrer einfachsten Version mit nur zwei Transistoren realisiert werden kann, ist der Stromspiegel. Später werden wir auch noch elementare 2-Transistorschaltungen mit MOSFETs betrachten.

8.1 Stromspiegel*

Ein Stromspiegel kopiert bzw. vervielfacht einen Eingangsstrom. Eine einfache Schaltung ist in Abbildung 8.1 dargestellt.

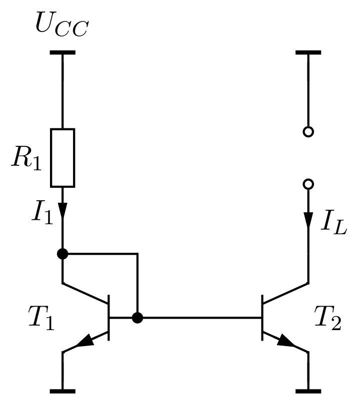


Abb. 8.1 Stromspiegel. Die nicht eingezeichnete Last R_L zwischen den offenen Anschlüssen kann die unterschiedlichsten Schaltungsblöcke darstellen.

Die Basis-Emitter-Spannung U_{BE} der Transistoren T_1 und T_2 ist identisch, da die Basis und die Emitter von T_1 und T_2 direkt verbunden sind. Der Transistor T_1 ist leitend. Durch den Kurzschluss von Kollektor C und Basis B entspricht er einer Diode. Der Strom durch den Transistor T_1 kann durch den Wert des Widerstands R_1 eingestellt werden. Sind T_1 und T_2 baugleich, so gilt $I_1 \approx I_L$.

Wichtig für das Verständis der Schaltung ist, dass der Kollektor von T_1 auf einem relativ konstanten Potential von etwa 0,7 V liegt. Die Kollektorspannung von T_2 ist dagegen variabel und kann sich relativ frei an den Spannungsabfall des Lastwiderstands R_L anpassen ($U_L + U_{CE} = U_{CC}$).

In integrierten Schaltkreisen stammen die Transistoren T_1 und T_2 aus demselben Halbleiterwafer und liegen nahe beieinander. Dadurch sind sie denselben Schwankungen der Fabrikationsparameter unterworfen und haben nahezu identische Eigenschaften. Ist die Geometrie von T_1 und T_2 unterschiedlich, kann man auch Stromverhältnisse ungleich 1 einstellen. Das Verhältnis der Ströme entspricht dem Verhältnis der Sperrströme $\frac{I_{S_2}}{I_{S_1}}$ und damit der Fläche des Basis-Emitter-Übergangs [1].

Die Temperaturabhängigkeit des Stromes durch R_L ist klein, da der Eingangsstrom I_1 durch

$$I_1 \approx \frac{U_{CC} - 0,7\,\mathrm{V}}{R_1} \tag{8.1}$$

gegeben ist und damit kleine Änderungen von U_{BE} keine große Rolle spielen.

Eine Darstellung des Stromspiegels mit einer Stromquelle statt einem einfachen Widerstand ist in Abb. 8.2 dargestellt.

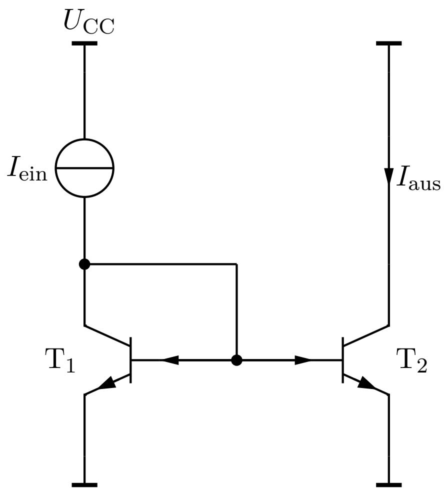


Abb. 8.2 Einfacher Stromspiegel aus zwei BJTs T_1 und T_2.

[1] Für MOSFETs gilt $I_{DS} \sim \frac{W}{L}$, wobei L und W die Kanallänge und -breite sind. Durch unterschiedliche Wahl von $\frac{W}{L}$ für T_1 und T_2 kann man das Verhältnis der Ströme zu $\frac{I_1}{I_2} = \frac{W_1\,L_2}{W_2\,L_1}$ einstellen.

$$\begin{aligned} I_{\text{aus}} &= I_{\text{CT}_2} = I_{\text{ein}} - 2I_{\text{B}} \\ &= I_{\text{ein}} - 2\frac{I_{\text{C}}}{\beta} \end{aligned} \tag{8.2}$$

I_{ein} und I_{aus} sind für BJTs also selbst bei identischen Transistoren nicht ganz gleich.

Mit einem weiteren Transistor (Abb. 8.3) reduziert man dieses „Problem“. Jetzt liefert T_3 den Basisstrom

$$I_{\text{ET}_3} = 2I_{\text{BT}_1} \tag{8.3}$$

T_3 reduziert den Strom I_{ein} aber nur minimal, da

$$I_{\text{BT}_3} = \frac{2I_{\text{B}_{\text{T}_1}}}{\beta_3} . \tag{8.4}$$

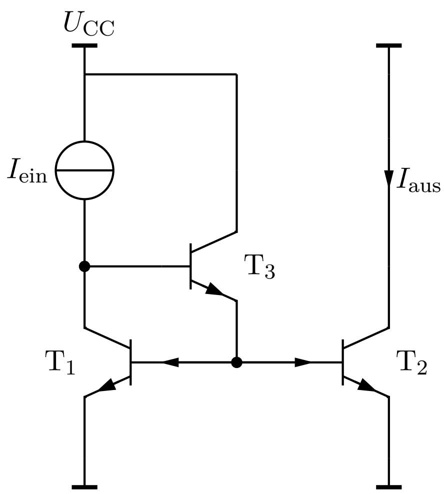


Abb. 8.3 Stromspiegel mit Basisstromkompensation.

Eine Variante des Stromspiegels nach Robert Widlar ist in Abb. 8.4 dargestellt.

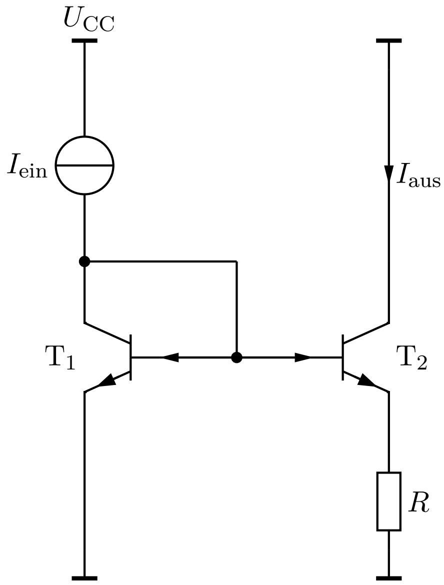


Abb. 8.4 Stromspiegel nach Robert Widlar.

Aus

$$U_{BET_1} = U_{BET_2} + I_{aus} R \tag{8.5}$$

und der Shockleygleichung

$$U_{BE} = U_T \ln\left(1 + \frac{I_C}{I_S}\right) \tag{8.6}$$

für T_1 und T_2 folgt

$$I_{aus} = \frac{U_T}{R} \ln \frac{I_{ein}}{I_{aus}} \,. \tag{8.7}$$

Diese Gleichung kann iterativ gelöst werden.

Die häufigere Stromquelle nach Wilson ist in Abb. 8.5 dargestellt. Die Grundidee dieser Stromquelle entspricht der Schaltung 8.3. Der Basistrom von T_1 und T_2 wird mithilfe des Transistors T_3 geliefert.

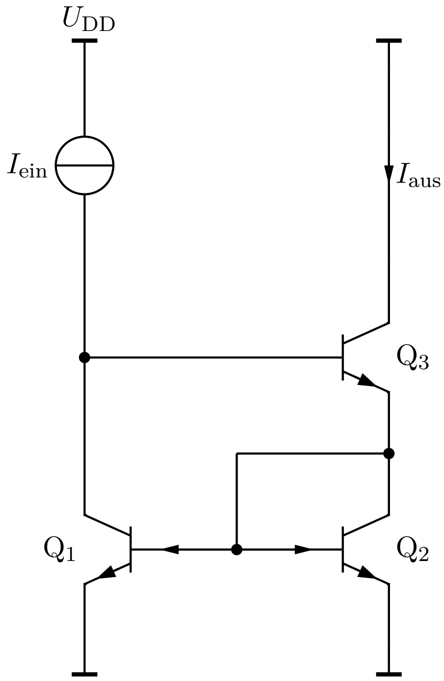


Abb. 8.5 Stromquelle nach Wilson.

8.2 Differenzverstärker

Der Differenzverstärker wird aus zwei Emitterverstärkern zusammengebaut, die an ihren Emittern verbunden sind (Abb. 8.6). Der Widerstand R_E kann auch durch eine Stromquelle ersetzt werden.

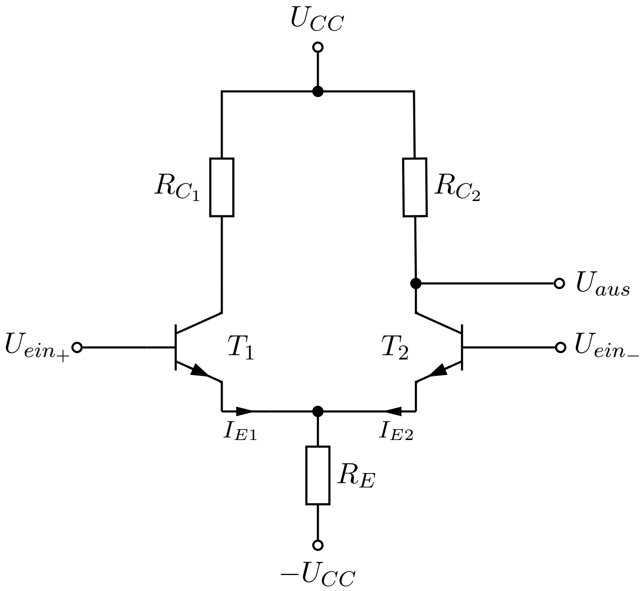


Abb. 8.6 Differenzverstärker.

Die Widerstände R_{C_1} und R_{C_2} werden gleich groß gewählt. Auch die beiden Transistoren T_1 und T_2 sollten im Idealfall identisch sein. Ihr Arbeitspunkt ist gleich eingestellt (nicht eingezeichnet). Unter vernünftigen Betriebsbedingungen ist der Gesamtstrom $I_{E_1} + I_{E_2}$ durch den gemeinsamen Emitterwiderstand relativ konstant. Dies gilt bei großen R_E mit einem Spannungsabfall, der groß gegen die Variationen der Eingangsspannung ist. Es gilt trivialerweise bei Ersatz von R_E durch eine Stromquelle mit Strom I_0. Wir betrachten nur den Fall einer hochohmigen Last, die keinen nennenswerten Strom zieht.

Den Differenzverstärker versteht man am besten, wenn man die Eingangsspannungen U_{ein_+} und U_{ein_-} als Linearkombination einer Differenzspannung

$$U_D = U_{ein_+} - U_{ein_-} \tag{8.8}$$

und einer Gleichtaktspannung

$$U_G = \frac{U_{ein_+} + U_{ein_-}}{2} \tag{8.9}$$

schreibt wie wir es auch beim OPV gemacht haben. Dann gilt

$$U_{ein_+} = U_G + \frac{U_D}{2} \quad \text{und} \quad U_{ein_-} = U_G - \frac{U_D}{2} \quad . \tag{8.10}$$

Ein schiefsymmetrisch zur Ruhespannung schwankendes Eingangssignal

$$dU_{ein_+} = \frac{dU_D}{2} \quad \text{und} \quad dU_{ein_-} = \frac{-dU_D}{2}$$

ändert das gemeinsame Emitterpotential, die Spannung oberhalb von R_E, nicht. Die Änderungen des Stroms im linken Zweig der Schaltung, dI_{E_1}, und im rechten Zweig, $-dI_{E_2}$, kompensieren sich.

Welchen Wert nimmt das Ausgangssignal an? Es entspricht der Emittergrundschaltung mit $dU_{aus}/dU_D = -\, R_C\, S$. Damit ergibt sich die Differenzspannungsverstärkung zu

$$V_D = \frac{dU_{aus}}{dU_D} = R_C\,\frac{S}{2} = R_C\,\frac{I_C}{4\,U_T} \quad . \tag{8.11}$$

Der Faktor $\frac{1}{2}$ erklärt sich dadurch, dass wir das Eingangsdifferenzsignal zu gleichen Teilen auf den positiven und den negativen Eingang aufgeteilt haben. Das positive Vorzeichen ergibt sich dadurch, dass wir das Ausgangssignal von T_2 und nicht T_1 abgreifen. Die Steilheit S ist hier durch $I_C/2\,U_T$ gegeben, da durch den Transistor T_2, an dem der Ausgang liegt, nur die Hälfte des Stromes fließt.

Was passiert bei einem symmetrischen Eingangssignal $dU_{ein_+} = dU_{ein_-} = \frac{dU_G}{2}$? Hier ergibt sich die Gleichtaktverstärkung zu

$$V_G = \frac{dU_{aus}}{dU_G} \approx \frac{-\,R_C}{2\,R_E} \quad . \tag{8.12}$$

Auch das ist plausibel, da jetzt die Spannung oberhalb von R_E nicht mehr konstant ist. Dann entspricht die Formel dem stromgegengekoppelten Emitterverstärker.

Je größer R_E, desto kleiner ist die unerwünschte Gleichtaktsverstärkung. Am geringsten ist die Gleichtaktverstärkung bei Ersatz des Widerstands R_E durch eine Stromquelle (genauer gesagt durch eine Stromsenke). Dann muss man R_E in Gleichung 8.12 durch den sehr großen Innenwiderstand R_i der Stromquelle ersetzen.

Für die Gleichtaktunterdrückung [2] gilt

$$\left|\frac{V_D}{V_G}\right| = S\,R_E \quad .$$

Es gibt viele Spielarten dieser Grundschaltung. Manchmal bevorzugt man nur ein Ausgangsignal wie in Abbildung 8.6. Man kann aber auch an beiden Kollektoren ein Ausgangssignal abgreifen und an die nächste (differentielle) Verstärkerstufe weiterreichen. Den Ersatz des Widerstands R_E durch eine Stromquelle haben wir schon erwähnt.

Man kann auch den Widerstand R_C durch eine Stromquelle ersetzen und damit deutlich vergrößern. Dann steigt die Differenzverstärkung (allerdings auch die Gleichtaktsverstärkung) stark an. Sehr einfach ist auch die Verwendung eines Stromspiegels, da ja zwei Kollektorwiderstände ersetzt werden sollen.

Diese Schaltung ist in Abbildung 8.7 gezeigt. Dieser Stromspiegel ist aus pnp-Transistoren aufgebaut!

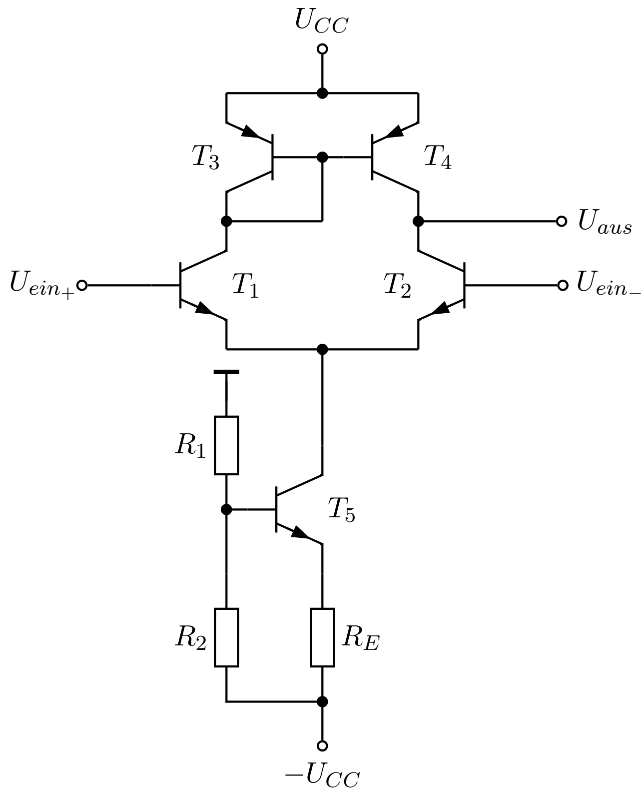


Abb. 8.7 Differenzverstärker (T_1 und T_2) mit Stromquelle (T_5) und Stromspiegel (T_3 und T_4).[2]

[2]Engl. *common mode rejection ratio* (CMRR)

Der Vergleich des Stromspiegels in Abbildung 8.1 mit dem in Abbildung 8.7 ist interessant. Der Widerstand R_1 in Abbildung 8.1, der den Strom einstellt, entspricht der Stromquelle mit Transistor T_5 im unteren Ast von Abb. 8.7. Wir werden diese Schaltung in der Implementierung mit FETs in Kapitel **??** genauer behandeln.

Der Differenzverstärker bildet die Eingangsstufe eines Operationsverstärkers. Schon allein dadurch hat er eine große Bedeutung. In der Detektorinstrumentierung ist der Differenzverstärker allerdings nicht so wichtig, wie man es vielleicht erwarten würde. Für die kritische Eingangsstufe eines ladungsempfindlichen Verstärkers wird meistens nur ein Eingang verwendet, da durch die zwei Eingänge des Differenzverstärkers das Rauschen um einen Faktor $\sqrt{2}$ ansteigen würde.

Die obigen Ausdrücke kann man auch mithilfe der Kleinsignalersatzschaltung ableiten. Eine gute Herleitung findet sich z.B. in
colorbox[rgb]0,0,0citereisch2006elektronische Kapitel 17.6.2.

8.3 Darlingtontransistor

In der Darlingtonschaltung werden zwei Transistoren so hintereinander geschaltet, dass sich eine besonders hohe Stromverstärkung ergibt (Abb. 8.8). Der Transistor T_1 ist ein Emitterfolger und steuert die Basis von T_2. Die Kollektoren von T_1 und T_2 sind verbunden. Manchmal wird der Widerstand R, der Werte im Bereich von $100\,\Omega$ bis einigen $k\Omega$ annimmt, auch weggelassen.

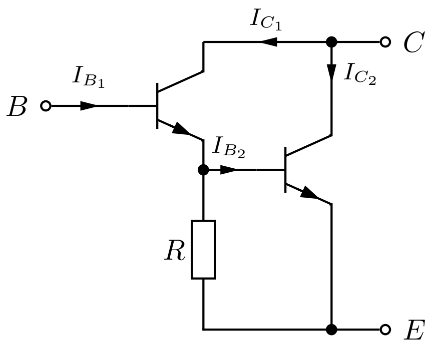


Abb. 8.8 Darlingtonschaltung.

Im Normalbetrieb leiten beide Transistoren. Der Strom I_R durch den Widerstand R ist durch $I_R \approx 0,7\,\mathrm{V}/R$ beschränkt. Für die Ströme gilt

$$I_C = I_{C_1} + I_{C_2} \quad \text{mit} \quad I_{C_1} = \beta_1\, I_{B_1} \quad \text{und} \quad I_{C_2} = \beta_2\, I_{B_2} = \beta_2\,(I_{C_1} + I_{B_1} - I_R) \quad . \tag{8.13}$$

Die Gesamtstromverstärkung β kann man hier als I_C/I_{B_1} definieren. Leiten beide Transistoren so gilt $\beta \approx \beta_1\,\beta_2$.

Bei kleinen Kollektorströmen I_C sperrt T_2 und nur T_1 leitet. Die Stromverstärkung $\beta \approx \beta_1$. Dies gilt für Ströme $I_R < 0,7\,\mathrm{V}/R$.

Die Darlingtonschaltung wird mitunter als eigenständiger Transistortyp aufgefasst. Entsprechend gibt es ein eigenes Schaltungssymbol (Abb. 8.9).

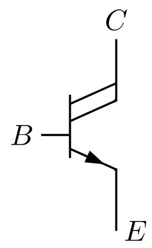


Abb. 8.9 Schaltbild des Darlingtontransistors.

Die Kennlinien des Darlingtontransistors ähneln denen des Einzeltransistors. Relevante Unterschiede zum Einzeltransitor sind:

- der weitaus größerer Gesamtstrom I_C und
- die verdoppelte Basisemitterspannung $U_{BE} = U_{BE_1} + U_{BE_2}$.

Wie ein einzelner Transistor kann auch der Darlingtontransistor auf vielfältige Weise eingesetzt werden, so z. B. als Emitterfolger mit einem zusätzlichen Widerstand R_E am Ausgang E oder als Spannungsverstärker mit einem Lastwiderstand R_C am Kollektor. Die Darlingtonschaltung ist langsamer als ein einzelner Transistor, denn es müssen bei entsprechendem Eingangssignal zwei Transistoren, T_1 und T_2, abgeschaltet werden, und T_2 ist vom Eingang durch T_1 abgekoppelt. T_1 lässt sich schnell abschalten. T_2 schaltet sich erst ab, wenn die in der Basis gespeicherte Ladung über R abgeflossen ist. Durch die Verkettung zweier Transistoren erhöht sich auch die Phasenverschiebung. Dies muss man bei negativer Rückkopplung beachten, damit die Schaltung nicht anfängt zu schwingen. Der Wert von R darf nicht zu klein sein, damit die Basisspannung von T_1 nicht zu klein wird bzw. damit der Kollektorstrom von T_1 nicht an T_2 vorbeifließt.

Eine schöne Beschreibung des Kleinsignalverhaltens des Darlingtontransistors findet sich in [10].

8.4 Kaskode

Die Kaskodenschaltung (Abb. 8.10) ist eine verbreitete und wichtige 2T-Schaltung. Sie ergibt sich aus der Emitterschaltung des Transistors T_1 und dem Kollektorwiderstand R_C durch Einfügen eines Transistors T_2 in Basisschaltung zwischen T_1 und dem Kollektorwiderstand R_C.

Der Transistor T_2 unterdrückt den Millereffekt, der bei der einfachen Emitterschaltung die Bandbreite begrenzt. Gleichzeitig bleiben die guten Eigenschaften der Emitterschaltung, wie hohe Eingangsimpedanz, Strom- und Spannungsverstärkung, erhalten.

Wie funktioniert die Kaskodenschaltung? Der Transistor T_2 ist als Basisschaltung mit konstantem Basispotential arrangiert. Dadurch ist die Emitterspannung von T_2 und die Kollektorspannung von T_1 relativ konstant. Das heißt, die Spannungsverstärkung von T_1 ist klein, was den Millereffekt auf Grund der Kapazität C_{BC} des Transistors T_1 eliminiert.

Anmerkung: Man kann die Verstärkung von T_1 berechnen, indem man den Eingangswiderstand von T_2 als den Kollektorwiderstand $R_{C'}$ von T_1 auffasst und die Formel $V_U = -R_{C'}\,S$ anwendet. Hier sind $R_{C'} = \frac{k\,T}{q_e\,I_C}$ (Gl. 5.28) und $S = \frac{q_e}{k\,T}\,I_C$, also $V_U(T_1) = dU_2/dU_{ein} \approx 1$.

Die Spannungsverstärkung der Kaskodenschaltung wird also nicht durch T_1, sondern durch T_2 bewirkt mit

$$V_U = \frac{dU_{aus}}{dU_{ein}} \approx \frac{dU_{aus}}{dU_2} = R_C\,S \quad . \tag{8.14}$$

Wieso tritt bei T_2 kein Millereffekt auf? Dies liegt daran, dass der Emitter von T_2 (Eingang der Basisschaltung) nicht direkt an den Kollektor (Ausgang der Basisschaltung) koppelt, sondern nur über die Kapazitäten C_{BE} und C_{BC}. Da aber die Basis auf einem festen Spannungspotential liegt, sind Emitter und Kollektor von T_2 kapazitiv entkoppelt. (Eine direkte Kopplung wäre zwar über C_{CE} möglich, aber C_{CE} ist deutlich kleiner als C_{BE} und C_{BC}.)

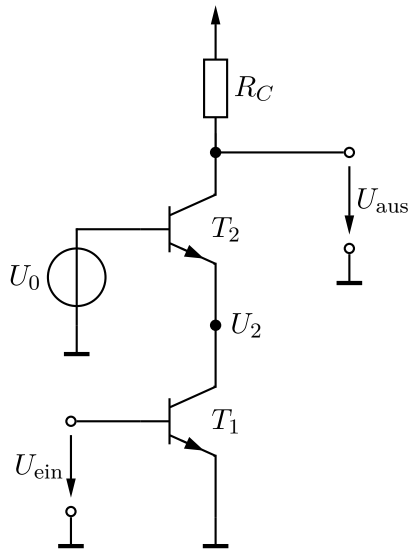


Abb. 8.10 Kaskode. Die Spannung am Kollektor von T_1 bzw. Emitter von T_2 bezeichnen wir als U_2.

8.5 Gleichspannungsregler*

Wir haben in Kapitel 3.13 schon eine einfache Spannungsregelung mittels einer Zenerdiode kennengelernt, die allerdings nur für kleine Lastströme geeignet ist. Die Schaltung stabilisiert die Ausgangsspannung U_{aus} gegen Schwankungen der Eingangsspannung U_{ein} oder des Laststroms. Diese Art der Schaltung nennt man auch Querregler. Zenerdiode und Last sind parallelgeschaltet. Falls sich das Lastprofil ändert und z.B. die Last momentan weniger Strom benötigt (der Lastwiderstand wäre dann größer), fließt der nicht benötigte Strom stattdessen durch die Diode.

Mit Hilfe eines Transistors kann man diese Schaltung verbessern (Abb. 8.11).

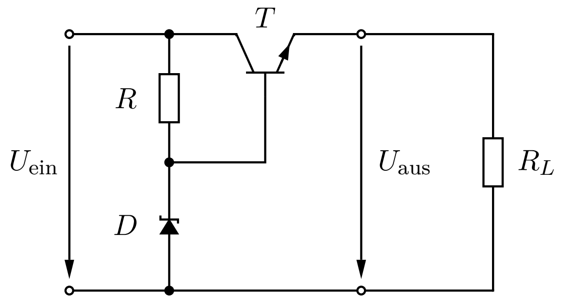


Abb. 8.11 Stabilisierte Spannungsquelle mit Zenerdiode und Transistor.

Der Laststrom fließt durch den Transistor T_1 und kann sehr groß sein. Der Strom durch den Widerstand R und die Diode D ist dagegen klein, was die Regelung effizienter macht, und er ist unabhängig vom Laststrom I_L. Die Ausgangsspannung ist um den Spannungsabfall U_{CE} an T_1 kleiner als die Eingangsspannung und ergibt sich einfach aus

$$U_{\text{aus}} = U_Z - U_{BE} \quad .$$

Sinkt U_{aus}, da I_L zu klein wird, steigt U_{BE} und damit I_L an. Entsprechend sinkt I_L, falls U_{aus} zu groß wird. Dieser Regler ist ein Längsregler. Der Transistor lässt genau den benötigten Laststrom zur Last fließen, ein Umlenken des Stroms wie im Querregler ist nicht nötig. Im Lastbetrieb muss T_1 leiten. Damit gilt

$$U_{BE} \approx 0{,}6\,\text{V} \quad .$$

Abhängig vom Laststrom wird U_{BE} aber um einige 100 mV schwanken. Dies begrenzt die Genauigkeit der Spannungsregelung.

In der folgenden Schaltung (Abb. 8.12) wird diese Schwankung durch eine negative Rückkopplung unterdrückt.

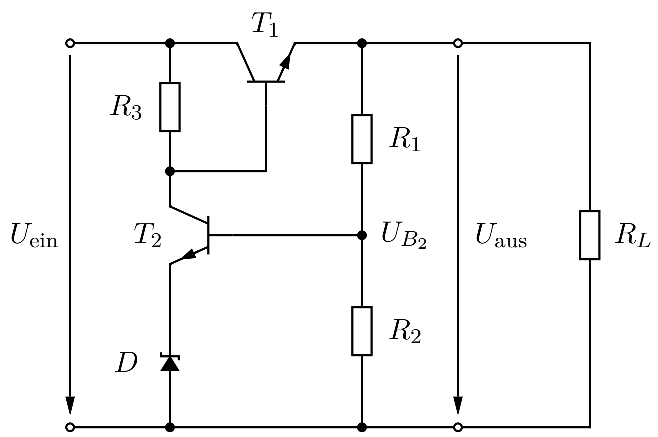


Abb. 8.12 Gleichspannungsquelle mit Rückkopplung.

Die Basisspannung von T_2 wird durch U_{aus} geregelt und beträgt $U_{B2} = U_{\text{aus}} \frac{R_2}{R_1+R}$. Die Basis-Emitterspannung U_{BE} von T_2 beträgt $U_{BE} = U_{B2} - U_Z$.

Falls U_{aus} vom Sollwert abweicht, z.B. durch eine Verringerung der Eingangsspannung, würde auch der Laststrom abnehmen. Die Schaltung wirkt dem entgegen, da beim Abnehmen von U_{aus} auch U_{BE} abnimmt. Damit sinkt der Strom durch T_2 und R_3, und U_{BE} nimmt zu. Der Transistor T_1 leitet also daher mehr Strom, U_{aus} nimmt wieder zu, und es stellt sich eine stabile Ausgangsspannung ein.

Um diese und die vorherige Schaltung zu verstehen, muss man sich klarmachen, dass I_L und U_{aus} voneinander abhängen. In unserem Beispiel durch $U_{\text{aus}} \approx R_L \cdot I_L$, da der Strom durch R_1 minimal ist. Die Zenerdiode erhöht die Emitterspannung von T_2, dadurch wird der Einfluss von temperaturbedingten Verschiebungen der Schwellenspannung im Verhältnis kleiner.

Wichtig für das Funktionieren der Schaltung ist auch, dass T_1 beim Anschließen der Eingangsspannung sofort leitet. Dies ist durch die Verbindung der Basis von T_1 über R_3 mit U_{ein} sichergestellt.

Dasselbe Konzept wird auch in Abb. 8.13 angewandt.

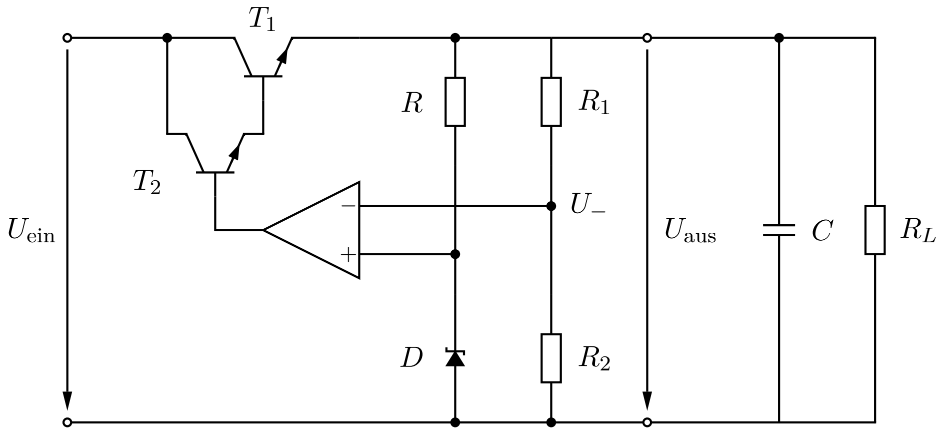


Abb. 8.13 Geregelte Spannungsquelle mit Zenerdiode, Darlingtonstufe T_1 und T_2 und OPV.

Hier wird der Transistor T_1 durch eine Darlington-Schaltung bestehend aus T_1 und T_2 ersetzt. Der Emitterstrom von T_2 dient als Basisstrom von T_1. Dadurch sind noch höhere Lastströme möglich. Die Spannungsregelung erfolgt nicht mehr über einen Transistor (T_2 in Abb. 8.11), sondern über einen Operationsverstärker. Der nichtinvertierende Eingang des Operationsverstärkers wird z.B. durch eine Zenerdiode oder besser eine noch genauere Referenzspannung konstant gehalten. Der invertierende Eingang des Operationsverstärkers ist über einen Spannungsteiler an die zu regelnde Ausgangsspannung U_{aus} gekoppelt. Dadurch entsteht eine negative Rückkopplung und

$$U_- = U_{\text{Ref}} = U_{\text{aus}} \frac{R_2}{R_1 + R_2} \quad . \tag{8.15}$$

Dies ist unabhängig von Temperatureffekten bei T_1 und T_2 und ist nur von der Stabilität der Referenzspannung abhängig.

Falls o.B.d.A. die Ausgangsspannung U_{aus} absinkt, nimmt auch die Spannung U_- ab. Damit nimmt die Ausgangsspannung U_{a} des Operationsverstärkers zu, und die Darlingtionstufe liefert einen höheren Strom. Dadurch erhöht sich U_{aus}. (Die Spannungserhöhung beträgt $U_{\text{aus}} \approx R_L \cdot I_C$, da R_L klein gegen $R_1 + R_2$ bzw. R ist.) Entsprechend wirkt die Gegenkopplung bei einer Abnahme des Laststroms.

Realistische Schaltungen sind noch etwas komplizierter, da zusätzliche Schaltungselemente ergänzt werden wie Überstrom- bzw. Kurzschlussschutz und eine Temperaturmessung am Leistungstransistor. Auch ist die Stabilität und die Regelgeschwindigkeit der Schaltungen wichtig.

Anmerkung: In dieser Schaltung wird der Ausgang des Operationsverstärkers nicht direkt, sondern über zwei Transistoren an die Last gekoppelt. Das ist ungewohnt, ändert aber nichts am Konzept der negativen Rückkopplung.

Der wesentliche Nachteil von linearen Gleichspannungsreglern ist ihre Ineffizienz. Da an T_1 die Kollektor-Emmiter-Spannung U_{CE} abfällt, entsteht eine Verlustleitung von $\approx U_{CE}\, I_L$. Damit ergibt sich die Effizienz der Regelung zu

$$\eta \approx \frac{U_{aus}\, I_L}{U_{CE}\, I_L + U_{aus}\, I_L} = \frac{U_{aus}}{U_{ein}} \quad .$$

Diese Gleichung vernachlässigt den kleinen Strom durch den OPV und die diversen Widerstände zur Rückkopplung. Um die Effizienz zu maximieren, sollte man möglichst nur kleine U_{CE} bzw. geringe Unterschiede von Eingangs- und Ausgangsspannung zulassen.

8.6 Verstärkerklassen

Wir unterscheiden Verstärker der Klasse A, B und AB, die unten kurz vorgestellt werden. Es gibt noch weitere Typen z. B. C, D, E, die aber für die Detektorinstrumentierung oder empfindliche analoge Messgeräte nicht relevant sind. Bei den folgenden Beispielen handelt es sich um Leistungsverstärker, die als als Stromverstärker geschaltet sind.

Klasse A

Wir haben bisher nur Verstärker der Klasse A behandelt. Der Arbeitsbereich wird so gewählt, dass immer ein nennenswerter Ruhestrom fließt, so dass das Eingangssignal nicht abgeschnitten und im gesamten Wertebereich linear verstärkt wird. Das einfachste Beispiel ist der Emitterverstärker. Soweit so gut, aber es gibt auch einen Haken. Da der Transistor immer „an“ ist, ist die Schaltung nicht sehr energieeffizient. Weiß man, dass z. B. nur unipolare positive Eingangsspannungen auftreten können, so ist dieses Arrangement nicht ideal, da der negative Bereich nicht genutzt wird.

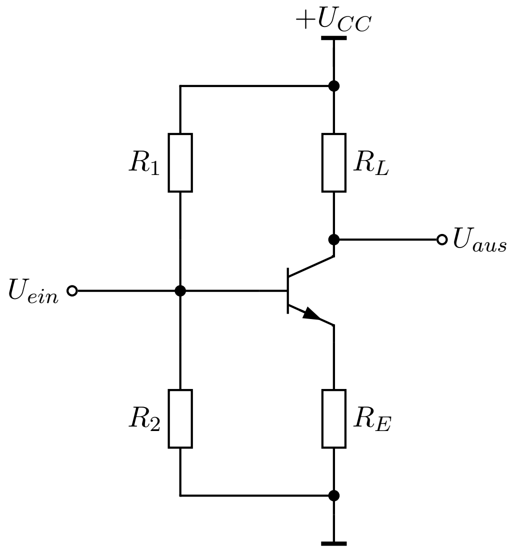


Abb. 8.14 Verstärker der Klasse A.

Klasse B

Alternativ könnte man den Arbeitspunkt der Schaltung in Abbildung 8.14 so einstellen, dass nur die positive Flanke des Eingangssignals durchkommt. Der Eingang liegt auf Emitterpotenzial, der Ruhekollektorstrom ist daher minimal, nur Eingangssignale größer 0,7 V werden verstärkt. Die Schaltung ist sehr effizient, aber nur für entsprechend positive Signale geeignet.

Interessanter ist der Verstärker der Klasse B aus Abbildung 8.15. Dieser Verstärker wird auch „push-pull amplifier" genannt.

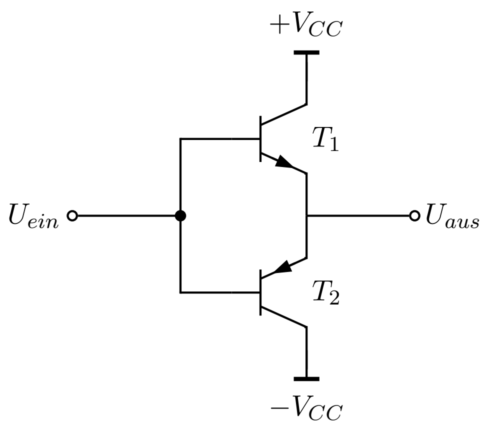


Abb. 8.15 Verstärker der Klasse B mit einem npn-Transistor (oben) und einem pnp-Transistor (unten) in Reihe.

Hier werden zwei komplementäre Transistoren, ein npn- und ein pnp-Transistor, benutzt, um jeweils die positive bzw. die negative Flanke zu verstärken. Der Arbeitspunkt ist wieder so eingestellt, dass ohne Eingangssignal kein Kollektorstrom durch die Transistoren fließt. Es leitet maximal ein Transistor, der andere sperrt. Dadurch ist der Verstärker recht effizient.

Anmerkung: Der pnp-Transistor leitet für $U_{BC} > 0,7$ V, $U_{EC} > 0$ V und $U_{EB} > 0$ V, analog zum npn-Transistor aber mit umgedrehten Vorzeichen.

Interessant ist das Verhalten bei kleinen Spannungen um 0 V. Hier sperren beide Transistoren, da der Betrag der Basisemitterspannungen beider Transistoren $< 0,7\,\text{V}$ ist. Dadurch wird das Signal verzerrt. Diesen Effekt wird mit der Verstärkerklasse AB vermieden.

Klasse AB

Eine Realisierung der Verstärkerklasse AB durch die Hinzunahme von Dioden ist in Abbildung 8.16 dargestellt. Dabei wird ein kleiner Arbeitspunktstrom hingenommen, um dafür auch kleine Signalhübe (unter $0,7\,\text{V}$) zu verstärken. Die Dioden haben dieselben Kennwerte wie die pn-Übergänge der Transistoren und sind unter Umständen aus entsprechenden Transistoren gebildet, deren Basis direkt mit dem Kollektor verbunden wurde. Die Transistoren sind dadurch an der Grenze, wo sie leitend werden und Signale verstärken, ohne dass sie gleichzeitig leiten (was einen Kurzschluss bewirken würde). Die Widerstände R_1 und R_2 bewirken eine Stromgegenkopplung und verringern die Temperaturabhängigkeiten der Schaltung.

Die Widerstände R_1 und R_2 dürfen nicht zu groß sein, damit an ihnen nicht zu viel Leistung abfällt. Andererseits wird die Stromkopplung mit kleinem R_1 und R_2 weniger effektiv. Dies begrenzt die Ausgangsströme der Schaltung. Für Ausgangsströme von mehreren Ampere setzt man Darlingtontransistoren ein.

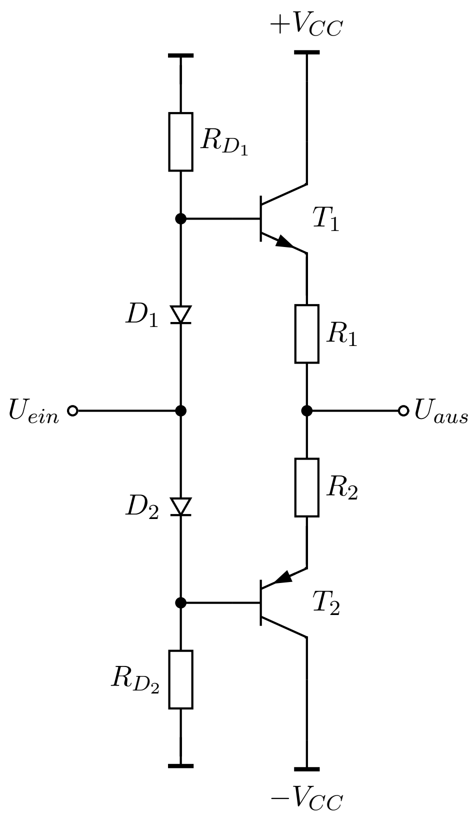


Abb. 8.16 Verstärkerklasse AB.

Anmerkung: Dieser Verstärkertyp nur für bipolare Signale wie z. B. Audiosignale wichtig. Für unipolare Signale ist Klasse A vollkommen ausreichend. Die meisten Detektoren haben unipolare Signale.

9 Feldeffekttransistor

Das Prinzip des Feldeffekttransistor wurde 1925 von Julius Lilienfeld entdeckt, weit vor der Erfindung und dem Bau des Bipolartransistors. Verwirklicht werden konnten die ersten Feldeffekttransistoren allerdings erst rund 30 Jahre später.

Der Feldeffekttransistor hat wie der Bipolartransistor „drei“ Anschlüsse, die „Source“, „Drain“ und „Gate“ genannt werden. Dabei spielen Drain und Source eine ähnliche Rolle wie Kollektor und Emitter beim Bipolartransistor, und Gate entspricht der Basis[1].

Es gibt im wesentlichen zwei Typen von Feldeffekttransistoren, den „Metal Oxide Semiconductor Field Effect Transistor“ (MOSFET) und den „Junction Field Effect Transistor“ (JFET). Beide Typen sind im Gegensatz zum Bipolartransistor unipolar, d. h. der Stromfluss basiert nur auf einer Sorte von Ladungsträgern, Elektronen oder Löchern. Bei beiden Typen ist die Eingangsimpedanz sehr hoch. Bei MOSFET fließt praktisch kein Strom in das Gate. Dies ist ein weiterer Unterschied zum BJT. Der Feldeffekttransistor ist also keine stromgesteuerte, sondern eine spannungsgesteuerte Stromquelle. Die Spannung zwischen Gate und Source bestimmt den Stromfluss zwischen Drain und Source.

Der MOSFET ist bei weiten der am häufigsten eingesetzte Transistortyp, insbesondere für integrierte digitale Bausteine. Der MOSFET kommt aber auch in „mixed-signal“ Schaltungen zum Einsatz, das heißt in Schaltungen mit analogen und digitalen Schaltungsblöcken. Von allen Transistortypen ermöglichen MOSFETs die bei weitem höchsten Integrationsdichten. Ein MOSFET hat einen eher geringen Stromverbrauch und eine extrem hohe Eingangsimpedanz.

9.1 MOSFET

Wir unterscheiden zwischen NMOS- und PMOS-Transistoren, ähnlich den beiden Typen des BJT, dem npn- und dem pnp-Transistor. Der Aufbau eines NMOS-Transistors ist in Abbildung 9.1 gezeigt.

[1] Den Einfluss einer Spannung an einem weiteren möglichen Anschluss, dem Substrat, werden wir in diesem Kapitel vernachläsigen

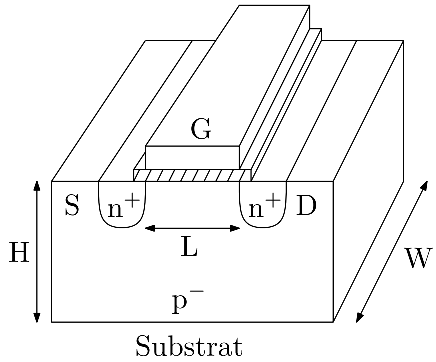


Abb. 9.1 Schematischer Aufbau eines NMOS-Feldeffekttranistors.

In das leicht dotierte p-Substrat sind zwei stark n-dotierte Kanäle eingebracht: Source und Drain. Der Strom zwischen Source und Drain hängt von der Gatespannung ab. Das Gate ist durch eine Isolatorschicht vom p-Substrat und den n-Kanälen getrennt[2].

Die Zeichnung ist nicht dimensionsgetreu. Der Abstand zwischen Drain und Source hängt von dem Fabrikationsprozess ab und kann rund 20 nm bis einige hundert nm betragen. Die Breite W wird vom Designer gewählt und ist meistens ein Vielfaches der Länge L. Die Substratdicke h beträgt einige 100 μm.

Als Schaltsymbol eines MOSFETs benutzen wir die in Abbildung 9.2 dargestellten Symbole. Hier gibt es in der Literatur zahlreiche ähnliche Varianten.

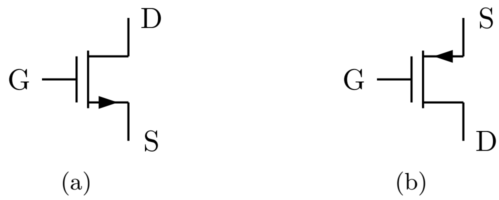


Abb. 9.2 (a) Schaltsymbol für den NMOS-Transistor und (b) für den PMOS-Transistor.

Die Kennlinien eines NMOS-Transistors sind in Abb. 9.3 dargestellt, die Kennlinien eines PMOS in Abb. 9.4.

[2]Der PMOS-Transistor hat zwei p-dotierte Kanäle in einem n-dotierten Substrat.

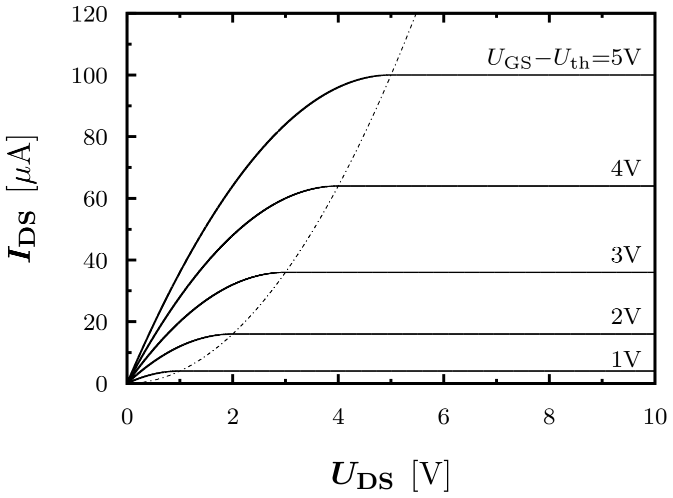


Abb. 9.3 Drainstrom eines NMOS-Transistors als Funktion von U_{DS} für verschiedene Werte von U_{GS}.

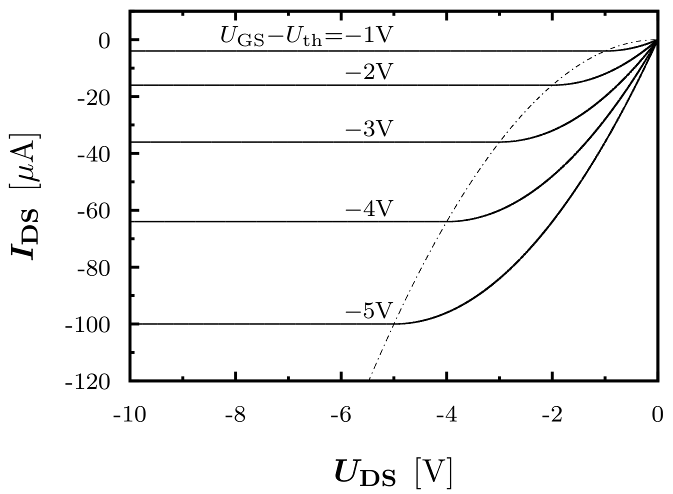


Abb. 9.4 Drainstromdiagramm eines PMOS-Transistors als Funktion von U_{DS} für verschiedene Werte von U_{GS}.

Der qualitative Verlauf der Kennlinien ähnelt dem eines BJT. Allerdings ist die Steilheit für große Werte von U_{DS} deutlich geringer. Die Steilheit ist analog zum BJT definiert und wird im folgendem durch das Symbol g_m bezeichnet. Die Steilheit g_m wird im englischen als „transconductance“ bezeichnet:

$$g_m = \frac{dI_{DS}}{dU_{GS}} \quad . \tag{9.1}$$

Um die Funktionsweise des MOSFET zu verstehen, betrachten wir die Abbildungen 9.5 - 9.7. In Abbildung 9.5 sind alle Anschlüsse des MOSFETs, also Gate, Drain, Source und Substrat, auf Masse gelegt. Da keine Spannung U_{DS} zwischen Drain und Source anliegt, kann auch kein Strom I_{DS} zwischen Drain und Source fließen. Zwischen Source und dem Substrat bzw.

Drain und Substrat bildet sich eine Sperrschicht. Dies entspricht exakt dem Verhalten einer Diode. Ein Stromfluss zwischen Drain und Source erfordert neben einer positiven Spannung U_{DS} eine positive Gatespannung U_{GS}. Man unterscheidet drei wesentliche Betriebsbereiche des MOSFET, abhängig vom Wert der Spannung U_{GS} zwischen Gate und Source und der Spannung U_{DS} zwischen Drain und Source:

- „Cut-off" oder „weak inversion"
- Linearer oder ohmscher Bereich
- „Strong inversion" oder Sättigungsbereich

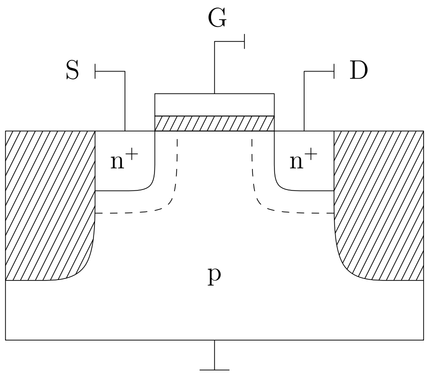


Abb. 9.5 Ein NMOS-Transistor mit Source S, Drain D, Gate G und Substrat auf Massenpotential. Die schraffierten Regionen sind SiO_2, zum einen das Gateoxid, zum anderen links und rechts das Feldoxid, das zur Isolation von benachbarten Transistoren dient und in Abb. 9.1 nicht eingezeichnet ist.

„Cut-off" oder „weak inversion"

Dieser Bereich ist durch $U_{GS} < U_{Th}$ definiert, wobei U_{Th} die transistorabhängige Schwellspannung ist. U_{Th} beträgt einige hundert mV. Hier ist der Transistor nahezu abgeschaltet, es fließt aber ein kleiner Leckstrom aus thermisch angeregten Elektronen. Der Strom

$$I_{DS} = I_{D_0}\, \mathrm{e}^{\frac{U_{GS}-U_{Th}}{n\, U_T}} \tag{9.2}$$

hängt exponentiell von U_{GS} ab. Wie bei der Diodengleichung gilt $n > 1$. Die Größe I_{DS_0} ist definiert als der Strom bei $U_{GS} = U_{Th}$:

$$I_{DS_0} = I_{DS}\,(U_{GS} = U_{Th}) \quad .$$

Die Steilheit ergibt sich zu

$$g_m = I_{DS}\, \frac{q_e}{n\, k\, T} \tag{9.3}$$

wie beim Bipolartransistor. Die Steilheit ist nur relativ zum Strom betrachtet groß; ihr absoluter Wert ist im Sättigungsbereich größer. Transistoren werden eher selten in weak inversion betrieben. Man erreicht große Steilheiten bei geringem Stromverbrauch, andereseits führen kleine produktionsbedingte Variationen der Schwellspannung zu deutlichen Unterschieden im Transistorstrom.

Anmerkung: Dieser Bereich liegt in Abbildung 9.3 in der Nähe des Nullpunktes und ist nur in logarithmischer Darstellung gut zu erkennen.

Linearer oder ohmscher Bereich

Dieser Bereich ist durch $U_{GS} > U_{Th}$ und $U_{DS} < U_{GS} - U_{Th}$ definiert. Er liegt links von der gepunkteten Linie in Abbildung 9.3. Dank der positiven Spannung U_{GS} bildet sich ein leitender Kanal, da die Löcher, die Majoritätsladungsträger des Substrats, vom Gate zurückgedrängt werden. Dieses Szenario ist in Abbildung 9.6 dargestellt.

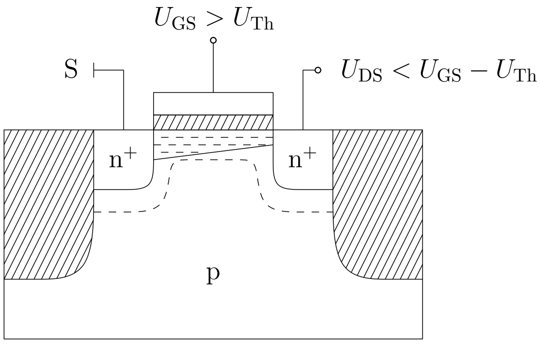


Abb. 9.6 Ein NMOS-Transistor im linearen Bereich. Die mit Elektronen angereicherte Region unter der Gate-Oxid-Schicht ist die leitende Inversionsschicht.

Die negative Inversionsladung im Kanal lässt sich über das Gaußsche Gesetz aus der Gatespannung und der Substratspannung berechnen. Einfacher ist es vielleicht, die Inversionsladung als Ladung eines Kondensators mit den Leiterplatten Gate und Substrat, die voneinander durch das Gateoxid isoliert sind, zu betrachten.

Im Fall $U_{DS} = 0$ ergibt sich die negative Flächenladung Q_F zu:

$$Q_F = C_{Ox} \cdot (U_{GS} - U_{Th}) \quad . \tag{9.4}$$

C_{Ox} ist die Kapazität des genannten Kondensators pro Fläche $W \cdot L$, W und L sind die Breite und Länge des Gates. Die Flächenkapazität C_{Ox} lässt sich gemäß der Formel für den parallelen Plattenkondensator berechnen

$$C_{Ox} = \epsilon_0 \, \epsilon_{SiO_2} \frac{1}{d_{Ox}} \quad , \quad \text{mit} \quad \epsilon_{SiO_2} \approx 11,9 \quad . \tag{9.5}$$

Werte von C_{Ox} werden typischerweise in $\frac{\mathrm{fF}}{\mu\mathrm{m}^2}$ angegeben.

Woher kommt die Größe U_{Th} in der Gleichung 9.4? Die Schwellspannung U_{Th} berücksichtigt, dass im Dielektrikum zwischen Gate und Source und auch in der Verarmungszone weitere Ladungsträger vorhanden sind.

Im Fall $U_{DS} > 0$ ergibt sich:

$$Q_F = C_{Ox} \cdot (U_{GS} - U_{Th} - U(x)) \quad . \tag{9.6}$$

Hier ist $U(x)$ die Spannung zwischen Source und der Position x im Kanal. Die Spannung zwischen dem Punkt x im Kanal und dem Gate ist damit $U_{GS} - U(x)$.

Im folgenden benötigen wir noch die Driftgeschwindigkeit v_D der Ladungsträger zwischen Source S und Drain D mit

$$\begin{aligned} v_{\mathrm{D}} &= \frac{dx}{dt} = \mu\, E_{\mathrm{Drift}} = \mu\, \frac{U_{\mathrm{DS}}}{L} \qquad (9.7) \\ \text{bzw.} \qquad v_{\mathrm{D}} &= \mu\, \frac{\mathrm{d}U(x)}{\mathrm{d}x}\,. \qquad (9.8) \end{aligned}$$

Dabei ist μ die Beweglichkeit der Ladungsträger. Für einen NMOS-Transistor ist μ_n die Beweglichkeit der Elektronen. Die Ladungsträgerbeweglichkeit hängt u. a. von der Temperatur, Defekten und der Dotierung ab, da die Wahrscheinlichkeit von Streuprozessen mit der Dotierung zunimmt. Bei Raumtemperatur und für schwach dotiertes Silizium beträgt die Beweglichkeit für Löcher, μ_p, bzw. Elektronen, μ_n,

$$\mu_p \approx 450\, \frac{\mathrm{cm}^2}{\mathrm{Vs}} \quad \text{bzw.} \quad \mu_n \approx 1400\, \frac{\mathrm{cm}^2}{\mathrm{Vs}} \quad .$$

Der Drainstrom ergibt sich aus $I_{DS} = dQ/dt = Q_F\, W\, dx/dt$ mit Gl. 9.6 und 9.8 zu

$$I_D = C_{Ox}\, W\, (U_{GS} - U_{Th} - U(x))\, \mu\, \frac{dU(x)}{dx} \quad . \qquad (9.9)$$

Integration über dx und dU ergibt

$$\begin{aligned} I_{DS} \int_0^L dx = I_{DS}\, L = \mu\, C_{Ox}\, W \int_0^{U_{DS}} (U_{GS} - U_{Th} - U(x))\, dU = \\ \mu\, C_{Ox}\, W\, \left[U_{GS} - U_{Th}\right] U_{DS} - \frac{U_{DS}^2}{2} \quad . \end{aligned}$$

Insgesamt gilt also

$$I_{DS} = \mu\, C_{Ox}\, \frac{W}{L}\, (U_{GS} - U_{Th})\, U_{DS} - \frac{U_{DS}^2}{2} \quad . \qquad (9.10)$$

Um Gleichung 9.10 etwas zu vereinfachen, definieren wir $\beta = \frac{W}{L}\, C_{Ox}\, \mu = \frac{W}{L}\, K$. Die Größe β nennt man auch den Steilheitskoeffizienten, sie darf nicht mit der Stromverstärkung des BJTs verwechselt werden. Damit gilt

$$I_{DS} = \beta \left[U_{GS} - U_{Th} - \frac{U_{DS}}{2}\right] U_{DS} \quad . \qquad (9.11)$$

Der Transistor verhält sich also in guter Näherung wie ein von U_{GS} gesteuerter Widerstand.

Anmerkung: Die Dicke der Inversionsschicht beträgt grob 5 nm.

Anmerkung: Die Größe $K_n = \mu_n C_{Ox}$ ist technologieabhängig und liegt in der Größenordnung von einigen 100 $\frac{\mu A}{V^2}$. Für 130 μm CMOS beträgt K_n rund 800 $\frac{\mu A}{V^2}$.

„Strong Inversion" oder Sättigungsbereich

Der Sättigungsbereich, auch *strong inversion* genannt, ist durch die Spannungen $U_{GS} > U_{Th}$ und $U_{DS} \geq U_{GS} - U_{Th} = U_{DS_{sat}}$ definiert. Hier ist also U_{DS} größer als im linearen Bereich. Für „strong inversion" sind die I_{DS}-U_{DS}-Kennlinien vergleichsweise flach. In diesem Bereich werden MOSFET-Verstärker meistens betrieben, entsprechend zu Bipolartransistoren im sogenannten aktiven Bereich[3]. Dieses Szenario ist in Abbildung 9.7 dargestellt.

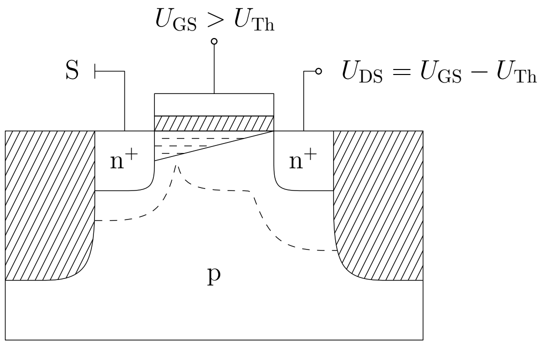


Abb. 9.7 Ein NMOS-Transistor im Sättigungsbereich für $U_{DS_{sat}} = U_{GS} - U_{Th}$.

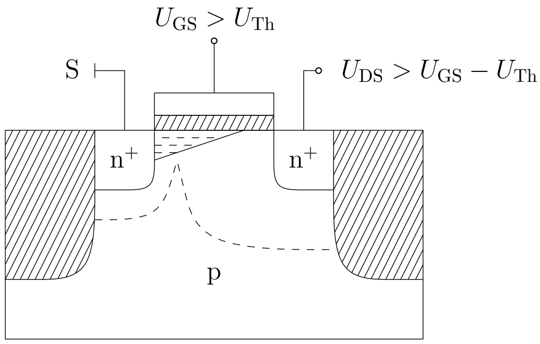


Abb. 9.8 Ein NMOS-Transistor im Sättigungsbereich für $U_{DS_{sat}} > U_{GS} - U_{Th}$.

Die Erhöhung von U_{DS} führt gegenüber Abbildung 9.6 zu zwei wesentlichen Änderungen.

1. Die Verarmungszone um den Drain vergrößert sich deutlich, da die mit U_{DS} auch die Spannung zwischen Drain und Substrat zugenommen hat.

2. Der Kanal beginnt sich am Drain abzuschnüren (*pinch-off*) Dies liegt an der verringerten Spannung oberhalb und unterhalb des Kanals in der Nähe des Drains im Vergleich zur

[3]Leider ist die Nomenklatur verwirrend. Bei MOSFETs entspricht der Sättigungsbereich großen U_{DS}, bei BJTs entspricht er kleinen U_{CE}

Region um die Source. In dem oben verwendeten Kondensatorbild nimmt die Kondensatorspannung mit zunehmender Nähe zum Drain ab und damit auch die gespeicherte Inversionsladung Q. Für $U_{DS} > U_{DS_{sat}}$ wird der Kanal noch weiter in Richtung Source zurückgedrängt.

Wichtig ist nun, dass das Abschnüren des Kanals den Stromfluss nicht verhindert. Elektronen werden aus dem Kanal in die Verarmungszone um den Drain injiziert und dort von der positiven Drainspannung in den Drain gezogen. Die Elektronen stammen teilweise aus thermischer Anregung der Substratatome und hauptsächlich aus der stark negativ dotierten Source. Eine Vergrößerung von U_{DS} verkürzt den Kanal, beeinflusst aber den nicht abgeschnürten Teil wenig. Damit tritt ein Sättigungseffekt auf. Der Drainstrom I_{DS} ergibt sich aus Gleichung 9.11 durch Einsetzen von $U_{DS} = U_{D_{sat}} = U_{GS} - U_{Th}$ zu

$$I_{DS} = \frac{\beta}{2}\,(U_{GS} - U_{Th})^2 \quad . \tag{9.12}$$

In erster Näherung ist I_{DS} also unabhängig von U_{DS} und hängt quadratisch, nicht exponentiell wie beim Bipolartransistor, von der Gate-Source-Spannung U_{GS} ab. Damit ergibt sich die Steilheit g_m im Sättigungsbereich zu

$$g_m = \beta\,(U_{GS} - U_{Th}) \quad . \tag{9.13}$$

Über

$$I_{DS} = \frac{\beta}{2}\,(U_{GS} - U_{Th})^2 \quad \text{und} \quad \sqrt{\frac{2\,I_{DS}}{\beta}} = U_{GS} - U_{Th}$$

ergibt sich

$$g_m = \sqrt{2\,\beta\,I_{DS}} \quad \text{und damit} \quad g_m \sim \sqrt{I_{DS}} \quad . \tag{9.14}$$

Eine andere Umformung ergibt die interessante Beziehung

$$\frac{g_m}{I_{DS}} = \frac{2}{U_{GS} - U_{Th}} \quad . \tag{9.15}$$

Der Steilheit g_m bezogen auf den Drainstrom I_{DS} ist im Sättigungsbereich also nicht geometrieabhängig.

Die Steilheit g_m liegt oft in der Größenordnung von $100\,\mu$S und damit wäre $\frac{1}{g_m} \approx 10\,\mathrm{k\Omega}$.

Die Ausgangsleitfähigkeit g_{DS} und der entsprechende differentielle Widerstand r_{DS} des Transistors berechnen sich nach Gleichung 9.12 zu

$$g_{DS} = \frac{\partial I_{DS}}{\partial U_{DS}} = 0 \quad , \quad \text{also} \quad r_{DS} = \frac{1}{g_{DS}} = \infty \quad .$$

Reelle Transistoren zeigen anders als in Abb. 9.3 dargestellt eine leichte Abhängigkeit des Stroms I_{DS} von U_{DS}. Dies wird gerne durch einen zusätzlichen Faktor $(1 + \lambda\,U_{DS})$ berücksichtigt. Der Parameter λ berücksichtigt den Early-Effekt, die Verkürzung der Kanallänge L mit steigendem U_{DS}.

$$I_{DS} = \frac{\beta}{2}\,(U_{GS} - U_{Th})^2\,(1 + \lambda\,U_{DS}) \quad \text{mit} \quad \lambda = 0{,}01 \ldots 0{,}1 \left[\frac{1}{\mathrm{V}}\right] \quad . \tag{9.16}$$

Damit ergibt sich der realistischere Ausdruck

$$g_{DS} = \frac{\partial I_{DS}}{\partial U_{DS}} = \frac{\lambda\, I_{DS}}{1+\lambda\, U_{DS}} \sim \lambda\, I_{DS} \quad .$$

Anmerkung: Die hier betrachteten Formeln sind didaktisch instruktiv und für Transistoren mit $L \approx 0,25\,\mu$m repräsentativ. Transitoren mit deutlich kleineren Abmessungen werden aber nicht mehr beschrieben.

Zusammenfassend gilt für MOSFETs das in Tabelle 9.1 aufgeführte Verhalten.

Bereich	**Weak Inversion** $U_{GS} < U_{Th}$	**Linearer Bereich** $U_{GS} > U_{Th}$, $U_{DS} < U_{GS} - U_{Th}$	**Strong Inversion** $U_{GS} > U_{Th}$, $U_{DS} \geq U_{GS} - U_{Th}$
I_{DS}	$I_{D_0}\, \mathrm{e}^{\frac{U_{GS}-U_{Th}}{n\, U_T}}$	$\beta\left[U_{GS} - U_{Th} - \frac{U_{DS}}{2}\right]\, U_{DS}$	$\frac{\beta}{2}\,(U_{GS}-U_{Th})^2\,(1+\lambda\, U_{DS})$
g_m	$I_{DS}\,\frac{q_e}{n\,k\,T}$	$\beta\, U_{DS}$	$\sqrt{2\,\beta\, I_{DS}}$

Tab. 9.1 Übersicht über die Betriebsmodi von MOSFETs

9.2 Grundschaltungen des FET

Wir können alle Schaltungen, die wir schon vom BJT kennen, auch mit MOSFETs realisieren. Die Grundschaltungen sind in den Abbildung 9.9 dargestellt. Die Eigenschaften der Schaltungen sind in Tab. 9.2 knapp zusammengestellt.

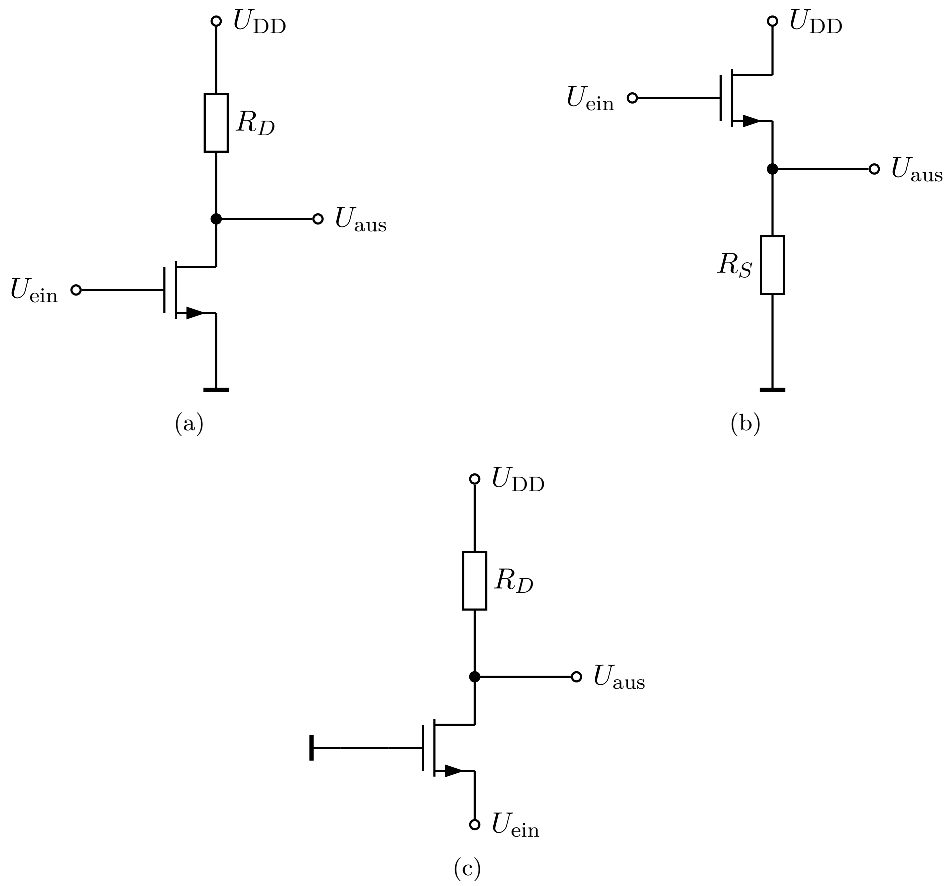


Abb. 9.9 Grundschaltungen des FET am Beispiel eines NMOS-Transistors: (a) Sourceschaltung, (b) Drainschaltung bzw. Sourcefolgerschaltung, (c) Gateschaltung.

Eigenschaften	**Sourceschaltung**	**Drainschaltung**	**Gateschaltung**
Eingangsimpedanz	sehr hoch	sehr hoch	klein, $Z_{ein} \approx \frac{1}{g_m}$
Ausgangsimpedanz	$Z_{aus} \approx R_D$	„klein“, $Z_{aus} \approx \frac{1}{g_m} \parallel R_S$	$Z_{aus} \approx R_D$
Spannungsverstärkung	$V_U = -\,R_D\,g_m$	$V_U \approx 1$	$V_U \approx R_D\,g_m$

Tab. 9.2 Eingangsimpedanz Z_{ein}, Ausgangsimpedanz Z_{aus} und Spannungsverstärkung V_U der Grundschaltungen eines FET.

Diese Schaltungen werden wir noch etwas genauer anschauen und führen aber zuerst die weiteren Grundbausteine, den FET als Diode und als Stromspiegel, ein.

9.3 FET als Diode

Diese Beschaltung eines Transistors als Diode kennen wir schon vom BJT. In der Standardkonfiguration mit positiver Versorgungsspannung gilt $U_{DS} = U_{GS} \geq U_{GS} - U_{Th}$. Der Transistor befindet sich in Sättigung und leitet.

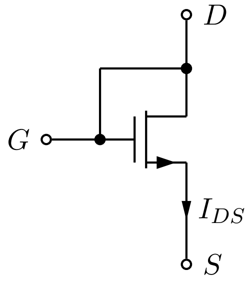


Abb. 9.10 NMOS in „Diodenschaltung".

Der differentielle Drain-Source-Widerstand im Arbeitspunkt r_{DS} errechnet sich zu

$$r_{DS} = \frac{\partial U_{DS}}{\partial I_{DS}} = \frac{\partial U_{GS}}{\partial I_{DS}} = \frac{1}{g_m} \ . \tag{9.17}$$

9.4 Stromspiegel

Der Stromspiegel in Abb. 9.11 ist analog zu der uns schon bekannten Schaltung mit Bipolartransistoren aufgebaut. Es gibt natürlich zahlreiche Varianten und Verbesserungen der Grundschaltung.

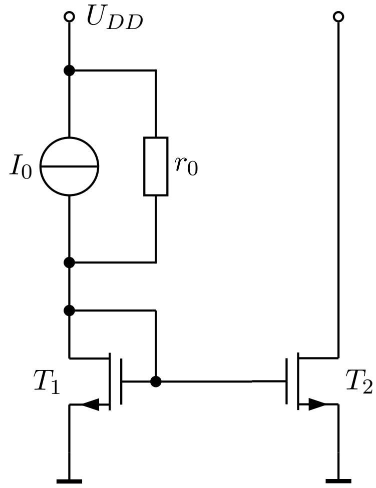


Abb. 9.11 Stromspiegel mit zwei NMOS-Transistoren und einer Stromquelle.

Der Transistor T_1 in Diodenschaltung muss den Strom I_0 der Stromquelle auf Grund des kleinen Widerstands der „Diode" durchleiten. Damit stellt sich die Spannung U_{GS_1} entsprechend ein. Für das Verhältnis der Ströme ergibt sich über $U_{\mathrm{GS}_1} = U_{\mathrm{GS}_2}$

$$\frac{I_1}{I_2} = \frac{\left(\frac{W}{L}\right)_1}{\left(\frac{W}{L}\right)_2} \, \frac{1 + \lambda U_{DS_1}}{1 + \lambda U_{DS_2}} \approx \frac{W_1}{W_2} \, \frac{L_2}{L_1} \quad . \tag{9.18}$$

Somit lässt sich der Strom I_2 über die Wahl des Stromes $I_0 = I_1$ und die Verhältnisse der Geometrieparameter W und L von T_1 und T_2 einstellen.

9.5 Common-Source-Verstärker

Die Schaltung in Abbildung 9.12 wird als Sourceschaltung oder Common-Source-Verstärker bezeichnet. Sie entspricht der Emitterschaltung des BJT.

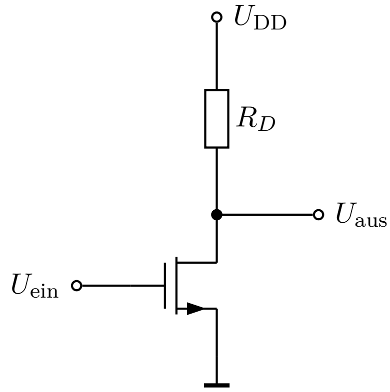


Abb. 9.12 Sourceschaltung eines NMOS-Transistors.

Die Kleinsignalersatzschaltung (Abb. 9.13) ist im Vergleich zum BJT einfacher, da der differenzielle Widerstand R_{BE} fehlt. Dadurch sind Eingangs- und Ausgangsbereich der Schaltung für niederfrequente Eingangssignale maximal voneinander entkoppelt.

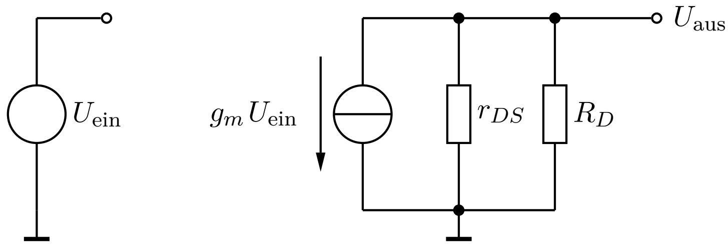


Abb. 9.13 Kleinsignalersatzschaltung des NMOS-Verstärkers.

Die Spannungsverstärkung V_U und die Ein- und Ausgangsimpedanzen R_{ein} und R_{aus} ergeben sich zu:

$$V_U = \frac{dU_{aus}}{dU_{ein}} = -\frac{(r_{DS} \parallel R_D)\, g_m\, dU_{ein}}{dU_{ein}} = -\,(r_{DS} \parallel R_D)\, g_m \approx -g_m\, R_D \tag{9.19}$$

$$R_{ein} = \frac{dU_{ein}}{dI_{ein}} \approx \infty \quad \text{und} \quad R_{aus} = \frac{dU_{aus}}{dI_{aus}} = R_D \parallel r_{DS} \approx R_D \quad , \text{da} \quad R_D \ll r_{DS}. \tag{9.20}$$

R_D kann durch einen zweiten Transistor, d. h. eine aktive Last, ersetzt werden. Warum ist dies sinnvoll?

1. Ein größeres R_D vergrößert die Verstärkung.

2. In integrierten Schaltungen benötigen ohmsche Widerstände viel Chipfläche und ihr Wert ist dadurch begrenzt.

In Abbildung 9.14 ist R_D deshalb durch eine Stromquelle ersetzt worden. Die Spannungsverstärkung ergibt sich aus Gleichung 9.19, wenn man für R_D den sehr großen Innenwiderstand R_i der Stromquelle einsetzt, der aber gegenüber r_{DS} vernachlässigt werden kann.

Mit $I_0 = I_{DS}$ folgt

$$V_U = -\,g_m\,(r_{DS} \,\|\, R_i) \approx -\,g_m\,r_{DS} = -\,\frac{g_m}{g_{DS}} \quad .$$

Für $I_{DS} = 100\,\mu\mathrm{A}$ und $\lambda \approx 0{,}1\frac{1}{\mathrm{V}}$ ergibt sich $g_{DS} \approx \lambda\,I_{DS} = 10\,\mu\mathrm{S}$. Für ein $g_m \approx 100\,\mu\mathrm{S}$ folgt $V_U \approx 10$.

Anmerkung: Das Verhältnis $\frac{g_m}{g_{DS}}$ ist ein wichtiger Technologieparameter. In modernen Transistortechnologien mit abnehmender Transistorgröße bzw. Gatelänge L sinkt auch $\frac{g_m}{g_{DS}}$ und damit die Verstärkung. Der analoge Schaltungsentwurf wird also komplexer.

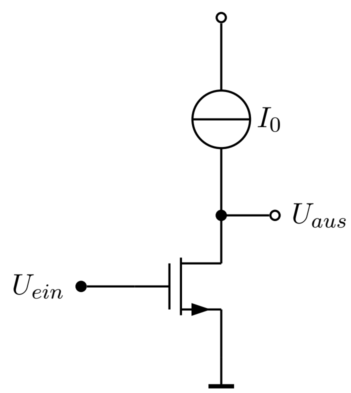


Abb. 9.14 NMOS-Verstärker mit „Stromquelle“.

Eine einfache Realisierung einer Stromquelle wäre zum Beispiel eine „Transistordiode“ (siehe Abb. 9.10) mit $r_{DS} = \frac{1}{g_{m_2}}$.

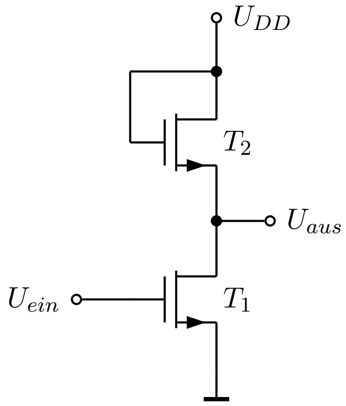


Abb. 9.15 NMOS-Verstärker mit Transistor als Stromquelle.

Es ergibt sich mit Gleichungen 9.17 und 9.19

$$V_U = -\,g_m\,R_D = -\,\frac{g_{m_1}}{g_{m_2}} = -\sqrt{\frac{\left(\frac{W}{L}\right)_1}{\left(\frac{W}{L}\right)_2}} \quad . \tag{9.21}$$

Auf Grund der obigen Wurzel ergeben sich Probleme. Für eine Verstärkung $V_U = -\,10$ wird eine Geometrieverhältnis $\frac{\left(\frac{W}{L}\right)_1}{\left(\frac{W}{L}\right)_2} = 100$ benötigt und damit nimmt T_1 viel Chipfläche in Anspruch. Die Schaltung aber hat bei Wahl kleiner Verstärkung für hohe Bandweiten ihre Berechtigung.

Um eine große Verstärkung zu erreichen, wird die Stromquelle aber besser nicht über eine Diode, sondern einen Stromspiegel realisiert (Abb. 9.16).

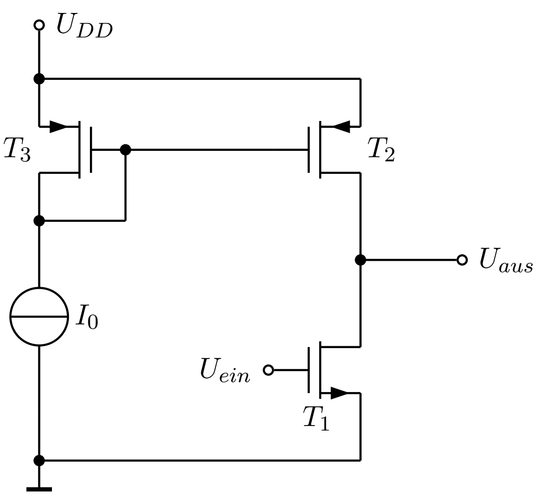


Abb. 9.16 NMOS-Verstärker mit PMOS-Stromspiegel.

Hier ist folgendes besonders interessant:

- Die Stromquelle I_0 beeinflusst die Schaltung nicht, sondern stellt nur den Strom ein.

- T_2 und T_3 sind PMOS Transistoren.

- Die Verstärkung $V_U = -\, g_{m_1}\,(r_{DS_1} \,\|\, r_{DS_2})$ nimmt zu, da hier $R_{\mathrm{DS}} = \frac{1}{g_{\mathrm{DS2}}} >> \frac{1}{g_{\mathrm{m2}}}$.

Beispiel: Für $g_m \approx 100\,\mu\mathrm{S}$ und $r_{DS_1} = r_{DS_2} \approx 1\,\mathrm{M}\Omega$ (wähle hierzu I_{DS} „klein“, z.B. $I_{DS} \approx 10\,\mu\mathrm{A}$) folgt $V_U = -\,50$.

Frage: Kann man die Verstärkung noch weiter vergrößern?
Nicht durch eine bessere Stromquelle, da nun auch r_{DS_1} die Verstärkung begrenzt.

Ersetzt man T_1 allerdings durch eine Kaskode, so kann man den Widerstand r_{DS_1} effektiv vergrößern. Dies ist in Abb. 9.17 illustriert. Hier ist der PMOS-Stromspiegel der Übersicht halber durch eine ideale Stromquelle dargestellt.

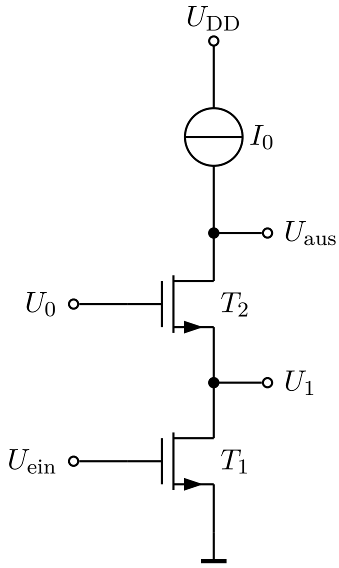


Abb. 9.17 Abwandlung von Abb. 9.16 mit einer Kaskode statt des einfachen Transistors T_1 und einer idealen Stromquelle statt des NMOS-Stromspiegels.

Das Kleinsignalersatzschaltbild ist in Abbildung 9.18 dargestellt. Die Stromquelle I_0 hat einen unendlich hohen Innenwiderstand und ist nicht extern gesteuert. Sie taucht also im Ersatzschaltbild nicht auf.

Anmerkung: Der Innenwiderstand der Spannungsquelle U_0 ist klein, U_0 liegt damit in der Kleinsignalanalyse auf Masse.

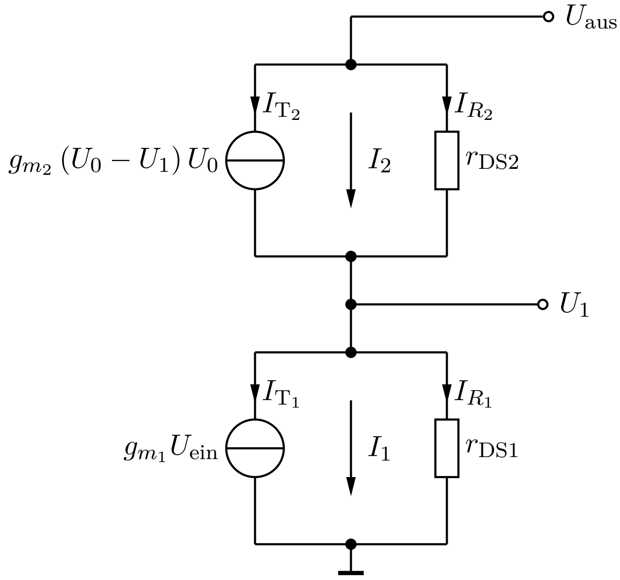


Abb. 9.18 Kleinsignalersatzschaltbild der Abb. 9.17.

Die Stromquelle I_0 legt den Strom durch die Transistoren T_1 und T_2 fest. Damit muss gelten

$$\mathrm{d}I_1 = g_{m1}\,\mathrm{d}U_{ein} + \frac{\mathrm{d}U1}{r_{DS1}} = 0 \quad \text{bzw.} \quad \mathrm{d}U_1 = -g_{m1}\,\mathrm{d}U_{ein} \cdot r_{DS1} \quad . \tag{9.22}$$

Gleichzeitig gilt analog

$$\mathrm{d}I_2 \;=\; g_{m2}\,(\mathrm{d}U_o - \mathrm{d}U_1) + \frac{\mathrm{d}U_{aus} - \mathrm{d}U_1}{r_{DS2}} = 0 \quad . \tag{9.23}$$

Der Strom durch die Transistoren hat jeweils zwei Anteile: den ohmschen Anteil I_R, der durch r_{DS} fließt und durch die Spannungsdifferenz an r_{DS} gegeben ist, und den „Transistoranteil“, der durch die Gatespannung gesteuert wird.

Mit $\mathrm{d}U_0 = 0$ und durch Einsetzen von Gl. (9.22) in Gl. (9.23) folgt

$$\mathrm{d}U_{aus} \;=\; \mathrm{d}U_1\,(g_{m2}\,r_{DS2} + 1) = -g_{m1}\mathrm{d}U_{ein} \cdot r_{DS1}\,(g_{m2}\,r_{DS_2} + 1) \tag{9.24}$$

$$\text{und} \qquad V_U = \frac{\mathrm{d}U_{aus}}{\mathrm{d}U_{ein}} \;\approx\; -g_{m1}\,r_{DS1}\,g_{m2}\,r_{DS_2} \quad . \tag{9.25}$$

Die Gesamtverstärkung ist damit um ein Vielfaches größer als in der Schaltung aus Abbildung 9.16.

Anschaulich lässt sich die Schaltung leicht verstehen, wenn man sich klar macht, dass $\mathrm{d}I_1 = 0$. Dies gilt, da der durch U_{ein} gesteuerte Strom dI_T, den durch $\mathrm{d}U_1$ gesteuerten Strom $\mathrm{d}I_R$ kompensiert. Mit zunehmenden U_{ein} steigt der Stromanteil $\mathrm{d}I_{T_1}$ gemäß $\mathrm{d}I_{T_1} \approx g_{m_1}\,\mathrm{d}U_{ein}$ an. Damit gilt

$$\mathrm{d}U_{ein}\,g_{m_1} \;=\; \mathrm{d}I_{T_1}\,, \tag{9.26}$$

$$\mathrm{d}I_{R_1}\,r_{DS_1} \;=\; \mathrm{d}U_1 \tag{9.27}$$

$$\text{also} \qquad \mathrm{d}U_1 \;\approx\; -r_{DS_1}\,g_{m_1}\,\mathrm{d}U_{ein} \quad . \tag{9.28}$$

Derselbe Effekt tritt für T_2 auf. Zwar liegt U_0 auf konstantem Potential, aber über $U_{\mathrm{GS}} = U_0 - U_1$ ändert sich der Stromanteil $\mathrm{d}I_{\mathrm{T}_2} = -\mathrm{d}U_1\, g_{\mathrm{m}_2}$ und damit

$$\begin{aligned} & \mathrm{d}I_{\mathrm{T}_2} &=& -\mathrm{d}U_1\, g_{\mathrm{m}_2} \\ \text{und}\quad & \mathrm{d}U_{\mathrm{aus}} &=& \mathrm{d}I_{R_2}\, r_{\mathrm{DS}_2} = -\mathrm{d}I_{\mathrm{T}_2}\, r_{\mathrm{DS}_2} \\ \text{also}\quad & \mathrm{d}U_{\mathrm{aus}} &=& -\mathrm{d}U_1\, g_{\mathrm{m}_2}\, r_{\mathrm{DS}_2} = \mathrm{d}U_{\mathrm{ein}}\, g_{\mathrm{m}_1}\, r_{\mathrm{DS}_1}\, g_{m_2}\, r_{\mathrm{DS}_2} \quad . \end{aligned} \tag{9.29}$$

Durch weiteres Kaskadieren könnte man die Verstärkung noch weiter erhöhen. Allerdings reduziert man den maximalen Spannungsausschlag von U_{aus}.

9.6 Sourcefolger

Die Drainschaltung[4] bzw. der Sourcefolger (Abb. 9.19) entspricht dem Emitterfolger des Bipolartransistors.

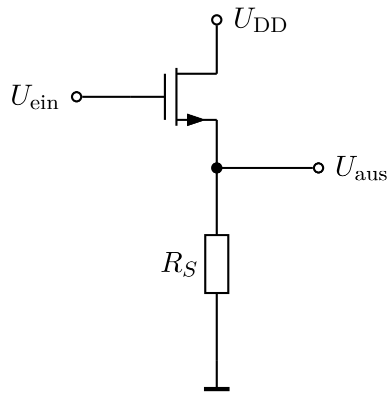
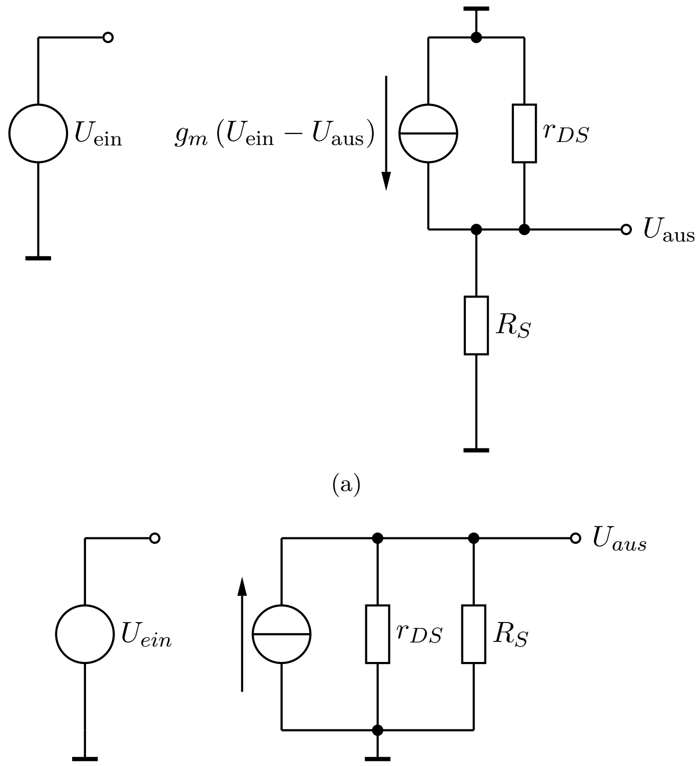


Abb. 9.19 Drainschaltung eines NMOS-Transistors.

In Abbildung 9.20(a) und 9.20(b) sind die Kleinsignalersatzschaltungen dargestellt.

[4]engl. *common drain amplifier*



(a)

(b)

Abb. 9.20 (a) Kleinsignalersatzschaltbild des NMOS-Sourcefolgers. (b) Vereinfachtes Kleinsignalersatzschaltung des NMOS-Sourcefolgers.

Aus der Knotengleichung in Abbildung 9.20(b) ergibt sich

$$I_{DS} = g_m\, dU_{GS} = g_m\, d\,(U_{ein} - U_{aus}) = \frac{dU_{aus}}{R_S \parallel r_{DS}} = \frac{dU_{aus}}{R_S^{'}} \quad \text{mit} \quad R_S^{'} = R_S \parallel r_{DS} \quad .$$

Daraus folgt

$$dU_{aus}\,(1 + g_m\, R_S^{'}) = dU_{ein}\, g_m\, R_S^{'}.$$

Die Spannungsverstärkung

$$V_U = \frac{dU_{aus}}{dU_{ein}} = \frac{g_m\, R_S^{'}}{1 + g_m\, R_S^{'}} < 1$$

Außerdem gilt

$$R_{ein} \to \infty \quad \text{und} \quad R_{aus} = \frac{dU_{aus}}{dI_{aus}} = \frac{1}{g_m} \parallel R_S^{'} \quad .$$

In der Praxis ist $R_s \approx 1\ldots10\,\text{k}\Omega$, g_m sollte groß sein (mS). Damit gilt $V_U \approx 1$ und die Ausgangsimpedanz R_{aus} ist eher klein.

9.7 Selbstleitende MOSFETs

Bisher haben wir nur den NMOS FET kennengelernt, und zwar den selbstsperrenden oder „enhancement“ NMOS-Transistor. Diese Bezeichnung beruht darauf, dass für $U_{GS} < U_{Th}$ selbst bei positiver Source-Drain Spannung $U_{DS} > 0\,\mathrm{V}$ kein nennenswerter Strom I_{DS} fließt.

Es gibt aber auch einen weiteren NMOS-Typ, den selbstleitenden oder depletion NMOS-Transistor. Hier ist eine dünne Schicht unterhalb des Gates negativ dotiert (Abb. 9.21), dadurch entsteht auch für $U_{GS} = 0$ eine leitende Verbindung zwischen Source und Gate. Der Strom I_{DS} verringert sich mit zunehmend negativer Spannung U_{GS}, da dann die Elektronen aus dem leitenden Kanal weggedrängt werden oder anders ausgedrückt sich eine Sperrschicht zwischen der n-dotierten Schicht unter dem Gate und dem p-bulk ausbildet.

Selbstleitende MOSFETs haben eine negative Schwellspannung. Sie können beim Schaltungsentwurf nützlich sein, werden aber deutlich seltener eingesetzt als selbstsperrende.

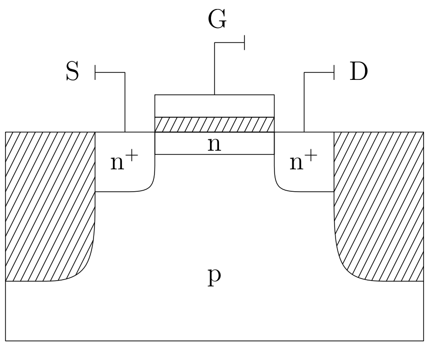


Abb. 9.21 Selbstleitender NMOS-Transistor.

9.8 Gefaltete Kaskode*

Die gefaltete Kaskode (Abb. 9.22) ist eine Variante der Schaltung aus Abb. 9.23, bei der der Strom nicht durch einen Ast und zwei NMOS-Transistoren fließt, sondern stattdessen ähnlich dem Differenzverstärker auf zwei Äste aufgeteilt wird. Es werden unterschiedliche Transistortypen, NMOS und PMOS, verwendet. Ein Vorteil der gefalteten Kaskode ist der größere Ausgangsspannungsbereich.

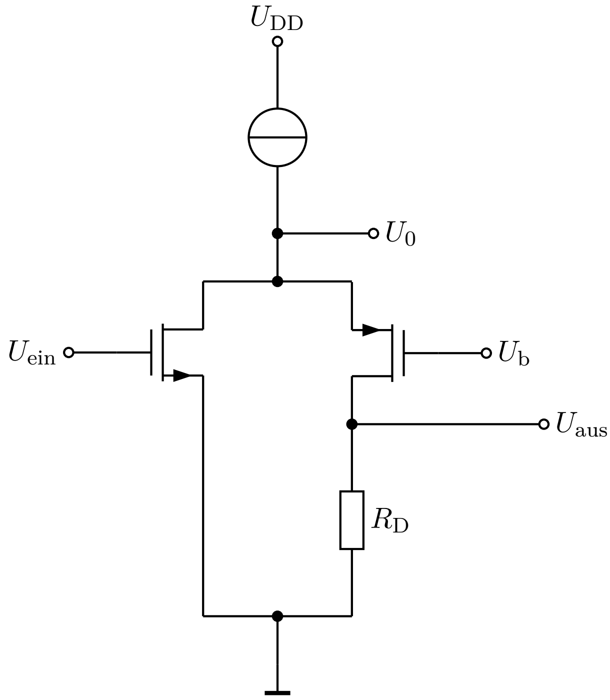


Abb. 9.22 Gefaltete Kaskode.

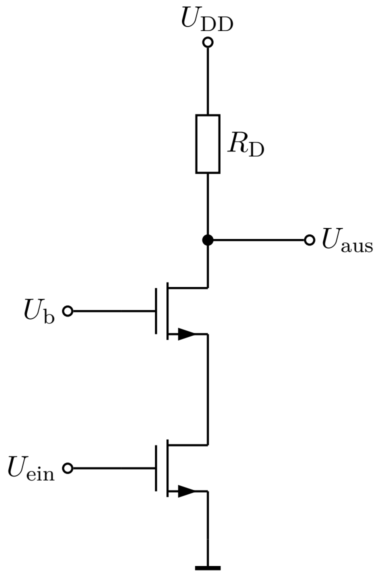


Abb. 9.23 Kaskodenschaltung mit zwei NMOSs.

9.9 CMOS - Differenzverstärker

Die Grundversion des CMOS-Differenzverstärkers ist in Abb. 9.24 dargestellt. Der gezeigte Verstärker hat zwei Ausgänge, von denen man aber nicht beide benutzen muss. Die Grundidee dieser Schaltung ist dieselbe wie beim Differenzverstärker mit Bipolartransistoren.

Die Verstärkung ergibt sich zu

$$V_U = \frac{\mathrm{d}U_{\mathrm{aus}}}{\mathrm{d}U_{\mathrm{ein}}} = \frac{\mathrm{d}U_{\mathrm{aus}+} - \mathrm{d}U_{\mathrm{aus}-}}{\mathrm{d}U_{\mathrm{ein}+} - \mathrm{d}U_{\mathrm{ein}-}} = -g_{\mathrm{m}}\, R_D \tag{9.30}$$

Betrachtet man nur den positiven Ausgang, so gilt

$$\frac{\mathrm{d}U_{\mathrm{aus}+}}{\mathrm{d}U_{\mathrm{ein}}} = \frac{g_m R_D}{2}\;. \tag{9.31}$$

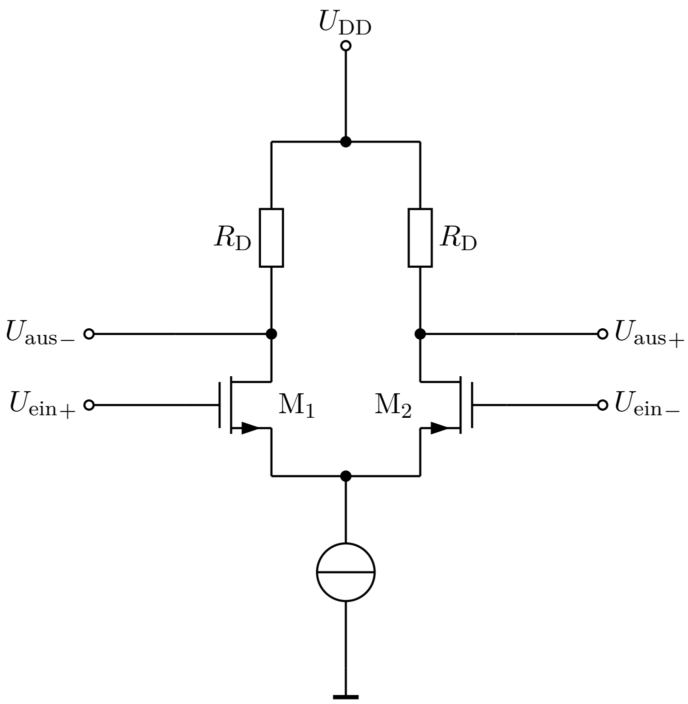


Abb. 9.24 Schaltplan eines CMOS-Differenzverstärkers.

Wie beim Differenzverstärker aus Bipolartransistoren gibt es zahlreiche Varianten dieser Schaltung. Die Stromquelle kann zum Beispiel durch einen Widerstand (Abb. 9.25 (a)) und einem Transistor mit konstanter Gatespannung (Abb. 9.25 (b)) realisiert sein.

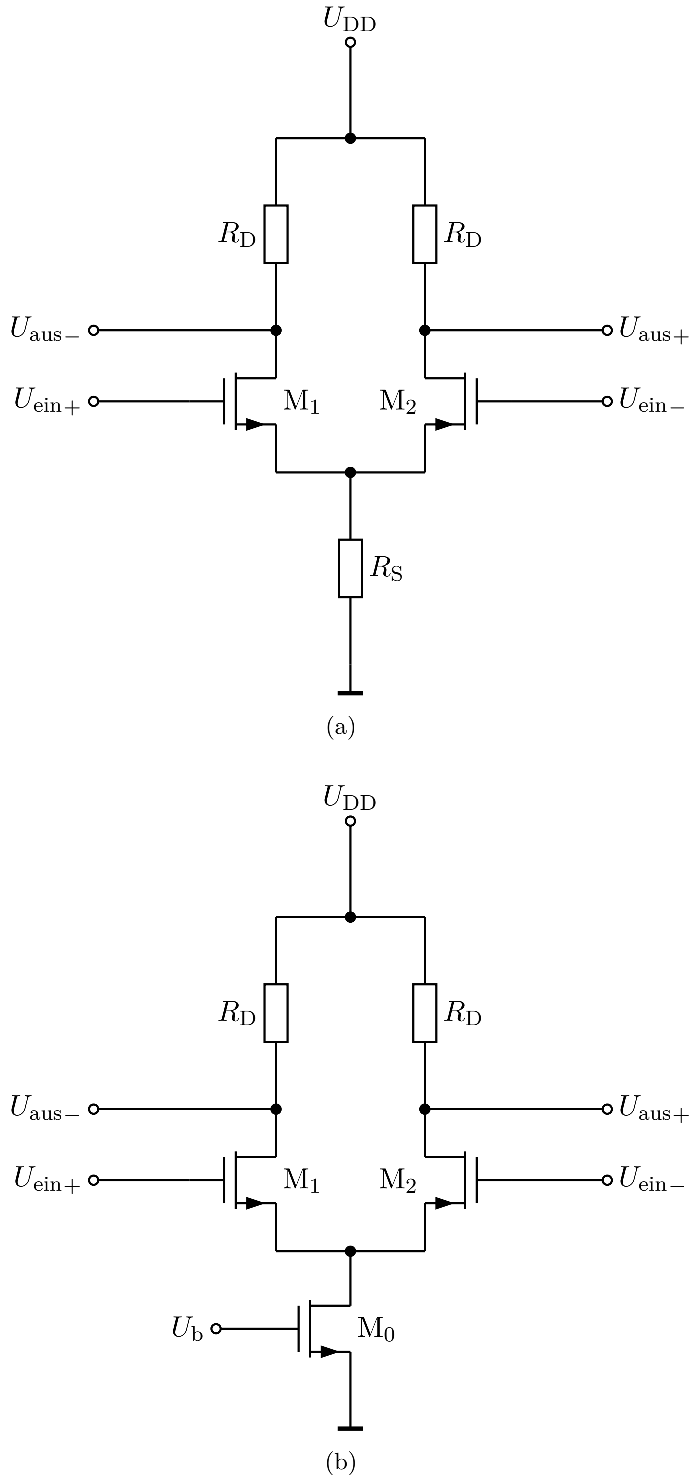


Abb. 9.25 Variation der Abb. 9.24 mit (a) einem Widerstand R_S und (b) mit einem Transistor M_0 als „Stromquelle".

Entsprechend können die Lastwidestände R_D durch eine aktive Last ersetzt werden, zum Beispiel einen einzelnen PMOS-Transistor (Abb 9.26), eine Kaskode (Abb. 9.27) beziehungsweise einen Stromspiegel (Abb. 9.28).

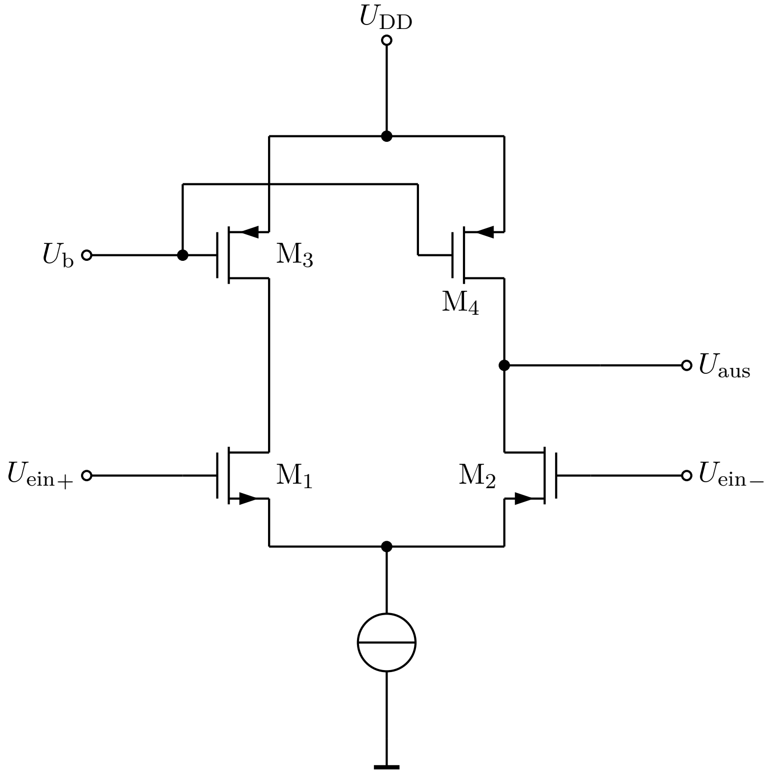


Abb. 9.26 CMOS-Differenzverstärker mit zwei PMOS-Stromquellen.

Wir gehen im folgenden davon aus, dass gegenüberliegende Transistoren baugleich sind. Also

$$g_{m_1} = g_{m_2} = g_{m_{12}} \,, \tag{9.32}$$
$$r_{\mathrm{DS}_1} = r_{\mathrm{DS}_2} = r_{\mathrm{DS}_{12}} \tag{9.33}$$
$$\text{und} \quad r_{\mathrm{DS}_2} = r_{\mathrm{DS}_4} = r_{\mathrm{DS}_{24}} \,. \tag{9.34}$$

Die Verstärkung der Schaltung in Abb. 9.26 ist

$$V_U = -g_{m12} \left(R_{DS12} || R_{DS24} \right) \,. \tag{9.35}$$

Durch die PMOS-Transistoren M_3 und M_4 lassen sich also große Impedanzen R_{DS} und dadurch große Verstärkungen bei relativ geringen Platzverbrauch bewirken. Gleichzeitig kann U_{DS} der PMOS relativ klein gehalten werden. Dadurch ist der verfügbare Bereich für U_{aus} größer, und der Verstärker geht nicht so leicht in Sättigung.

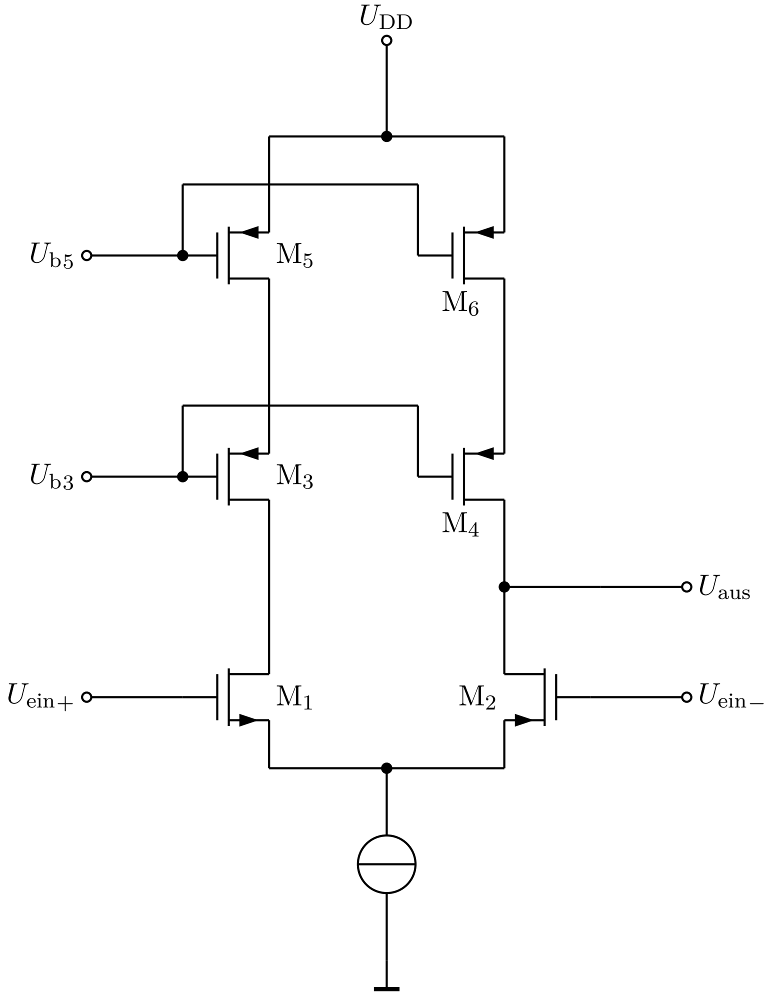


Abb. 9.27 CMOS-Differenzverstärker mit zwei kaskadierten PMOS-Stromquellen.

Die Schaltung in Abb. 9.27 ist noch raffinierter. Die Impedanz der gegenüberliegenden Transistoren M_3 und M_4 bzw. M_5 und M_6 ist jeweils

$$\frac{1}{g_{mP}}||R_{DS} \approx \frac{1}{g_{mP}} . \tag{9.36}$$

Ignoriert man die Biastransistoren M_3 und M_6, deren Impedanz R_{DS} groß ist, ergibt sich die Verstärkung also zu

$$V_U \approx -\frac{g_{mN}}{g_{mP}} . \tag{9.37}$$

Wie kann man g_{mP} klein halten, um eine große Verstärkung zu erzielen. Dies gelingt, indem man W und L der parallelen Transistoren M_3 und M_4 so dimensioniert, dass M_3 den Großteil des Stromes führt und M_4 den Rest.

Schaltung 9.27 ermöglicht eine große Ausgangimpedanz im oberen Ast der Schaltung von

$$R_{out} = r_{DS_3} + r_{DS_4} + g_{m_4} r_{DS_4} r_{DS_6} \approx g_{m_4} r_{DS_4} r_{DS_6} . \tag{9.38}$$

(Hier müsste man auch noch den unteren Ast kaskodieren, um dies auszunutzen, da $R_{\mathrm{DS12}} \,||\, g_{\mathrm{m34}}\, R_{\mathrm{DS34}}\, R_{\mathrm{DS56}} \approx R_{\mathrm{DS12}}$.)

In Abb. 9.28 ist der Differenzverstärker mit einem Stromspiegel als aktive Last dargestellt. Der Vorteil ist, dass man im Gegensatz zu Abb. 9.26 und 9.27 die Spannung U_{b} nicht einstellen muss.

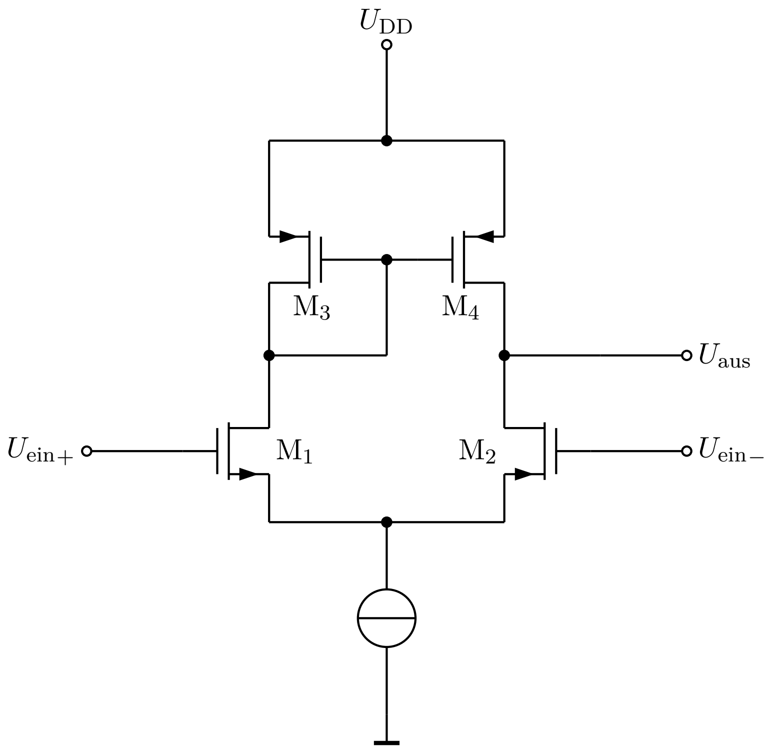


Abb. 9.28 CMOS-Differenzverstärker mit PMOS-Stromspiegel.

Der Differenzverstärker aus Abb. 9.28 ist qualitativ leicht zu verstehen, wenn man sich folgendes klar macht:

- Die gemeinsame Stromquelle am Fuß der Schaltung legt den Gesamtstrom durch den linken und rechten Ast fest

$$I_{\mathrm{L}} + I_{\mathrm{R}} \quad = \quad I_0 \; . \tag{9.39}$$

- Der Stromspiegel (M_3 und M_4) stellt nur dann einen identischen Strom durch M_3 und M_4 ein, wenn beide Transistoren in Sättigung sind und gilt $U_{\mathrm{GS}_1} = U_{\mathrm{GS}_2}$. M_3 ist aufgrund der Diodenschaltung immer in Sättigung, M_4 aber nicht.

- Die Spannung der Stromquelle ist nicht notwendigerweise konstant.

Falls die Eingangspannung $U_{\mathrm{ein+}} >> U_{\mathrm{ein-}}$ so gilt

$$I_{\mathrm{M}_1} \quad = \quad I_{\mathrm{M}_3} \approx I_0 \; . \tag{9.40}$$

Dies ist möglich, wenn M_4 sich im ohmschen Bereich befindet, $U_{SD_4} = -U_{DS_4}$ sehr klein ist und damit $U_{aus} \approx U_{DD}$. Der Wert von $U_{S_1} = U_{S_2}$ der NMOS M_1 und M_2 stellt sich so ein, dass M_2 keinen Strom I_2 führt.

Ist dagegen $U_{ein_+} << U_{ein_1}$, so kann im linken Ast kein Strom fließen, $U_{SG_3} = -U_{GS_3}$ ist klein, und damit ist auch U_{SD_3} klein.

Wieso fließt durch den rechten PMOS-Transistor so ein großer Strom $I_{M_4} = I_{M_2} \approx I_0$? Dies liegt daran, dass sich U_{aus} stark absenkt und damit der Strom $I_4 = (U_{DD} - U_{aus})\, R_{DS_4}$ durch R_{DS_4} fließt.

Analysiert man die Schaltung im Kleinsignalersatzschaltbild quantitativ, was nicht schwierig aber etwas aufwendig ist, so ergibt sich die Spannungsverstärkung zu

$$V_U \;=\; g_{m_{12}} \cdot r_{DS_4} \,||\, r_{DS_2} \,. \qquad (9.41)$$

Der Ausgangswiderstand ist

$$R_{out} \;=\; r_{DS_4} \,||\, r_{DS_2} \,. \qquad (9.42)$$

9.10 Beispiel eines CMOS-Schaltkreises

In manchen Experimenten sind die instrumentellen Anforderungen so hoch oder die Umgebungsbedingungen so ungewöhnlich, dass nicht auf kommerzielle Elektronikbausteine zurückgegriffen werden kann. Ein gutes Beispiel ist der Detektorbau in der Teilchenphysik, wo die große Anzahl elektronischer Kanäle (1 - 100 Millionen), die hohen Kollisionsraten der Strahlpakete des Beschleunigers (z. B. 40 MHz), die Strahlenbelastung der Elektronik und weiteres in der Regel Eigenentwicklungen erfordern. Ein Beispiel ist der Schaltkreis in Abbildung 9.29, entworfen von J. Kaplon (CERN).

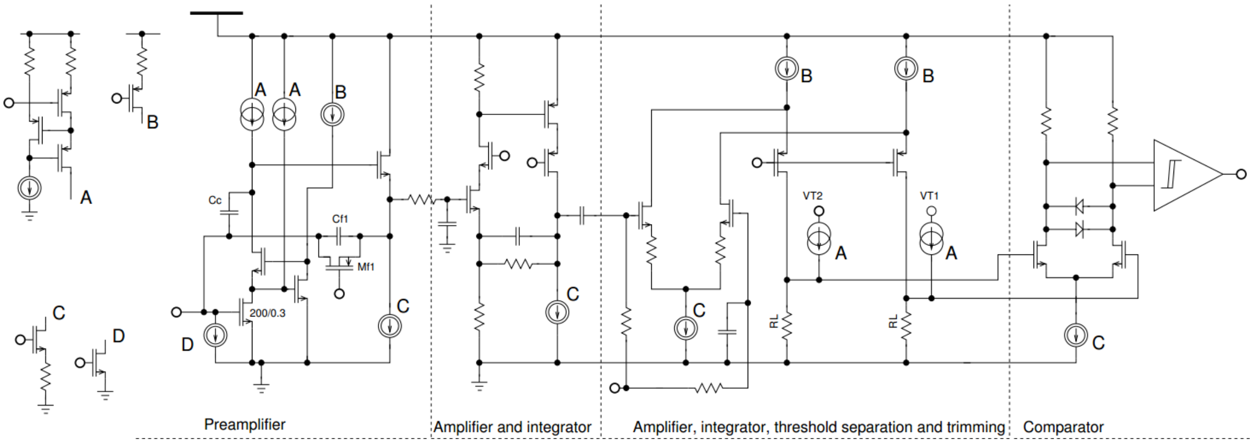


Abb. 9.29 Schaltkreis eines Frontend-Kanals eines Siliziumspurdetektors für LHC mit Vorverstärker, Integrator, Differenzverstärkerstufen.

Es wurde zur Auslese eines zukünftigen Siliziumstreifendetektors für ATLAS in einer 90 nm CMOS-Technologie konzipiert.

Der Schaltkreis besteht aus vier Blöcken, die in Abb. 9.29 durch gestrichelte Linien angedeutet sind: Vorverstärker (Block 1), Verstärker und Integrierer (Blöcke 2 und 3), Komparator (Block 4).

Grundbausteine, die wir teilweise schon kennengelernt haben sind:

a) Der Vorverstärker ist nicht differenziell. Diese Struktur, eine sogenannte „regulierte“ Kaskode kennen wir noch nicht. Der Kondensator C_{f_1} zwischen Eingang und Ausgang des Vorverstärkers macht diese Schaltung zu einem ladungsempfindlichen Vorverstärker. Die Verstärkung beträgt 5.5 mV/fC. Parallel zum Kondensator ist ein Transistor M_{f_1} geschaltet (active feed-back), der den Kondensator kontinuierlich, aber langsam, entlädt.

b) Der 2. Block besteht aus zwei kaskadierten Kaskoden. Die Kaskode im linken Ast der Schaltung aus zwei NMOS, die rechte aus zwei PMOS.

c) Der 3. Block enthält einen Differenzverstärker. Die aktive Last des Differenzverstärkers ist allerdings komplizierter als unsere bisherigen Beispiele.

d) Es werden verschiedene Stromquellen A, B, C und D unterschiedlicher Komplexität verwendet.

Die Versorgungsspannung beträgt 1,2 V. Der dargestellte Kanal hat einen Stromverbrauch von 240 μA. Davon entfällt ein gutes Drittel auf nur einen Transistor, den Eingangstransistor der mit $\frac{W}{L}$ von 200 μm/0,3 μm dimensioniert ist. Die Blöcke sind teilweise DC-gekoppelt und teilweise AC-gekoppelt. So sind Block 1 und 2 über einen Widerstand und Block 3 und 4 direkt miteinander verbunden. Block 2 und 3 dagegen über einen Kondensator.

10 Weitere Halbleiterbauelemente*

Neben den Halbleiterbauteilen Dioden, Zenerdiode, Schottkydiode, Bipolartransistor und MOSFET gibt es noch eine Vielzahl weiterer wichtiger Bauteile, die häufig verwendet werden. Im Folgenden wollen wir kurz auf den JFET, IGBT, HEMT und Thyristor eingehen (später noch VECSEL, Pin-Diode, elektrooptischer Modulator).

10.1 JFET

Die Junction FET (JFET) oder Sperrschicht-Feldeffekttransistor ähnelt im Aufbau dem MOSFET (siehe Abb. 10.1).

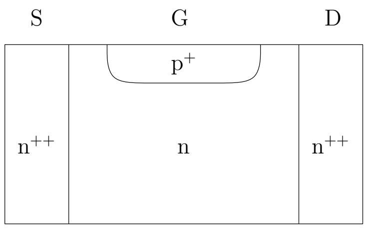


Abb. 10.1 Schematischer Aufbau eines n-Kanal-JFET

Als Schaltsymbol eines JFETs benutzen wir die in Abbildung 10.2 dargestellten Symbole.

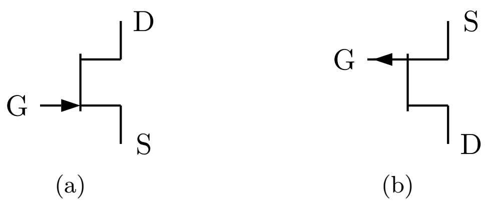


Abb. 10.2 (a) Schaltsymbol für den n-Kanal JFET und (b) für den p-Kanal JFET.

Die Kennlinien eines JFET sind in Abb. 10.3 dargestellt.

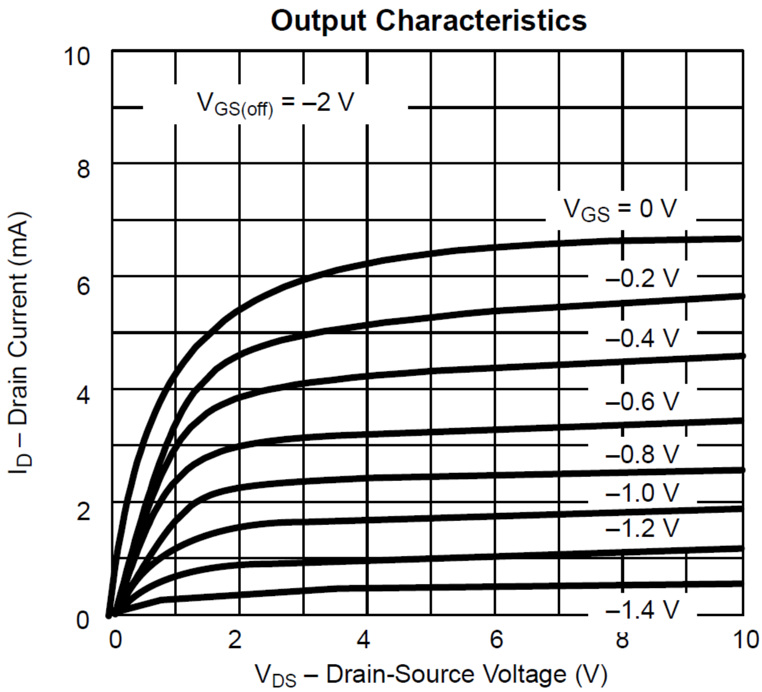


Abb. 10.3 Drainstromdiagramm und Schaltung eines JFET (aus dem Datenblatt eines 2N4416 von Vishay).

Die gestrichelte Linie entspricht der Sättigungsspannung U_{DS}.

Der JFET ist wie auch der MOSFET eine spannungsgesteuerte Stromquelle. Das Gate ist allerdings nicht durch einen Isolator vom Kanal getrennt, sondern bildet mit dem Kanal eine pn-Diodenstruktur. Liegt am Gate keine Spannung an, $U_{GS} = 0$, so sind Source und Drain über den n-Kanal leitend verbunden, und es fließt abhängig von UDS ein Strom IDS. Liegt dagegen eine negative Spannung U_{GS} an, so bildet sich eine Sperrschicht (Abb. 3).

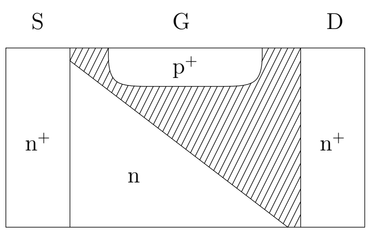


Abb. 10.4 JFET mit $U_{DS} > 0$ und $U_{GS} < 0$

Der Kanal wird „abgeschnürt“. Die Sperrschicht ist zwischen Gate und Drain größer als zwischen Gate und Source, da die Spannungsdifferenz zu Gate und Drain größer ist ($U_{DS} > 0$). Mit zunehmender Abschnürung sinkt der Strom I_{DS} durch den JFET.

Es gibt zwei Betriebsbereiche des JFET. Im linearen Bereich ist der Kanal zwischen Source und Drain noch nicht abgeschnürt. Der JFET verhält sich dann in Näherung wie ein spannungsgesteuerter Widerstand mit der Steuerspannung U_{GS}.

Im Sättigungsbereich ist der Kanal bereits abgeschnürt und die Abhängigkeit des Drainstroms I_{DS} von U_{DS} ist klein. In diesem Bereich hängt der Strom in Näherung quadratisch von U_{GS} als $I_{DS} = I_{DSS}\,(1 - \frac{U_{GS}}{U_P})^2$.

I_{DSS} ist sog. Sättigungsstrom bei $U_{GS} = 0$ und großen U_{DS}, U_P ist die Abschnürspannung bei $U_{GS} = 0$. Dies entspricht dem Schnittpunkt der gestrichelten Linie in Abb. 2 mit der Kennlinie für $U_{GS} = 0$.

JFETs haben eine hohe Eingangsimpedanz (> 100 MΩ), die allerdings kleiner als die Eingangsimpedanz von MOSFETs ist. Das Rauschen von JFETs ist bei kleinen Frequenzen nur sehr gering, JFETs werden daher als Eingang von empfindlichen diskreten Verstärkerschaltungen verwendet. JFETs sind relativ strahlenhart.

Es ist wichtig, den JFET nicht in Vorwärtsrichtung $U_{GS} > 0$ zu betreiben, dann brennt er durch. Wählt man U_{DS} zu groß, kommt es trotz Sperrbetrieb ($U_{GS} < 0$) zu einem Durchbruch, der den JFET auch zerstört.

10.1.1 Thyristor

Ein Thysristor oder gesteuerter Gleichrichter[1] ist ein Bauteil der Leistungselektronik. Thyristoren können Sröme im kA-Bereich, bei sehr geringem Spannungsabfall, leiten und Sperrspannungen im kV-Bereich aushalten.

Ein Thyristor hat drei Anschlüsse: Anode, Kathode und Gate. Der Thyristor besteht aus einer Folge von unterschiedlich dotierten Siliziumschichten: pnpn (siehe Abb. 10.5).

[1] engl.: Silicon controlled rectifiers (SCR)

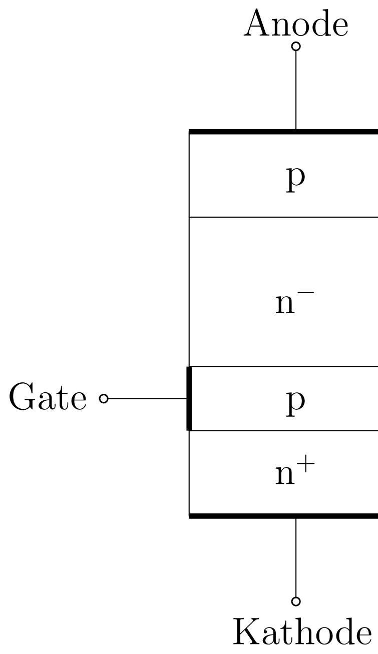


Abb. 10.5 Schematischer Aufbau eines p-Gate-Thyristors.

Das Schaltsymbol des Thyristors ist in Abb. 10.6 zu sehen.

Abb. 10.6 Schaltungssymbol eines Thyristors.

Man kann den Thyristor als zwei hintereinander geschaltete Dioden verstehen, es entstehen drei Übergänge (pn, np und pn). Liegt an der Anode eine positive Spannung gegenüber der Kathode an, dann sind die beiden äusseren pn-Übergänge zwar in Durchlassrichtung gepolt aber der mittlere np-Übergang sperrt. Hier fällt der Großteil der Anoden-Kathoden-Spannung U_{AK} ab, der Thyristorstrom ist sehr klein.

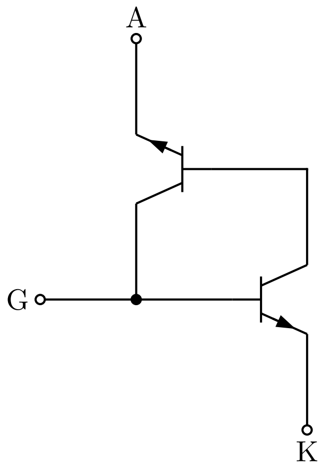


Abb. 10.7 Ersatzschaltbild eines p-Gate-Thyristors.

10.2 Der Leistungs-MOSFET

In der Leistungselektronik kommen MOSFETs für mittlere Spannungen von bis zu einigen hundert Volt zum Einsatz. Ihr geometrischer Aufbau (siehe Abb. 10.8) unterscheidet sich von den im Kapitel 9 vorgestellten MOS-Strukturen durch die vertikale Anordnung von Source und Drain.

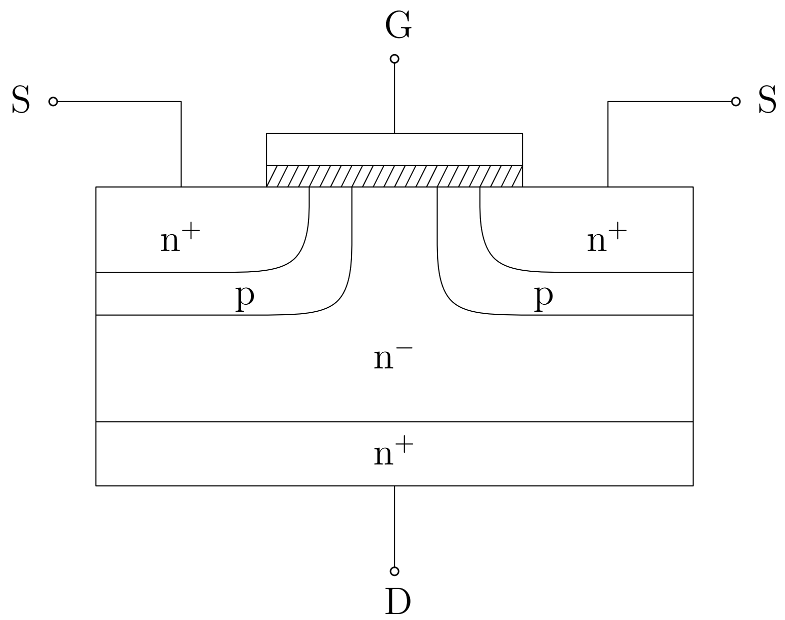


Abb. 10.8 Aufbau eines vertikalen n-Kanal-Leistungs-MOSFET.

Der Drainstrom fließt bei positiver Gatespannung vertikal durch den Transistor.

Die Gatestruktur selbst ist im gezeigten Beispiel planar und befindet sich an der Chipoberfläche. Es gibt aber auch Leistungs-MOSFETs mit einer vertikalen Gattestruktur.

Die vertikale Struktur hat zwei Vorteile:

- Auf der Chipoberfläche wird weniger Platz benötigt.
- Man kann relativ große Driftregionen (n^--Regionen) realisieren, so dass große Spannungen ermöglicht werden.

Oft werden viele benachtbarten Leistungs-MOSFETs parallel geschaltet und dadurch besonders große Ströme ermöglicht.

Dies macht sich den positiven Temperaturkoeffizienten von MOSFETs zunutze. Die Erhöhung der Temperatur führt zu einer Abnahme des Stromes. Dadurch wird der Erwärmung entgegengewirkt und „thermal run-away“ verhindert.

10.3 Der IGBT

Der „Insulated Gate Bipolartransistor“ (IGBT) ist wie der Thyristor eines Halbleiters mit (im wesentlichen) vier dotierten Halbleiterschichten, wie z.B. ($n^+pn^-p^+$) für die Leistungselektronik. Im Gegensatz zu einem klassischen Thyristor lässt sich der Strom durch einen IGBT durch die Gatespannung nicht nur anschalten, sondern auch wieder ausschalten.

IGBTs können Sperrspannungen von mehreren kV und Durchlassströme von mehreren kA aushalten! Die maximalen Schaltfrequenzen liegen im Bereich von 10-100 kHz. Dabei ist die minimale Anschaltzeit deutlich höher als die minimale Abschaltzeit.

Der IGBT enthält Elemente eines MOSFET und eines Bipolatransistors. Entsprechend bezeichnet man seine drei Pole als Kollektor, Emitter und Gate. Der schematische Aufbau eines n-Kanal-IGBTs ist in Abb. 10.9 dargestellt. Die Regionen 1, 2/3 und 4 entsprechen einem pnp-Bipolartransistor. Der Transistor hat eine vertikale Struktur. Der Emitter ist hier auf der Oberseite und der Kollektor auf der Unterseite, dem Substrat, angeschlossen. Die Regionen 3, 4 und 5 bilden einen NMOS-Transistor, der durch das Gate gesteuert wird. Dabei entsprechen die Regionen 3 und 5 Drain und Source des NMOS.

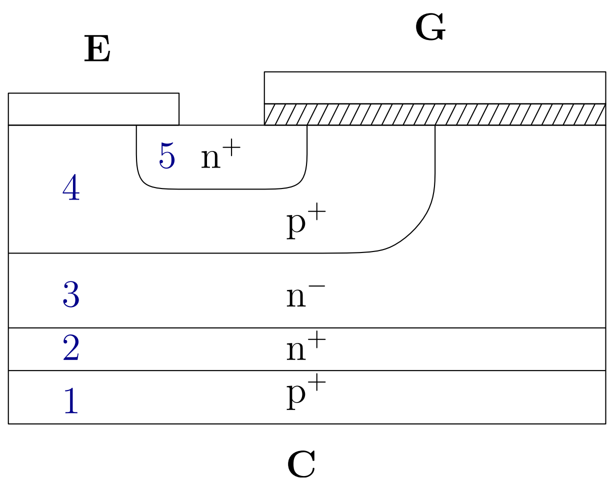


Abb. 10.9 Aufbau eines vertikalen n-Kanal-IGBTs.

10.3.1 Wie funktioniert der IGBT ?

Im Vorwärtsbetrieb ist $U_{\mathrm{CE}} > 0\,\mathrm{V}$. Dadurch ist der pn-Übergang 1, 2/3 in Durchlassrichtung gepolt. Da allerdings der np-Übergang 3, 4 in Sperrichtung gepolt ist, fließt für $U_{\mathrm{G}} \approx 0\,\mathrm{V}$ kein Kollektorstrom. Legt man eine positive Gatespannung von einigen Volt an, so bildet sich unter dem Gate zw. Zone 5, 4 und 3 ein leitender Kanal. Der IGBT leitet.

Das Kennliniendiagramm eines IGBT ist in Abb. 10.10 dargestellt.

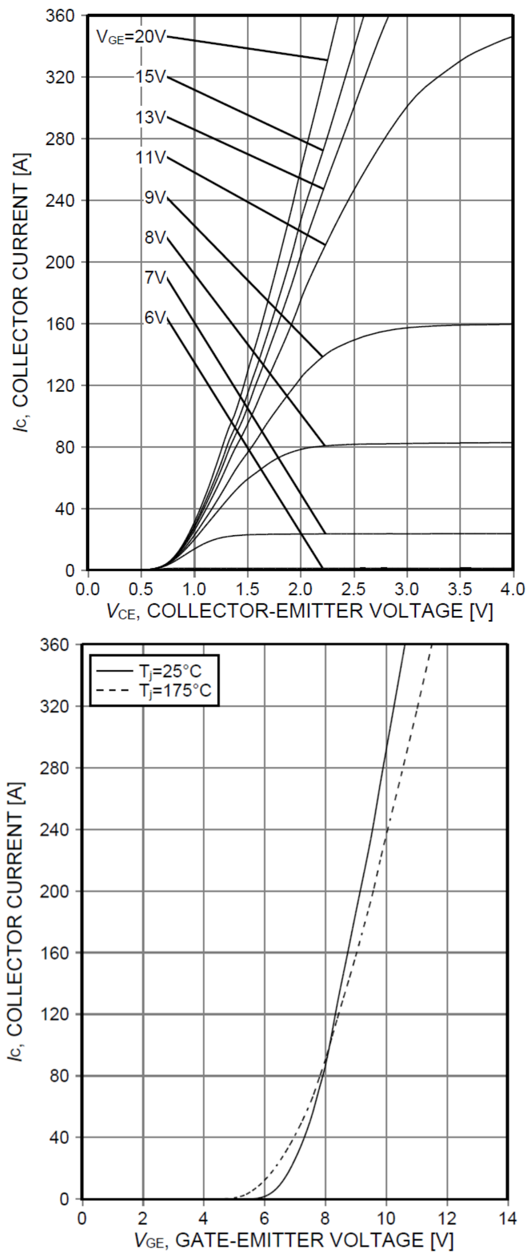


Abb. 10.10 Auszug aus dem Datenblatt eines IGBT (IKQ120N60TA)[11].

Man erkennt die großen Stöme und Spannungen und (relativ) geringe Minimalspannung U_{CE} im leitenden Zustand.

Das Ersatzschaltbild eines IGBTs ist in Abb. 10.11 dargestellt.

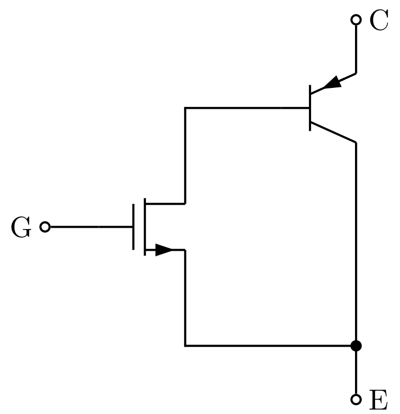


Abb. 10.11 Ersatzschaltbild eines IGBT.

Schaut man den Aufbau eines IGBT genauer an (Abb. 10.12) erkennt man noch eine weitere parasitäre npn-Struktur. Der npn- und pnp-Transistor können im Prinzip einen Thyristor im Kurzschlussbetrieb bilden (Latch-Up). Dies muss man bei Design bzw. Betreiben eines IGBT beachten.

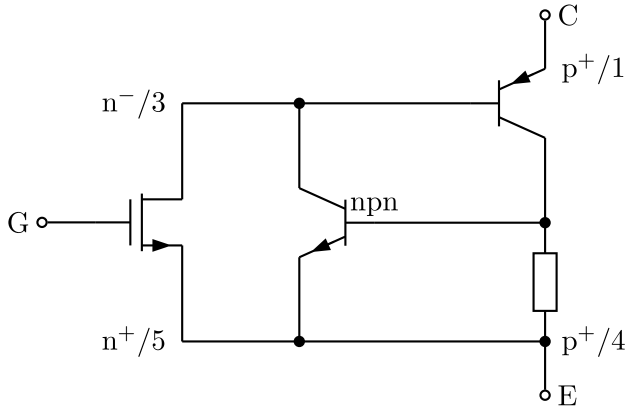


Abb. 10.12 Ersatzschaltbild eines IGBT mit parasitären npn-Transistor.

11 Rauschen und Techniken zur Rauschunterdrückung*

Dieses Kapitel ist noch in Arbeit.

Die Messgenauigkeit eines Detektors oder Verstärkers ist durch elektronisches Rauschen und externe Störungen begrenzt. Es ist äußerst wichtig, diese Effekte zu verstehen und zu minimieren. Das Eigenrauschen eines Vorverstärkers ist beim Nachweis kleiner Signale oft die entscheidende Eigenschaft, wichtiger als seine Linearität und Bandbreite. Beim Entwurf eines Detektorsystems spielt das Rauschen oder das Signal-zu-Rauschverhältnis[1] eine große Rolle. Oft muss ein Kompromiss zwischen minimalem Rauschen, der Leistung, Bandbreite und der Zahl der elektronischen Kanäle gefunden werden. Auch die Wahl der Transistortechnologie ist oft durch ihr Rauschverhalten bestimmt.

Besonders ärgerlich, aber leider keineswegs selten, ist es, wenn in einem Detektorsystem mit vielen Modulen die intrinsisch mögliche Auflösung der einzelnen Module nicht erreicht wird. Ähnliches gilt auch für komplexe Leiterplatten mit Analog- und Digitalfunktionen[2] und vielen Bausteinen. Dies kann viele Ursachen haben. Meistens liegen nicht externe Störeinflüsse, sondern hausgemachte Schwierigkeiten zugrunde. Hier ist es ratsam grundlegende Techniken wie die Abschirmung, Erdung und Filtern, zu kennen und anzuwenden.

Wir wollen zuerst das Eigenrauschen elektronischer Bauteile verstehen.

11.1 Rauscharten

Betrachten wir mit beliebig hoher Auflösung den Strom durch ein Bauteil, z. B. durch einen Widerstand, eine Diode oder einen Transistor, ergibt sich folgendes exemplarisches Bild (Abb. 11.1).

[1]engl. *signal-to-noise ratio: S/N oder SNR*
[2]engl. *mixed-signal*

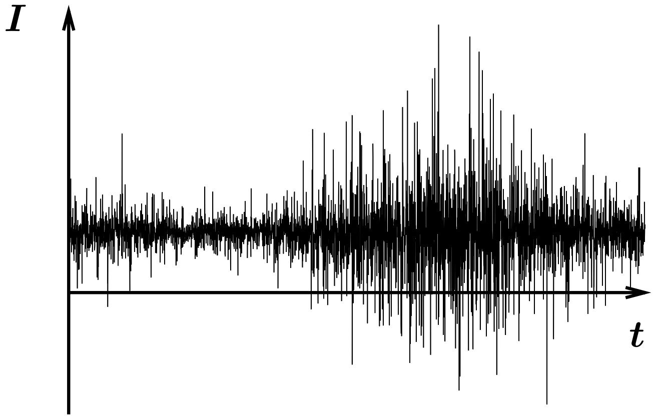


Abb. 11.1 Illustration des elektrischen Rauschens. Strom als Funktion der Zeit in willkürlichen Einheiten.

Der Strom schwankt statistisch um einen Mittelwert I_0. Dies können folgende Rauscharten bewirken:

1. Thermisches Rauschen
2. Schrotrauschen[3]
3. $\frac{1}{f}$-Rauschen.

Allen Rauscharten ist gemeinsam, dass ihr zeitliches Mittel über längere Zeiten verschwindet

$$\overline{I}_R = \lim_{T\to\infty} \frac{1}{2\,T} \int_{-T}^{T} I_R(t)\,\mathrm{d}t = 0 \quad .$$

Es geht also um den Effektivwert

$$I^2_{R\,RMS} = \overline{I}^2_R = \lim_{T\to\infty} \frac{1}{2\,T} \int_{-T}^{T} I^2_R(t)\,\mathrm{d}t$$

und das Frequenzverhalten des Rauschens. Wir definieren die spektrale Rauschstromdichte $S_i(f)$ mit

$$I^2_{R\,RMS}(f) = \int_{f_1}^{f_2} S^2_i(f)\,\mathrm{d}f \quad .$$

Der Index i steht für den Strom, S für spektrale Dichte. $S_i(f)$ hat die Einheit $\frac{\mathrm{A}}{\sqrt{Hz}}$. Entsprechend definiert man die spektrale Leistungsdichte S_p und die spektrale Spannungsdichte S_u mit den Einheiten $\frac{\mathrm{W}}{\sqrt{\mathrm{Hz}}}$ und $\frac{\mathrm{V}}{\sqrt{\mathrm{Hz}}}$.

[3]engl. *shot noise*

Vorsicht: Hier gehen die Bezeichnungen in der deutschen und der angelsächsischen Literatur und zwischen unterschiedlichen Autoren auseinander. Oft wird statt $S_i(f)$ die Größe $\mathrm{dI_R}/\sqrt{\mathrm{d}f} = I_R$ verwendet, wobei man leicht die spektrale Stromdichte und mit dem Strom verwechselt. Reisch definiert die Größe $S_i(f)$ als die quadratische Rauschstromdichte (aus Sicht der obigen Definition).

Zum Verständnis des thermischen Rauschens und des Schrotrauschens betrachten wir einen Strom zwischen zwei Elektroden im Abstand L mit

$$I = \frac{\mathrm{d}Q}{\mathrm{d}t} = \frac{n\,q_e\,v}{L} \quad . \tag{11.1}$$

Hier ist n die Zahl der Ladungsträger, q_e die Elementarladung in Coulomb und v die mittlere Geschwindigkeit der Elektronen. Schwankungen des Stroms können durch statistische Fluktuationen der Zahl der Ladungsträger oder durch Geschwindigkeitsschwankungen hervorgerufen werden.

11.2 Thermisches Rauschen

Thermisches Rauschen oder Johnsonrauschen tritt in Widerständen auf. Es wird durch Geschwindigkeitsschwankungen der Elektronen, die aufgrund ihrer statistischen Wärmebewegung (Brownsche Bewegung) auftreten, verursacht. Die Rauschstromdichte des thermischen Rauschens in einem Widerstand R der Temperatur T ist

$$S_i(f) = \sqrt{\frac{4\,k\,T}{R}} \quad . \tag{11.2}$$

Die Rauschspannungsdichte ist

$$S_u(f) = \sqrt{4\,k\,T\,R} \quad . \tag{11.3}$$

In beiden Ausdrücken tritt die Frequenz f nicht explizit auf. Thermisches Rauschen ist sogenanntes „weißes Rauschen“. Dies gilt für alle praktisch relevanten Frequenzbereiche und Temperaturen, genaugenommen für $h\,f \ll k\,T$[4].

Damit ergibt sich für das Gesamtrauschen in einem Frequenzbereich $B = f_2 - f_1$ der Wert

$$I^2_{R\,RMS} = \int_{f_1}^{f_2} S_i^2(f)\,\mathrm{d}f = \frac{4\,k\,T\,B}{R} \quad . \tag{11.4}$$

Die Wahrscheinlichkeit W, das bei gegebenen mittleren Rauschen $I^2_{R\,RMS}$ zu einem willkürlichen Zeitpunkt eine Rauschstromamplitude I_R auftritt, entspricht einer Gaußverteilung

$$W(I_R) = \frac{\mathrm{e}^{-\frac{I_R^2}{2\,I^2_{R\,RMS}}}}{\sqrt{2\,\pi\,I^2_{R\,RMS}}} \quad . \tag{11.5}$$

[4] Für beliebig hohe Frequenzen gilt $S_p(f) = \frac{h\,f}{\mathrm{e}^{\frac{h\,f}{k\,T}} - 1}$, entsprechend der Plankschen Formel für Schwarzkörperstrahlung.

Es können mit abnehmender Wahrscheinlichkeit also beliebig große Rauschströme auftreten.

Durch das Einfügen von Spannungs- oder Stromrauschquellen in den Schaltkreis kann man die Wirkung einer Rauschquelle in einer Schaltung berücksichtigen. Man schaltet zu einem Widerstand R eine entsprechende Spannungsquelle

$$U_R = \sqrt{4\,k\,T\,R} \quad . \tag{11.6}$$

in Serie. Manchmal ist es praktischer, aber mathematisch äquivalent, eine Stromquelle

$$I_R = \sqrt{\frac{4\,k\,T}{R}} \tag{11.7}$$

parallel zu schalten (Abb. 11.2).

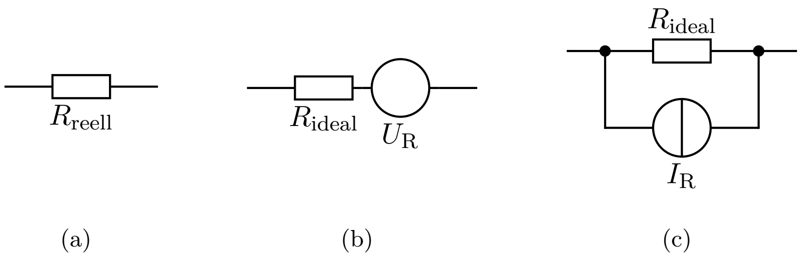


Abb. 11.2 Ersatzschaltung eines rauschenden Widerstands (a) durch einen rauschlosen Widerstand und entsprechender Spannungs- oder Stromrauschquelle (b) und (c).

Bemerkung: In Gleichung 11.6 und 11.7 ist die Integration über die Frequenz noch nicht durchgeführt.

Beispiel: Wie groß ist die mittlere Rauschspannung eines Widerstands von $1\,\mathrm{M\Omega}$ bei Raumtemperatur?

Mit $k = 1{,}38 \cdot 10^{-23}\,\frac{\mathrm{J}}{\mathrm{K}}$, $R = 1\,\mathrm{M\Omega}$, $T = 300\,\mathrm{K}$ und einer hypothetischen Bandbreite $B = 10\,\mathrm{MHz}$ ergibt sich $U_R = 0{,}4\,\mathrm{mV}$[5]. Ist dies nun groß oder klein? Das hängt von den Details der Schaltung und bei einem Vorverstärker auch von der Größe des Eingangsignals ab.

11.3 Schrotrauschen

Schrotrauschen tritt in Röhren, Halbleiterdioden und Transistoren auf und beruht auf der statistischen Fluktuation der Zahl der Ladungsträger (Elektronen oder Löcher), z. B. beim Durchqueren einer Potentialschwelle oder bei der Erzeugung oder Rekombination von Ladungsträgern.

Die Rauschstromdichte ist

$$S_i(f) = \sqrt{2\,q_e\,I} \quad , \tag{11.8}$$

wobei I der Durchschnittsstrom in Ampere ist. I kann der Sperrstrom oder der Vorwärtsstrom einer Diode oder der Basisstrom eines bipolaren Transistors sein. Auch das Schrotrauschen

[5] *Nebenrechnung:* $U_R = \sqrt{4 \cdot 1{,}38 \cdot 3 \cdot 10^{-23+2+6+7}}\,\mathrm{V} \approx \sqrt{5{,}6 \cdot 3 \cdot 10^{-8}}\,\mathrm{V} \approx 4 \cdot 10^{-4}\,\mathrm{V} = 0{,}4\,\mathrm{mV}$.

ist frequenzunabhängig, also „weiß“.

Wie groß ist das Schrotrauschen?

Mit $q_e = 1{,}6 \cdot 10^{-19}\,\mathrm{C}$ gilt, $\sqrt{2\,q_e\,I} = 5{,}7 \cdot 10^{-10}\,\sqrt{I[\mathrm{A}]}$. Dieser Wert wirkt auf den ersten Blick sehr klein. Lässt man aber diesen Rauschstrom durch einen Widerstand von $1\,\mathrm{M}\Omega$ fließen und berücksichtigt man einen Frequenzbereich von z.B. 1 MHz, so gilt $U_{RMS} = 1\,\mathrm{M}\Omega \cdot \sqrt{10^6}\,\mathrm{Hz} \cdot 5{,}7 \cdot 10^{-10}\,\sqrt{I[\mathrm{A}]}\mathrm{V}$. Für einen Strom von $1\,\mu\mathrm{A}$ wäre $U_{RMS} \approx 6\,\mathrm{mV}$. Hier sieht man, dass es sich lohnt, eine Rauschanalyse einer Schaltung durchzuführen.

11.4 $\frac{1}{f}$-Rauschen

Das $\frac{1}{f}$-Rauschen tritt in vielen Bauelementen auf und hat unterschiedliche Ursachen, so z. B. Fluktuationen in der Beweglichkeit der Ladungsträger, Erzeugung und Rekombination von Ladungsträgern mit unterschiedlichen, aber festen, Zeitkonstanten, Einfangen[6] von Ladungen in der Oxidschicht von Feldeffekttransistoren. Die Rauschstromdichte ist

$$S_i(f) = \sqrt{\frac{A}{f}} \quad . \tag{11.9}$$

Der Parameter A ist abhängig vom Bauteil und ist auch eine Funktion des Stromes.

Im Frequenzbereich zwischen f_1 und f_2 trägt das $\frac{1}{f}$-Rauschen

$$I_{RMS}^2 = \int_{f_1}^{f_2} \frac{A}{f}\,\mathrm{d}f = A\,\log\frac{f_2}{f_1} \quad . \tag{11.10}$$

Wichtig ist hier, dass relativ kleine Frequenzbereiche, für die das thermische Rauschen und Schrotrauschen vernachlässigbar ist, über das $\frac{1}{f}$-Rauschen schon einen substantiellen Beitrag zum Gesamtrauschen liefern können.

11.5 Rauschquellen und Verstärker

Welchen Einfluss haben Rauschquellen am Eingang eines Verstärkers oder im Verstärker selbst auf das Ausgangssignal. Dazu betrachten wir Abbildung 11.3.

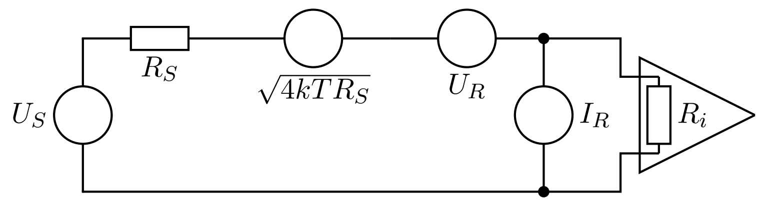


Abb. 11.3 Vorverstärker mit Eingangssignalquelle U_S, Innenwiderstand R_S, Stromrauschquelle I_R und Spannungsrauschquelle U_R.

[6] engl. *Trapping*

U_R und I_R können entweder externe oder interne Rauschquellen des OPV sein.

Die Rauschspannung am Eingang des OPV ergibt sich dann zu

$$S^2_{u_{\mathrm{ein}}}(f) = S^2_s(f) + S^2_u(f) + S^2_i(f) \cdot R^2_S = 4\,k\,T\,R_S + U^2_R + (\bar{I}_R \cdot R_S)^2 \quad .$$

Da die drei Rauschterme statistisch unabhängig voneinander sind, muss man sie quadratisch addieren. Man kann daraus durch Multiplikation des Eingangsrauschens mit der Verstärkung V_u des OPV das Rauschen am Ausgang berechnen

$$S^2_{u_{\mathrm{aus}}}(f) = V^2_u(f)\, S^2_{u_{\mathrm{ein}}}(f) \quad .$$

Aussagekräftiger als der Absolutwert des Rauschens ist oft das Verhältnis der Signalamplitude U_S zum Rauschen $S_{u_{\mathrm{ein}}}$, also

$$\frac{V^2_u(f)\, U^2_S}{V^2_u(f)\ \left[S^2_s(f) + S^2_u(f) + S^2_i(f) \cdot R^2_S\right]} \quad .$$

Diese Größe, integriert über alle Frequenzen, bezeichnet man auch als das Signal-zu-Rauschverhältnis S/N[7]. Es ist ist unabhängig von der Verstärkung V_u.

$$\left(\frac{S}{N}\right)^2 = \frac{\int\limits_0^\infty U^2_S(f)\, |A_u|^2(f)\, \mathrm{d}f}{\int\limits_0^\infty S^2_{u_{\mathrm{ein}}}(f)\, |A_u|^2(f)\, \mathrm{d}f} \tag{11.11}$$

Hier tritt ein neuer Effekt auf, da der Frequenzgang der Übertragungsfunktion des Verstärkers eingeht. Falls die Übertragungsfunktion des Verstärkers nicht an das Signalspektrum angepasst ist und die Bandbreite zu groß ist, ist das S/N unnötig klein. Dies ist in Abbildung 11.4 dargestellt.

[7] In dieser Herleitung haben wir den Wert von R_i nicht berücksichtigt. Das Verhältnis $\frac{S}{N}$ ist aber unabhängig von der Eingangimpedanz des Verstärkers.

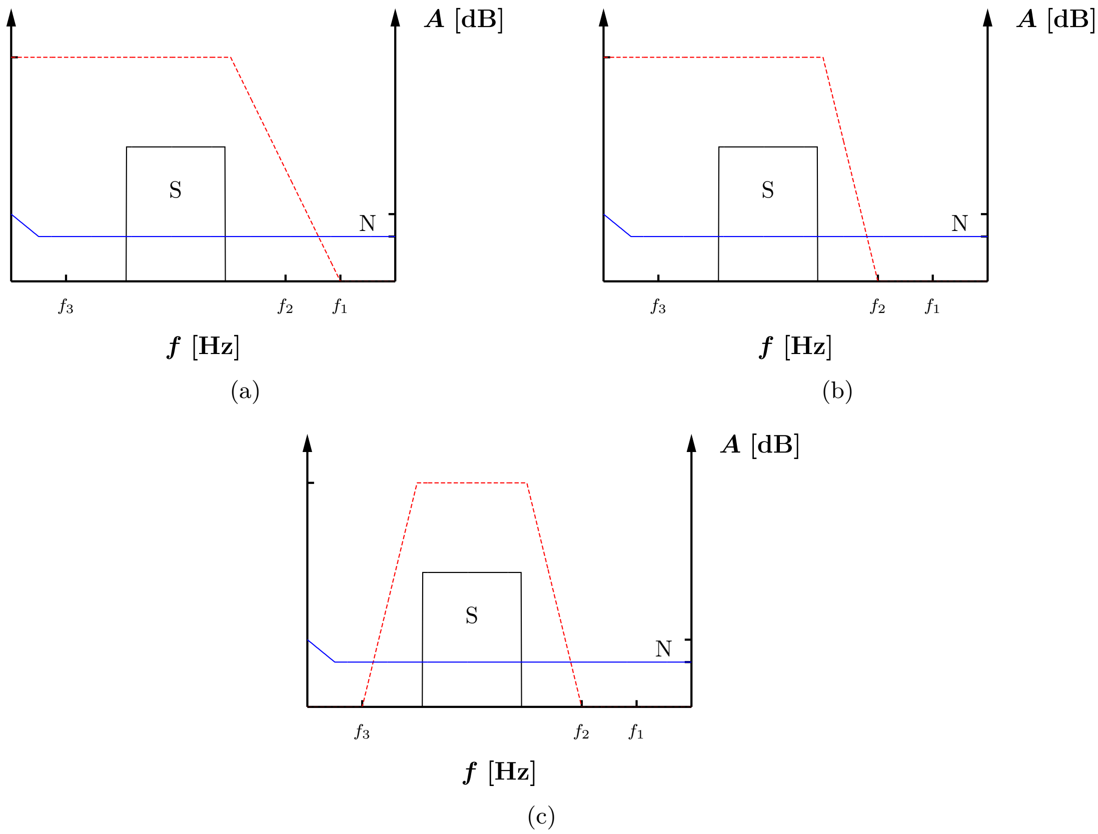


Abb. 11.4 Rausch- und Signalanteile N und S, die den Verstärker durchlaufen, für drei Übertragungsfunktionen (gestrichelt): (a) großer Verstärkerbereich bis zur Frequenz f_1, (b) reduzierter Bereich bis zur Frequenz $f_2 < f_1$ und (c) Bandpass im Bereich f_3 bis f_1.

Das Signal-Rausch-Verhältnis S/N ist ungünstig für das Szenario in Abbildung 11.4(a). Der Verstärker lässt in dem Bereich hoher Frequenzen, bei dem keine Signalkomponenten auftreten, nur Rauschanteile durch. S/N wird zu klein.

Das Szenario in Abbildung 11.4(b) ist viel vorteilhafter. Hier ist es natürlich wichtig, dass das Signal bekannt ist, so dass man durch zu frühes Abschneiden von hohen Frequenzanteilen die Signalform nicht verfälscht bzw. Signal verliert.

Wenn man es besonders gut machen will, kann man sowohl die hohen und tiefen Frequenzkomponenten, die nicht relevant sind, abschneiden wie in Szenario 11.4(c).

11.6 Äquivalente Rauschbandbreite

Die Rauschbandbreite Δf_R, die für „weiße“ Rauschquellen eine hilfreiche physikalische Größe ist, ist über folgende Beziehung definiert:

$$U_R^2 = \int_0^\infty A_u^2(f)\, S_{u^2_{\text{ein}}}\, \mathrm{d}f = A_0^2\, S^2_{u_{\text{ein}}} \int_0^\infty G(f)\, \mathrm{d}f = A_0^2\, S^2_{u_{\text{ein}}}\, \Delta f_R \quad ,$$

wobei $A_u(f) = A_0\,G(f)$ gilt und S_{u_ein}frequenzunabhängig ist, so dass es vor dem Integral gezogen werden kann.

Für einen Verstärker mit einem Pol

$$A_u(f) = \frac{A_0}{1 + \mathrm{j}\,\frac{f}{f_G}}$$

ergibt sich

$$\int\limits_0^\infty \frac{1}{1 + \mathrm{j}\,\frac{f^2}{f_G^2}}\,\mathrm{d}f = \left[\frac{\arctan\left(\frac{f}{f_G}\right)}{\frac{1}{f_G}}\right]_0^\infty = \frac{\pi}{2}\,f_G \quad .$$

Die äquivalente Rauschbandbreite Δf_R ist also um $\frac{\pi}{2}$ kleiner als die Bandbreite f_G des Signals. Bei Schaltungen mit mehreren identischen Polen und damit steilerem Abfall der Übertragungsfunktion verringert sich der Unterschied zwischen Δf_R und f_G.

11.7 Equivalent Noise Charge

Eine vielbenutzte Größe für ladungsempfindliche Vorverstärker ist die *equivalent noise charge* (ENC). Hier drückt man das Signal und das Rauschen als Ladung bzw. als Zahl der Ladungsträger aus. Die ENC berechnet sich aus $\frac{S}{N}$ und der Signalladung Q_S zu

$$\mathrm{ENC} = \frac{Q_S}{S/N} \quad .$$

Beispiel aus der Detektorinstrumentierung: Ein geladenes, relativistisches Teilchen deponiert beim Durchqueren eines $300\,\mu$m dicken Siliziumdetektors rund 24 000 Elektron-Loch-Paare im Detektor. In der Regel werden nur die Löcher oder die Elektronen ausgelesen:

$$Q_n \approx 4\,f\,C \quad .$$

Mit einem $\frac{S}{N} = 10$ ergibt sich die ENC zu $2\,400\,e$. Meistens ist das $\frac{S}{N}$ aber deutlich höher und ENC entsprechend kleiner.

S/N hängt wie wir schon gesehen haben von der Bandbreite des Verstärkers ab.

11.8 Systeme mit mehreren Verstärkerstufen

Wie beeinflusst die Zahl der Verstärkerstufen das Rauschen des Gesamtsystems?

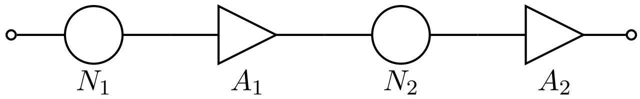


Abb. 11.5 Rauschen in einem System mit zwei Verstärkerstufen.

Wir betrachten Abbildung 11.5. Am Eingang der ersten Verstärkungsstufe mit Verstärkung A_1 liegt eine Rauschquelle mit Rauschen N_1 an. Dieses Rauschen mag aus äußeren Quellen stammen oder nur vom Verstärker selbst verursacht worden sein. Genauso liegt am Eingang des folgenden Verstärkers mit der Verstärkung A_2 eine Rauschquelle mit Rauschen N_2 an. N_2 ist allerdings rein durch den Verstärker bedingt. Das Signal des Gesamtsystems ist $S\,A_1\,A_2$. Das Rauschen des Gesamtsystem hat zwei Anteile:

- $N_1\,A_1\,A_2$, das Rauschen N_1 verstärkt in beiden Verstärkerstufen
- $N_2\,A_2$, das Rauschen N_2 der zweiten Stufe.

Diese Anteile müssen quadratisch addiert werden.

Zusammenfassend ergibt sich das Signal-Rausch-Verhältnis

$$\left(\frac{S}{N}\right)^2 = \frac{(S\,A_1\,A_2)^2}{(N_1\,A_1\,A_2)^2 + (N_2\,A_2)^2} = \frac{S^2}{N_1^2 + \left(\frac{N_2}{A_1}\right)^2} = \left(\frac{S}{N_1}\right)^2 \frac{1}{1 + \left(\frac{N_2}{A_1\,N_1}\right)^2} \quad .$$

Dies bedeutet, dass das Signal-Rausch-Verhältnis im Wesentlichen durch die Eingangsstufe bestimmt wird, falls die Verstärkung A_1 diese Stufe ausreichend groß ist. Die Aneinanderreihung mehrerer Verstärkerstufen ist also unproblematisch.

11.9 KTC-Rauschen

Mit KTC-Rauschen bezeichnet man die Rauschspannung an einem parallel geschalteten Kondensator C und Widerstand R (Abb. 11.6).

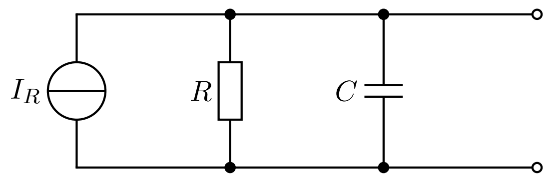


Abb. 11.6 Parallelschaltung aus Kondensator C und Widerstand R.

Der Widerstand R verursacht einen Rauschstrom I_n mit $S_i(f) = \sqrt{\frac{4\,k\,T}{R}}$. Dies führt in der Impedanz

$$Z = R\,\|\,Z_C = \frac{R}{\sqrt{1 + \omega^2\,R^2\,C^2}}$$

zu einer Rauschspannung

$$\overline{U_n^2} = \int_0^\infty Z^2\,S_i^2(f)\,\mathrm{d}f = 4\,k\,T\,R \int_0^\infty \frac{1}{1 + (2\,\pi\,f\,R\,C)^2} = \frac{k\,T}{C} \quad .$$

Die Rauschspannung hängt also vielleicht überraschenderweise nicht mehr von R ab. Dies liegt daran, dass sich die jeweilige Abhängigkeit der Stromrauschdichte und der „Bandbreite“ von R aufhebt.

Anmerkung: Die Anwendung aus Abbildung 11.6 entspricht der eines typischen ladungsempfindlichen Detektors mit der Detektorkapazität C und dem spannungsempfindlichen Widerstand R. Der Rauschstrom entspricht dem Rauschen des Detektors. Für einen Siliziumdetektor, der als Diode in Sperrrichtung geschaltet ist, ist der Leckstrom die Rauschquelle.
Die ENC ergibt sich aus der Rauschspannung aus $Q = CU$. Die ENC aufgrund von KTC-Rauschen beträgt damit kTC, daher auch der Name.

11.10 Transistorrauschen

Die oben eingeführten grundlegenden Rauschquellen treten auch in aktiven Bausteinen wie BJTs, MOSFETs und JFETs auf.

BJT:

1. Schrotrauschen in der Basis-Emitter-Diode $S_i(f) = \sqrt{2\,q_e\,I_B}$
2. Schrotrauschen im Kollektor $S_i(f) = \sqrt{2\,q_e\,I_C}$
3. Das thermische Rauschen des Basisbahnwiderstads mit $S_i(f) = \sqrt{\frac{4\,k\,T}{R_B}}$

MOSFET:

1. Thermisches Rauschen im Kanal dominiert.
2. Zusätzlich tritt 1/f-Rauschen auf.

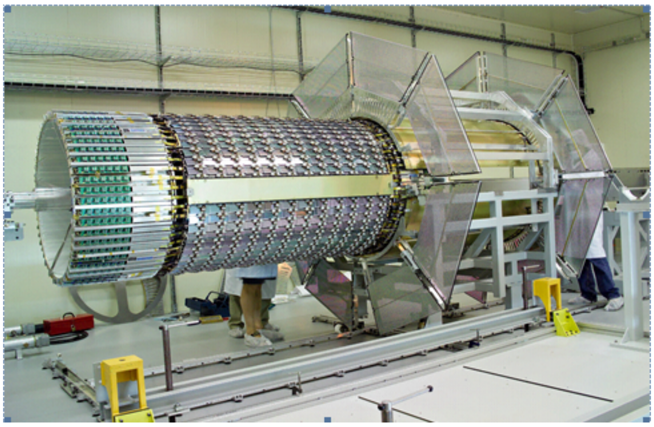
JFET:

1. Thermisches Rauschen im Kanal
2. 1/f-Rauschen, allerdings deutlich kleiner als in MOSFETs
3. Schrotrauschen in der Gate-Substrat-Diode

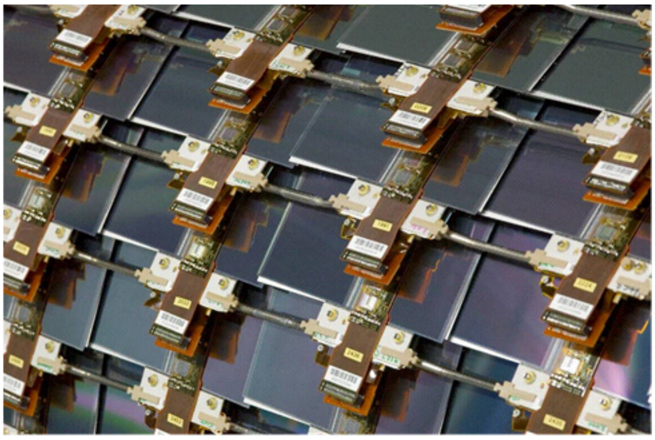
11.11 Rauschen in Systemen

In einem komplexen System ist es oft schwierig, das intrinische Rauschniveau eines einzelnen Moduls oder Kanals zu erreichen.

Zur Illustration betrachten wir Abbildung 11.7. Hier ist der Semiconductor Tracker (SCT) des ATLAS-Experiments am Large Hadron Collider in Genf gezeigt. Dies ist ein Teildetektor mit 4 000 Detektormodulen und 6 Millionen elektronischen Auslesekanälen. Jedes Modul benötigt drei unterschiedliche Versorgungsspannungen, diverse Steuersignalleitungen, Datenleitungen, etc.



(a)



(b)

Abb. 11.7 Die äußere Lage des SCT [12] des ATLAS-Experiments: (a). Detailansicht mit Siliziumsensoren, Auslesechips, Multichipmodulen und Kühlrohren (b).

Der Detektor ist von zahlreichen weiteren Detektoren von vergleichbarer Komplexität umgeben. Hier gibt es vielfältige Möglichkeiten der Wechselwirkungen zwischen den Modulen untereinander oder mit den anderen Detektoren, die zu erhöhtem Rauschen oder sogar Störsignalen führen.

Ein ganz anders geartetes elektronisches System ist in Abbildung 11.8 dargestellt. Die beidseitig beschichtete Platine ist Hauptbestandteil eines komplexen Datennahmesystems zur Verstärkung von Radiosignalen. Das Board trägt zahlreiche analoge und digitale Komponenten wie mit Analogverstärker, ADC, DRAM, Memory, GPS-Uhr, FPGA zur digitalen Datenverarbeitung und Uhr, CPU, LVDS, Tranceiver für I/O. Auch hier gibt es zahlreiche Möglichkeiten zur unerwünschten Wechselwirkung von Schaltungskomponenten und damit erhöhtes Rauschen.

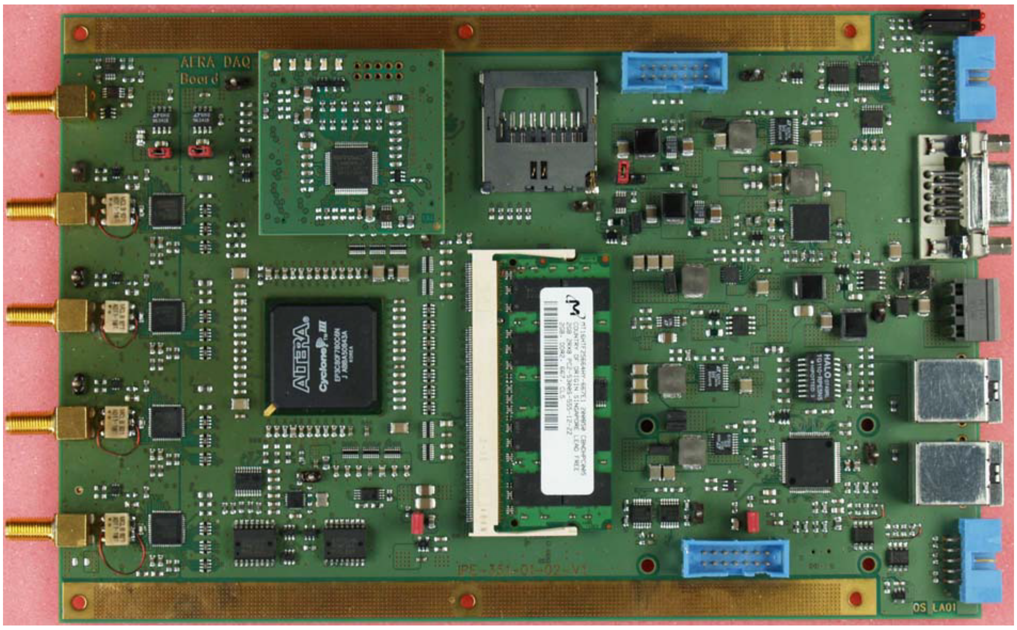


Abb. 11.8 Ein komplexes Datanahmesystem zur Verstärkung von Radiosignalen aus der kosmischen Strahlung.

Ein letztes Beispiel für ein empfindliches elektronisches System ist in Abbildung 11.9 zu sehen. Dies ist ein komplexer „mixed-signal“ ASIC. Der integrierte Schaltkreis dient der Auslese eines zukünftigen Siliziumdetektors à la Abbildung 11.7.

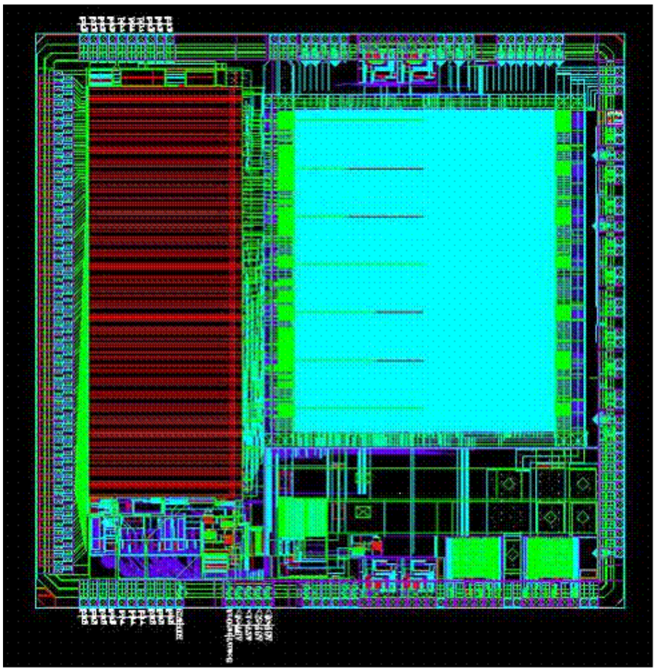


Abb. 11.9 Layout des ABCN-Chips.

Zwar ist der Chip nur rund $6 \times 7\,\mathrm{mm}^2$ groß, er besteht aber aus einer Vielzahl unterschiedlicher

Funktionsblöcke, wie rauscharme Vorverstärker für 128 Kanäle, Datenpipeline, Konfigurationsspeicher, Spannungsregler, Filter, Diskriminatoren, Kontrollogik, Datenleitungsgtreiber, DACs u. v. m.

Trotz aller äußerlichen Unterschiede zwischen einem Großdetektor, eine Leiterplatte und einem ASIC gelten die grundlegenden Prinzipien zur optimalen Systemauslegung für alle drei Beispiele. Man kann Rauscheffekte an drei Stellen bekämpfen (Abb. 11.10):

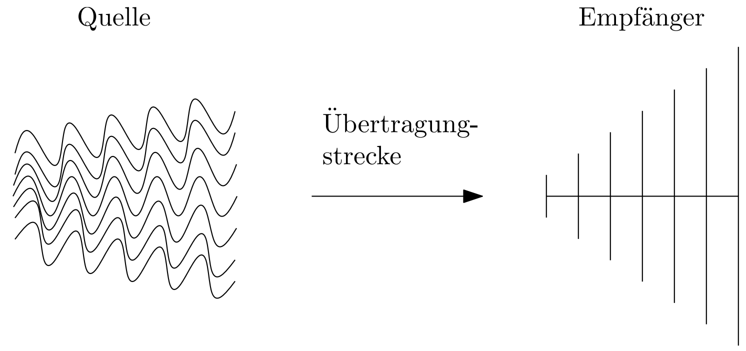


Abb. 11.10 Rauschmodell.

1. durch Elimination der Rauschquelle selbst,
2. durch Unterbrechung der Übertragungsstrecke oder
3. durch Verringern der Empfindlichkeit des Empfängers auf das Rauschsignal.

Welche Übertragungs- oder Einkopplungsmöglichkeiten der Rauschquelle gibt es?

Wir betrachten die folgenden Wege:

- Galvanisch (über die Zuleitung)
- Kapazitiv
- Induktiv
- Elektromagnetische Strahlung

Eine wichtige Regel ist also, das Signal vor der Verstärkung sorgfältig vor Störeinflüssen zu schützen.

Beispiel für den Signalpfad

Ein geladenes Teilchen oder ein Röntgenquant erzeugt in der Diode Elektron-Loch-Paare. In der in Abb. 11.11(a) gezeigten Konfiguration wird der Löcherstrom, kapazitiv von dem p-dotierten Streifen in darüberliegenden den Aluminiumstreifen gekoppelt. Der Aluminiumstreifen ist direkt mit dem Vorverstärker verbunden. Der Stromrückfluss erfolgt über die Masse des Vorverstärkerchips, den Hybrid und Hochspannungskoppelkondensator und den n-dotierten Substratanschluss des Sensors. In Abbildung 11.11(b) ist diese Konfiguration als Schaltkreis dargestellt.

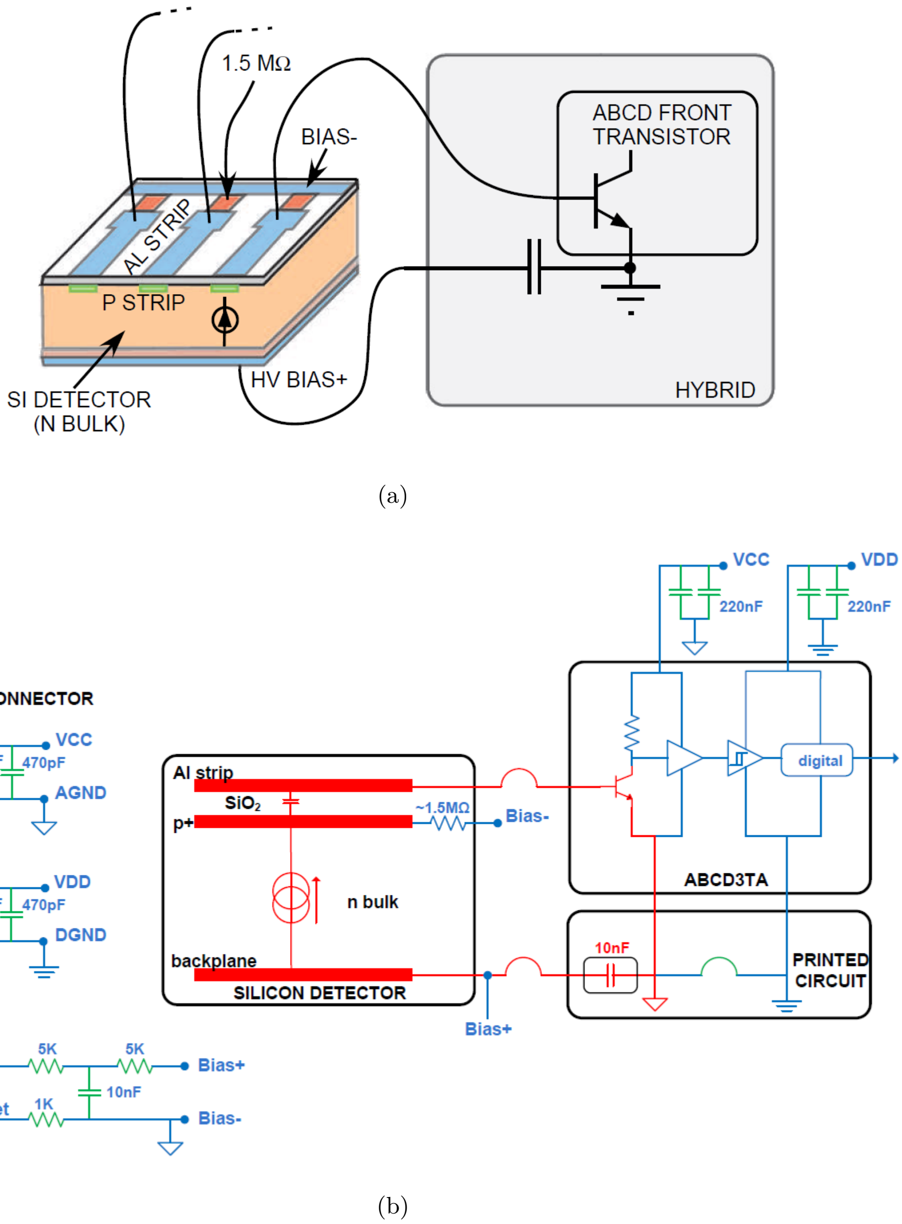


Abb. 11.11 (a) Signalpfad in einem Streifendetektor. (b) Schaltkreis zum Signalpfad.

Wie schützt man den Signalpfad? Dies wird anhand der folgenden Kapitel erklärt.

11.12 Kopplung über die Stromversorgung (conductive noise)

Der einfachste Weg, viel Rauschen einzufangen, verläuft über die Spannungsversorgung. Eine rauschende Spannungsquelle leitet die unsaubere Spannung durch die Stromkabel, die eventuell sehr lang sein können, direkt an den Verbraucher. Der Verbraucher kann ein empfindlicher Vorverstärker in einem Auslesechip sein. Abhängig vom Verstärkerdesign wirkt sich die unsaubere Spannungsversorgung des Chips als ein erhöhtes Rauschen am Ausgang des Verstärkers aus.

Hier kann man wie im vorherigen Abschnitt erläutert die Spannungsquelle austauschen, das Rauschsignal auf den Versorgungsleitungen unterdrücken (z. B. durch einen Tiefpass) oder den Verstärker so entwerfen, dass er unempfindlich gegenüber Schwankungen der Versorgungsspannung im relevanten Frequenzbereich ist. Die ersten beiden Methoden sind in der Praxis relevanter.

Ein Spezialfall der rauschenden Einkopplung über Zuleitungen ist das Einkoppeln von Rauschen aufgrund gemeinsamer Zuleitungen mehrerer Verbraucher. Es ist oft bequem, die Rückleitung (Erde) zweier Module auf eine Leitung zu legen (siehe Abb. 11.12). Über die Zuleitung kann man die Module dann immer noch individuell ein- oder ausschalten und spart bei der Rückleitung ein Kabel ein. Die Masseleitung hat einen kleinen, aber nicht verschwindenden Widerstand R.
Deshalb wirkt sich eine Änderung des Stroms dI_M eines der Module auf den Wert von U_0 mit $dU_0 = R\,dI_M$ aus. Dies reduziert (auch) die Versorgungsspannung $U_1 - U_0$ des anderen Moduls.

Eine einfache Methode dies zu minimieren, ist die Wahl eines kleinen Widerstands R des Rückleitungskabels oder der Einsatz eines Spannungsreglers im Modul.

Welche Auswirkung hat die Reduktion auf einen fehlenden Verstärker? Dies ist in Abbildung 11.13 illustriert. Warum stört diese Kombination überhaupt?

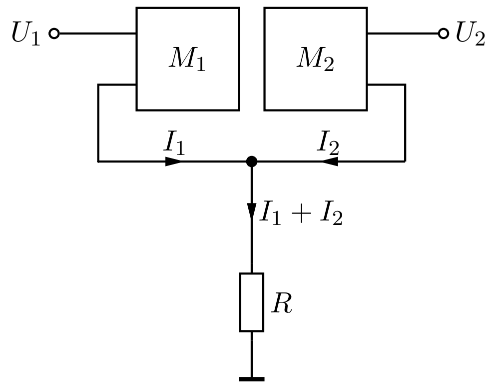


Abb. 11.12 Illustration zweier Module mit gemeinsamer Masseleitung der Modulspannungen.

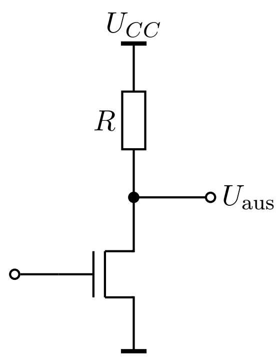


Abb. 11.13 Schematische Erklärung für Störungen bei einer Kopplung über die Stromversorgung.

Eine Variation von U_{CC} wirkt wie ein Signal am Eingang des Emitterverstärkers, da $\mathrm{d}U_{\text{aus}} = \mathrm{d}U_{CC} - g_m\, R\, \mathrm{d}U_{\text{ein}}$ ist.

$$PSRR \quad \equiv \quad \frac{\mathrm{d}U_{CC}}{\mathrm{d}U_{\text{aus}}} \tag{11.12}$$

Der $PSRR$ wird in der Regel in dB ausgedrückt und ist eine Funktion der Frequenz. PSRR ist nicht notwendigerweise sehr groß! Manchmal liegt PSRR bei ≈ 1, d.h. die Schwankung von U_{CC} wird im relevanten Frequenzbereich kaum abgeschwächt.

Dies ist nur eine einfache Architektur. Mit Spannungsregler und Rückkopplung sieht es in weiten Frequenzbereichen besser aus!

Wichtig ist auch der Entkopplungskondensator C (Abb. 11.14). C wirkt als AC-Kurzschluss. Der Kondensatortyp ist zu beachten!

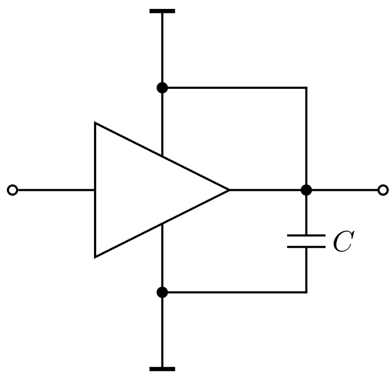
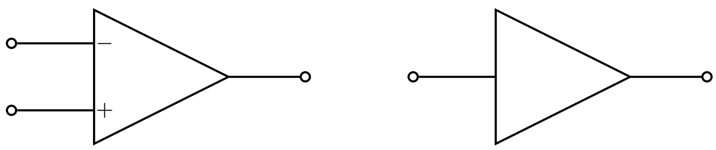


Abb. 11.14 Entkopplungskondensator.

Ein OPV mit zwei differentiellen Eingängen ist besser als ein OPV mit einem Eingang. Aufgrund der zwei Eingänge ist das thermische Rauschen dann aber um einen Faktor $\sqrt{2}$ größer als beim OPV mit einem einzigen Eingang.



Wie kann eine Störung auftreten?

- Powersupply rauscht! Lösungen sind:
  a) Besseres kaufen.
  b) Lineares Power supply statt „switched-mode mode“ power supply.
  c) Filtern mit L und C
- Schwankung des Stromverbrauchs!
- Analoge Verstärker haben meistens einen fast konstanten Strom I_L. Die Kopplung von Analog- und Digitalteil der Schaltung ist oft kritisch. Häufig tritt das in Abbildung 11.15

beschriebene Problem bei Kopplung rauschender digitaler mit empfindlichen analogen Systemen auf.

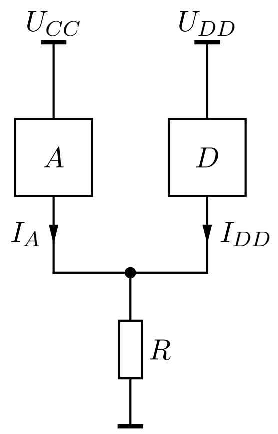


Abb. 11.15 Kopplung der I_D-Variation auf U_A.

$$
\begin{aligned}
U_A &= U_{CC} - (I_A + I_D)\,R \\
U_D &= U_{CC} - (I_A + I_D)\,R
\end{aligned}
$$

Besser wäre die Anordnung aus Abbildung 11.16.

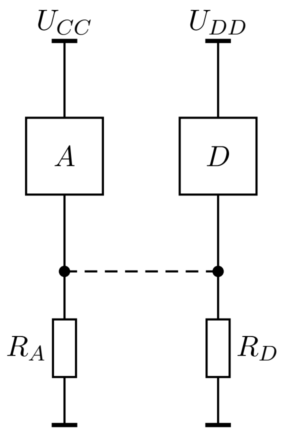


Abb. 11.16 Bessere Analog-Digital-Kopplung.

Meistens aber müssen die Masse der Analogspannung und der Digitalspannung gekoppelt sein.

Es ist also wichtig, R_A und R_D klein zu halten. Alternativ kann I_D *konstant* gehalten werden, selbst wenn es mit Energievelust einhergeht.

11.13 Kapazitive Kopplung

In Abbildung 11.17 ist eine Art der kapazitiven Einkopplung von Störsignalen über eine benachbarte Leitung illustriert [13].

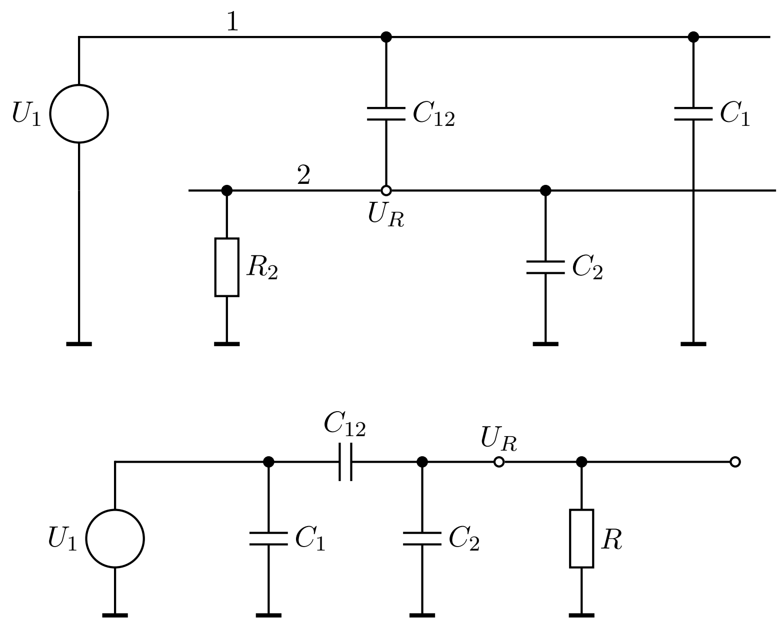


Abb. 11.17 Zwei parallele Leitungen 1 und 2 mit kapazitiver Kopplung C_{12}.

$$
\begin{aligned}
\frac{U_R}{U_{\text{ein}}} &= \frac{Z_2 \parallel R}{(Z_2 \parallel R) + Z_{12}} = \frac{\frac{R\,Z_2}{R+Z_2}}{\frac{R\,Z_2}{R+Z_2} + Z_{12}} = \frac{1}{1 + Z_{12}\,\frac{R+Z_2}{R\,Z_2}} = \frac{1}{1 + \frac{1}{\mathrm{j}\,\omega\,C_{12}}\,\frac{R+\frac{1}{\mathrm{j}\,\omega\,C_2}}{\frac{R}{\mathrm{j}\,\omega\,C_2}}} \\
&= \frac{\mathrm{j}\,\omega\,C_{12}\,R}{\mathrm{j}\,\omega\,C_{12}\,R + \mathrm{j}\,\omega\,C_2\,R + 1} = \frac{\mathrm{j}\,\omega\,C_{12}\,R}{\mathrm{j}\,\omega\,R\,(C_{12} + C_2) + 1}
\end{aligned}
\tag{11.13}
$$

Falls $R \ll \frac{1}{\mathrm{j}\,\omega\,(C_{12}+C_2)}$ gilt, folgt $U_R = \mathrm{j}\,\omega\,R\,C_{12}\,U_{\text{ein}}$.

Was kann man tun, um die Übersprechspannung U_R kleinzuhalten?

- Die Frequenz ω klein wählen; ω ist aber oft durch das analoge Signal bzw. die nötigen Übertragungsraten vorgegeben.

  Falls das Signal digital ist, ist die Signaländerung sprunghaft. Besser ist aber eine langsamere Signaländerung. Hier könnte man statt einem Bus im LVDS-Standard M-LVDS wählen. Dies bewirkt Stromänderungen von $I = 10\,\mathrm{mA}$ mit reduzierter An- und Abstiegszeit von mehreren ns, statt Zeiten im sub-Nanosekundenbereich (siehe Abbildung 11.18).

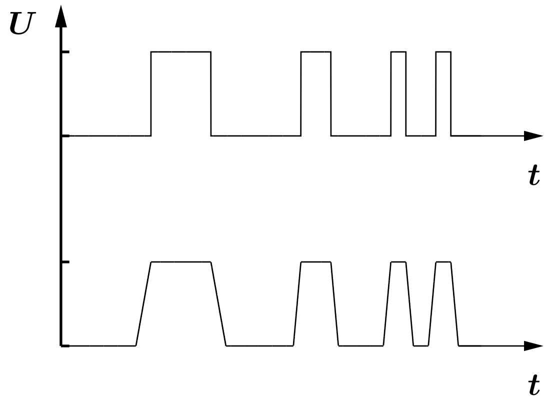


Abb. 11.18 Illustration eines digitalen Signals mit LVDS (oben) und M-LVDS (unten).

- Die Wahl eines kleinen Spannungspegels U ist beim digitalen Bus vorteilhaft, solange der Rauschabstand groß genug bleibt.

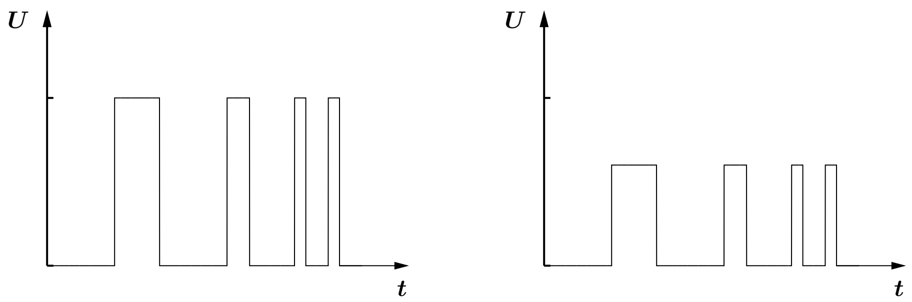


Abb. 11.19 Illustration eines digitalen Signals unterschiedlicher Spannungspegel.

- Halte C_{12} klein.

  Wie kann man die Kapazität C_{12} klein halten?
  - Je kleiner der Abstand zwischen den Kabeln gewählt wird, desto kleiner ist die Kapazität C_{12}.
  - Abschirmung

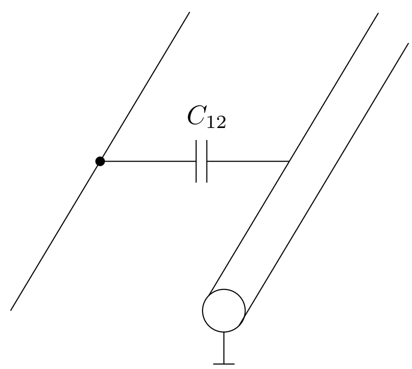


Abb. 11.20 Abschirmung durch Koaxialkabel mit Außenleiter auf Masse.

$\Rightarrow U_n = 0.$

- Man kann einen kleinen Widerstand R wählen. Dieser kann aber in den meisten Fällen nicht ausgesucht werden. Liegt R nahe an Masse, gibt es kein Problem.

11.14 Induktive Kopplung

In Abbildung 11.21 wird die induktive Kopplung zweier Leiter betrachtet [13].

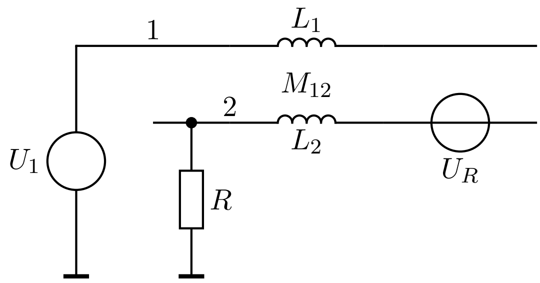


Abb. 11.21 Induktive Einkopplung über benachbarte Leitungen.

Mit M = Gegeninduktivität gilt, analog zu $U = L\,\frac{dI}{dt}$

$$U_n = M_{12}\,\frac{dI_1}{dt} \quad \text{bzw.} \quad U_n = \mathrm{j}\,\omega\,M_{12}\,I_1 \quad . \tag{11.14}$$

Um eine minimale induktive Kopplung zu realisieren, sollten die Ströme klein bzw. die Anstiegszeiten langsam sein. Ferner ist eine kleine Gegeninduktivität M_{12} durch die Wahl einer kleinen Schleifenfläche vorteilhaft.

Induktive Kopplung koplanarer Leiterschleifen:

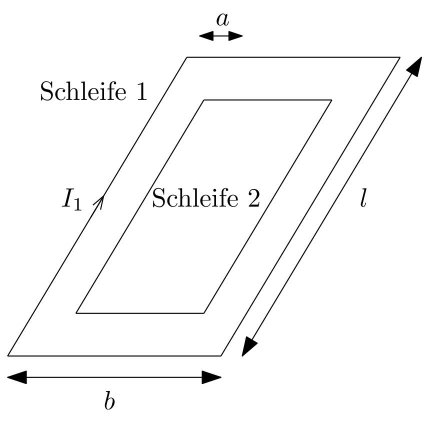


Abb. 11.22 Gegeninduktivität koplanarer Leiterschleifen.

Die Gegeninduktivität koplanarer Leiterschleifen ist durch die folgende Gleichung gegeben.

$$M_{12} = \frac{\mu}{2\,\pi}\,\ln\left(\frac{b}{a}\right) = 4\cdot 10^{-7}\,\ln\left(\frac{b}{a}\right)\frac{\mathrm{H}}{\mathrm{m}} \quad .$$

Daraus folgt: Die Seiten der Schleifenfläche müssen so klein wie möglich gewählt werden.

Beispiel aus der Praxis:

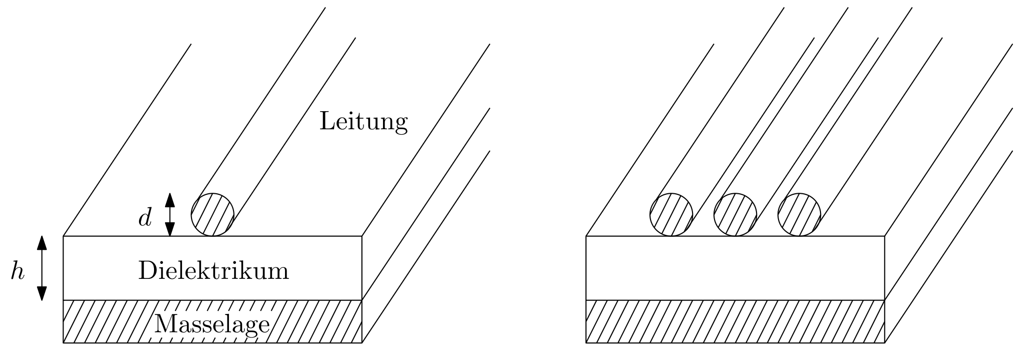


Abb. 11.23 Leiter des Durchmessers d im Abstand h über einer geerdeten Leiterfläche (links). Alternativanordnung mit drei identischen Leitern (rechts).

Jeder Leiter bildet nicht nur einen ohmschen Widerstand, sondern auch eine Induktivität. Die Induktivität eines runden Leiters mit Durchmesser d im Abstand h über einer „geerdeten" Leiterplatte (Abb. 11.23) beträgt für $h > 1,5\,d$:

$$\frac{L}{l} = \frac{\mu}{2\,\pi} \ln \frac{4\,h}{d} \frac{\mathrm{H}}{\mathrm{m}} \quad . \tag{11.15}$$

Mit $\mu = 4\,\pi \cdot 10^{-7}\,\frac{\mathrm{H}}{\mathrm{m}}$ folgt

$$\frac{L}{l} = 2 \cdot 10^{-7} \ln\left(\frac{4\,h}{d}\right) \frac{\mathrm{H}}{\mathrm{m}} \quad .$$

Es ist generell sinnvoll L klein zu halten. Die Variation von h und d ist wegen des Logarithmus' ineffektiv, dagegen ist die Verwendung mehrerer Leitungen gut, kurze Leitungen sind vorteilhaft. Interessanterweise ist die Induktivität nicht zu groß. Eine Massenlage (*ground plane*) führt also für vernünftige Werte zu einer kleinen Induktivität. Die Impedanz dieser Anordnung beträgt

$$Z_0(\Omega) \quad = \quad \frac{60}{\sqrt{\epsilon_R}} \ln \frac{4\,h}{d} \quad . \tag{11.16}$$

Oft ist der Leiter kein Draht, sondern eine Leiterbahn (siehe Abb.11.24).

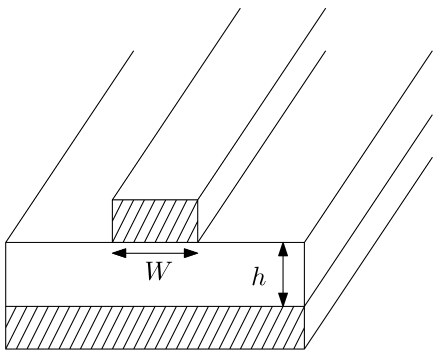


Abb. 11.24 Darstellung einer Mikrostreifenleitung auf einem leitenden Substrat.

Oft liegt auch ein Szenario à la Abbildung 11.25 vor. Hier gibt es keine Massenlage, sondern Zuleiter und Rückleiter liegen nebeneinander. Hier gilt Gleichung 11.17.

$$\frac{L}{l} = 4 \cdot 10^{-7} \ln\left(\frac{2\,h}{d}\right) \frac{\mathrm{H}}{\mathrm{m}} \tag{11.17}$$

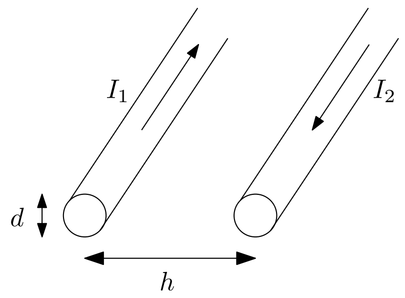


Abb. 11.25 Zwei parallele Leiter des Durchmessers d im Abstand h.

Zahlenbeispiele

In welcher Größenordnung liegen realistische Induktivitäten?

Wir betrachten einen Leiter von 10 cm Länge mit einem Höhen-zu-Abstandsverhältnis von $\frac{4\,h}{d} = 8$ (z. B. 25 μm dicken Draht in einem Abstand von 100 μm über Masse).

$$l\,L = 0,1\mathrm{m} \cdot 2 \cdot 10^{-7} \cdot \ln 8\,[\mathrm{H}] \approx 4 \cdot 10^{-8}\,[\frac{\mathrm{H}}{m}] \quad .$$

Zu welchen Störungen könnte dies führen?

Schon digitale Signale könnten leicht Stromänderungen von $\approx \frac{3}{10}\frac{\mathrm{mA}}{\mathrm{ns}}$ bewirken. Dies führt dann mit Gleichung 11.14 zu einer induzierten Spannung $U_N = 12\,\mathrm{mV}$.

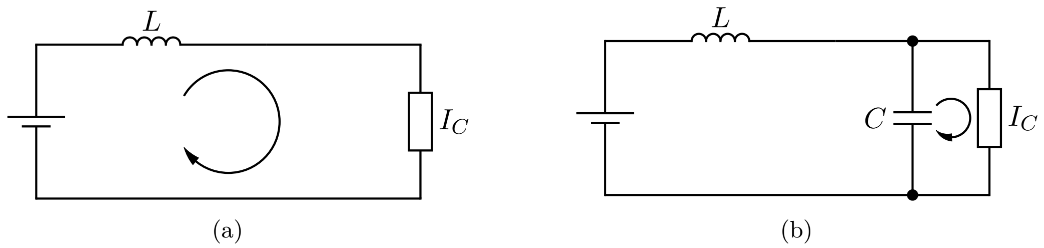


Abb. 11.26 (a) Spannungsquelle mit Leitung der Induktivität L und einem IC als Verbraucher. (b) Dieselbe Schaltung mit einem zusätzlichen Entkopplungskondensator C.

Schnelle Ströme kommen von der Kapazität, mit kurzen Wegen und daher kleiner Induktivität L_p, danach langsames Aufladen des Kondensators durch die externe Spannungsquelle über die Induktivität. Die Kapazität C muss wegen

$$\mathrm{d}U = \frac{I\,\mathrm{d}t}{C} \tag{11.18}$$

groß genug sein, damit die Spannungsänderung dU trotz der Stromänderung klein bleibt.

Beispiel

Der IC braucht 50 mA für 2 ns. Mit d$V < 0{,}01$ V folgt $C \geq \frac{50 \cdot 2}{0{,}01} 10^{-3-9}\,\mathrm{F} = 10\,\mathrm{nF}$.

Typisch sind Kondensatoren mit Kapazitäten von 10 bis 100 nF direkt neben dem IC.

Es ist besser die Kapazität C nicht zu groß wählen, sonst kommt es zum Resonanzeffekt. Mit $f = \frac{1}{2\,\pi\,\sqrt{L\,C}} = \frac{1}{2\,\pi\,\sqrt{1\,\mathrm{nH}\,1\,\mathrm{nF}}} \approx 30\,\mathrm{MHz}$ ist es OK! Oft ist die Induktivität der Drahtverbindung eher 10 nH.

11.15 Elektromagnetische Strahlung

- Elektromagnetische Wellen (z. B. Funkwellen[8], Blitze, Motoren etc.) können mit der Schaltung wechselwirken.

  Von diesen Effekten wird oft berichtet.Viel wichtiger sind aber meistens Störungen aus dem eigenen System. Am empfindlichsten ist der Analogteil und das Signal vor der Verstärkung.

- Eine Abschirmung durch eine 25 bis 50 μm leitenden dicke Folie ist sehr effektiv für elektrische Felder. Die Folie muss auf einem festen Potential liegen.
- Magnetische Felder sind für kleine Frequenzen (1…10 kHz) problematisch.
- Weniger Löcher und kürzere Schlitze im Schirm sind besser.

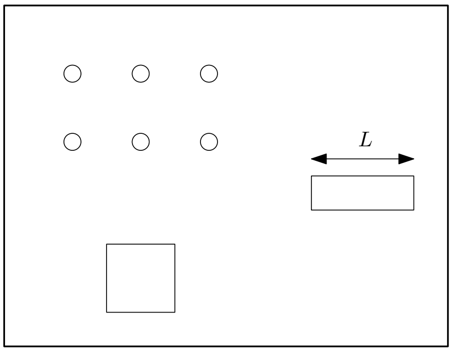


Abb. 11.27 Abschirmung elektromagnetischer Strahlung durch eine Folie mit kurzen Schlitzen und Löchern.

[8] Dies umfasst Radiowellen, Mikrowellen, Infrarotwellen, etc.

12 Grundlagen der Digitalelektronik

Digitalelektronik wird zur elektronischen Verarbeitung von diskreten Größen eingesetzt. Dies können digitalisierte Signale, logische Zustände oder Binärzahlen sein.

Grundelemente der Digitalelektronik sind logische Gatter, Speicher, Flip-Flops und programmierbare Logikbausteine. Zur komplexeren Digitalelektronik zählen Mikrokontroller, Mikroprozessoren, CPUs oder FPGAs.

Die Alltagstechnologie, wie z. B. Smartphones, Computer und elektronische Haushaltsgeräte, beinhaltet Digitalelektronik. Die Programmierbarkeit und dadurch die große Flexibilität standardisierter Digitalbausteine führen dazu, dass zunehmend Digitalelektronik in klassischen Bereichen der Analogelektronik angewandt wird. Beispiele sind die digitale Signalverarbeitung und digitale Filter. Andererseits werden das Verständnis analoger Effekte für digitale Schaltungen bei höchsten Frequenzen und für moderne Transistortechnologie immer wichtiger!

In diesem Kapitel beschäftigen wir uns mit den Grundlagen der Digitalelektronik:

1. Zahlensysteme: Binärzahlen, Hexadezimalzahlen, Graycode, etc.
2. Logische Verknüpfungen und Schaltalgebra
3. Logikfamilien und Implementierung logischer Verknüpfungen auf Transistorniveau.

12.1 Zahlensysteme

Computer und andere elektronische Spielzeuge rechnen im **Binärsystem**. Das heißt, sie benutzen nur zwei Zahlensymbole nämlich „0“ und „1“, statt 0, 1, 2 bis 9 wie im Dezimalsystem.

In Tabelle 12.1 sind die Dezimalzahlen von 1 bis 16 und die entsprechenden Darstellung im Binärsystem und Hexadezimalsystem (Sechzehnersystem) zusammengestellt. Die rechte Ziffer einer Binärzahl wird auch das „least significant bit“ (LSB) genannt, also die Ziffer, die der Potenz 2^0 entspricht. Die linke Ziffer das „most significant bit“ (MSB). „Bit“ steht für binary digit. (Mitunter wird die Reihenfolge auch umgekehrt.) Die Umrechnung einer Binärzahl in die entsprechende Dezimalzahl ist einfach. Jede Ziffer der Binärzahl muss mit der Potenz der Basis 2, die ihrer Position in der Zahl entspricht multipliziert werden und alle Terme aufsummiert werden.

Beispiel:

$$1101_2 = 1 \cdot 2^3 + 1 \cdot 2^2 + 0 \cdot 2^1 + 1 \cdot 2^0 = 8 + 4 + 0 + 1 = 13_{10} \quad .$$

Die Umrechnung von Dezimalzahlen in Binärzahlen ist auch einfach und kann nach dem folgenden iterativen Algorithmus durchgeführt werden.

Dezimal	**Binär**	**Hexadezimal**
0	0	0
1	1	1
2	10	2
3	11	3
4	100	4
5	101	5
6	110	6
7	111	7
8	1000	8
9	1001	9
10	1010	A
11	1011	B
12	1100	C
13	1101	D
14	1110	E
15	1111	F
16	1 0000	10

Tab. 12.1 Die Zahlen 0 bis 16 dargestellt im Dezimal-, Binär- und Hexadezimalsystem.

Die Dezimalzahl D wird durch 2 geteilt. Es ergibt sich eine kleinere natürliche Zahl N und ein Rest R, „0" oder „1":

$$\frac{\mathrm{D}}{2} = \mathrm{N} + \mathrm{R} \quad .$$

Der Rest entspricht dem LSB der Binärdarstellung. Die natürliche Zahl N wird wiederum durch 2 geteilt und der neue Rest stellt die nächste Stelle der Binärdarstellung dar, usw.

Beispiel: Dezimalzahl 13_{10}.

$$\frac{13}{2} = 6\,\mathrm{R}1, \ \frac{6}{2} = 3\,\mathrm{R}0, \ \frac{3}{2} = 1\,\mathrm{R}1, \ \frac{1}{2} = 0\,\mathrm{R}1, \ \frac{0}{2} = 0\,\mathrm{R}0 \quad .$$

Damit gilt

$$13_{10} = 1101_2 \quad .$$

Eine Binärzahl hat schnell viele Stellen. Das kann unpraktisch werden und deshalb verwendet man in der Datenverarbeitung sehr gerne das **Hexadezimalsystem**.

Das Hexadezimalsystem beruht auf der Basis 16. Als einstellige Symbole für die zweistelligen Dezimalzahlen 10, 11, 12, 13, 14 und 15, werden im Hexadezimalsystem einfach die Buchstaben A, B, C, D, E und F verwendet (Tab. 12.1). Durch die Mischung von Zahlen mit Buchstaben sind Hexadezimalzahlen in der Regel leicht also solche zu erkennen.

Das Prinzip zur Umrechnung von Binärzahlen in Dezimalzahlen und umgekehrt funktioniert auch für die Umrechnung von Hexadezimalzahlen.

Beispiel:

$$3\mathrm{EF} = 3 \cdot 16^2 + 14 \cdot 16^1 + 15 \cdot 16^0 = 3 \cdot 256 + 14 \cdot 16 + 15 \cdot 1 = 1007_{10} \quad .$$

Die Quadratzahl $16 \cdot 16 = 256\,(= 1\,0000\,0000_2 = 100_{16})$ darf man sich merken. Auch den Inhalt von Tabelle 12.1.

Hexadezimalzahlen sind sehr kompakt. Das praktische an ihnen ist nun, dass eine 4-stellige Binärzahl, auch „**Nibble**“ genannt, gerade einer einstelligen Hexadezimalzahl entspricht.

Beispiele:

$$255_{10} = 1111\ 1111_2 = \mathrm{FF}_{16}$$

$$11\ 0111\ 1100\ 0101_2 = 37\mathrm{C}5_{16} \quad ,$$

da

$$11_2 = 3 = 3_{16} \quad , \quad 0111_2 = 7 = 7_{16} \quad , \quad 1100_2 = 12 = \mathrm{C}_{16} \quad \text{und} \quad 0101_2 = 5 = 5_{16} \quad .$$

Gray Code

Der Gray Code benutzt nur zwei Symbole, „0“ und „1“. Der Gray Code entspricht in etwa der Binärdarstellung mit der wesentlichen Modifikation, dass sich benachbarte Zahlen nur in einer Stelle unterscheiden dürfen.

Für eine Zahl mit 3 bits ergibt sich das in Tabelle 12.2 dargestellte Bild.

Man kann eine n-stellige Binärzahl, $b_n ... b_3\, b_2\, b_1$, einfach nach folgender Vorschrift in Gray Code, $g_n ... g_3\, g_2\, g_1$, umrechnen.

1) Binärcode und der zugehörige Gray Code haben dieselbe Anzahl von Stellen n. Die führende Stelle in Binärcode und Graycode ist identisch: $g_n = b_n$.

2) Alle folgenden Stellen g_i hängen von den zwei Stellen b_{i+1} und b_i ab.

Falls $b_{i+1} = 0 \Rightarrow g_i = b_i$.

Falls $b_{i+1} = 1 \Rightarrow g_i = \overline{b_i}$. (Also $g_i = 1$ für $b_i = 0$, und $g_i = 0$, für $b_i = 1$.)

Dezimal	**Binär**	**Gray**
0	000	000
1	001	001
2	010	011
3	011	010
4	100	110
5	101	111
6	110	101
7	111	100

Tab. 12.2 Die Zahlen 0 bis 7 als Dezimalzahl, Binärzahl und Gray Code.

Stimmt diese Vorschirft mit Tabelle 12.2 überein?

Vorschrift 1) ist erfüllt, $b_3 = g_3$. Vorschrift 2) ist auch erfüllt. Wir überprüfen dies für die Zahlen $b_3\, b_2\, b_1 = 010$ und 101.

Beispiel $b_3\, b_2\, b_1 = 010$

$g_3 = 0$, nach 1).

$g_2 = b_2 = 1$, da $g_3 = 0$.

$g_3 = \overline{b_3} = 1$, da $b_2 = 1$.

Insgesamt gilt also $g_3\, g_2\, g_1 = 011$ entsprechend der Tabelle.

Beispiel $b_3\, b_2\, b_1 = 101$

$g_3 = 1$, nach 1).

$g_2 = \overline{b_2} = 1$, da $g_3 = 1$.

$g_3 = b_3 = 1$, da $b_2 = 0$.

Insgesamt gilt also $g_3\, g_2\, g_1 = 111$ entsprechend der Tabelle.

Der Gray Code hat verschiedene Vorzüge insbesondere bei der Datenübertragung und bei automatischer Übertragungsfehlerkorrektur.

BCD Code

BCD steht für *binary coded decimal.* In diesem Code wird jede Stelle einer Dezimalzahl durch eine entsprechende 4-stellige Binärzahl dargestellt. Also

$$129_{10} = 0001\,0010\,1001_{BCD}$$

Die Binärzahlen 1010_2 bis 1111_2 werden nicht benötigt, daher ist dieser Code aus moderner Sicht nicht sehr effizient und nicht besonders relevant.

Negative Binärzahlen*

Es gibt verschiedene Möglichkeiten negative Binärzahlen darzustellen. Die naheliegende Variante ist durch ein zusätzliches Vorzeichenbit, wie in der zweiten Spalte der Tabelle 12.3 illustiert. Hier wird jeder 3-stelligen Binährzahl eine „0“ oder „1“ vorangestellt. „0“ steht für positive und „1“ für negative Zahlen.

Anders als bei Dezimalzahlen mit Vorzeichen (+, -) wird hier also kein zusätzliches Symbol als Vorzeichen benutzt. Deshalb ist eine klare Definition, die die Stelle des Vorzeichenbits regelt, nötig, damit es nicht zu Verwirrungen kommt.

Dezimal	Vorzeichen	Offset	IEEE 754	Zweierkomplement
+ 8	-	-	1111	-
+ 7	0111	1111	1110	0111
+ 6	0110	1110	1101	0110
+ 5	0101	1101	1100	0101
+ 4	0100	1100	1011	0100
+ 3	0011	1011	1010	0011
+ 2	0010	1010	1001	0010
+ 1	0001	1001	1000	0001
0	0000	1000	0111	0000
- 1	1001	0111	0110	1111
- 2	1010	0110	0101	1110
- 3	1011	0101	0100	1101
- 4	1100	0100	0011	1100
- 5	1101	0011	0010	1011
- 6	1110	0010	0001	1010
- 7	1111	0001	0000	1001
- 8	-	0000	-	1000
(- 0)	1000	-	-	-

Tab. 12.3 Darstellung der negative Zahlen im Binärsystem.

In der Offsetdarstellung führt man kein explizites Vorzeichen ein, sondern man setzt die kleinste (negative) Zahl gleich 0_2, die nächsthöhere Zahl gleich 1_2 und zählt entsprechend nach oben weiter (Tab. 12.3, 3. Spalte). Dabei ergibt es sich, dass ein MSB von „1“ Zahlen ≥ 0 entspricht. Ein MSB von „0“ entspricht negativen Zahlen. Es gibt eine alternative Offsetdarstellung, die nicht bei -8, sondern -7 beginnt (Tab. 12.3, 4. Spalte).

Eine andere Darstellung von Binärzahlen mit Vorzeichen ist das sogenannte Zweierkomplement[1] (Tab. 12.3, rechte Spalte). Positive Zahlen werden in der Zweierkomplementdarstellung, wie gehabt, ohne „Vorzeichen“, sondern mit einer vorangestellten „0“ dargestellt. Die Darstellung einer negativen Binarzahl erhält man aus der entsprechenden positiven Binärzahl durch

1. Ersetzen aller Nullen durch Einsen und umgekehrt. (Bilden eines Einerkomplements.)

2. Summe des Einerkomplements und 1_2

Beispiel 1

Gesucht: $x_2 = -7_{10}$.

1. Bestimme die entsprechende positive Binärzahl: $+7_{10} = 0111_2$.

2. Bilden des Einerkomplements: $0111_2 \Rightarrow 1000_2$.

[1] Engl. *two's complement*

3. Addition mit 0001_2:

$$\begin{array}{ccccc} & 0 & 1 & 1 & 1 \\ + & 0 & 0 & 0 & 1 \\ & {\scriptstyle 1} & {\scriptstyle 1} & {\scriptstyle 1} & \\ \hline & 1 & 0 & 0 & 0 \end{array}$$

$\Rightarrow -7_{10} = 1001_2$.

Beispiel 2

Gesucht: $x_2 = -5_{10}$.

1. Bestimme die entsprechende positive Binärzahl: $+5_{10} = 0101_2$.
2. Bilden des Einerkomplements: $0101_2 \Rightarrow 1010_2$
3. Addition mit 0001_2:

$$\begin{array}{ccccc} & 1 & 0 & 1 & 0 \\ + & 0 & 0 & 0 & 1 \\ \hline & 1 & 0 & 1 & 1 \end{array}$$

$\Rightarrow -5_{10} = 1011_2$.

Ein Vorteil der Zweierkomplementdarstellung ist, dass sich negative und positive Zahlen mit geringem elektrischen Aufwand addieren/subtrahieren und multiplizieren lassen.

Beispiel 3

$$-5_{10} + 5_{10} = 5_{10} + (-5_{10}) = 0_{10} \quad \overset{\textit{Zweierkomplement}}{\rightarrow} \quad \begin{array}{ccccc} & 0 & 1 & 0 & 1 \\ + & 1 & 0 & 1 & 1 \\ {\scriptstyle 1} & {\scriptstyle 1} & {\scriptstyle 1} & {\scriptstyle 1} & \\ \hline 1 & 0 & 0 & 0 & 0 \end{array}$$

Beispiel 4

$$5_{10} - 7_{10} = 5_{10} + (-7_{10}) = -2_{10} \quad \overset{\textit{Zweierkomplement}}{\rightarrow} \quad \begin{array}{ccccc} & 0 & 1 & 0 & 1 \\ + & 1 & 0 & 0 & 1 \\ & & & {\scriptstyle 1} & \\ \hline & 1 & 1 & 1 & 0 \end{array}$$

Beispiel 5

$$-2_{10} \cdot 3_{10} = -6_{10} \quad \overset{\textit{Zweierkomplement}}{\rightarrow} \quad \begin{array}{cccccccccc} & 1 & 1 & 1 & 0 & \cdot & 0 & 0 & 1 & 1 \\ \hline & & & 0 & 0 & 0 & 0 & & & \\ & & & & 0 & 0 & 0 & 0 & & \\ & & & & & 1 & 1 & 1 & 0 & \\ + & & & & & & 1 & 1 & 1 & 0 \\ & & & & {\scriptstyle 1} & {\scriptstyle 1} & {\scriptstyle 1} & & & \\ \hline & & & 0 & 1 & 0 & \mathbf{1} & \mathbf{0} & \mathbf{1} & \mathbf{0} \end{array}$$

$$-6_{10} = 1010_2 \Leftarrow$$

Beispiel 6

$$-4_{10} \cdot 2_{10} = -8_{10} \quad \xrightarrow{Zweierkomplement} \quad \begin{array}{ccccccccc} 1 & 1 & 0 & 0 & \cdot & 0 & 0 & 1 & 0 \\ \hline & & 0 & 0 & 0 & 0 & & & \\ & & & 0 & 0 & 0 & 0 & & \\ & & & & 1 & 1 & 0 & 0 & \\ + & & & & & 0 & 0 & 0 & 0 \\ \hline & & 0 & 0 & 1 & \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{array}$$

$$-8_{10} = 1000_2 \Leftarrow$$

Anmerkung: Hier werden nur die letzten vier Stellen berücksichtigt.

Binärzahlen kann man analog zu Dezimalzahlen als Bruch schreiben.

Beispiel 7

$$11,10_2 = 1 \cdot 2^1 + 1 \cdot 2^0 + 1 \cdot 2^{-1} + 0 \cdot 2^{-2}$$

Welchem Dezimalbruch entspricht dies?

Um den Dezimalbruch zu bestimmen, multiplizieren wir den Bruch mit 2^2, was einer Verschiebung des Kommas um zwei Stellen entspricht[2].

$$11,01_2 \cdot 100_2 = 1101_2 \quad = \quad 8_{10} + 4_{10} + 1_{10} = 13_{10}$$

$$11,01_2 \quad \widehat{=} \quad \left.\frac{13}{4}\right|_{10} = 3,25_{10}$$

Weitere Beispiele:

$$\begin{aligned} 0,10_2 &= \left.\tfrac{1}{2}\right|_{10} &= 0,5_{10} \\ 0,01_2 &= \left.\tfrac{1}{4}\right|_{10} &= 0,25_{10} \\ 0,11_2 &= \left.\tfrac{3}{4}\right|_{10} &= 0,75_{10} \\ 0,1011_2 &= \left.\tfrac{11}{16}\right|_{10} &= 0,6875_{10} \end{aligned}$$

Gleitkommadarstellung

Neben der Festkommadarstellung gibt es auch die Gleitkommadarstellung mit Vorzeichen S (engl. *sign*), Exponent E und Mantisse M in der Form $(-1)^S\, M\, 2^E$. So gilt

$$11,01_2 = 0,1101_2 \cdot \mathrm{E}100_2$$

In der Standarddarstellung für einfach genaue[3] Binärzahlen (IEEE-P754) werden 32 Bits verwendet, nämlich

[2]Die Multiplikation einer Binärzahl mit 2 entspricht dem Anfügen einer „0“ an die ursprüngliche Zahl, z. B. $101_2 \cdot 10_2 = 1010_2 \;\widehat{=}\; 5_{10} \cdot 2_{10} = 10_{10}$.

[3]Engl. *single precision*

- 1 Vorzeichenbit
- 8 Exponentenbits (hier wird der Exponent in Offsetdarstellung angegeben, entsprechend der 5. Spalte von Tabelle 12.3)
- 23 Mantissenbits

Doppelt genaue[4] Binärzahlen werden mit 64 Bits dargestellt. Hier wird wieder ein Vorzeichenbit verwendet, dazu 11 Bits für den Exponenten und 52 Bits für die Mantisse.

Die gößte darstellbare Zahl in „single precision“ beträgt rund $3,4 \cdot 10^{38}|_{10}$, die größte darstellbare Zahl in „double precision“ beträgt rund $2 \cdot 10^{308}|_{10}$.

Die Darstellung der Mantisse M hat die Form

$$\mathrm{M} = 1 + \sum_{i=1}^{23} m_i \, 2^{-i} = 1 + \frac{m_1}{2} + \frac{m_2}{4} + \cdots + \frac{m_{23}}{2^{23}} \quad ,$$

das heißt, die erste Stelle m_0 wird $m_0 = 1$ gesetzt. Sie wird nicht explizit angegeben.

Beispiel:

$$\underbrace{0}_{Vorzeichen} \;|\underbrace{011\ 1111\ 1}_{Exponent}|\underbrace{110\ 0000\ 0000\ 0000\ 0000\ 0000}_{Mantisse} = \mathrm{3FE00000}_{16}$$

Mit einem Offset von 127 für single precision und einem Offset von 255 für double precision ergibt sich

$$+\,2^{127-127}\,(1+0,75) = 1,75 \quad .$$

Die größte einfachgenaue Zahl ist 7F7 FFFFF, also

$$\underbrace{0}_{Vorzeichen} \;|\underbrace{111\ 1111\ 0}_{Exponent}|\underbrace{111\ 1111\ 1111\ 1111\ 1111\ 1111}_{Mantisse\, m_1 - m_{23}} \Rightarrow +\,2^{254-127}\,(2-2^{-23}) \approx 2^{128} \quad . \tag{12.1}$$

Warum ist nicht 7FF FFFFF die größte einfachgenaue Zahl?

Am einfachsten beginnt man vom Ende der Zahl. Die FFFFF ist offensichtlich. Die nächsten 4 Stellen in Gl. 12.1 (nach links gehend) sind 0111. Hier würde man eine 1111 erwarten, dies ist aber nicht zugelassen. Es würde einem Exponenten von 1111 1111 entsprechen, der eine Sonderfunktion hat. Die erste Stelle (7) ist offensichtlich, da am Anfang eine „0“ stehen muss (positives Vorzeichen) und nur 32 Bits zur Verfügung stehen.

Etwas unbequem an dieser Darstellung ist, dass die Mantisse und der Exponent nicht jeweils durch von einander getrennte Hexadezimalzahlen ausgedrückt werden.

12.2 Logische Verknüpfungen

Ein wichtiges Gebiet der digitalen Elektronik ist die kombinatorische Logik oder Schaltalgebra. Hier geht es darum, mehreren logischen Eingangsvariablen eine oder mehrere logische

[4] Engl. *double precision*

Ausgangsvariablen zuzuordnen.

Es gibt nur zwei logische Zustände: „wahr“ (true) oder „falsch“ (false). Für „wahr“ schreiben wir hier „1“ und für „falsch“ schreiben wir „0“.

Für logische Variablen gibt es drei elementare Verknüpfungen: die Negation, Konjunktion und Disjunktion.

Wir bezeichnen im Folgenden die Eingangsvariablen meistens als a, b und c und die Ausgangsvariablen als x und y.

Negation

$$\mathrm{x} = \overline{\mathrm{a}} \tag{12.2}$$

Die Wahrheitstabelle der Negation ist besonders einfach, da nur eine Eingangsvariable auftritt.

Die Negation der Variable a drückt man durch Querstrich über a aus, also $\overline{\mathrm{a}}$.

Anmerkung: Alternative Symbole für $\overline{\mathrm{a}}$ sind ¬a und !a.

a	**x**
0	1
1	0

Tab. 12.4 Wahrheitstabelle der Negation (NOT).

Die zweifache Anwendung der Negation ergibt wieder die Ausgangsvariable, also $\overline{\overline{\mathrm{a}}} = \mathrm{a}$.

Konjunktion oder 0→0-Verknüpfung

$$\mathrm{x} = \mathrm{a} \cdot \mathrm{b} = \mathrm{a\,b} \tag{12.3}$$

Die Konjunktion entspricht der UND-Verknüpfung. Nur wenn alle Eingangswerte, hier a und b, wahr sind, ist auch die Ausgangsvariable x wahr. Dies entspricht der Wahrheitstabelle 12.5.

a	**b**	**x**
0	0	0
0	1	0
1	0	0
1	1	1

Tab. 12.5 Wahrheitstabelle der Konjunktion oder UND-Verknüpfung (AND).

Man kann dies auch anders ausdrücken. Ist mindestens eine Eingangsvariable „0“, so ist die Ausgangsvariable „0“. Ansonsten ist die Ausgangsvariable „1“.

Das hier benutzte Symbol „·“ für die Konjunktion entspricht dem angelsächsischen Gebrauch. Es ist sehr suggestiv, da das Ergebnis der Konjunktion ab einfach der Multiplikation der natürlichen Zahlen „0“ und „1“ entspricht.

Anmerkung: Das deutsche Symbol für $a \cdot b$ ist $a \wedge b$. Man sollte es tunlichst vermeiden.

Disjunktion oder 1→1-Verknüpfung

$$x = a + b \tag{12.4}$$

Die Disjunktion entspricht der ODER-Verknüpfung. Wenn mindestens einer der Eingangswerte, hier a und b, wahr ist, ist auch die Ausgangsvariable x wahr. Dies entspricht der Wahrheitstabelle 12.6.

a	**b**	**x**
0	0	0
0	1	1
1	0	1
1	1	1

Tab. 12.6 Wahrheitstabelle der Disjunktion oder ODER-Verknüpfung (OR).

Das Symbol + für Disjunktion ist zwar gewöhnungsbedürftig aber auch suggestiv, da bis auf die $1 + 1 = 1$ die Disjunktion der Addition der natürlichen Zahlen „0“ und „1“ entspricht.

Die deutschen Symbole $\vee$[5] statt „+“ und $\wedge$ statt „·“ sind Sonderzeichen und dadurch unpraktisch und unintuitiv.

Konjunktion und Disjunktion kann man auch für mehr als zwei Eingangsvariablen definieren.

Neutrales und komplementäres Element

Für Konjunktion wie Disjunktion gibt es je ein neutrales Element n mit der Eigenschaft

$$a \cdot n = a \quad \text{und} \quad a + n = a \quad . \tag{12.5}$$

Den Wahrheitstabellen entnimmt man, dass das neutrale Element der Konjunktion die „1“ und das neutrale Element der Disjunktion die „0“ ist.

Es gibt auch zu jedem a ein komplementäres Element $\bar{a}$, so dass gilt

$$a \cdot \bar{a} = 0 \quad \text{und} \quad a + \bar{a} = 1 \quad . \tag{12.6}$$

Alle diese Gesetzmäßigkeiten und auch die folgenden Gesetze, deren Anwendung das Rechnen mit logischen Ausdrücken vereinfachen kann, lassen sich leicht durch einfaches Einsetzen überprüfen, da es nur die beiden Werte „0“ und „1“ gibt.

Kommutativgesetze

$$a\,b = b\,a \quad \text{und} \quad a + b = b + a \tag{12.7}$$

[5] Abkürzung des lateinischen Worts *vel*, zu Deutsch „oder“.

Assoziativgesetze

$$(\mathrm{a\,b})\,\mathrm{c} = \mathrm{a}\,(\mathrm{b\,c}) \quad \text{und} \quad (\mathrm{a}+\mathrm{b})+\mathrm{c} = \mathrm{a}+(\mathrm{b}+\mathrm{c}) \tag{12.8}$$

Die Reihenfolge der Anwendung der Konjunktion oder Disjunktion bei mehreren Eingangsvariablen ist beliebig.

Achtung: Auch der digitalen Algebra gilt *Punkt vor Strich*, also **UND vor ODER**.

Distributivgesetz

$$\mathrm{a}\,(\mathrm{b}+\mathrm{c}) = \mathrm{a\,b}+\mathrm{a\,c} \quad \text{und} \quad \mathrm{a}+\mathrm{b\,c} = (\mathrm{a}+\mathrm{b})(\mathrm{a}+\mathrm{c}) \tag{12.9}$$

Absorptionsgesetz

$$\mathrm{a}\,(\mathrm{a}+\mathrm{b}) = \mathrm{a\,a}+\mathrm{a\,b} = \mathrm{a}+\mathrm{a\,b} = \mathrm{a} \tag{12.10}$$

Diese Identität ist sehr hilfreich, wenn man komplexe Ausdrücke vereinfachen will. Wieso gilt sie? Ist $a = 0$, so ist auch $ab = 0$. Ist $a = 1$, so ändert der Wert ab in der Summe $a + ab$ nichts mehr.
Außerdem gilt

$$\mathrm{a\,a} = \mathrm{a} \quad \text{und} \quad \mathrm{a}+\mathrm{a} = \mathrm{a} \quad .$$

Das heißt die Verknüpfung von a mit sich selbst ergibt wieder a. Höhere Potenzen von a treten deshalb in logischen Ausdrücken nicht auf.

De Morgansche Gesetze

$$\overline{\mathrm{a\,b}} = \overline{\mathrm{a}}+\overline{\mathrm{b}} \quad \text{und} \quad \overline{\mathrm{a}+\mathrm{b}} = \overline{\mathrm{a}}\,\overline{\mathrm{b}} \tag{12.11}$$

Die De Morganschen Gesetze drücken die Tatsache aus, dass man die Wertetabelle der UND-Verknüpfung durch mehrfache Anwendung der Negation in die Wertetabelle der ODER-Verknüpfung umformen kann und umgekehrt (siehe Tabelle 12.7).

a	b	ab	$\overline{\mathrm{ab}}$	$\overline{\mathrm{a}}$	$\overline{\mathrm{b}}$	$\overline{\mathrm{a}}+\overline{\mathrm{b}}$
0	0	0	1	1	1	1
0	1	0	1	1	0	1
1	0	0	1	0	1	1
1	1	1	0	0	0	0

Tab. 12.7 Wahrheitstabelle zu den de Morganschen Gesetzen.

Die linke und die rechte Seite der Gleichung 12.11 sind nicht unabhängig voneinander.

$$\overline{\mathrm{a\,b}} = \overline{\mathrm{a}}+\overline{\mathrm{b}} \quad \Leftrightarrow \quad \overline{\overline{\mathrm{a\,b}}} = \overline{\overline{\mathrm{a}}+\overline{\mathrm{b}}} \quad \Leftrightarrow \quad \mathrm{a\,b} = \overline{\overline{\mathrm{a}}+\overline{\mathrm{b}}}$$

Durch Ersetzen von $\overline{\mathrm{a}}$ durch a und von $\overline{\mathrm{b}}$ durch b folgt

$$\overline{\mathrm{a}+\mathrm{b}} = \overline{\mathrm{a}}\,\overline{\mathrm{b}} \quad .$$

Entsprechend den Wertetabellen für die Konjunktion und die Disjunktion können 14 weitere Wertetabellen gebildet werden. (Es gibt 16 verschiedene vierstellige binäre Zahlen von 0000 bis 1111, die als Ergebnisvariable x dargestellt werden können.)

Drei weitere wichtige Verknüpfungen sind NAND, NOR und XOR.

NAND und NOR entsprechen der Negierung der Konjunktion $\overline{\mathrm{a\,b}}$ und der Disjunktion $\overline{\mathrm{a + b}}$. XOR steht für Exclusive OR. Dies ist eine Abart der Disjunktion, bei der der Ausgang nur dann wahr ist, wenn genau eine Eingangsvariable wahr ist. Die Wertetabellen für NAND und NOR sind in 12.8 und 12.9 dargestellt.

a	**b**	**x**
0	0	1
0	1	1
1	0	1
1	1	0

NAND ($0 \rightarrow 1$)

a	**b**	**x**
0	0	1
0	1	0
1	0	0
1	1	0

NOR ($1 \rightarrow 0$)

Tab. 12.8 Wahrheitstabelle der Verknüpfungen NAND und NOR.

a	**b**	**x**
0	0	0
0	1	1
1	0	1
1	1	0

XOR

a	**b**	**x**
0	0	1
0	1	0
1	0	0
1	1	1

Exclusive-NOR

Tab. 12.9 Wahrheitstabelle der Verknüpfungen XOR und Exclusive-NOR.

Die Verknüpfung XOR wird auch Antivalenz genannt. Fallunterscheidungen:

$$x = \begin{cases} 0, & \text{wenn a = b} \\ 1, & \text{wenn a} \neq \text{b} \end{cases}$$

Das Symbol für die Verknüpfung XOR ist $\oplus$.

Entsprechend definiert man die Verknüpfung Äquivalenz oder Exclusive-NOR:

$$x = \begin{cases} 1, & \text{wenn a = b} \\ 0, & \text{wenn a} \neq \text{b} \end{cases}$$

XOR und Exclusive-NOR sind in Tabelle 12.9 dargestellt.

Es ist sehr leicht logische Funktionen für bekannte Werte der Eingangsvariablen schnell auszuwerten.

Beispiel

$$\bar{a}(b + c) + bca + b^2(b + c) = 0 \cdot (1) + 1 + 1 \cdot (1) = 1 \quad \text{für} \quad a = b = c = 1 \quad .$$

12.3 Normalformen

Wie kommt man von einer Wertetabelle zu der entsprechenden algebraischen Funktion?

Es gibt zwei Standarddarstellungen algebraischer Funktionen, die disjunktive Normalform und die konjunktive Normalform.

Will man die disjunktive Normalform (DNF), auch **Summe von Produkten** genannt, aus einer Wertetabelle (z.B. 12.10) aufstellen, geht man folgendermaßen vor:

Eingangsvariablen				Ausgangsvariable
a_1	a_2	$\cdots$	a_n	x
a_{11}	a_{12}	$\cdots$	a_{1n}	x_1
a_{21}	a_{22}	$\cdots$	a_{2n}	x_2
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
a_{m1}	a_{m2}	$\cdots$	a_{mn}	x_m

Tab. 12.10 Allgemeingültige Wahrheitstabelle mit einer Ausgangsvaraiblen x und n Eingangsvariablen a_1 - a_n zur Erläuterung der DNF.

1. Berücksichtige nur die Zeilen i der Wahrheitstabelle mit Ergebnis $x_i = 1$.

2. Aus jeder berücksichtigten Zeile wird genau ein Produktterm gebildet. Jeder Produktterm besteht aus n Variablen und zwar taucht entweder die Variable a_j oder $\bar{a}_j$ (für alle Werte von j) auf. Ein möglicher Produktterm hat also die Form $a_1\,\bar{a}_2 \dots a_n$. Ist in der Wahrheitstabelle der Eintrag a_{ij} = „1“, so taucht im Produktterm der Zeile i die Variable a_j auf. Ist a_{ij} = „0“, die negierte Eingangsvariable $\bar{a}_j$.

3. Bilde die Summe der Produktterme aus allen Zeilen unter Berücksichtigung von 1) und 2).

Beispiel UND

a	**b**	**x**
0	0	0
0	1	0
1	0	0
1	1	1

1. Hier gibt es nur eine Zeile mit Ergebnis „1“.
2. a = b = „1“, damit ergibt sich die DNF zu x = ab.

Beispiel ODER

a	b	x
0	0	0
0	1	1
1	0	1
1	1	1

1. Hier müssen wir die drei letzten Zeilen berücksichtigen.
2. Es treten sowohl Eingangsvariablen mit Wert „0“ wie mit Wert „1“ auf.

Die DNF ergibt sich zu

$$x = a\,\overline{b} + \overline{a}\,b + a\,b \quad . \tag{12.12}$$

Dieser Ausdruck ist noch etwas unhandlich, aber er lässt sich vereinfachen zu

$$a\,\overline{b}+\overline{a}\,b+a\,b = \overline{a}\,b+a\left(\overline{b}+b\right) = \overline{a}\,b+a\cdot 1 = \overline{a}\,b+a\,(a+b) = (\overline{a}+a)\;b+a^2 = b+a \quad . \tag{12.13}$$

Hier wurde das Absorptionsgesetz angewendet.

Die Produkte nennt man auch Minterme.

Man kann auch die konjunktive Normalform (KNF), das **Produkt der Summen**, bilden. Diese Summenterme nennt man auch Maxterme. Dazu betrachtet man nur die Zeilen i der Wertetabelle mit Ergebnis „0“. Jeder Summenterm besteht aus n Variablen, z.B. $a_1 + \overline{a}_2 \ldots + a_n$. Falls $a_{ij} = 0$ taucht die Eingangsvariable a_j in dem Summeterm zur Zeile i auf, falls $a_{ij} = 1$, die negierte Eingangsvariable $\overline{a_j}$.

Soll man die Wahrheitstabelle mit Hilfe der DNF oder der KNF in algebraischer Form bringen?

Die Normalform mit den wenigsten Termen ist die geeigneteste!

Beispiel ODER

a	b	x
0	0	0
0	1	1
1	0	1
1	1	1

Hier ergibt sich mit der KNF nur ein Term, nämlich x = a + b! Hier sollte also diese Methode gewählt werden.

Beispiel UND

a	**b**	**x**
0	0	0
0	1	0
1	0	0
1	1	1

Mit Hilfe der KNF, die hier allerdings eine umständliche Wahl ist, ergibt sich

$$\mathrm{x} = (\mathrm{a}+\mathrm{b})\left(\mathrm{a}+\overline{\mathrm{b}}\right)(\overline{\mathrm{a}}+\mathrm{b}) = \left(\mathrm{a}^2+\mathrm{a}\,\overline{\mathrm{b}}+\mathrm{b}\,\mathrm{a}+\mathrm{b}\,\overline{\mathrm{b}}\right)(\overline{\mathrm{a}}+\mathrm{b}) = \left(\mathrm{a}+\mathrm{a}\left[\mathrm{b}+\overline{\mathrm{b}}\right]+0\right)(\overline{\mathrm{a}}+\mathrm{b}) =$$

$$(\mathrm{a}+\mathrm{a})(\overline{\mathrm{a}}+\mathrm{b}) = \mathrm{a}\,(\overline{\mathrm{a}}+\mathrm{b}) = \mathrm{a}\,\overline{\mathrm{a}}+\mathrm{a}\,\mathrm{b} = \mathrm{a}\,\mathrm{b} \quad .$$

Natürlich gibt es auch komplexere Wahrheitstabellen wie

a	**b**	**c**	**x**	**y**
0	0	0	0	0
0	0	1	1	0
0	1	0	1	0
0	1	1	0	1
1	0	0	1	0
1	0	1	0	1
1	1	0	0	1
1	1	1	1	1

Mit der DNF ergibt sich

$$\mathrm{x} = \left(\overline{\mathrm{a}}\,\overline{\mathrm{b}}\,\mathrm{c}\right)+(\overline{\mathrm{a}}\,\mathrm{b}\,\overline{\mathrm{c}})+\left(\mathrm{a}\,\overline{\mathrm{b}}\,\overline{\mathrm{c}}\right)+\mathrm{a}\,\mathrm{b}\,\mathrm{c} \quad \text{und}$$

$$\mathrm{y} = \overline{\mathrm{a}}\mathrm{bc}+\mathrm{a}\overline{\mathrm{b}}\mathrm{c}+\mathrm{ab}\overline{\mathrm{c}}+\mathrm{abc} \quad .$$

Aufgabe: Welcher algebraische Ausdruck entspricht dem XOR-Gatter? Welcher dem Exclusive-NOR?

12.4 Logikbausteine

Die eingeführten Verknüpfungen kann man mit kommerziellen Standardbausteinen, sogenannten Gattern, realisieren. Ein Gatter entspricht einer Verknüpfung. Diese Bausteine sind dann natürlich eher sperrig. Deshalb sind integrierte komplexe und eventuell auch programmierbare Logikbausteine heute viel häufiger als einzelne Gatter. Leider gibt es unterschiedliche Schaltsmbole für ein und dieselbe Gatterfunktion. Da keine der üblichen Varianten besonders suggestiv sind, empfehle ich keine.

In Abbildung 12.1 sind die ANSI-Versionen dargestellt. Die US- und die IEC-Versionen finden Sie im Appendix.

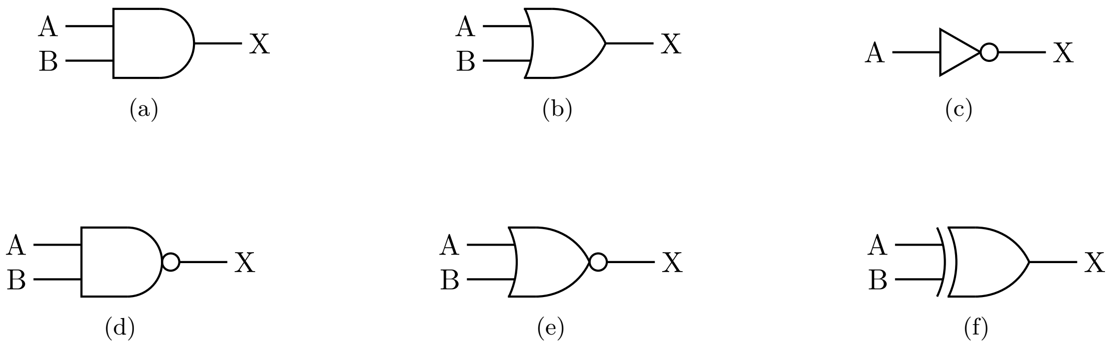


Abb. 12.1 (a) UND-Gatter. (b) ODER-Gatter. (c) NOT-Gatter. (d) NAND-Gatter. (e) NOR-Gatter. (f) XOR-Gatter.

Ein Beispiel für eine Schaltung aus mehreren Gattern ist in Abbildung 12.2 dargestellt.

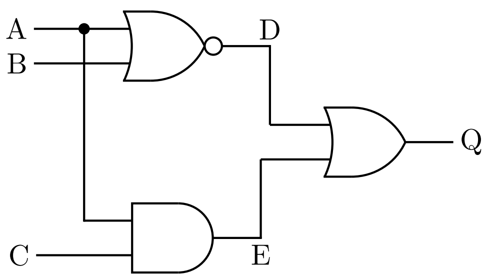


Abb. 12.2 Einfache Schaltung aus einem NOR-, AND- und OR-Gatter [14].

Es ergibt sich:

$$\Rightarrow \left.\begin{array}{rcl} D & = & \overline{A+B} \\ E & = & AC \end{array}\right\} \quad Q = D + E = \overline{A+B} + AC \tag{12.14}$$

Die Wahrheitstabelle 12.11 bildet man iterativ. Erst D, E und dann die Ausgangsvariable Q.

A	**B**	**C**	**D**	**E**	**Q**
0	0	0	1	0	1
0	0	1	1	0	1
0	1	0	0	0	0
0	1	1	0	0	0
1	0	0	0	0	0
1	0	1	0	1	1
1	1	0	0	0	0
1	1	1	0	1	1

Tab. 12.11 Wahrheitstabelle zu Abb. 12.2.

NAND und NOR als Universalbausteine

Sind diese vorgestellten Gatter bzw. Verknüpfungen alle gleichwertig? Nein! Aus mehreren NAND- bzw. NOR-Gattern kann man alle anderen logischen Gatter aufbauen. Dies gilt für die anderen Gattertypen nicht.

Frage: Warum nicht?

Es kann praktische Vorteile haben, nur einen Gattertyp zu verwenden, z. B. weil es gerade besonders schnelle oder günstige kommerzielle Realisierungen gibt.

Beispiel NOT

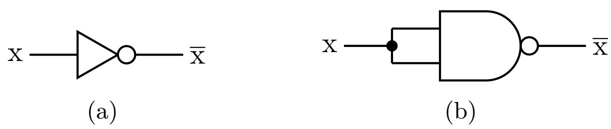


Abb. 12.3 (a) NOT-Gatter. (b) Äquivalente Schaltung mit einem NAND-Gatter.

Es gilt

$$\overline{X} \hat{=} \underbrace{\overline{XX}}_{X} = \overline{X}$$

Beispiel UND

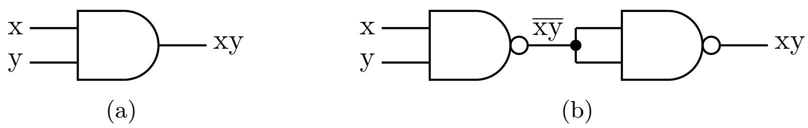


Abb. 12.4 (a) UND-Gatter. (b) Äquivalente Schaltung mit NAND-Gattern.

$$XY \quad \hat{=} \quad \overline{\overline{XY}\,\overline{XY}} = \overline{\overline{XY}} = XY$$

Beispiel ODER

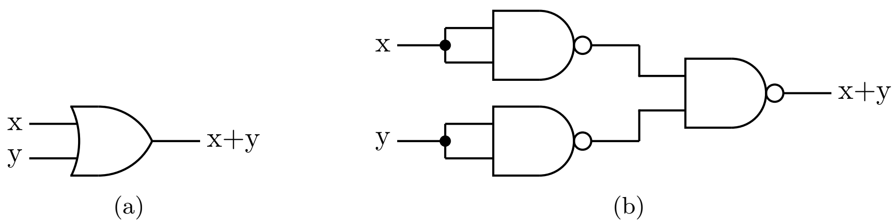


Abb. 12.5 (a) ODER-Gatter. (b) Äquivalente Schaltung mit NAND-Gattern.

$$
\begin{aligned}
X + Y \quad \widehat{=} \quad \overline{\overline{XX}\,\overline{YY}} &= \overline{\overline{X}\,\overline{Y}} && |\text{De Morgan'sche Regel} \\
&= \overline{\overline{X}} + \overline{\overline{Y}} \\
&= X + Y
\end{aligned}
$$

12.5 Elementare Schaltkreise für Logikfunktionen

Es gibt zahlreiche Logikfamilien, das heißt integrierte Schaltkreise, die die obigen Logikfunktionen realisieren. Die Gatter einer Familie kann man, wie im obigen Beispiel, sehr flexibel, baukastenartig miteinander kombinieren und fast beliebige Schaltungen herstellen. Die Logikfamilien unterscheiden sich zuerst in den verwendeten Bausteinen, z.B. NAND oder NOR, und auch in den Spannungspegeln, die die Zustände „1“ und „0“ repräsentieren.

Daraus ergeben sich dann die relevanteren Unterschiede in der Verlustleistung, der Schaltgeschwindigkeit, bzw. der Verzögerung zwischen Eingangs- und Ausgangssignal, dem Preis, et cetera.

Ein Beispiel für die Zuordnung von logischen Zuständen und Spannungspegeln in der sogenannten Transistor-Transistor-Logik (TTL) ist in Abb. 12.6 dargestellt. Die Zuordnung ist für Gattereingänge und -ausgänge leicht unterschiedlich. Eingangsspannungen zwischen 2 V und 5 V werden als logische „1“ gewertet. Eingangsspannungen zwischen 0 V und 0,8 V als logische „0“. Dazwischen liegt ein nicht zulässiger Spannungsbereich, der zu einem undefinierten Logikzustand führen würde.

Anmerkung Die erlaubten Ausgangsspannungen sind restriktiver definiert als die Eingangsspannung. Warum?

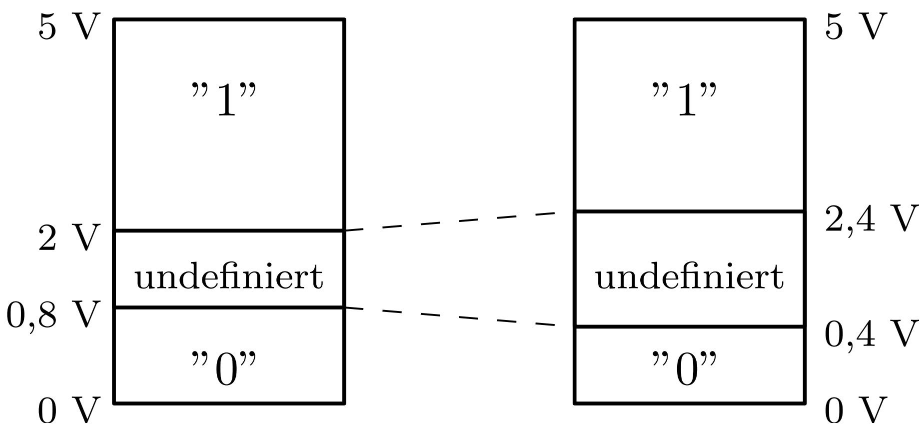


Abb. 12.6 Spezifikation der Spannungspegel für TTL.

Die drei wichtigsten Logikfamilien sind

- CMOS
- TTL
- ECL

Die Schaltkreise in komplementärer MOSFET-Technologie (CMOS) sind am leichtesten zu verstehen. TTL und ECL sind Familien, die auf Bipolartransistoren basieren. Es gibt von jeder Familie zahlreiche Varianten, die z.B. auf hohe Schaltgeschwindigkeiten oder geringen Stromverbrauch optimiert sind.

Im Folgenden sind exemplarisch einige Schaltkreise dieser Familien erklärt.

CMOS-Logikgatter

Die logische „1“ entspricht in der CMOS-Logik einer Spannung nahe der Versorgungsspannung U_{DD}, während die logische „0“ der Masse entspricht. Man versteht die folgenden CMOS-Schaltungen sehr leicht, wenn man den FET als einfachen spannungsgesteuerten Schalter auffasst. In den folgenden Schaltungen liegt die Source der NMOS-Transistoren immer auf Masse und die Source der PMOS-Transistoren auf U_{DD}. Dadurch ist die Steuerspannung U_{GS} festgelegt.

Für U_G = „1“ ist der NMOS-Schalter geschlossen (U_{GS} = 1 V), für U_G = „0“ (U_{GS} = 0 V) geschlossen (siehe Tabelle 12.12). Für U_G = „1“ ist der PMOS-Schalter offen (U_{GS} = 0 V), für U_G = „0“ (U_{GS} = -1 V) geschlossen.

	$\mathbf{U}_{DS}$	$\mathbf{U}_G$	$\mathbf{U}_{GS}$	$\mathbf{I}_{DS}$	Region
NMOS	1 V	„1“	1 V	maximal	A
	1 V	„0“	0 V	minimal	B
	0 V	„0“	0 V	minimal	C
	0 V	„1“	1 V	minimal	C
PMOS	-1 V	„0“	-1 V	maximal	A’
	-1 V	„1“	0 V	minimal	B’
	0 V	„1“	0 V	minimal	C’
	0 V	„0“	-1 V	minimal	C’

Tab. 12.12 Das Schaltverhalten von NMOS- und PMOS-Transistoren für einen generischen FET mit Versorgungsspannung U_{DD} = 1 V.

Inverter

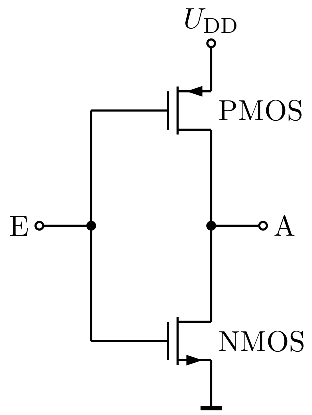


Abb. 12.7 Inverter in CMOS-Logik.

Falls am Eingang E der High-Pegel anliegt, also E = „1“, dann ist der NMOS-Transistor leitend mit Masse verbunden und der PMOS-Transistor sperrt. Der Ausgang A liegt also auf Masse. Allerdings fliesst kein Strom, da für den NMOS gilt $U_{DS} = 0$. Falls dagegen am Eingang E der Low-Pegel anliegt, also E = „0“, dann leitet der PMOS-Transistor und der NMOS sperrt. Damit liegt A auf der Versorgungsspannung. Auch hier fliesst kein Strom.

Anmerkung: Wir gehen immer davon aus, dass kein Strom in den Ausgang A fließt, der typischerweise mit einem weiteren hochohmigen Gattereingang verbunden ist. Dies gilt für Gatter aus FETs.

In Tabelle 12.13 sind auch die hypothetischen Zustände E = A = „0“ und E = A = „1“ dargestellt. Diese Zustände sind nicht zulässig, da nicht gleichzeitig ein Transistor einen Maximalstrom leiten und der andere sperren kann (es fliesst kein Strom in den Ausgang A).

E/A	FET	$\mathbf{U}_{DS}$	$\mathbf{U}_{GS}$	$\mathbf{I}_{DS}$	Kommentar
E=0; A=0	NMOS	0 V	0 V	min.	C
	PMOS	-1 V	-1 V	max.	A’
E=0; A=1	NMOS	1 V	0 V	min.	B
	PMOS	0 V	-1 V	min.	C’
E=1; A=0	NMOS	0 V	1 V	min.	C
	PMOS	-1 V	0 V	min.	B’
E=1; A=1	NMOS	1 V	1 V	max.	B
	PMOS	0 V	0 V	min.	C’

Tab. 12.13 Das Schaltverhalten der NMOS- und PMOS-Transistoren des Inverters für hypothetische Werte des Ausgangs A.

Den Inverter kann man auch mit nur einem Transistor, einem NMOS oder einem PMOS, und einem Widerstand realisieren (Abb. 12.8). Dann fließt allerdings immer für einen der Zustände Strom durch den Transistor.

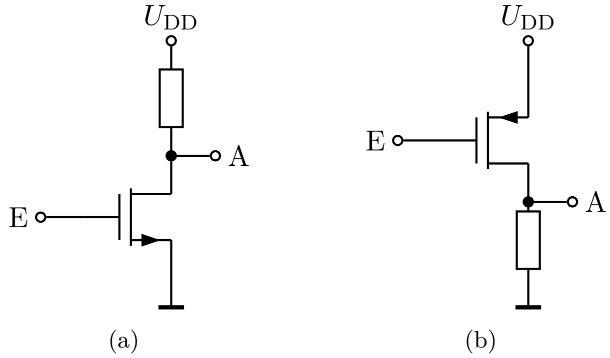


Abb. 12.8 (a) Inverter in NMOS-Logik, (b) Inverter in PMOS-Logik.

NAND-Gatter

Die NAND-Gatterschaltung (Abb. 12.9) besteht aus vier Transistoren, zwei NMOS und zwei PMOS. Die PMOS-Transistoren sind parallel geschaltet, die NMOS-Transistoren in Serie. Wie beim CMOS-Inverter ist jeder Eingang mit einem NMOS- und einem PMOS-Transistor verbunden. Auch hier ist der Drain der PMOS-Transistoren mit dem Drain der NMOS-Transistoren verbunden, so dass nur ein minimaler Ruhestrom fließt. Nur wenn beide Eingänge „1“ sind, ist der Ausgang über die Kette der NMOS-Transistoren mit der Masse verbunden. Ist auch nur ein Eingang „0“, so leitet der mit ihm verbundene PMOS und der Ausgang A liegt auf U_{DD}.

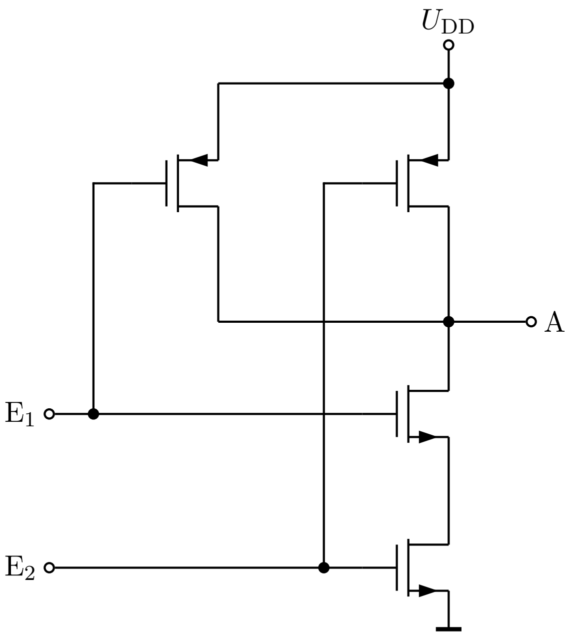


Abb. 12.9 NAND in CMOS-Logik.

$\mathbf{E_1}$	$\mathbf{E_2}$	**A**
0	0	1
0	1	1
1	0	1
1	1	0

Tab. 12.14 Wahrheitstabelle eines NAND oder 0→1-Gatters (Abb. 12.9).

NOR-Gatter

Das Konzept des NOR-Gatter (Abb. 12.10) entspricht dem des NAND-Gatters, nur sind jetzt die PMOS in Serie und die NMOS parallel geschaltet. Daraus folgt: Wenn nur an einem Eingang der High-Pegel anliegt, sperrt ein PMOS-Transistor und es besteht keine leitende Verbindung zu U_{DD}. Der Ausgang A liegt auf Masse, da die NMOS parallel geschaltet sind und mindestens ein NMOS leitet. Sind aber beide Eingänge „0", so sperren beide NMOS und leiten beide PMOS. Der Ausgang A liegt auf „1".

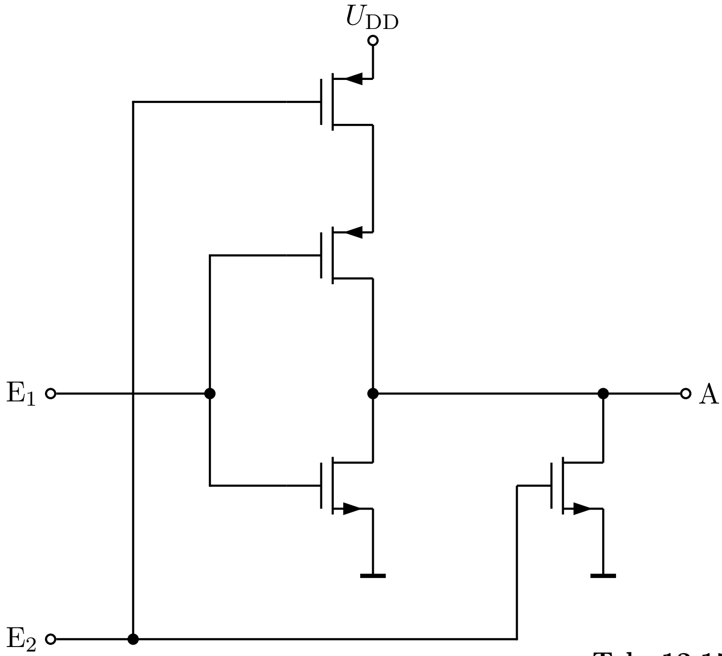


Abb. 12.10 NOR in CMOS-Logik.

$\mathbf{E_1}$	$\mathbf{E_2}$	**A**
0	0	1
0	1	0
1	0	0
1	1	0

Tab. 12.15 Wahrheitstabelle eines NOR oder 1→0-Gatters (Abb. 12.10).

Tristate

Es gibt neben den Zuständen „1" und „0" einen dritten Zustand, den sogenannte *tristate*. Hier ist der Ausgang des Bausteins hochohmig. Das heißt, er belastet die Ausgangsleitung bzw. weitere an ihr angeschlossene Ein- oder Ausgänge nicht und zieht die Leitung nicht auf ein definiertes Potential. Dies ist z. B. dann wichtig, wenn mehrere Gatterausgänge parallelgeschaltet sind oder mehrere digitale Datensender auf einer Busleitung liegen (Abb. 12.11).

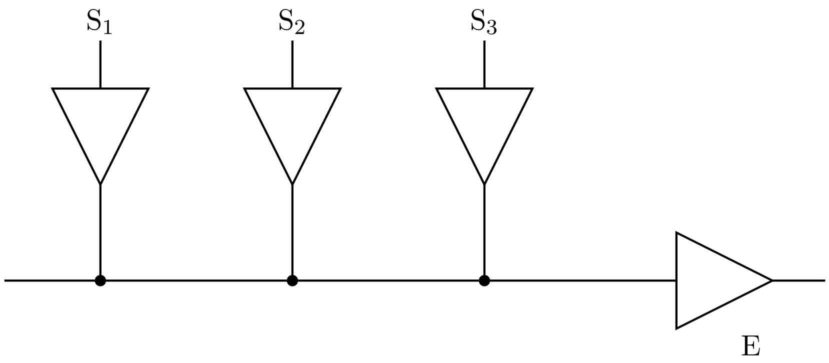


Abb. 12.11 Illustration eines Bussystems mit einem Empfänger E und drei Sendern S_1, S_2 und S_3 auf einer gemeinsamen Datenleitung.

Transmissionsgatter

Das Transmissionsgatter (Abb. 12.12) schaltet, abhängig vom Steuersignal C, den Eingangszustand zum Ausgang durch oder nicht. Im letzteren Fall ist der Ausgang hochohmig. Ein CMOS Transmissionsgatter besteht nur aus zwei Transistoren, einem NMOS- und einem PMOS-Transistor und einem zusätzlichen Inverter mit zwei weiteren Transistoren.

Die Schaltung ist raffinierter als sie auf den ersten Blick aussieht. Man versteht sie am besten, wenn man alle möglichen Eingangszustände einzeln durchgeht. Für C = „1“ am Gate des NMOS, liegt $\overline{\mathrm{C}}$ = „0“ am Gate des PMOS. In diesem Fall leiten entweder NMOS oder PMOS (oder beide) je nach dem Wert der Eingangsspannung. Der Eingangszustand wird also auf den Ausgang übertragen. Für C = „0“, sperren beide Transistoren und der Ausgang wird hochohmig (Tab. 12.16). Die Schaltung funktioniert bi-direktional, man kann Eingang und Ausgang vertauschen. Es gibt keine einfache äquivalente Schaltung mit Bipolartransistoren.

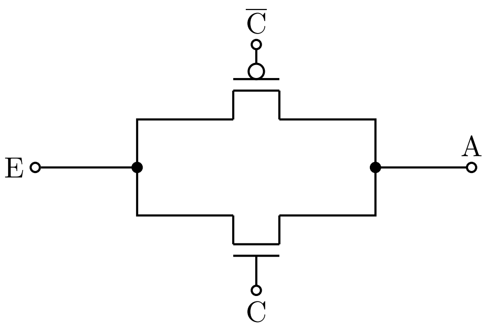


Abb. 12.12 CMOS-Transmissionsgatter: Der Inverter des Steuereingangs C ist nicht dargestellt.

E	**C**	**A**
0	0	Z
1	0	Z
1	1	1
0	1	0

Tab. 12.16 Wahrheitstabelle zu Abb. 12.12. Z steht für hochohmige Ausgänge.

Leistungsverbrauch von CMOS-Logik

Da MOSFETs einen sehr hohen Eingangswiderstand haben, kann man viele Gatter parallelschalten, ohne den Ausgang nennenswert zu belasten. Abbildung 12.13 entspricht aus Sicht des linken Gatters einer kapazitiven Last C am Ausgang.

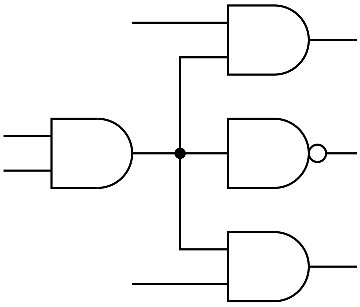


Abb. 12.13 Eingangsgatter mit drei parallelgeschalteten Gattern am Ausgang [15].

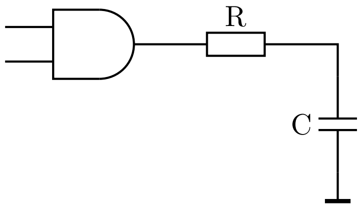


Abb. 12.14 Ersatz der parallelgeschalteten Gatter in 12.13 durch eine Kapazität C.

Durch den Leistungswiderstand R und die kapazitive Last C kommt es zu einer Signalverzögerung und zu einer Verlustleitung durch das Umladen von C. Die Werte von R und C haben erheblichen Einfluss auf das Schaltverhalten. Zum einen begrenzt $\tau = RC$ die Schaltgeschwindigkeit des Gatters; sind die Werte RC für unterschiedliche Pfade, die zu einem gemeinsamen Gatter führen, verschieden, so kann dies zu Logikfehlern führen. Zum anderen bestimmt die kapazitive Last die Verlustleistung.

Wie groß ist die Verlustleistung? Wir betrachten zuerst den Fall einer einmaligen Entladung des Kondensators mit $U(t) = U_0\, e^{-\frac{t}{RC}}$. Dies entspricht einem Energieverbrauch E

$$E = \int_0^\infty U(t)\, I(t)\, dt = \int_0^\infty \frac{U^2}{R}\, dt = \frac{U_0^2}{R} \int_0^\infty e^{-\frac{2t}{RC}}\, dt = -\frac{U_0^2}{R} \frac{RC}{2} \left[e^{-\frac{2t}{RC}}\right]_0^\infty = \frac{1}{2} C\, U_0^2 \quad .$$

Der Energie zum Aufladen des Kondensators beträgt auch $\frac{1}{2}\, C\, U_0^2$. Ein Lade-Entlade-Zyklus erfordert eine Energie von $C\, U_0^2$. Bei mehrfachen Zyklen nimmt die Verlustleistung P deshalb mit der Rate f der Schaltvorgänge zu. Es gilt

$$P = f\, C\, U_0^2 \quad . \tag{12.15}$$

Die Verlustleistung nimmt also linear mit der Frequenz f, linear mit der Lastkapazität C und quadratisch mit der Versorgungsspannung U_0 zu.

Will man eine kleine Verlustleistung erzielen, muss man entweder

1. die Taktfrequenz f klein halten.
   Dies geht aber oft nicht, da man meistens auch schnell sein will.

2. die Versorgungsspannung U_0 klein halten.
   Dies erfordert die Nutzung moderner Transistortechnologien.

3. oder die Lastkapazität C klein halten.
   Dies schränkt die Anzahl der parallel geschalteten Gatter ein. Mit der Wahl der Technologie wird auch C bestimmt (Kap. 9).

Frage: Warum ist die Verlustleistung unabhängig von R?

TTL-Logikgatter*

Die Familie TTL ist fast obsolet wird aber im Praktikum und in kleineren Laboraufbauten noch verwendet. TTL-Logik ist wie CMOS-Logik positiv, d. h., der Zustand „1“ entspricht einer positiven Spannung nahe der Versorgungsspannung. Der Zustand „0“ entspricht einer kleinen Spannung nahe Masse (Abb. 12.6).

TTL-Logik ist mit ~ 1 bis 2 ns Laufzeitverzögerung[6] eher schnell. Der Stromverbrauch ist nicht klein. Typisch sind ~ 2 bis 10 mW pro Gatter. Dafür ist aber der Stromverbrauch relativ konstant und unabhängig vom Eingangs- oder Ausgangszustand. TTL-Bausteine können substantiellen Strom liefern[7] oder schlucken[8]: 8 mA in „0“ (z. B. zum Betrieb einer Leuchtdiode), 400 μA in „1“.

Beispiel: TTL Buffer

Wozu ist ein Buffer (Abb. 12.15) gut?

Der Buffer erneuert die Spannungspegel, so dass diese gut in die Spezifikation von TTL passen.

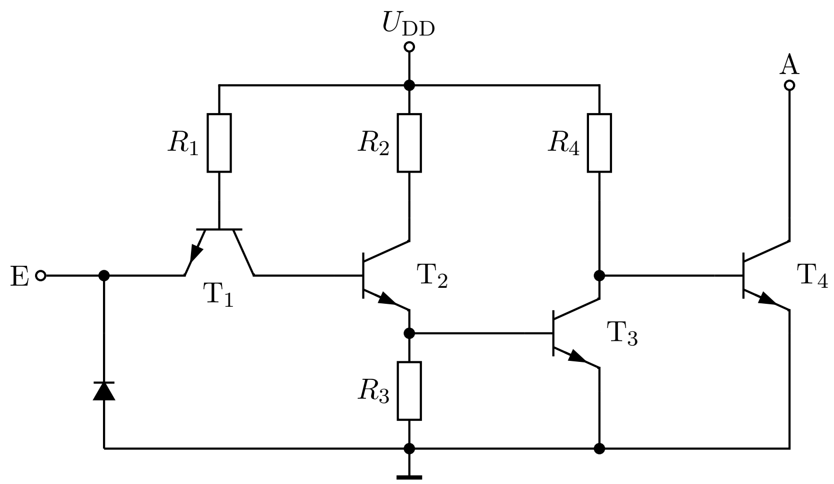


Abb. 12.15 Aufbau eines TTL-Buffers.

Im Vergleich mit CMOS-Logik sind hier folgende Eigenarten bemerkenswert:

1. Es werden nur npn- und keine pnp-Transistoren verwendet.

[6]Engl. *propagation delay*
[7]Engl. *source*
[8]Engl. *sink*

2. Der Ausgang A ist ein sogenannter Open-Collector-Ausgang. Hier muss weitere Logik angehängt werden bzw. A noch über einen Widerstand an Versorgungsspannung angeschlossen werden, damit das Gatter funktioniert.

3. Liegt der Eingang E auf Masse, so fließt Strom in die Eingangsschaltung hinein.

   Anmerkung: Der Eingang E entspricht dem Ausgang A eines vorangehenden Gatters.

4. Der Transistor T_1 wird entweder in Vorwärts- oder Rückwärtsrichtung betrieben. Er ersetzt hier zwei antiseriell geschaltete Dioden.

D ist eine Schutzdiode gegen negative Eingangsspannung und entsprechende Ströme.

Diese Schaltung ist deutlich komplizierter als der entsprechende Buffer in CMOS-Logik.

Wir unterscheiden die zwei Fälle E = „0“ und E = „1“:

E = „0“:
Hier liegt E auf Masse und Strom fließt von der Versorgungsspannung U_{DD} durch R_1 in E und nicht in die Basis von T_2. Also sperrt T_2 und damit sperrt auch T_3. Strom fließt über R_4 in die Basis von T_4 und nicht durch T_3. Damit stellt A eine Stromsenke dar und A = „0“.

Im Prinzip könnte gleichzeitig die Basis-Kollektor-Diode von T_1 leiten, aber da dann U_{BE} von $T_2 \geq 0,7$ V sein müsste (Die Basis von T_1 liegt auf 0,7 V.), geht das nicht.

E = „1“:
Die Basis-Emitter-Diode sperrt. Strom fließt durch R_1 in die Basis von T_2. T_1 wird also in Rückwärtsrichtung betrieben und die Basis-Kollektor-Diode von T_1 ist durchlässig. Also leitet T_2 und entsprechend T_3. Der Spannungsabfall an R_4 ist groß. Damit wird U_{BE} von T_4 so klein, dass T_4 sperrt. Damit schwimmt das Potential von A, und A geht bei Anschluss an U_{DD} über einen Widerstand auf „1“.

In beiden Fällen fließt durch die Diode D kein Strom. Der Transistor T_4 fungiert als Inverter.

Beispiel: TTL NAND

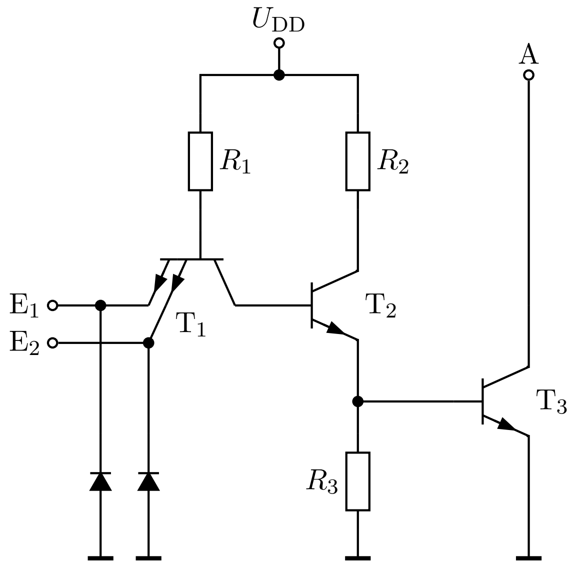


Abb. 12.16 Aufbau eines NAND-Gatters (0→1-Gatter) in TTL.

Das NAND-Gatter in TTL-Logik (Abb. 12.16) zeigt eine weitere Besonderheit, einen Transistor T_1 mit mehreren Emittern. T_1 leitet nur dann Strom, wenn mindestens ein Emitter auf Masse liegt bzw. wenn $U_{BE} \approx 0,7\,\mathrm{V}$ ist.

Falls E_1 oder E_2 gleich „0“ ist, fließt Strom von U_{DD} durch R_1 in E_1 bzw. E_2 und nicht in die Basis von T_2. Damit sperren T_2 und T_3. Der Ausgang A schwimmt und ist nicht leitend mit Masse verbunden. (Es kann hier kein Strom aus der nächsten Stufe abfließen.)

Falls umgekehrt E_1 oder E_2 gleich „1“ ist, fließt Strom von U_{DD} über R_1 in die Basis von T_2. Damit leiten T_2 und T_3. Der Ausgang A liegt auf Masse. Hier kann Strom aus der nächsten Stufe abfließen.

Open-Collector

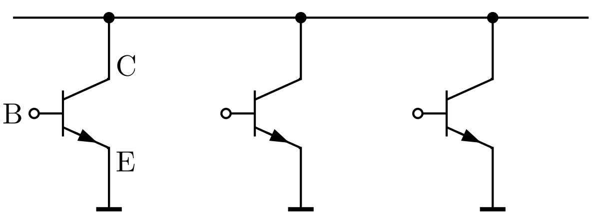


Abb. 12.17 Drei Transistoren in einer Open-Collector-Anordnung.

In Abbildung 12.17 sind drei Transistoren in einer Open-Collector-Anordnung dargestellt. Nur wenn der gemeinsame Kollektorausgang (über einen Widerstand) an eine Spannungsquelle angeschlossen ist, ist die Schaltung funktionstüchtig.

ECL*

Die schnellste Logikfamilie ist die **Emittergekoppelte** Logik (ECL). Mit ihr können Schaltgeschwindigkeiten unter einer Nanosekunde verwirklicht werden. Dies erreicht man dadurch, dass die Transistoren nicht in Sättigung gehen. Der zulässige Arbeitsbereich ist dementsprechend relativ klein. Es gilt für leitende Transistoren $U_{CE} \geq 0,6\,\mathrm{V}$ bzw. $U_C > U_B$. Auch der Potentialunterschied zwischen dem High-Pegel, also „1“, und dem Low-Pegel, also „0“, ist mit $\approx 0,9\,\mathrm{V}$ gering. Der Stromverbrauch ist zwar hoch, aber dafür, wie auch bei TTL, relativ unabhängig von der Schaltungsfrequenz.

ECL wird mit einer negativen Versorgungsspannung von $-5,2\,\mathrm{V}$ betrieben. Der High-Pegel beträgt $-0,9\,\mathrm{V}$, der Low-Pegel entspricht $-1,8\,\mathrm{V}$.

Ein NOR-Gatter in ECL ist in Abbildung 12.18 dargestellt. Die Schaltung hat zwei Ausgänge A_1 und A_2. Der Ausgang A_1 ist der NOR-Ausgang, der Ausgang A_2 ist der ODER-Ausgang.

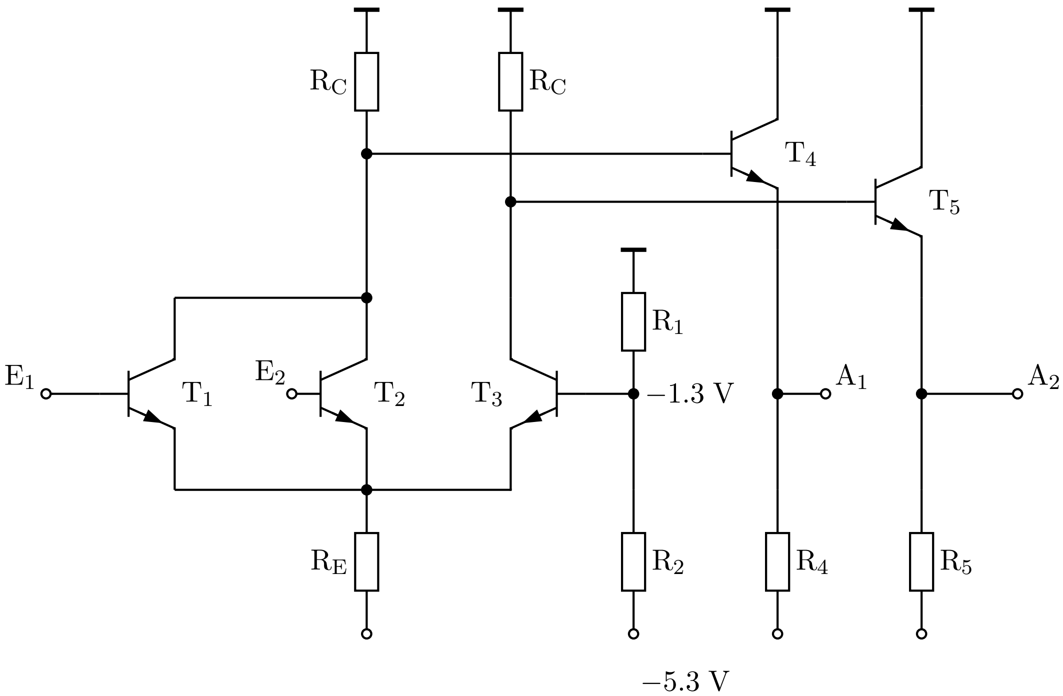


Abb. 12.18 ODER-NOR-Gatter in ECL.

$\mathbf{E}_1$	$\mathbf{E}_2$	$\mathbf{A}_1$ NOR (1→0)	$\mathbf{A}_2$ ODER (1→1)
0	0	1	0
0	1	0	1
1	0	0	1
1	1	0	1

Tab. 12.17 Wahrheitstabelle zu Abb. 12.18.

Die Transistoren T_1 bzw. T_2 und T_3 sind an ihren Emittern miteinander verbunden, daher die Bezeichnung *Emittergekoppelte Logik.*

Der Differenzverstärker, bestehend aus T_1 bzw. T_2 und T_3, ist so dimensioniert, dass der Strom im Wesentlichen nur durch einen Ast - also T_1 bzw. T_2 oder T_3 - fließt. Der andere Ast sperrt. Mögliche Werte von R_C, R_E und $R_4 = R_5$ sind 220 Ω, 780 Ω und 510 Ω.

T_4 und T_5 sind Emitterfolger, die durch die Ausgänge des Differenzverstärkers T_1 bzw. T_2 und T_3 gesteuert werden. Die Basisspannung von T_3 ist über den Spannungsteiler R_1, R_2 fest eingestellt.

Die Eingänge E_1 und E_2 entscheiden, wohin der Strom I_1 fließt und zwar durch R_1 und R_2.

Falls $E_1 = E_2 =$ „0“, sperren T_1 und T_2, T_3 schaltet durch. Entsprechend fällt an R_C im $T_1 - T_2-$Ast keine Spannung ab, an R_C im T_3-Ast dagegen schon. Damit leitet T_4 und T_5 sperrt. Der Ausgang A_1 liegt entsprechend auf „1“ und A_2 auf „0“.

Falls dagegen $E_1 = E_2 =$ „1“, leiten T_1 und T_2. T_3 sperrt. Die Basisspannung U_B an T_4 sinkt und A_1= „0“. Entsprechend gilt A_2= „1“.

Die Fälle $E_1 =$ „0“, $E_2 =$ „1“ und $E_1 =$ „1“, $E_2 =$ „0“ entsprechen der Bedingung $E_1 = E_2$ = „1“, da T_1 und T_2 parallel geschaltet sind.

Die wesentlichen Eigenschaften von CMOS, TTL und ECL sind in Tabelle 12.18 zusammengestellt.

	CMOS	**TTL**	**ECL**
„1“	U_{DD}	5 V	-0,9 V
„0“	0 V	0 V	-1,8 V
Geschwindigkeit	gering: 3 - 30 ns	hoch: 1 - 10 ns	sehr hoch: 0,5 - 2 ns
Leistung	hoch: 0,5 mW/MHz	hoch: 1 - 10 mW	sehr hoch [8]: 35 - 70 mW
Technologie	CMOS	bipolar	bipolar
Versorgungsspannung	5 V, 3,3 V, ...	5 V	-5,2 V

Tab. 12.18 Eigenschaften von CMOS-, TTL- und ECL-Logik.

[8] Unabhängig von der Schaltfrequenz!

CMOS-Gatter in Technologien mit Strukturgrößen $\leq 130\,\mathrm{nm}$ sind deutlich schneller als oben angegeben und brauchen aufgrund ihrer geringeren Versorgungsspannungen deutlich weniger Leistung.

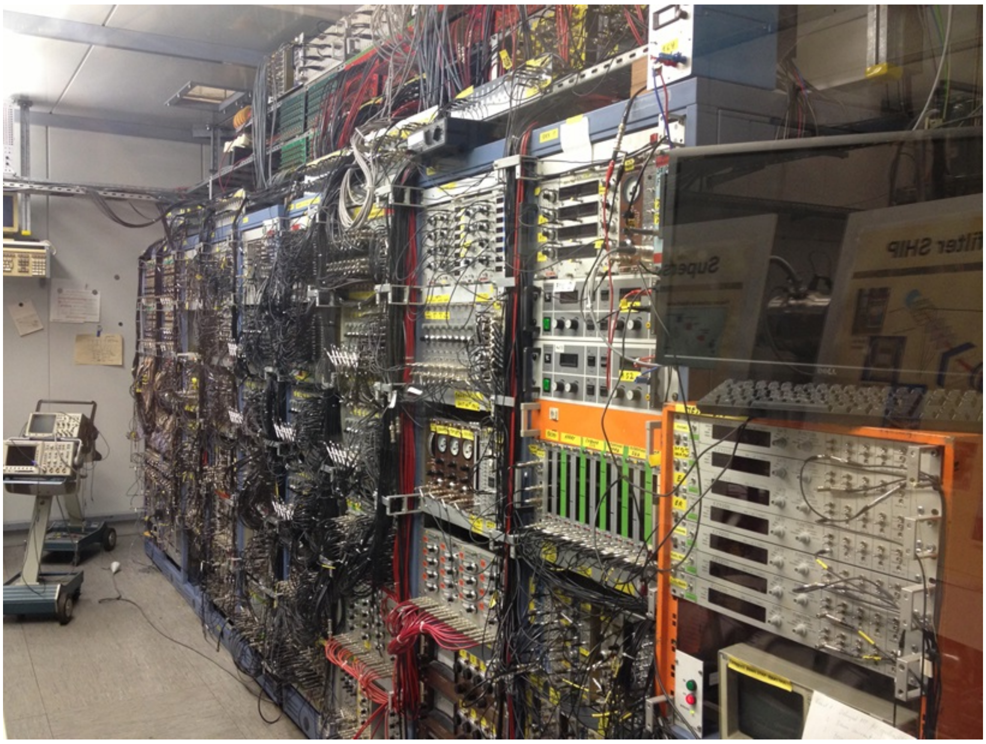


Abb. 12.19 Elektronikmodule und Logikbausteine zur Suche nach schweren Elementen an der GSI.

13 Kombinatorische Logik

Aus den grundlegenden Logikgattern, die wir im vorigen Kapitel kennengelernt haben, lassen sich allgemeine und komplexere Logikbausteine zusammenstellen. Dies sind z. B. Addierer und Multiplizierer, Dekodierer und Komparatoren. Allen diesen Bausteinen ist gemeinsam, dass der Ausgangszustand („das Ergebnis“) nur von den momentanen Eingangszuständen abhängt, genau wie bei den elementaren UND-, ODER-Gattern, etc. Solche Bausteine werden auch *Schaltnetze* genannt.

13.1 Addierer

Zur Addition von Binärzahlen beliebiger, aber fester Stellenzahl verwendet man Halbaddierer und Addierer.

Ein Halbaddierer ist ein Logikgatter mit zwei Eingängen und zwei Ausgängen entsprechend der Wahrheitstabelle 13.1.

a	**b**	**c**	**s**
0	0	0	0
0	1	0	1
1	0	0	1
1	1	1	0

Tab. 13.1 Wahrheitstabelle des Halbaddierers.

Der Halbaddierer addiert die einstelligen Binärzahlen a und b. Die Ausgänge s und c stehen für die Summe s und den Übertrag c (*carry*). Dieser Wahrheitstabelle entspricht die Verknüpfung

$$\mathrm{s} = \mathrm{a} \oplus \mathrm{b} \quad \text{und} \quad \mathrm{c} = \mathrm{a}\,\mathrm{b} \quad .$$

Die Schaltung aus den entsprechenden Logikgattern ist in Abbildung 13.1 dargestellt.

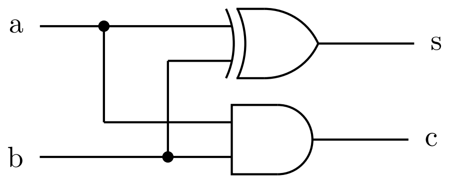


Abb. 13.1 Halbaddierer aus einem XOR- und einem UND-Gatter.

Will man mehrstellige Binärzahlen addieren, ist der Halbaddierer nicht ausreichend, da der Übertrag c_v einer vorangehenden Stelle nicht berücksichtigt werden kann. Es gilt ja nur zwei Eingänge. Ein Volladdierer hat dagegen drei Eingänge a, b, c_v und zwei Ausgänge s unc c, die entsprechend Tabelle 13.2 miteinander verbunden sind.

a	**b**	$\mathbf{c}_v$	**c**	**s**
0	0	0	0	0
0	0	1	0	1
0	1	0	0	1
0	1	1	1	0
1	0	0	0	1
1	0	1	1	0
1	1	0	1	0
1	1	1	1	1

Tab. 13.2 Wahrheitstabelle des Volladdierers.

Aufgabe: Bilden Sie die Summe der Produkte aus Tabelle 13.2 und weisen Sie nach, daß sie der folgenden Schaltung entspricht.

Ein Volladdierer lässt sich aus zwei Halbaddierern und einen ODER-Gatter zusammenbauen (Abb. 13.2).

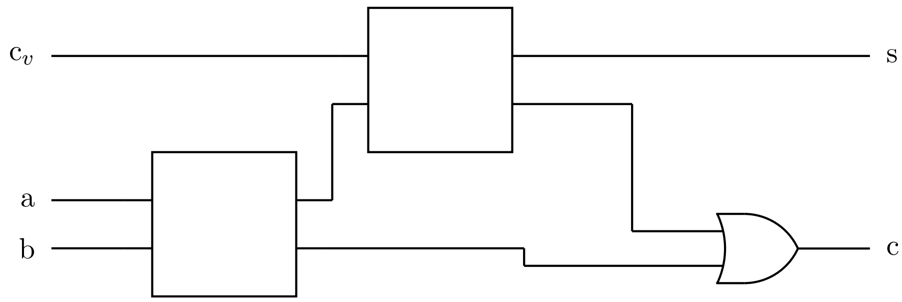


Abb. 13.2 Volladdierer aus zwei Halbaddierern und einem ODER-Gatter.

Zur Addition von vielstelligen Binärzahlen benötigt man für das LSB einen Halbaddierer, für alle anderen Stellen einen Volladdierer. Dies ist in Abbildung 13.3 am Beispiel zwischen einer 3-stelligen Binärzahl illustriert.

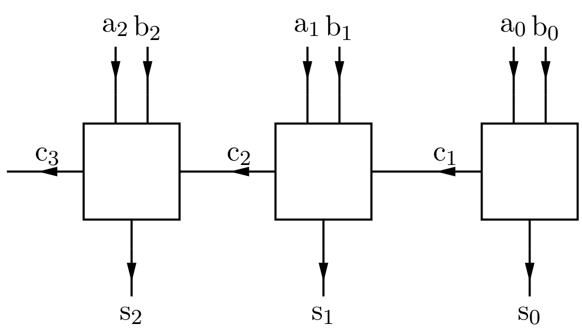


Abb. 13.3 Schaltung aus zwei Volladdierern und einem Halbaddierer zur Summation zweier 3-stelliger Binärzahlen a_2, a_1, a_0 und b_2, b_1, b_0. Eingangs- und Ausgangsvariablen werden durch die Pfeilrichtung unterschieden.

13.2 Multiplexer

Ein Multiplexer ist ein Baustein, der abhängig von einem Steuersignal eine von mehreren möglichen Eingangsleitungen mit seinem Ausgang verbindet (Abb. 13.4).

Abb. 13.4 Multiplexer mit drei Eingängen E_1, E_2, E_3, einer Steuerleitung S und einem Ausgang A (links). Demultiplexer mit einem Eingang E, einer Steuerleitung S und drei Ausgängen (rechts).

Anmerkung: Ein Demultiplexer ist der dazu komplementäre Baustein.

Ein Multiplexer (MUX) ist ein wichtiger und häufig verwendeter Baustein. Er kommt z. B. dann zum Einsatz, wenn die Datenraten der Eingänge nicht sehr groß sind oder die Eingänge nicht gleichzeitig senden, aber wenig Platz für mehrere paralle Datenleitungen besteht. Man kann die Eingangsdaten mittels des MUX nacheinander über eine Leitung übertragen.

Im einfachsten Fall besteht der MUX nur aus zwei Eingängen E_0, E_1 (Abb. 13.5). Selbst für diesen Fall ist die Wahrheitstabelle (Tab. 13.3) schon recht lang.

Anmerkung: Dieses Blockdiagramm werden wir noch mehrfach benötigen.

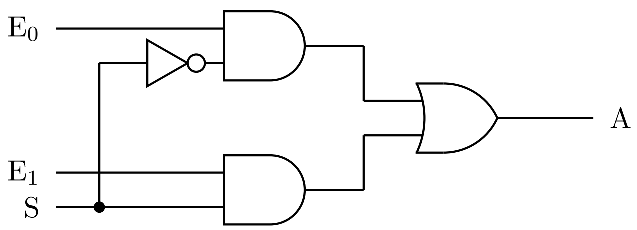


Abb. 13.5 Multiplexer aus zwei UND-Gattern, einem Inverter und einem ODER-Gatter.

$\mathbf{E_0}$	$\mathbf{E_1}$	**S**	**A**
0	0	0	0
0	1	0	0
1	0	0	1
1	1	0	1
0	0	1	0
0	1	1	1
1	0	1	0
1	1	1	1

Tab. 13.3 Wahrheitstabelle des MUX.

Ich finde es lohnend, sich einmal zu Fuß durch alle Verzweigungen dieser einfachen Schaltung zu bewegen:

Die Wahrheitstabelle entspricht der offensichtlichen Gleichung

$$\mathrm{A} = \mathrm{S}\,\mathrm{E}_1 + \overline{\mathrm{S}}\,\mathrm{E}_0 \quad . \tag{13.1}$$

Auch Gl. 13.1 ergibt sich über die DNF (Summe von Produkten).

Nebenrechnung: Hier erweist sich die Schreibweise der UND-Verknüpfung als Multiplikation und die der ODER-Verknüpfung als Addition wieder als sehr praktisch.

$$
\begin{aligned}
\overline{\mathrm{S}}\,\mathrm{E}_0\,\overline{\mathrm{E}}_1 + \overline{\mathrm{S}}\,\mathrm{E}_0\,\mathrm{E}_1 + \mathrm{S}\,\overline{\mathrm{E}_0}\,\mathrm{E}_1 + \mathrm{S}\,\mathrm{E}_0\,\mathrm{E}_1 = {} & \overline{\mathrm{S}}\,\left(\mathrm{E}_0\,\overline{\mathrm{E}}_1 + \mathrm{E}_0\,\mathrm{E}_1\right) + \mathrm{S}\,\left(\overline{\mathrm{E}}_0\mathrm{E}_1 + \mathrm{E}_0\,\mathrm{E}_1\right) \\
= {} & \overline{\mathrm{S}}\,\mathrm{E}_0 + \mathrm{S}\,\mathrm{E}_1
\end{aligned}
$$

Aufgabe: Berechnen Sie diesen Ausdruck über das Produkt der Summen.

Was macht man, falls alle Eingangsdatenleitungen gleichzeitig und kontinuierlich Daten übertragen, aber nur eine lange Übertragungsstrecke zur Verfügung steht?

Damit in keinem Fall Daten verloren gehen, muss man hier anders vorgehen und die Taktfrequenz der Ausgangsleitung erhöhen. Dann kann man die entsprechend verkürzten Eingangssignale nacheinander auf die Ausgangsleitung legen. Dies ist in Abb. 13.6 für zwei Eingangsleitungen und eine Ausgangsleitung illustiert.

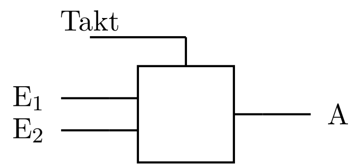


Abb. 13.6 Schaltung mit zwei Eingängen, einem Taktgeber und einem Ausgang.

Eine zufällige Signalfolge für die Eingänge E_1 und E_2 und das resultierende Ausgangssignal A ist in Tabelle 13.4 und Abb. 13.4 dargestellt. Die Eingangssignale seien z. B. mit 40 MHz getaktet. Das erfordert eine Taktung der Ausgangsleitung mit 80 MHz. Bei mehr als zwei Eingangssignalen erhöht sich die erforderliche Ausgangsfrequenz entsprechend. Die Schaltung erfordert einen PLL (*phase locked loop*) zur Verdopplung der Eingangstaktes und einen DLL zur Verschiebung der Eingangssignale gegen Ende um einen halben Takt. Sowohl DLL als auch PLL sind aufwendige und stromverbrauchende Blöcke.

Eingangstakt (40 MHz)	1	2	3	4	5	6	7	8	9	10
$\mathbf{E}_1$	0	0	1	0	0	0	1	1	0	1
$\mathbf{E}_2$	1	0	1	0	0	1	0	0	1	1
Ausgangstakt (80 MHz)	12	34	56	78	90	12	34	56	78	90
A	01	00	11	00	00	01	10	10	01	11

Tab. 13.4 Wahrheitstabelle zur Schaltung 13.6.

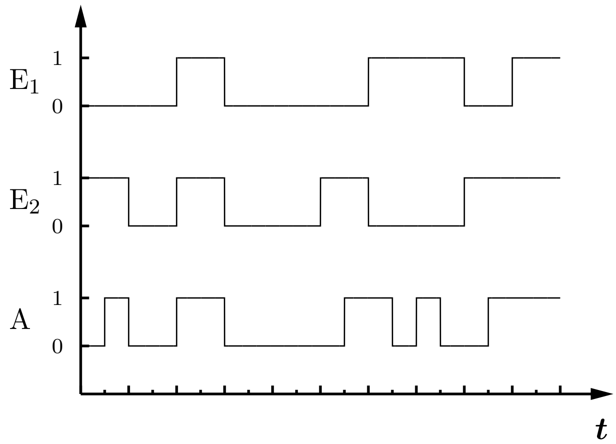


Abb. 13.7 Zeitfunktion zur Schaltung 13.6.

Einen Multiplexer kann man auch als Speicher benutzen (Abb. 13.8). Hierzu legt man die Eingänge E_i auf den gewünschten Zustand und hält sie konstant. Durch Variation der Adresseneingänge a, b kann man am Ausgang A die entsprechende Speicherinformation abrufen.

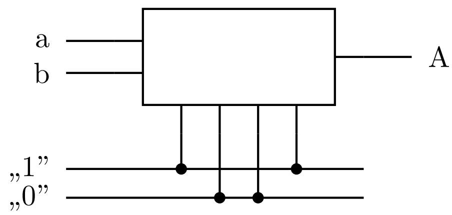


Abb. 13.8 Multiplexer als Speicher.

13.3 Decoder

Ein Decoder ist ein Logikbaustein mit wenigen Eingängen und vielen Ausgängen, typisch sind n Eingänge und 2^n Ausgänge (Binärdecoder). Man spricht von einem 3-nach-8-Decoder, 2-nach-4-Decoder etc.

Abhängig von dem Zustand der Eingangsleitungen wird genau eine der Ausgangsleitungen auf „1" gesetzt, die anderen Ausgangsleitungen liegen auf „0". Die Eingangsleitungen entsprechen also der „Adresse" der Ausgangsleitungen, die der Decoder dekodiert. Ein Blockdiagramm eines 2-nach-4-Decoders ist in Abbildung 13.9 dargestellt, die Wahrheitstabelle in Tabelle 13.5. Dieser Decoder besitzt einen zusätzlichen Eingange E („enable").

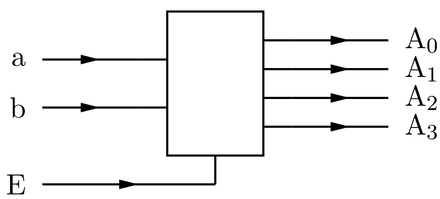


Abb. 13.9 2-nach-4-Decoder mit einem zusätzlichen *enable*-Eingang E.

a	**b**	$\mathbf{A_0}$	$\mathbf{A_1}$	$\mathbf{A_2}$	$\mathbf{A_3}$
0	0	1	0	0	0
0	1	0	1	0	0
1	0	0	0	1	0
1	1	0	0	0	1

Tab. 13.5 Wahrheitstabelle zu Abb. 13.9 für E = „1".

Die Wahrheitstabelle 13.5 gilt für E = „1“, für E = „0“ sind alles Ausgänge gleich „0“, unabhängig von a und b.

Der Schaltkreis ohne „enable“ist in Abbildung 13.10 dargestellt.

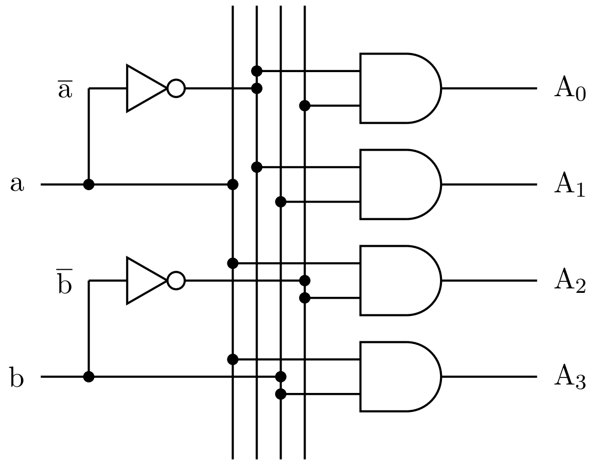


Abb. 13.10 Schaltung eines 2-nach-4-Decoders mit NICHT- und UND-Gattern.

Decoder sind vielverwendete und wichtige Bausteine. Oft wird die selektierende Ausgangsleitung zur Aktivierung eines weiteren mit ihr verbundenen Chips (*chip select*) verwendet oder sie ermöglicht den Zugriff auf einen Speicherplatz (z. B. *row select* eines RAM).

Einem Decoder mit *enable* kann man als Demultiplexer verwenden, indem man das *enable* als (die einzige) Eingangsdatenleitung verwendet. Die Adresseingänge (entsprechend der Leitung S in Abbildung 13.4) dienen der Auswahl der Ausgangsleitung.

13.4 Encoder

Ein Encoder (siehe Abb. 13.11) ist ein umgedrehter Decoder. Die Wahrheitstabelle eines 4-nach-2-Encoders ist in Tabelle 13.6 dargestellt. Dabei wird vorausgesetzt, dass genau eine Eingangsleitung auf „1“ liegt. Dieser Eingangsleitung wird am Ausgang eine entsprechende Adresse zugeordnet. Eingangszustände mit mehr als einer „1“ auf den Eingangsleitungen sind nicht zulässig bzw. führen zu nicht definierten Ausgangszuständen a, b.

Der **Prioritätsencoder** ist ein Baustein, der beliebige Werte auf den Eingängen zulässt. Hier wird die Adresse derjenigen Eingangsleitung auf „1“ ausgegeben, die die höchste Eingangsnummer hat. Liegen also z.B. sowohl $\mathbf{E}_2$ wie $\mathbf{E}_3$ auf „1“, so würde die Adresse von $\mathbf{E}_3$ ausgegeben.

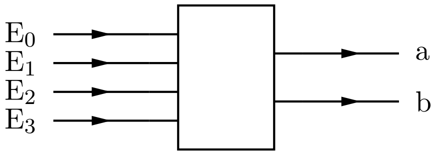


Abb. 13.11 4-nach-2-Encoder.

$\mathbf{E}_0$	$\mathbf{E}_1$	$\mathbf{E}_2$	$\mathbf{E}_3$	**a**	**b**
1	0	0	0	0	0
0	1	0	0	0	1
0	0	1	0	1	0
0	0	0	1	1	1

Tab. 13.6 Wahrheitstabelle eines 4-nach-2-Encoders.

Der **Prioritätsencoder** ist ein Baustein, der beliebige Werte auf den Eingängen zulässt. Hier wird die Adresse derjenigen Eingangsleitung auf „1“ ausgegeben, die die höchste Eingangsnummer hat. Liegen also z.B. sowohl $\mathbf{E}_2$ wie $\mathbf{E}_3$ auf „1“, so würde die Adresse von $\mathbf{E}_3$ ausgegeben.

13.5 Komparator

Ein Komparator (Abb. 13.12) vergleicht den Wert zweier Binärzahlen a, b an seinem Eingang und setzt die entsprechenden Ausgänge „$a > b$“, „$a = b$“, „$a < b$“.

Vorsicht: Hier sind a und b Zahlen und nicht logische Zustände!

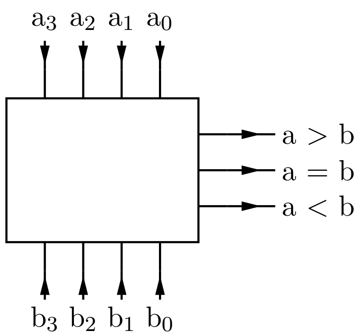


Abb. 13.12 4-Bit-Komparator.

Die Wahrheitstabelle ist für nur einstellige Zahlen (Tab. 13.7) schon umfangreich. Sie entspricht den Gleichungen

$$„a = b“ = \overline{\mathrm{a}}\,\overline{\mathrm{b}} + \mathrm{a}\,\mathrm{b} \quad , \quad „a > b“ = \mathrm{a}\,\overline{\mathrm{b}} \quad , \quad „a < b“ = \overline{\mathrm{a}}\,\mathrm{b} \quad .$$

a_0	b_0	$a > b$	$a = b$	$a < b$
0	0	0	1	0
0	1	0	0	1
1	0	1	0	0
1	1	0	1	0

Tab. 13.7 Wahrheitstabelle eines 1-Bit-Komparators mit drei Ausgängen.

Die Erweiterung auf mehrstelligen Binärzahlen ist offensichtlich. So muss man bei der Äquivalenzfeststellung $a = b$ jedes Bit a_i, b_i überprüfen und die UND-Verknüpfung aller Ergebnisse

$a_i = b_i$ bilden.

Den digitalen Komparator darf man nicht mit dem analogen Komparator (Abs. 4.3) verwechseln. Der analoge Komparator vergleicht eine beliebige Eingangsspannung mit einer extern eingestellten Schwellenspannung.

13.6 Das Karnaugh-Diagramm*

Umfangreiche Wahrheitstabellen lassen sich mithilfe von Karnaugh-Diagrammen oft schneller auf Gleichungen zurückführen als beim direkten Ausrechnen der Produktsummen oder Summenprodukte.

Wir betrachten als erstes Beispiel die Wahrheitstabelle 13.8

A	**B**	**C**	**Q**
0	0	0	1
0	0	1	1
0	1	0	0
0	1	1	1
1	0	0	0
1	0	1	0
1	1	0	0
1	1	1	1

Tab. 13.8 Vereinfachte Wahrheitstabelle zu Abb. 12.2.

1. Zuerst wird das Karnaugh-Diagramm gemalt. Dazu gibt es verschiedene Möglichkeiten. Bei mehr als zwei Eingangsvariablen muss man Variablen zusammenfassen. Wichtig ist es, die Zeilen und Spalten im Gray Code anzuordnen, also in der Reihenfolge 00, 01, 11, 10 statt 00, 01, 10, 11.

$AB\backslash C$	0	1
00	1	1
01	0	1
11	0	1
10	0	0

Abb. 13.13 Karnaugh-Diagramm der Wahrheitstabelle 13.8

2. Die Werte der Ergebnisvariablen werden in die so gebildete Matrix eingetragen. Hier Q.

3. Jetzt werden alle Rechtecke gefunden und markiert, die vollständig mit „1” gefüllt sind. Wichtig ist, die Rechtecke müssen die Größe 1, 2, 4, 8, ..., 2^n haben. Drei „1” in einer Reihe bilden kein gültiges Rechteck. Rechtecke dürfen überlappen. Es dürfen keine „1”, die nicht in einem Rechteck liegen, übrigbleiben. Die gegenüberliegenden Seiten des Gitters sind dabei topologisch anliegend. So können im folgenden Beispiel die folgenden Rechtecke gebildet werden.

$AB\backslash CD$	00	01	11	10
00	0	1	1	0
01	0	0	0	0
11	0	0	0	0
10	0	1	1	0

$AB\backslash CD$	00	01	11	10
00	1	0	0	1
01	0	1	1	0
11	0	1	1	0
10	1	0	0	1

$AB\backslash CD$	00	01	11	10
00	1	0	0	1
01	0	0	0	0
11	0	1	1	0
10	1	0	0	1

Abb. 13.14 Beispiele für gültige Rechtecke in Karnaugh-Diagrammen

Je größer die Rechtecke sind, desto einfacher werden die resultierenden Gleichungen.

4. Die den Rechtecken entsprechenden Produkte werden gebildet und summiert. Das blaue Rechteck in Abb. 13.13 entspricht dem Produkt $\overline{AB}$.

   Warum? Die „1" tritt sowohl für $C = 0$ und $C = 1$ auf. Daher muss beim Term C ein Produkt verwendet werden. Gleichzeitig ist der Wert von A und B konstant, daher gibt es einen Term A und B im Produkt. Da aber $A = 0$ und $B = 0$, werden A und B im Produkt invertiert. Das grüne Rechteck führt zu dem Term BC. Beide Variablen nehmen nur den Wert „1" an. A tritt mit „0" und „1" auf und kann ignoriert werden. Das Ergebnis entspricht Gleichung 12.14, da $\overline{A+B} + BC = \overline{AB} + BC$.

Als ein weiteres Beispiel betrachten wir Tabelle 13.3 Das Karnaugh-Diagramm ist

$E_0E_1\backslash S$	0	1
00	0	0
01	0	1
11	1	1
10	1	0

Abb. 13.15 Karnaugh-Diagramm der Wahrheitstabelle 13.3

Die beiden grauen Rechtecke entsprechen den Produkten $A = \overline{S}E_0 + SE_1$, in Übereinstimmung mit Gleichung 13.1.

Man kann das Karnaugh-Diagramm auch alternativ auswerten. Dazu modifizert man die Regel 4).

Es werden nur Rechtecke mit „0” betrachtet und Summenprodukte gebildet. Das blaue Rechteck entspricht der Summe $(S+E_0)$, da hier die Variablem mit Wert „0” in der Summe auftauchen. Variablen mit dem Wert „1” werden invertiert. Dies erklärt den Summenterm für das grüne Rechteck $(\overline{S}+E_1)$. Hier wird das Rechteck durch die beiden Felder $(E_0E_1, S) = (00, 0)$ und $(10, 1)$ auf gegenüberliegenden Rändern des Karnaugh-Diagramms gebildet. Es ergibt sich der alternative Ausdruck $A = (S + E_0)(\overline{S} + E_1)$.

Als letztes einfaches Beispiel betrachten wir eine ODER-Verknüpfung. Das Karnaugh-Diagramm (Abb. 13.16) hat zwei überlappende (graue) Rechtecke mit Inhalt „11“. Das horizontale Rechteck entspricht dem Ausdruck a, das vertikale b, insgesamt also $a + b$ wie es sein sollte.

$b\backslash a$	0	1
0	0	1
1	1	1

Abb. 13.16 Karnaugh-Diagramm der ODER-Verknüpfung

Alternativ könnte man die Produktsume $(a + b)$ bilden, indem man das rote Rechteck mit Inhalt „0“ betrachtet. Hier gibt es nur ein Summenterm, da es nur ein Rechteck gibt.

14 Sequentielle Logik

Sequentielle Logikbausteine sind Flip-Flops, Zähler und Schieberegister. Hier bestimmen nicht nur die Eingangszustände den Ausgangszustand, sondern auch vorausgehende Zustände. Sequentielle Logikbausteine speichern vorangegangene Zustände. Man bezeichnet sie auch als *Schaltwerke* oder *Zustandsmaschinen*[1]. Man unterscheidet zwischen getakteten und ungetakteten Schaltwerken.

14.1 RS-Flipflop

Ein Flipflop oder eine bistabile Kippstufe ist ein Logikbaustein mit Rückkopplung. Durch die Rückkopplung ist der Ausgangszustand nicht ausschließlich durch die Eingangszustände bestimmt, sondern auch durch den Ausgang bzw. die Ausgänge. Dadurch wird die Analyse der möglichen Ausgangszustände nicht nur komplizierter, es treten auch zwei neue Phänomene auf:

- Man kann Ausgangszustände speichern, selbst wenn sich der Eingangszustand ändert.
- Es gibt auch instabile Zustände.

Zur Illustration betrachten wir zuerst eine einfache Schaltung mit Invertern (Abbildung 14.1).

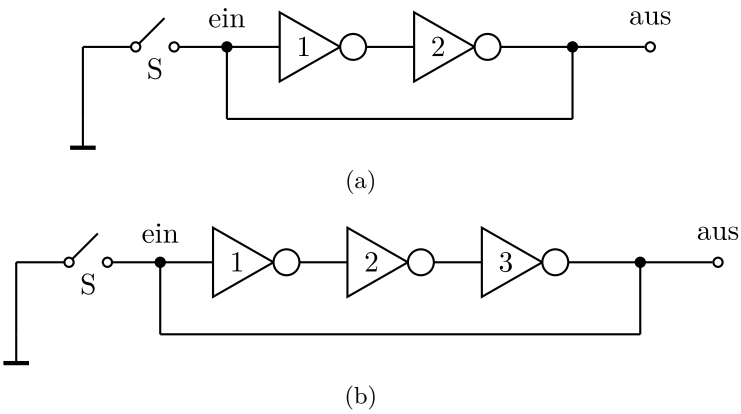


Abb. 14.1 (a) Kette aus zwei Invertern mit Rückkopplung. (b) Kette aus drei Invertern mit Rückkopplung.

Im ersten Beispiel (Abb. 14.1(a)) sind zwei Inverter hintereinander geschaltet. Der Ausgang des zweiten Inverters ist an den Eingang des ersten Inverters rückgekoppelt. Legen wir den Eingang des ersten Inverters über den Schalter S kurz auf Masse, also „0“, so tritt am Ausgang von 1 der Zustand „1“ und am Ausgang von 2 der Zustand „0“ auf. Dies entspricht dem

[1] engl. *finite state machine (FSM)*

ursprünglichen Eingangszustand beim Schließen des Schalters. Der Schalter wird nicht mehr benötigt. Die Schaltung ist stabil und speichert den Eingangswert „0“ Entsprechend würde auch ein Eingangswert „1“ stabil gespeichert.

Die Schaltung in Abbildung 14.1(b) verhält sich anders. Hier ist die Zahl der hintereinander geschalteten Inverter ungerade. Bei kurzem Schließen des Schalters S wird der Eingang des ersten Inverters in den Zustand „0“ gebracht. Der Ausgangszustand von 3 ist dagegen „1“ Der Ausgangszustand ist nicht identisch mit dem Eingangszustand und ist auch nicht zeitlich konstant. Er oszilliert zwischen „0“ und „1“ hin und her. Die Periode entspricht der sechsfachen Durchlaufzeit eines Inverters.

Das RS-Flipflop ist ein Logikbaustein aus zwei verschränkten NOR- oder NAND-Gattern (Abb. 14.2). Im Gegensatz zu den oben vorgestellten Inverterschaltungen gibt es also zwei Ausgänge und jeder Ausgang ist zu dem Eingang des anderen Gatters rückgekoppelt. Die Bezeichnungen R und S stehen für *Reset* und *Set*.

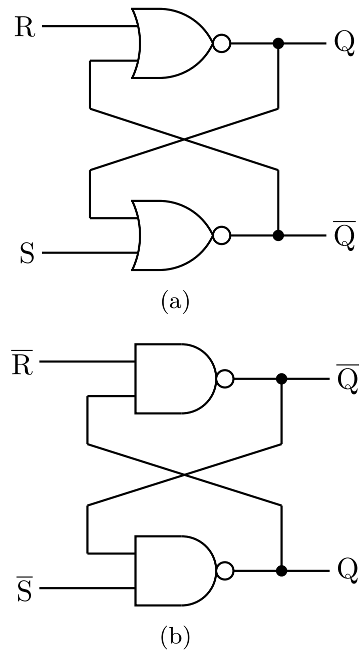


Abb. 14.2 (a) RS-Flipflop aus zwei NOR-Gattern. (b) RS-Flipflop aus zwei NAND- bzw. „$0 \to 1$“-Gattern.

Wir betrachten im Folgenden hauptsächlich die NAND-Version, die wir noch mehrfach weiterverwenden und modifizieren werden. Das RS-Flipflop hat zwei stabile Zustände, nämlich die Zustände mit $\overline{\mathrm{R}}$, $\overline{\mathrm{S}}$ = „0“, „1“ und $\overline{\mathrm{R}}$, $\overline{\mathrm{S}}$ = „1“, „0“. Hier nehmen $\overline{\mathrm{Q}}$ und Q die Werte „1“, „0“ bzw. „0“, „1“ an. Wir betrachten zuerst den Eingangszustand $\overline{\mathrm{R}}$, $\overline{\mathrm{S}}$ = „0“, „1“. Für $\overline{\mathrm{R}}$ = „0“ gilt $\overline{\mathrm{Q}}$ = „1“ unabhängig vom Wert des anderen Eingangs des oberen Gatters 1. Damit liegen am Gatter 2 die Werte $\overline{\mathrm{S}}$, $\overline{\mathrm{Q}}$ = „1“, „1“ an und deshalb wird Q = „0“.

Ist dieser Zustand stabil oder oszilliert er wie die Kette aus drei rückgekoppelten Invertern? Dazu betrachten wir noch einmal das Gatter 1: Wir kennen jetzt beide Eingangswerte $\overline{\mathrm{R}}$, Q = „0“, „0“ und wieder gilt $\overline{\mathrm{Q}}$ = „1“. Damit ist dieser Zustand stabil.

Erinnerung: Eine Null an einem der Eingänge bestimmt eindeutig den Ausgangswert eines NAND-Gatters zu „1“. Daher auch die suggestive Bezeichnung „0 → 1“-Gatter.

Entsprechend erhalten wir für den Eingangszustand $\overline{\mathrm{R}}$, $\overline{\mathrm{S}}$ = „1“, „0“ den stabilen Ausgangszustand $\overline{\mathrm{Q}}$, Q = „0“, „1“. Hier erklärt sich auch die Namensgebung $\overline{\mathrm{Q}}$ und Q. Die beiden Ausgänge sind in der Tat komplementär zueinander.

Es gibt allerdings noch weitere Zustände ($\overline{\mathrm{R}}$, $\overline{\mathrm{S}}$ = „0“, „0“ und $\overline{\mathrm{R}}$, $\overline{\mathrm{S}}$ = „1“, „1“), deren Verhalten vielleicht weniger offensichtlich ist.

$\overline{\mathrm{R}}$, $\overline{\mathrm{S}}$ = „1“, „1“:
Wir betrachten zuerst den Zustand $\overline{\mathrm{R}}$, $\overline{\mathrm{S}}$ = „1“, „1“. Dabei setzen wir der Einfachheit halber voraus, dass das Flipflop sich im Ausgangszustand $\overline{\mathrm{R}}$, $\overline{\mathrm{S}}$, $\overline{\mathrm{Q}}$, Q = „1“, „0“, „0“, „1“ befindet und wir jetzt die Leitung $\overline{\mathrm{S}}$ von „0“ auf „1“ ziehen. Dies ändert den Ausgangszustand nicht! Gehen wir von dem stabilen Zustand $\overline{\mathrm{R}}$, $\overline{\mathrm{S}}$, $\overline{\mathrm{Q}}$, Q = „0“, „1“, „1“, „0“ aus und setzen wir $\overline{\mathrm{R}}$ auf „1“, so ändert sich der Zustand auch nicht. Der vorherige Ausgangszustand ($\overline{\mathrm{Q}}_{-1}$, Q_{-1}) bleibt erhalten.

Man kann also das Flipflop als Speicher benutzen. Dazu legt man die Eingänge $\overline{\mathrm{R}}$ und $\overline{\mathrm{S}}$ hochohmig auf „1“. Durch einen kurzen Puls mit Zustand „0“ über eine externe $\overline{\mathrm{R}}$- oder $\overline{\mathrm{S}}$-Leitung kann man das Flipflop in den Zustand $\overline{\mathrm{Q}}$, Q = „1“, „0“ oder „0“, „1“ bringen. Danach sind diese Leitungen zum Erhalt des Speicherzustands nicht mehr nötig.

$\overline{\mathrm{R}}$, $\overline{\mathrm{S}}$ = „0“, „0“:
Der Zustand $\overline{\mathrm{R}}$, $\overline{\mathrm{S}}$ = „0“, „0“ ist unerwünscht. Dies hat zwei Gründe. Zwar ist der Ausgangszustand stabil ($\overline{\mathrm{Q}}$, Q = „1“), aber $\overline{\mathrm{Q}}$ und Q sind nicht komplementär zueinander. Außerdem ist beim Übergang von $\overline{\mathrm{R}}$, $\overline{\mathrm{S}}$ = „0“, „0“ zu $\overline{\mathrm{R}}$, $\overline{\mathrm{S}}$ = „1“, „1“ (dem „Ruhezustand“) nicht eindeutig definiert, welcher Wert $\overline{\mathrm{Q}}$, Q sich einstellt.

Die entsprechende Wahrheitstabelle ist in Tabelle 14.1 dargestellt.

R	**S**	$\overline{\mathbf{R}}$	$\overline{\mathbf{S}}$	$\overline{\mathbf{Q}}$	**Q**
1	1	0	0	1	1
1	0	0	1	1	0
0	1	1	0	0	1
0	0	1	1	$\overline{\mathrm{Q}}_{-1}$	Q_{-1}

Tab. 14.1 Wahrheitstabelle eines RS-Flipflops.

Eine primitive Implementierung eines RS-Flipflops auf Transistorniveau ist in Abbildung 14.3 dargestellt. Hier sind die Eingänge R, S ähnlich wie oben beschrieben im Ruhezustand über einen Widerstand von 100 kΩ auf Masse gelegt. Dieses Potential wird bei Anlegen einer positiven Spannung an R oder S „überschrieben“. (Dies ist möglich, weil R und S über einen deutlich kleineren Widerstand von 1 kΩ eingekoppelt werden.) Diese Schaltung nennt man auch bistabile Kippstufe oder bistabiler Multivibrator.

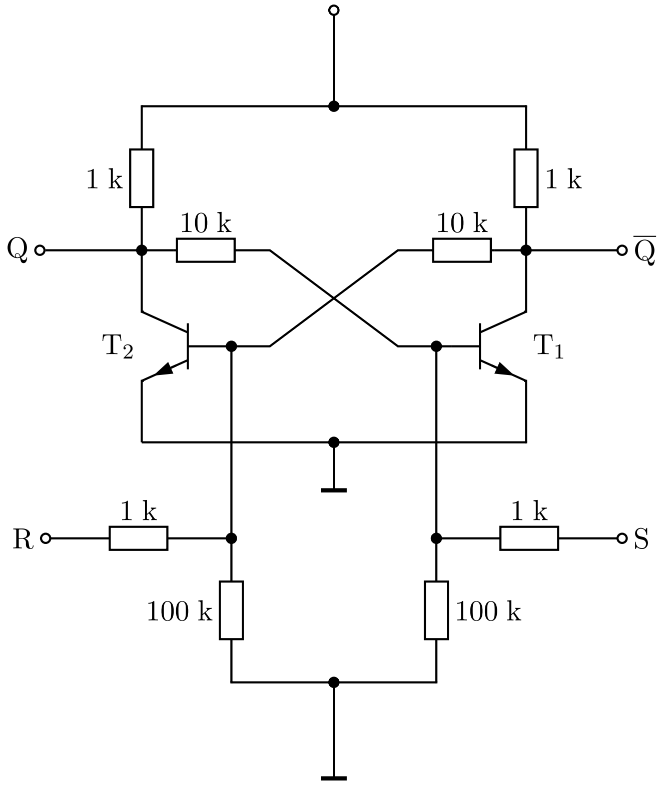


Abb. 14.3 RS-Flipflop (als bistabile Kippstufe) aufgebaut aus zwei Bipolartransistoren.

Es gibt zwei wesentliche und naheliegende Erweiterungen des elementare RS-Flipflops:

- Das Hinzufügen eines Takteingangs C (*clock*).
- Das Hinzufügen von Logik zum Vermeiden des Zustands $\overline{\text{R}}$, $\overline{\text{S}}$ = „0“.

14.2 Getaktetes RS-Flipflop

Das getaktete RS-Flipflop ist in Abbildung 14.4 dargestellt.

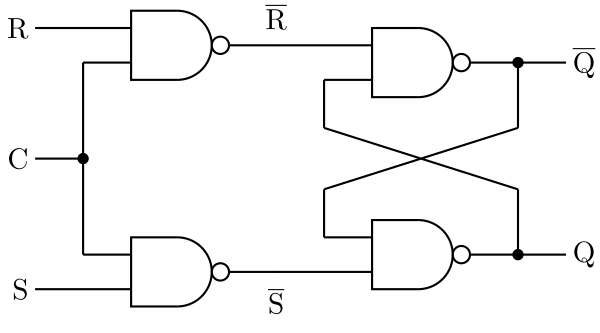


Abb. 14.4 Taktzustandgesteuertes RS-Flipflop.

Im Vergleich zu Abbildung 14.2(b) werden hier zwei zusätzliche NAND-Gatter am Eingang hinzugefügt.

Es treten drei (nicht zwei) Eingangszustände R, S und C auf. Damit hängt die Funktionsweise jetzt auch von dem Zustand des Takteingangs C ab.

Für C = „0“ sind die Ausgänge der ersten NAND-Stufe „1“ und damit gilt $\overline{\mathrm{R}}$, $\overline{\mathrm{S}}$ = „1“. Hier wird entsprechend Tabelle 14.1 des gespeicherten Zustand $\overline{\mathrm{Q}}_{-1}$, Q_{-1} erhalten.

Für C = „1“ sind die Eingänge R und S wie ihre Bezeichnung suggeriert komplementär zu $\overline{\mathrm{R}}$ bzw. $\overline{\mathrm{S}}$. Hier gilt Tabelle 14.2.

C	**R**	**S**	$\overline{\mathbf{R}}$	$\overline{\mathbf{S}}$	$\overline{\mathbf{Q}}$	**Q**
0	x	x	1	1	$\overline{\mathrm{Q}}_{-1}$	Q_{-1}
1	0	0	1	1	$\overline{\mathrm{Q}}_{-1}$	Q_{-1}
1	0	1	1	0	0	1
1	1	0	0	1	1	0
1	1	1	0	0	1	1

Tab. 14.2 Wahrheitstabelle eines taktzustandgesteuerten RS-Flipflops. „x“ steht für einen beliebigen Zustand.

14.3 D-Flipflop

Wichtiger als das RS-Flipflop ist das D-Flipflop, bei dem der Zustand R = S = „0“ bzw. $\overline{\mathrm{R}} = \overline{\mathrm{S}}$ = „1“ vermieden wird. Dies erreicht man einfach, indem man den externen Eingang R abschafft und stattdessen über einen zusätzlichen Inverter $\mathrm{R} = \overline{\mathrm{S}}$ setzt (Abb. 14.5). Der Takteingang wird beigehalten. Dieses Flipflop nennt man auch taktzustandgesteuertes D-Flipflop.

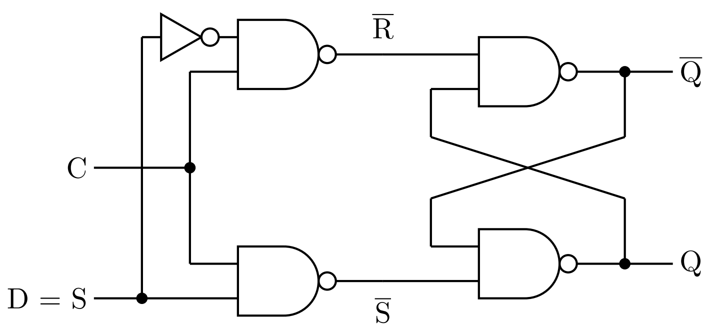


Abb. 14.5 Taktzustandgesteuertes D-Flipflop.

Die Wahrheitstabelle 14.3 vereinfacht sich gegenüber Tabelle 14.2.

C	**D**	$\overline{\mathbf{R}}$	$\overline{\mathbf{S}}$	$\overline{\mathbf{Q}}$	**Q**
0	x	1	1	$\overline{Q}_{-1}$	Q_{-1}
1	0	0	1	1	0
1	1	1	0	0	1

Tab. 14.3 Wahrheitstabelle eines taktzustandgesteuerten D-Flipflops. x steht für einen beliebigen Zustand.

Anmerkung: Man kann den Inverter aus Abbildung 14.5 auch vermeiden, indem man das D-Flipflop wie in Abbildung 14.6 aufgebaut. Hier wird das NAND-Gatter selbst als Inverter genutzt. Die beiden Alternativen ergeben dieselbe Zustandstabelle.

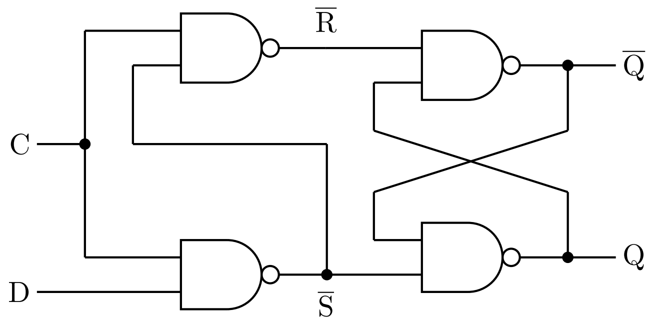


Abb. 14.6 Alternative Schaltung des taktzustandgesteuerten D-Flipflops.

Das Schaltsymbol des taktzustandgesteuerten[2] D-Flipflops ist in Abbildung 14.7 dargestellt.

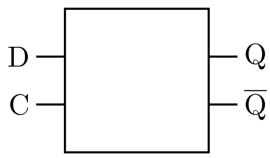


Abb. 14.7 Schaltsymbol des taktzustandgesteuerten D-Flipflops.

14.4 Flankengesteuerte Flipflops

Das oben eingeführte D-Flipflop überträgt seinen Eingangszustand bei C = „1“ sofort auf den Ausgang. Es wird deshalb auch transparentes D-Flipflop genannt. Diese Transparenz ist für viele Anwendungen ein Nachteil. So könnten kurze Spannungsstörungen („glitches“) unerwünschterweise den Zustand des Flipflops ändern.

Bei den flankengesteuerten[3] Flipflops wird der Zustand stattdessen nur während der positiven (oder negativen) Taktflanke am Eingang des Flipflops auf den Ausgang übertragen. Eventuell folgende Änderungen des Eingangszustands während C= „1“ werden ignoriert. Der Ausgang kann sich erst mit der nächsten positiven Taktflanke wieder ändern.

[2]engl. *level triggered*
[3]engl. *edge triggered*

Ein flankengesteuertes D-Flipflop kann man leicht aus einem Inverter und zwei taktzustandgesteuerten D-Flipflops aufbauen (Abb. 14.8). Es handelt sich hierbei um eine sogenannte Master-Slave-Konfiguration. Der Datenzustand D muss erst das Masterflipflop (links) durchqueren, bevor er auf den Ausgang des Slave übertragen werden kann. Beide Flipflops der Schaltung werden synchron getaktet, dabei liegt der Master am invertierten Takt $\overline{\mathrm{C}}$ und der Slave an C.

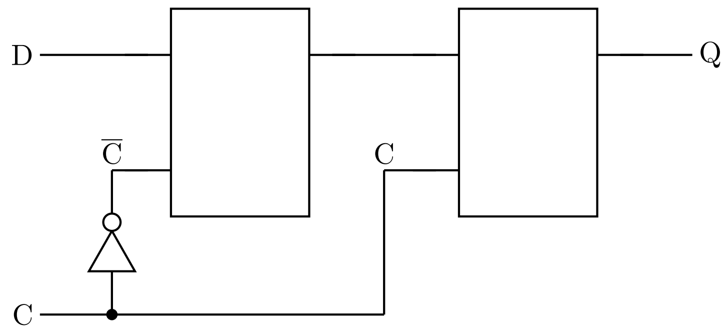


Abb. 14.8 Flankengesteuertes D-Flipflop aus zwei taktzustandgesteuerten D-Flipflops und einem Inverter.

Wie funktioniert das flankengesteuerte D-Flipflop? Zum Verständnis betrachten wir Abbildung 14.9.

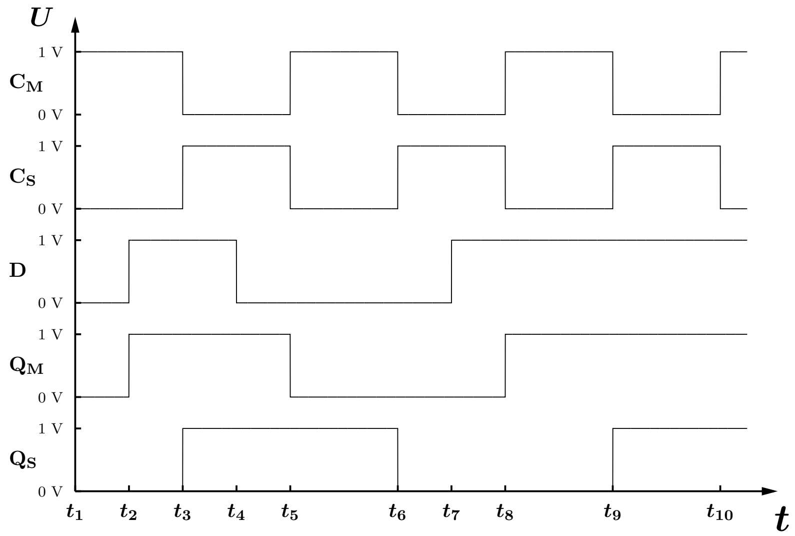


Abb. 14.9 Eingangs- und Ausgangszustände eines flankengesteuerten D-Flipflops als Funktion der Zeit.

Die Signale C_S, C_M und D sind vorgegeben mit $\mathrm{C}_M = \overline{\mathrm{C}}_S$. C_M ist der Takteingang des Masterflipflops, C_S der Takteingang des Slave. Wir machen noch die willkürliche Annahme, dass zum Zeitpunkt t_1, der Ausgang Q des Slave auf „0“ liegt.

Wie erklärt sich der Verlauf der Signale Q_M und Q_S?

t_2: D wechselt von „0“ auf „1“. Q_M folgt D, da $\mathrm{C}_M(t_2)$ = „1“. Q_S ändert sich nicht, da $\mathrm{C}_S(t_2)$ = „0“.

t_3: Jetzt wechselt C_S von „0“ auf „1“ und damit wird $Q_S(t_3) = Q_M(t_3) =$ „1“.

t_4: D springt auf „0“. Da $C_M(t_4) =$ „0“, kann Q_M nicht folgen und ändert seinen Zustand nicht. Q_S könnte Q_M folgen, da $C_S(t_4) =$ „1“. Da Q_M aber „eingefroren“ ist, ändert sich auch Q_S nicht.

t_5: Q_M wird wegen $C_M(t_5) =$ „1“ wieder freigegeben, sieht $D(t_5) =$ „0“ und folgt D. Gleichzeitig (eventuell vielleicht sogar eine Gatterlaufzeit früher), wird aber $C_S =$ „0“. Der Slave sieht den Wechsel von Q_M nicht mehr und verharrt in $Q_S =$ „1“.

t_6: Q_M wird ab t_6 auf „0“ eingefroren. Q_S kann wegen $C_S(t_6) =$ „1“ jetzt Q_M folgen, also wird $Q_S(t_6) =$ „0“.

t_7: D wechselt wieder auf „1“. Da Q_M eingefroren ist, kann sich auch Q_S nicht ändern.

t_8: Der Master wird wieder aktiv und $Q_M(t_8) =$ „1“.

t_9: Der Slave wird wieder aktiv und folgt Q_M, also $Q_S(t_9) =$ „1“.

Wenn wir nur C_S, D und Q_S betrachten ergibt sich ein sehr einfaches Bild: **Nur wenn C_S bzw. C von „0“ auf „1“ übergeht, nimmt Q_S den aktuellen Zustand von D an.** Danach bleibt der Zustand Q_S bis zur nächsten positiven Flanke von $C = C_S$ erhalten.

Das Schaltsymbol für das flankengesteuerte D-Flipflop ist in Abbildung 14.10 dargestellt

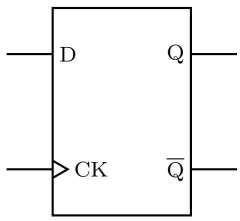


Abb. 14.10 Schaltsymbol des flankengesteuerten D-Flipflops.

Man kann das Verhalten der Schaltung auch anschaulich erklären. Sobald der Takt des Slave von „0“ auf „1“ wechselt, aktualisiert der Slave seinen Ausgangszustand Q_S. Würde sich jetzt der Eingang des Slave ändern können, so würde diese Änderung auf den Ausgang Q_S weitergegeben. Da bei $C_S =$ „1“ auch $C_M =$ „0“ gilt, ist allerdings der Zustand des Master eingefroren und eine Änderung am Eingang D des Masters hat keinen Effekt. Änderungen von D treten also verzögert auf und nur beim Übergang C_S von „0“ auf „1“ . Erst muss der Zeitpunkt $C_M =$ „1“ erreicht werden und dann $C_S =$ „1“.

14.5 Toggle-Flipflop

Ein Toggle-Flipflop bildet man aus einem flankengesteuerten D-Flipflop, indem man den Ausgang $\overline{Q}$ auf den Dateneingang D zurückkoppelt (Abb. 14.11).

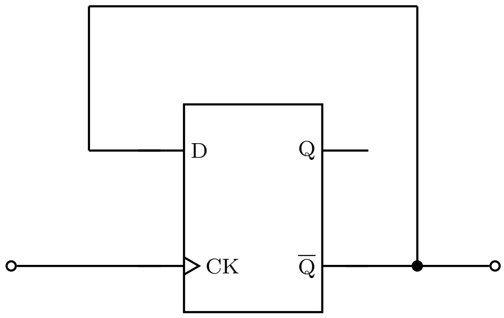


Abb. 14.11 Ein Toggle-Flipflop gebildet aus einem flankengesteuerten D-Flipflop.

Der Trick der Schaltung ist die Rückkopplung des invertierten Ausgangs $\overline{\mathrm{Q}}$ an den Eingang D. Wenn $\overline{\mathrm{Q}}$ = „1“, dann schaltet das Flipflop während einer positiven Flanke Q = „1“ und damit $\overline{\mathrm{Q}}$ = „0“.

Das Toggle-Flipflop entspricht einem Frequenzteiler. Dies ist in Abbildung 14.12 dargestellt.

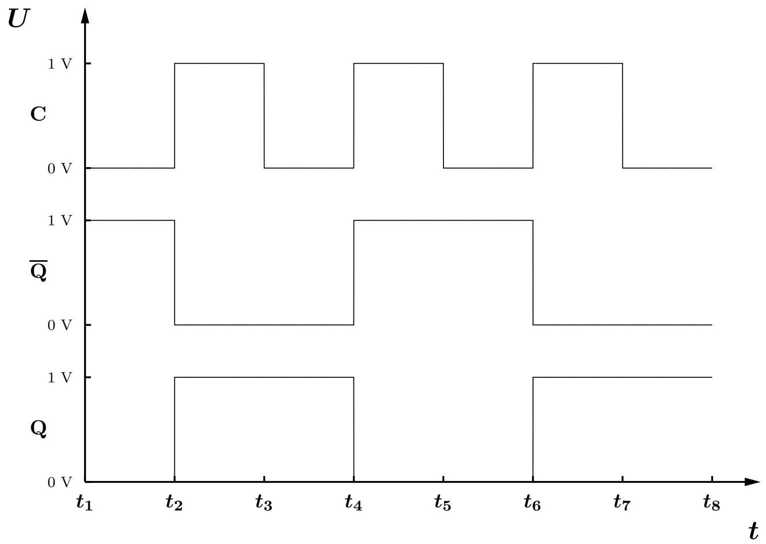


Abb. 14.12 Zeitlicher Verlauf des Zustände C, Q und $\overline{\mathrm{Q}}$ des Toggle-Flipflops aus Abb. 14.11.

Eine Erweiterung dieser Schaltung ist in Abbildung 14.13(a) dargestellt. Der Toggle-Flipflop wurde um einen zusätzlichen Toggleeingang T erweitert. Das Schaltsymbol ist in Abbildung 14.13(b) dargestellt. Über den Toggleeingang T wird gesteuert, ob intern der Ausgang Q (T=0) oder der Ausgang $\overline{\mathrm{Q}}$ (T=1) mit dem Eingang D verbunden ist.

Ist der Ausgang Q auf den Eingang D zurückgekoppelt, so ändert sich der Zustand des Flipflops nicht, D und Q sind konstant.

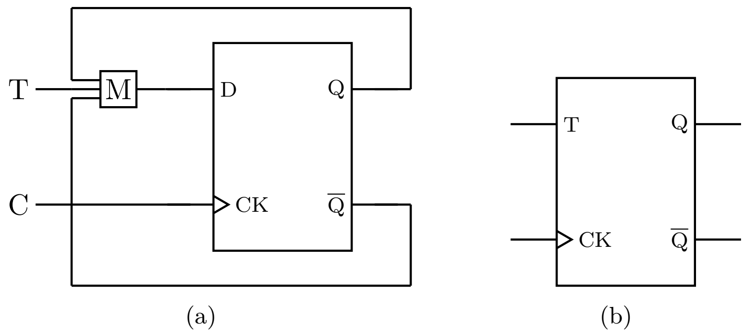


Abb. 14.13 (a) Interne Schaltung eines Toggle-Flipflops mit einem Multiplexer. (b) Schaltsymbol eines Toggle-Flipflops.

Dieses Flipflop werden wir in Abb. 14.19 zum Aufbau eines synchronen Binärzählers benutzen.

14.6 Schieberegister

Ein Schieberegister ist aus einer Kette von nichttransparenten Flipflops aufgebaut (Abb. 14.14), die an einer gemeinsamen Taktleitung (*Clock*) hängen. Der Baustein ist komplexer als er aussieht.

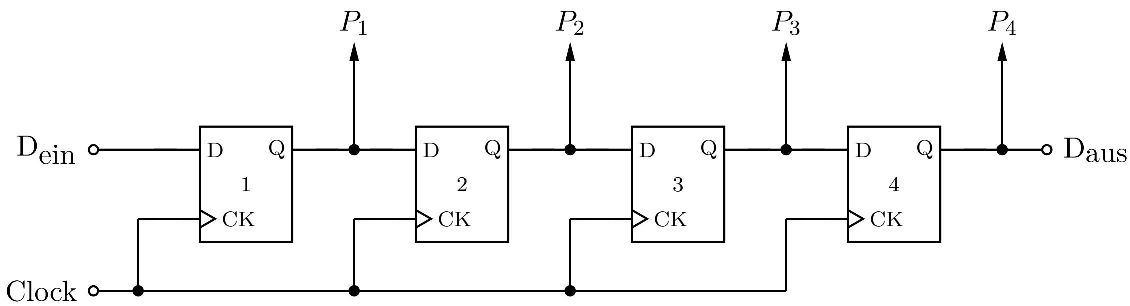


Abb. 14.14 Ein Schieberegister aus vier flankengesteuerten D-Flipflops.

Mit jeder positiven Taktflanke werden der Speicherinhalt des Flipflops i in das folgende Flipflop $i+1$ übertragen und dort gespeichert. Entsprechend wird der Inhalt des i-ten Flipflops durch den Wert des vorangehenden Flipflops überschrieben.

Oft kann man den Inhalt des Schieberegisters auch parallel durch die Ausgänge P_1 - P_4 auslesen.

Wir betrachten eine Datenfolge „1101“ am Eingang des Flipflops 1. Das LSB liegt zuerst am Eingang an, das MSB zuletzt. Es braucht vier Takte, um diese Bitfolge vollständig in das Schieberegister Flipflop 1 - 4 zu übertragen. Dies ist in Tabelle 14.4 dargestellt. Hier gehen wir davon aus, dass der Anfangszustand des Schieberegisters „0000“ ist.

Takt	$\mathbf{P}_1$	$\mathbf{P}_2$	$\mathbf{P}_3$	$\mathbf{P}_4$
0	0	0	0	0
1	1	0	0	0
2	0	1	0	0
3	1	0	1	0
4	1	1	0	1

Tab. 14.4 Wahrheitstabelle eines Schieberegisters aus vier flankengesteuerten D-Flipflops (Abb. 14.14).

Warum wird der Ausgang eines Gatters während der positiven Flanke nur auf den direkten Nachfolger und nicht alle folgenden Gatter übertragen? Dies liegt daran, dass jedes Flipflop nur mit einer gewiessen Verzögerung den Ausgangzustand ändert. Diese Verzögerung muß so groß sein, dass die positive Flanke schon vorbei ist, bevor der Ausgang das kommende Flipflop beeinflußt.

Hausaufgabe: Schauen Sie das Datenblatt eines Flipflops an, z.B. des 74VHCT74A.

Das Schieberegister ist ein erstaunlich nützlicher Baustein. So ist es eine einfache Verzögerungsschaltung und Zwischenspeicher. Solange ein Bit das Schieberegister nicht verlassen hat, kann man seinen Wert über die Ausgänge P_i auslesen. Man muss nur wissen, an welcher Position i es sich gerade befindet.

Beispiel: Datenzwischenspeicher

Der Zwischenspeicher des Semiconductor Tracker (SCT)-Detektors des Experiments ATLAS am Large Hadron Collider (LHC) ist 132 Plätze tief und ist mit 40 MHz getaktet. Das entspricht $25\,\mathrm{ns} \cdot 132 \approx 3,3\,\mu\mathrm{s}$ Speicherzeit. In dieser Zeit muss eine Triggerentscheidung (*L1 accept*) des Gesamtexperiments vorliegen, die ein gespeichertes Ereignis als interessant klassifiziert und die Auslese des Ereignisses veranlasst. Die Zwischenspeicherung ist deshalb sinnvoll, weil die meisten Ereignisse verworfen werden. Das Datenvolumen ist viel zu groß ($\approx 1/4$ Pb/s), um die Daten ohne Vorselektion auslesen und speichern zu können.

Beispiel: Deserialiser

Ein Schieberegister kann auch zur Umwandlung eines seriellen in einen parallelen Datenstroms benutzt werden. Dazu liest man die Daten nicht seriell am Ausgang D_{aus}, sondern parallel an den Ausgängen P_1, P_2, P_3 und P_4 aus. Für die Auslese kann man mit einer vierfach kleineren Taktfrequenz auskommen.

Es gibt auch Schieberegister, die man parallel einlesen kann.

Wichtig ist vielleicht auch, dass man durch Anhalten des Taktes, die im Schieberegister enthaltenen Daten dauerhaft speichern kann.

In Abbildung 14.14 ist ein 4-stufiges 1-Bit-Schieberegister dargestellt, aber es gibt auch Schieberegister mit mehreren parallel gespeicherten Bits.

Beispiel: Schnelle Rechenoperationen

Einen Schieberegister kann man auch für Rechenoperationen einsetzen. So ist die Multiplikation einer Binärzahl mit Zweierpotenzen z. B. $0101_2 \cdot 10_2 = 1010_2$, entsprechend $5_{10} \cdot 2_{10} = 10_{10}$, schnell durch Anhängen einer „0“ und Verschieben der 0101_2 um eine Position in einem Schieberegister zu realisieren.

Beispiel: Mustererkennung

Manchmal ist es nötig, aus einen Datenstrom eine festgelegte Abfolge von Bits zu erkennen. Oft markiert diese Bitfolge z. B. 111011 den Beginn (Header) oder das Ende eines in sich abgeschlossenen Datensatzes. Auch hier ist ein Schieberegister hilfreich (Abb. 14.15). Der Ausgang des UND-Gatters ist nur dann „1“, wenn das gesuchte Bitmuster auftritt.

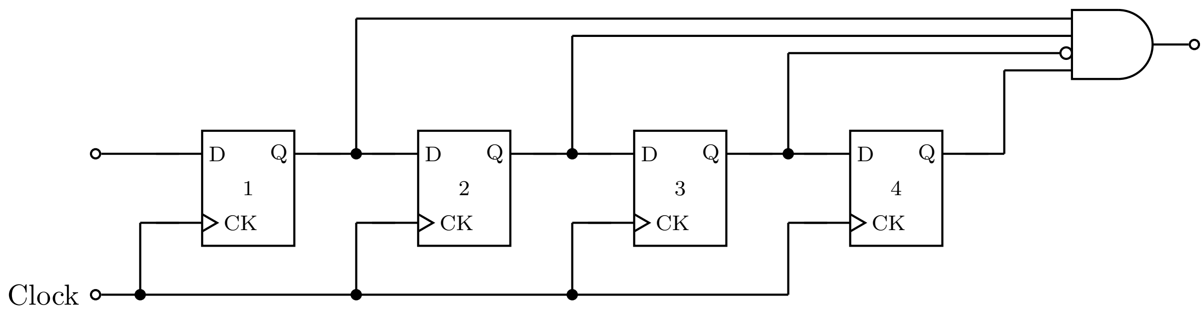


Abb. 14.15 Ein Schieberegister mit einem UND-Gatter und einem Invertierer (hier symbolisiert durch den Kreis am UND-Gatter) zur Erkennung der Bitfolge 1101.

14.7 Zähler

Ein Zähler registriert die Anzahl der Zustände „1“, die mit der Zeit an seinem Eingang anliegen, und verändert dementsprechend seinen Ausgang.

Ein anderer Zahlentyp, der **Ringzähler** ähnelt dagegen mehr einer Stoppuhr. Er zählt auf ein Startkommando kontinuierlich nach oben. Erreicht er das Ende seines Wertebereichs, beginnt er wieder von vorne.

Ein einfacher Ringzähler besteht nur aus einem Schieberegister mit einer gemeinsamen Takt- und Clearleitung (Abb. 14.16).

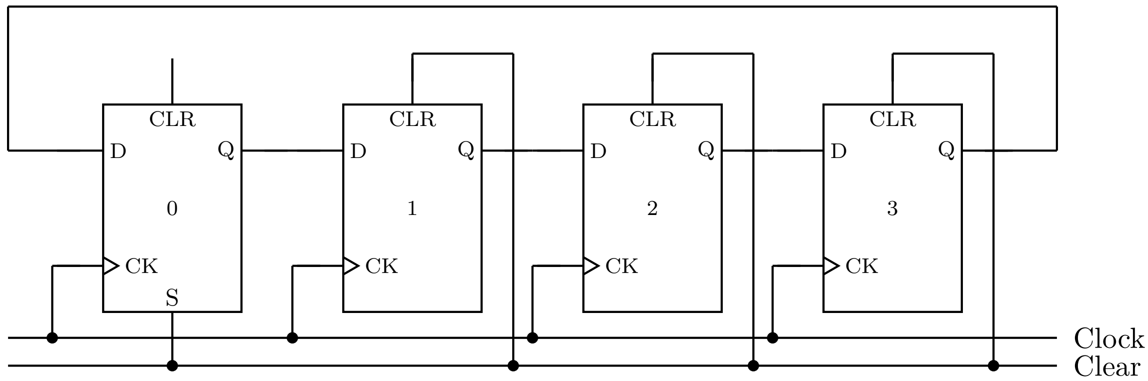


Abb. 14.16 Ein Schieberegister als Ringzähler. Hier werden noch zwei weitere nützliche Eingänge benutzt, nämlich Clear- und Set (CLR und S).

Es ist offensichtlich, daß man das Schieberegister initialisieren möchte. Da der Eingang D den Flipflop 0 mit dem Ausgang Q des Flipflop 3 verbunden ist, ist D dazu nicht geeignet. Deshalb benutzt man hier einen zusätzlichen Eingang Set (S).

Anmerkung: Eine „1“ auf Set setzt den Ausgang Q des Flipflops auf „1“, unabhängig von C. Eine „1“ auf Clear setzt Q = 0. Eine „0“ auf Set oder Clear hat keinen Effekt.

Der Ringzähler wird durch eine kurze „1“ auf der Clearleitung initialisiert. In Abbildung 14.16 entspricht dies dem Zustand 1000. Dies liegt daran, dass der Flipflop 0 die „1“ der Clearleitung auf seinem Seteingang führt, die Flipflops 1 - 3 aber auf den Cleareingang.

Mit jedem Taktzyklus wird die „1“ des Flipflops 0 an die folgenden Flipflops weitergereicht. Durch die Rückkopplung von Flipflop 3 an Flipflop 0 wird in den ersten drei Zyklen Flipflop 0 wieder auf „0“ gesetzt. Die Wahrheitstabelle ist in Tabelle 14.5 dargestellt.

Takt	$\mathbf{P}_0$	$\mathbf{P}_1$	$\mathbf{P}_2$	$\mathbf{P}_3$
0	1	0	0	0
1	0	1	0	0
2	0	0	1	0
3	0	0	0	1

Tab. 14.5 Wahrheitstabelle des Ringzählers in Abb. 14.16.

Nach dem Taktzyklus 3 springt der Ringzähler wieder in seinen Ausgangszustand 1000.

Von den 16 möglichen Ausgangszuständen P_0, P_1, P_2, P_3 nutzt der Ringzähler nur 4. Das ist etwas ungeschickt. Eine raffiniertere Version des Ringzählers ist der **Johnsonzähler** (Abb. 14.17).

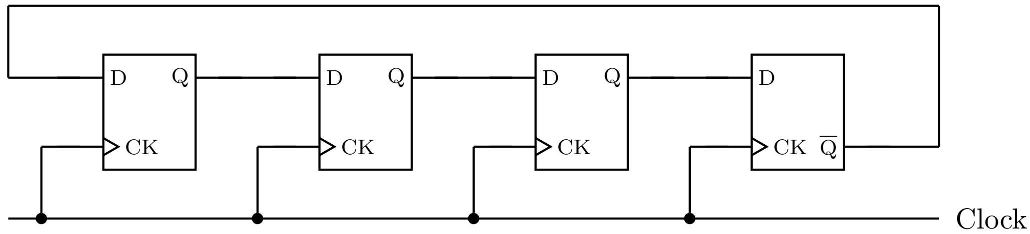


Abb. 14.17 Der Johnsonzähler. Die Clearleitung wurde übersichtshalber nicht gezeigt. Sie entspricht Abb. 14.16.

Hier wird nicht der Ausgang Q des letzten Flipflops des Schieberegisters an den Eingang D des ersten Flipflops gelegt, sondern der Ausgang $\overline{Q}$. Initialisiert man den Johnsonzähler in den Zustand 1000, so nimmt er nach dem folgenden Takt den Zustand 1100 an. Dies liegt daran, dass der Zustand Q_4 = „0“ nicht an den Flipflop 0 weitergereicht wird, sondern der Zustand $\overline{Q}_4$ = „1“. Die Wahrheitstabelle (Tab. 14.6) ergibt sich damit zu:

Takt	$\mathbf{P}_0$	$\mathbf{P}_1$	$\mathbf{P}_2$	$\mathbf{P}_3$
0	1	0	0	0
1	1	1	0	0
2	1	1	1	0
3	1	1	1	1
4	0	1	1	1
5	0	0	1	1
6	0	0	0	1
7	0	0	0	0

Tab. 14.6 Wahrheitstabelle des Johnsonzählers (Abb. 14.17).

Nach dem Taktzyklus 7 springt der Johnsonzähler wieder in seinen Ausgangszustand 1000.

Hier werden immerhin die Hälfte der möglichen 16 Zustände des Zählers genutzt.

Einen **Binärzähler** kann man mithilfe einer Kette von D-Flipflips aufbauen, die als Toggle-Flipflops geschaltet sind (Abb. 14.18).

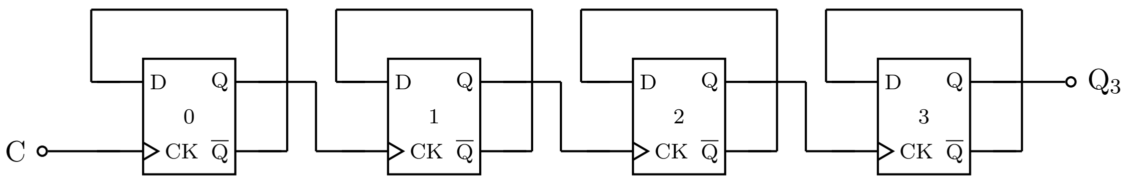


Abb. 14.18 Asynchroner Binärzähler aus vier Toggleflipflops.

Hier ist der Ausgang Q eines jeden Flipflops mit dem Takteingang C des folgenden Flipflops verbunden.

Dieser Aufbau entspricht einer Kette von Frequenzhalbierern. Die Zustände sind in Tabelle 14.7 angegeben. Die Schaltung ist leicht zu verstehen, wenn man sich die Funktionsweise des Toggleflipflop in Erinnerung ruft.

Takt	Q_3	Q_2	Q_1	Q_0
0	0	0	0	0
1	0	0	0	1
2	0	0	1	0
3	0	0	1	1
4	0	1	0	0
5	0	1	0	1
6	0	1	1	0
7	0	1	1	1
8	1	0	0	0
9	1	0	0	1
10	1	0	1	0
11	1	0	1	1
12	1	1	0	0
13	1	1	0	1
14	1	1	1	0
15	1	1	1	1

Tab. 14.7 Wahrheitstabelle des asynchronen Binärzählers (Abb. 14.18).

Dieser Zähler ist asynchron, da es keine gemeinsame Taktleitung gibt. Das heißt, die Flipflops ändern ihren Zustand sequentiell. Der Zustand des Flipflops 1 ändert sich erst nach der Schaltzeit von Flipflop 0, da der Ausgang des Flipflops i am Takteingang als Flipflop it1 liegt. Dies führt zu einer manchmal störenden Verzögerung.

Ein **synchroner Binärzähler** ist in Abbildung 14.19 dargestellt. Hier sind alle Flipflops mit einer Taktleitung verbunden. Dafür sind die Toggleeingänge der Flipflops jetzt nicht mehr parallel geschaltet und nur der erste Flipflop hat einen dauerhaft auf „1“ gesetzten Toggleeingang.

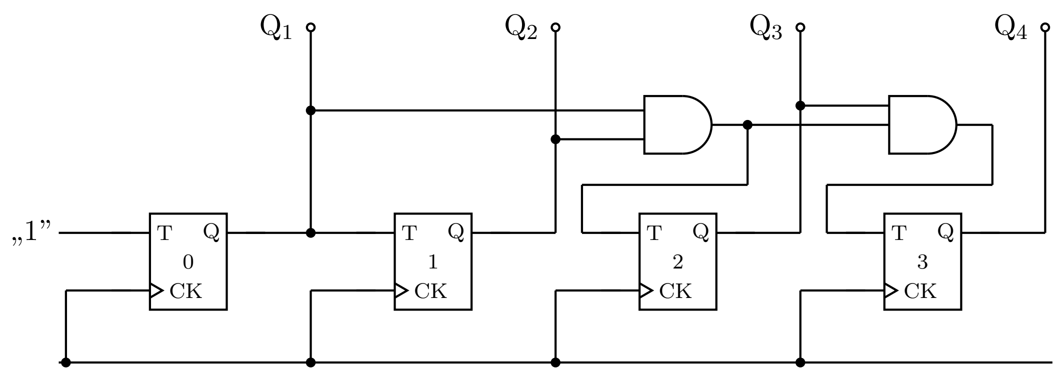


Abb. 14.19 Synchroner Binärzähler aus vier Toggleflipflops und zwei UND-Gattern (siehe Abb. 14.13).

Wie funktioniert diese Schaltung?

Flipflop 0 ist ein Frequenzhalbierer und schaltet unabhängig von allen anderen Flipflops mit jeder positiven Taktflanke zwischen „0“ und „1“ hin und her.

Das Toggle-Flipflop hat die Eigenschaft (siehe Abs. 14.5), seinen Ausgang Q nur dann zu ändern, wenn T = „1“ und falls eine positive Taktflanke anliegt. Bei T = „0“ bleibt Q unverändert.
Jeder folgende Flipflop schaltet genau dann auf „1“, wenn im vorherigen Taktzyklus alle vorangehenden Flipflops auf „1“ liegen, z. B. 0111 → 1000. Für Flipflop 1 bedeutet dies, Q_0 = „1“. Für Flipflop 2, $Q_0 = Q_1$ = „1“, also $Q_0 \cdot Q_1$ = „1“. Dementsprechend sind in Abbildung 14.19 zusätzlich zu den Flipflops zwei UND-Gatter eingebaut.

14.8 Zustandsdiagramme*

Zustandsdiagramme sind sehr hilfreich beim Entwurf von Schaltungen. Beispiele für Zustandsdiagramme sind in den Abbildungen 14.20 bis 14.23 zusammengestellt.

2-bit-Schieberegister

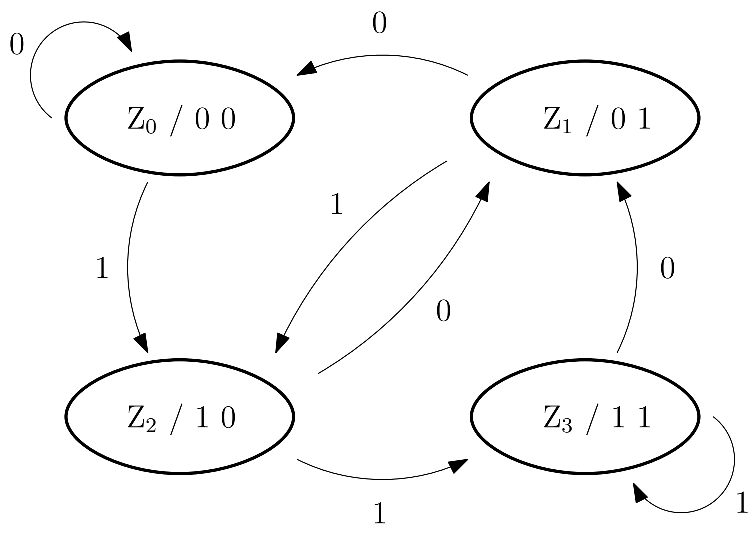


Abb. 14.20 Zustandsdiagramm eines 2-bit-Schieberegisters.

Das Diagramm stellt alle möglichen Zustände $Z_0 - Z_3$ als Ellipsen dar. Die Zustände entsprechen den Inhalten des Schieberegisters. Abhängig vom neuen Bit, das mit dem nächstem Taktzyklus in das Schieberegister geschoben wird, ändert sich der Zustand und Inhalt des Schieberegisters. Die möglichen Übergänge werden durch die Pfeile symbolisiert. Es gibt auch neue Bits, die den Zustand nicht ändern. Eine Zustandstabelle dieses Diagramms sieht folgendermaßen aus:

$\mathbf{A}_1$	$\mathbf{A}_0$	$\mathbf{N}_1$ nb=0	$\mathbf{N}_0$ nb=0	$\mathbf{N}_1$ nb=1	$\mathbf{N}_0$ nb=1	$\mathbf{Q}_1$	$\mathbf{Q}_0$
0	0	0	0	1	0	0	0
0	1	0	0	1	0	0	1
1	0	0	1	1	1	1	0
1	1	0	1	1	1	1	1

Tab. 14.8 Zustandstabelle eines 2-bit-Schieberegisters.

Hierbei stellen A_1A_0 den alten Inhalt des Schieberegisters dar und N_1N_0 den neuen. Die Ausgangsvariablen Q_1, Q_0 entsprechen hier einfach den alten Inhalt des Schieberegisters. (Wir hätten auch den neuen Inhalt des Schieberegisters als Ausgangsvariablen wählen können.) Als Zustandsgleichung ergibt sich

$$\begin{aligned}
Q_0 &= A_0\,, \qquad Q_1 = A_1 \\
N_0 &= A_1\overline{A}_0\,\overline{nb} + A_1A_0\,\overline{nb} + A_1\overline{A}_0\,nb + A_1A_0\,nb \\
&= \overline{nb}\,A_1 + nb\,A_1 = A_1 \\
N_1 &= \overline{A}_1\overline{A}_0\,nb + \overline{A}_1A_0\,nb + A_1\overline{A}_0\,nb + A_1A_0\,nb \\
&= nb\left[\overline{A}_1\left(\overline{A}_0 + A_0\right) + A_1\left(\overline{A}_0 + A_0\right)\right] = nb\left[\overline{A}_1 + A_1\right] = nb
\end{aligned}$$

also

$$Q_0 = A_0 \quad \text{und} \quad Q_1 = A_1 \tag{14.1}$$

und

$$N_0 = A_1 \quad \text{und} \quad N_1 = nb \quad , \tag{14.2}$$

was natürlich offensichtlich ist. Es ist klar, dass ein Schieberegister à la Abb. 14.4 diese Funktion erfüllt.

Mustererkennung der Bitfolge 101

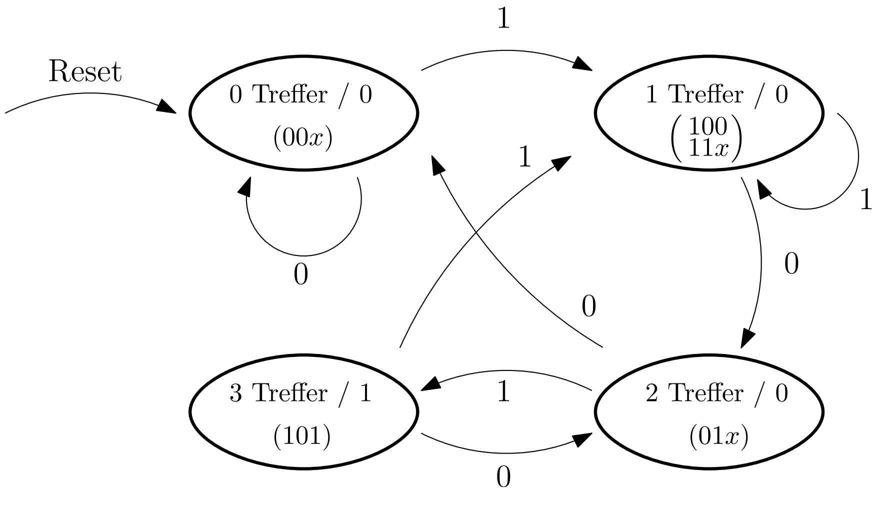


Abb. 14.21 Zustandsdiagramm zur Mustererkennung der Bitfolge 101.

Eine Ausgangsvariable Q_0 soll anzeigen, ob die Bitfolge 101 aufgetreten ist (Q_0 = „1“) oder nicht (Q_0 = „0“). Die Zustandstabelle lautet wie folgt mit einer etwas unorthodoxen aber hoffentlich suggestiven Notation.

Zuletzt registrierte Bits	Zahl der passenden Bits	Aktueller Zustand A_1A_0	Neuer Zustand N_1N_0 nb = 0	nb = 1	Q_0
000 001 } 00x	0	00	00	01	0
100	1	01	10	01	0
110 111 } 11x	1	01	10	01	0
010 011 } 01x	2	10	00	11	0
101	3	11	10	01	1

Tab. 14.9 Zustandstabelle zur Mustererkennung der Bitfolge 101.

Hier entspricht A_1A_0 der Zahl der Treffen im aktuellen Zustand und N_1N_0 der Zahl der Treffer des neuen Zustands. Die Zustandsgleichungen ergeben sich mit Tabellen 14.10 und 14.11 zu

$$Q_0 \;=\; A_1A_0 \quad , \quad N_0 = nb \quad , \quad N_1 = A_0\,\overline{nb} + A_1\,\overline{A}_0\,nb\,. \tag{14.3}$$

Die Schaltung ergibt sich damit zu Abb. 14.22. Diese Schaltung scheint deutlich aufwendiger als das Schieberegister in Abb. 14.15 auch wenn weniger Flipflops benötigt werden. Dies ist ein Anzeichen dafür, dass die Zustandsvariablen A_0, A_1 und n_b ungeschickt gewählt wurden. Besser wäre es, die Werte der Bits zu verwenden.

A_1A_0\nb	0	1
00	0	1
01	0	1
11	0	1
10	0	1

Tab. 14.10 Karnaugh-Diagramm für N_0 entsprechend der Tabelle 14.9 zur Mustererkennung.

A_1A_0\nb	0	1
00	0	0
01	1	0
11	1	0
10	0	1

Tab. 14.11 Karnaugh-Diagramm für N_1 entsprechend der Tabelle 14.9 zur Mustererkennung.

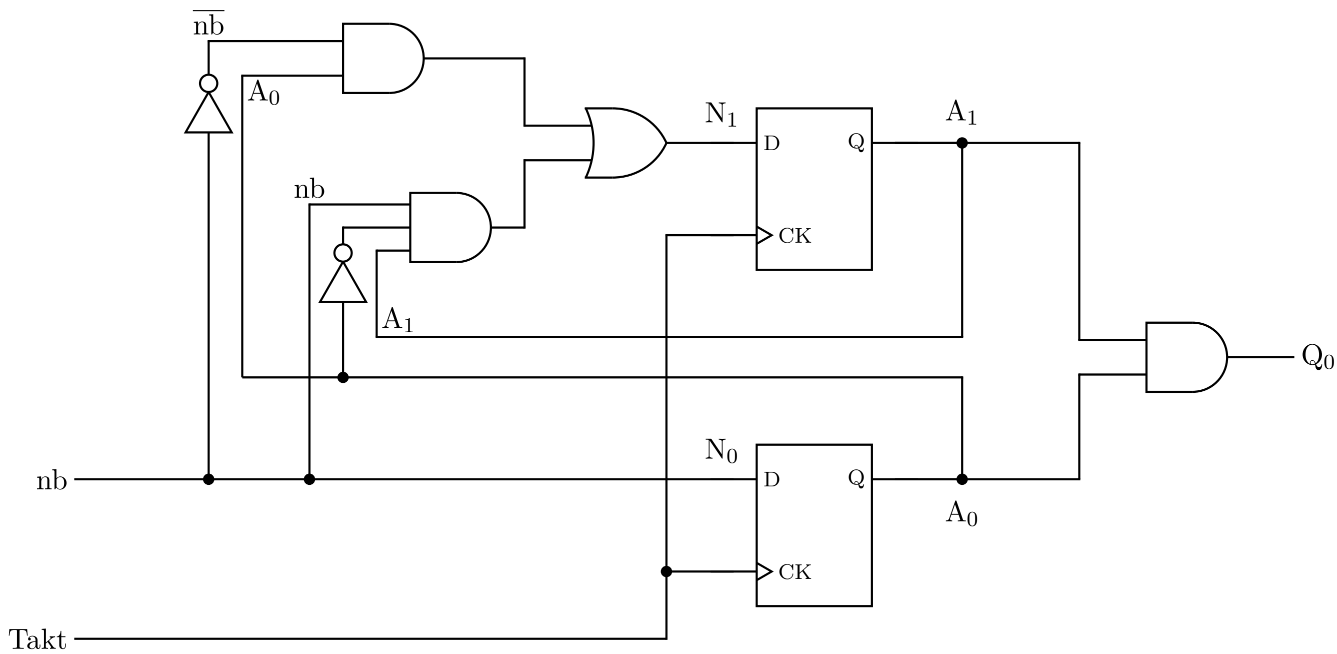


Abb. 14.22 Realisierung einer Schaltung zur Mustererkennung der Bitfolge 101.

Synchroner Zähler

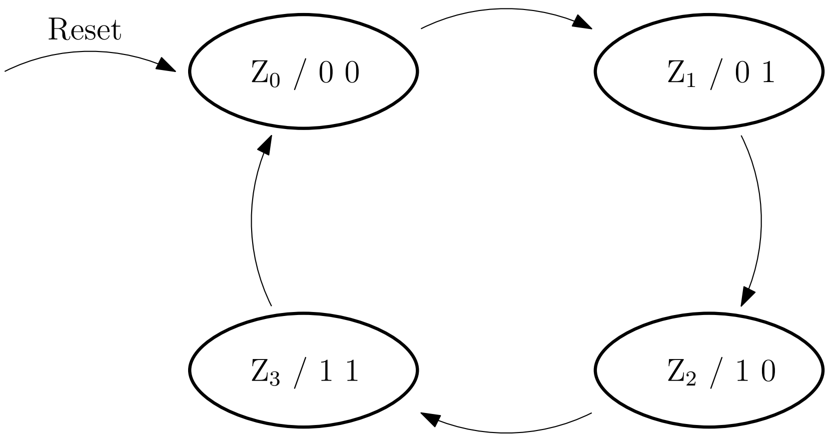


Abb. 14.23 Zustandsdiagramm eines synchronen Zählers.

Das Diagramm in Abbildung 14.23 beschreibt einen binären Ringzähler, der von 00 bis 11 zählt. Aus der Zustandstabelle (Tabelle 14.12)

A_1	A_0	N_1	N_0	Q_1	Q_0
0	0	0	1	0	0
0	1	1	0	0	1
1	0	1	1	1	0
1	1	0	0	1	1

Tab. 14.12 Zustandstabelle eines synchronen Zählers.

folgt

$$A_0 = Q_0 \quad , \quad A_1 = Q_1 \tag{14.4}$$

und

$$N_0 = \overline{A}_0 \quad , \quad N_1 = \overline{A}_1 A_0 + A_0 \overline{A}_1 = A_0 \oplus A_1 \ . \tag{14.5}$$

Dieser Zähler kann also mit zwei Flipflops, einem Inverter und einem XOR-Gatter aufgebaut werden (siehe Abb. 14.24).

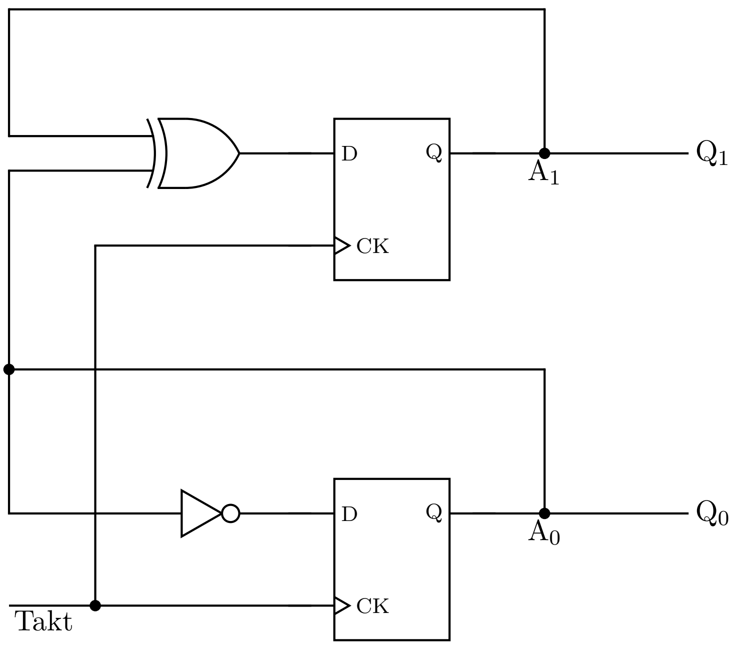


Abb. 14.24 Realisierung eines Synchronzählers mit zwei Flipflops, einem Inverter und einem XOR-Gatter.

15 Speicherbausteine

Die große Bedeutung leistungsfähiger Datenspeicher ist offensichtlich, nicht nur seit der Erfindung des Internets. Die Kapazität von Datenspeichern hat in den letzten Jahren noch schneller zugenommen als die Leistungsfähigkeit von Prozessoren. Datenspeicher basieren auf den unterschiedlichsten physikalischen und elektronischen Konzepten. Es gibt im wesentlichen fünf grundlegende Speichertypen:

- ROM,
- RAM,
- Flash memory,
- magnetische Speicher und
- optische Speicher.

Wir unterscheiden außerdem zwischen unterschiedlichen Speichermedien, zwischen flüchtigen[1] und permanenten[2] Speichern, sowie zwischen dynamischen und statischen Speichern. Neben der Speicherkapazität, den oben genannten Eigenschaften, sind auch die Zugriffszeit, die physikalische Größe des Speichers und der Preis wichtig. Eine grobe Einordnung häufiger Speicher ist in Tabelle 15.1 dargestellt.

permanent	CD DVD		Flash	Magnet- speicher
flüchtig	ROM	PROM		SRAM DRAM
	nicht	einmalig	beschränkt	beliebig oft
		beschreibbar		

Tab. 15.1 Klassifizierung von Speicherbausteinen

15.1 ROM

ROM steht für Read-only Memory. Ein ROM ist ein **nichtflüchtiger** Speicher. Das heißt bei Abschalten der Stromversorgung geht die Information nicht verloren. Der Name ROM suggeriert, dass der Inhalt eines ROM nicht modifiziert, sondern nur gelesen werden kann. Diese

[1] engl. *volatile*
[2] engl. *non-volatile*

Einschränkung ist für modernere Typen nur teilweise gültig. Grundsätzlich werden aber nur Informationen, die sich nicht oder selten ändern, z. B. boot firmware, auf ein ROM gelegt.

Ein ROM ist ein Tabellenspeicher. Jeder Adresse ist ein Speicherplatz mit festem Inhalt zugeordnet. Die Speicherplätze sind in einer Matrixstruktur angeordnet, die durch gleichzeitiges Anlegen einer Zeilen- und Spaltenadresse ausgelesen wird (Abb. 15.1).

Wie funktioniert das?

Die An- oder Abwesenheit einer Diode an einem Kreuzungspunkt bestimmt den Speicherinhalt. Für eine gegebene Adresse A wird die entsprechenden Adressleitung X_n auf eine positive Spannung „1“ gesetzt. Auf der Datenleitung D_i liegt genau dann eine positive Spannung an, wenn die Adressleitung und die Datenleitung am Kreuzungspunkt ni mit einer Diode verbunden sind. Sonst liegt die Datenleitung über den „Pull-down“-Widerstand auf Masse. Die Addressleitung wird auch „Worldline“ und die Datenleitung „Bitline“ genannt.

Legt man in dem in Abbildung 15.1 gezeigten Beispiel die Adresse „0“ an, so wird die Information 1010 ausgelesen.

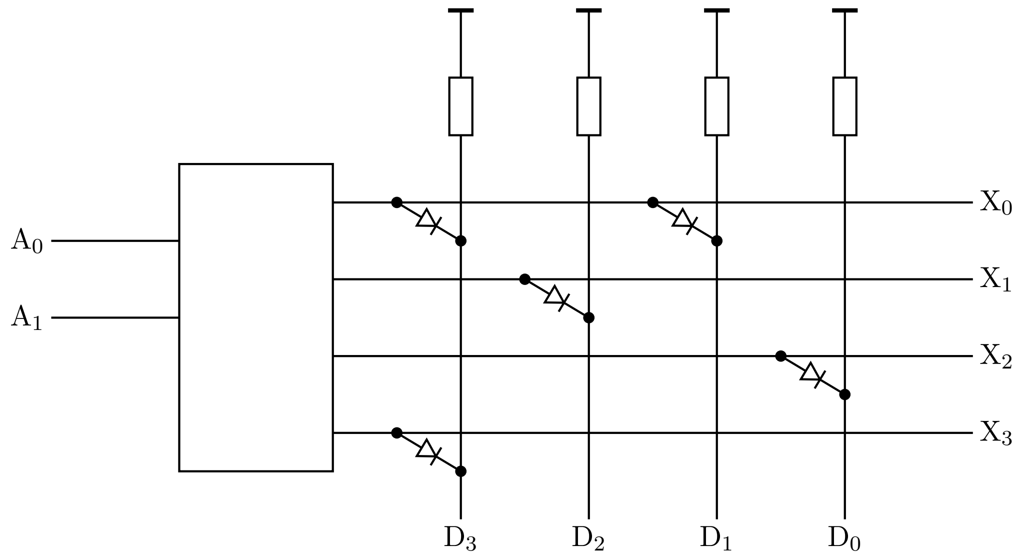


Abb. 15.1 ROM mit Addressdecoder, Speicherdioden und „Pull-down“-Widerständen.

Es gibt verschiedene Typen des ROM nämlich MROM, PROM, EPROM und EEPROM.

MROM: Das „M“ steht für maskenprogrammiert. Hier wird der Inhalt des ROM bei der Produktion des Halbleiters festgelegt und kann nicht mehr verändert werden.

PROM: Das „Programmable read-only memory“ (PROM) ist ein einmalig vom Nutzer programmierbares d. h. beschreibbares Speicherelement.
Bei der Fabrikation sind alle Transistoren durch Dioden in Sperrrichtung oder Schmelzsicherungen (dünnen Metallbrücken) verbunden (siehe Abbildung 15.2). Die Schmelzsicherungen sind so dimensioniert, dass sie durch Anlegen einer hohen Datenleitungsspannung gezielt durch gebrannt werden können. Dazu braucht man ein Spezialgerät. Entsprechend kann durch

das Anlegen einer Überspannung in den Dioden bzw. der Gateoxidschicht eines Transistors ein dauerhafter Kurzschluss erzeugt werden. Beide Vorgänge sind irreversibel.

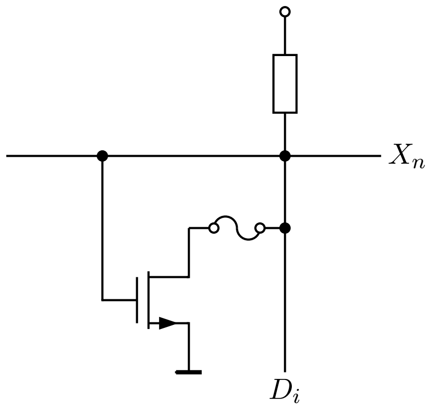


Abb. 15.2 PROM-Speicherzelle mit einem Transistor, einem „Pull-up"-Widerstand und einer Schmelzsicherung.

EPROM: Ein EPROM ist ein „Erasable programmable read-only memory". Hier kommen keine Schmelzsicherungen, sondern ein „floating-gate"-Transistor (FGT) zum Einsatz. Dies ist ein Feldeffekttransistor mit einem inneren Gate, welches nicht mit einer Elektrode mit der Aussenwelt verbunden ist (Abb. 15.3).

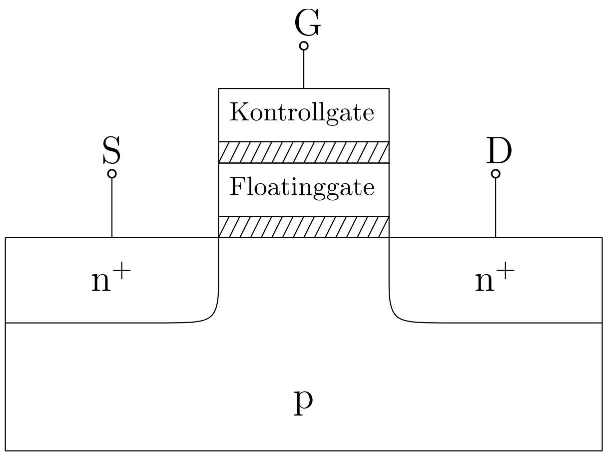


Abb. 15.3 Aufbau eines „floating-gate"-Transistors.

Das floating gate ist vorerst nicht geladen, es kann aber trotzdem über die sogenannte „hot carrier injection" aufgeladen werden. Dazu wird eine eher hohe positive Spannung zwischen Source und Drain angelegt, die die Elektronen im Kanal beschleunigt, und gleichzeitig eine noch höhere positive Spannung $U_G > U_D$ an das Gate angelegt.

Sind nun Elektronen auf dem Gate, so ist dies negativ geladen. Dies verschiebt wiederum die Schwellenspannung des FGT. Das floating gate entlädt man hier nicht durch Lichteinstrahlung, sondern durch Anlegen einer negativen Kontrolgatespannung und einer positiven

Spannung an Source bzw. Drain.

Das Gute am EPROM ist, dass sich das floating gate durch Einstrahlung von UV-Licht durch ein Quarzfenster wieder löschen lässt. Dafür muss man ein entsprechendes Fenster vorsehen, was den Preis erhöht. Abbildung 15.4 zeigt ein Bild eines EPROM mit Fenster. Dieses Bauteil hat einen 64 kB großen Speicher und Lesezeiten von rund 100 ns. Es kostet ein Paar Euro pro Stück.

Zur Bestrahlung muss das EPROM der Schaltung entnommen werden und rund 20 Minuten lang bestrahlt werden. Da für sehr kleine Transistorstrukturen ein Großteil der Oberfläche mit Metalllagen abgedeckt ist, kommt das UV-Licht hier nicht mehr gut durch.

EEPROM: Ein EEPROM ist ein „Electrically erasable programmable read-only memory“. Hier kann man statt durch die umständliche Einstrahlung von UV-Licht, die Speicherinformation elektrisch löschen, eventuall sogar ohne den Baustein auszubauen. Ein Beispiel eines EEPROM ist das beliebte flash memory, dass weiter unten beschrieben wird.

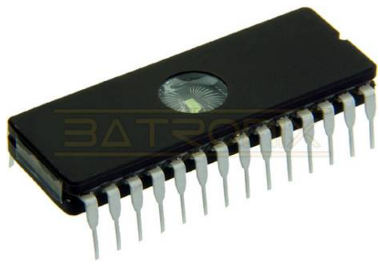


Abb. 15.4 Bild des EPROM M27C64A-10F6 mit Quarzfenster.

15.2 RAM

RAM steht für „Random access memory“. Dies soll ausdrücken, dass der Zugriff auf alle Daten gleichlang dauert. Dies ist also anders als beim Plattenspieler oder einer CD. Ein RAM kann man beliebig oft lesen und überschreiben. Der Speicherinhalt ist aber in der Regel flüchtig; das Abschalten der Versorgungsspannung löscht den gespeicherten Inhalt. Man unterscheidet zwei wesentliche Typen des RAM, das dynamische RAM (DRAM) und das statische RAM (SRAM).

15.2.1 DRAM

Die Speicherzellen eines DRAM bestehen aus je einem Kondensator und einem Transistor (Abb. 15.5). Ein Bit entspricht der auf dem Kondensator gespeicherten Ladung. Der Kondensator wird über den Transistor zugänglich.

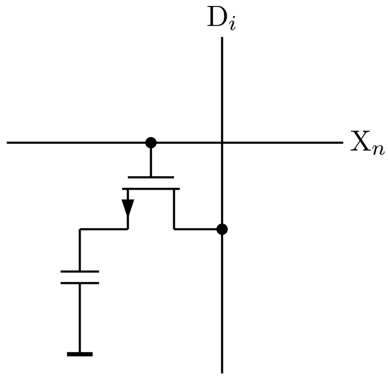


Abb. 15.5 Speicherzelle eines DRAMs.

Ein DRAM ist kompakt und dadurch billig. Der Stromverbrauch von DRAMs ist gering. Allerdings muss der Kondensator regelmäßig aufgeladen werden (alle 1 - 10 ms), da er mit der Zeit immer etwas Ladung verliert („leckt"). Das Aufladen erfordert zusätzliche Logik, die auch Platz kostet und sich daher erst ab einer Mindestgröße des Speichers lohnt.

Ein „Synchronous DRAM" ist ein DRAM, welches Werteänderungen nur bei Taktflanken umsetzt. Der Einsatz in Computern erfolgt oft mit mehreren Speicher-ICs auf einer zweiseitig bestückten Platine, welche auf beiden Seiten Kontaktpins für unabhängige Signale besitzt (DIMM, Dual Inline Memory, Speichermodule für Arbeitsspeicher). Die typische Breite des Datenbusses ist 64 Bit. Als Latenz bezeichnet man die Zugriffszeit, z. B. 10 ns. Diese begrenzt die Bandbreite.
Ein DDR SDRAM (Double Data Rate Synchronous Dynamic Random Access Memory) wird von beiden Flanken getaktet, der ansteigenden und der fallenden. Dadurch verdoppelt sich die effektive Taktfrequenz. Bei einem DDR3 ist die Taktrate gegenüber dem Standard DDR2 noch einmal verdoppelt. Es gibt DDR3-Speicher der Größe 2, 4 und 8 GB mit Spitzentransferraten von mindestens 6,4 GB/s. Ein DDR3 von 8 GB (z.B. 4 x 2 GB) kostet zwischen 60 € und 80 € (Stand 2014).

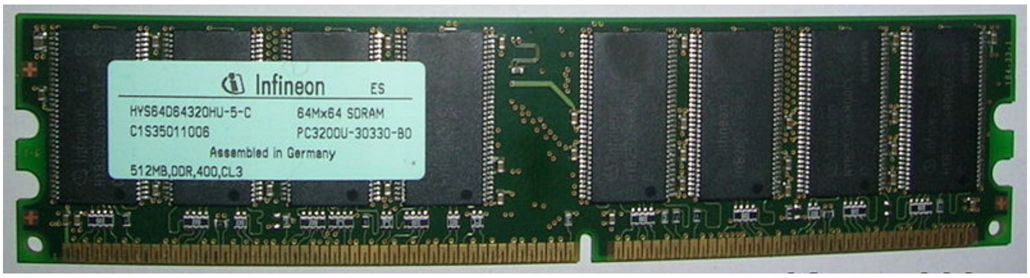


Abb. 15.6 Photo eines SDRAM von Infineon.

15.2.2 SRAM

Die Speicherzelle eines SRAM besteht dagegen aus Logikbausteinen z. B. Flipflops. Dadurch sind SRAMs sehr schnell, müssen nicht aufgefrischt werden, und haben einen geringeren Stromverbrauch als DRAMs.

Eine mögliche Realisierung mit sechs Transistoren ist in Abbildung 15.7 illustriert. Die inneren vier Transistoren bilden einen Speicher, ein Latch, aus zwei gegengekoppelten CMOS-Invertern. Die Transistoren M_1 und M_2 bilden den linken und die Transistoren M_3 und M_4 den rechten Inverter.

Um Daten zu schreiben, muss die Zelle mit der Aussenwelt verbunden werden. Dies geschieht durch Freischalten der Datenleitung (bit line) durch die „word line“ durch ein positives Potential auf das Gate von M_5 und M_6. Jetzt kann die Information der Zelle über die bidirektionale Bitline gelesen oder überschrieben werden.

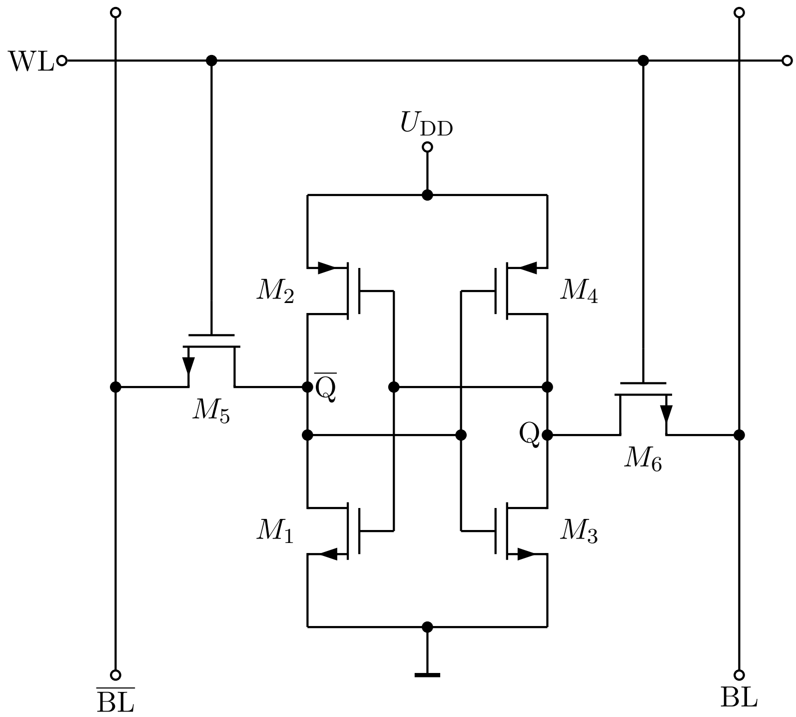


Abb. 15.7 6T-Speicherzelle eines SRAMs mit Wordline und Bitline.

Ein SRAM wird beschrieben, indem man die Bitleitungen geeignet vorspannt und dann über die Wordline den Zugang zu den gewünschten Speicherzelle ermöglicht: Soll in eine gegebene Zelle eine „1“ gespeichert werden, so wird die BL auf „1“ (also U_{DD}) gesetzt und $\overline{\mathrm{BL}}$ auf „0“. Bei der Dimensionierung der Bitlinetreiber muss man sicherstellen, dass die Treiber stark genug ausgelegt sind, die Speicherinformation zu überschreiben.

Ein SRAM wird gelesen, indem man beide Bitlines, BL und $\overline{\mathrm{BL}}$, auf U_{DD} vorspannt und sie dann über die Wordline mit der Speicherzelle verbindet. Die Zelle zieht dann abhängig von ihrem Inhalt eine der Leitungen in Richtung „0“. Die Abweichung vom vorgespannten Wert U_{DD} wird vom Leseverstärkern erkannt und dadurch der Inhalt der Speichzelle interpretiert.

Annmerkung: Wie können Bitlinetreiber einmal stark und einmal schwach sein? Dazu muss man W/L der sonst baugleichen Transistoren in Abb. 15.7 so dimensionieren, dass gilt

$$
\begin{aligned}
& \left(K\,\frac{W}{L}\right)_{M_5} < \left(\mu\,\frac{W}{L}\right)_{M_1} \\
\text{und}\qquad & \left(K\,\frac{W}{L}\right)_{M_5} \gg \left(\mu\,\frac{W}{L}\right)_{M_2} .
\end{aligned}
$$

Entsprechende Bedingungen gelten für M_3, M_4 und M_6.

Die erste Bedingung sichert das Lesen einer $\overline{Q}$ = „0“ bei vorgespanntem $\overline{\text{BL}}$, ohne gleichzeitig den Zustand zu ändern, weil dadurch die Verbindung von $\overline{Q}$ zu Masse über M_1 niederohmiger ist als die Verbindung zu U_{DD} über M_5.

Die zweite Bedingung sichert das Schreiben einer $\overline{Q}$ = „0“ bei gespeicherter $\overline{Q}$ = „1“.

Eine vereinfachte Darstellung der Zelle wird in Abb. 15.8 dargestellt.

Frage für sehr Interessierte: Warum werden beim Lesen die Bitleitungen vorgespannt und nicht auf Tristate gelegt (hochohmig abgeschlossen)?

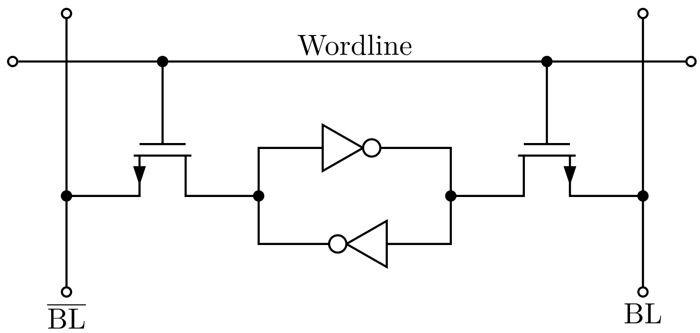


Abb. 15.8 Vereinfachte Darstellung einer SRAM-Zelle mit zwei gekoppelten Invertern.

Eine schematische Darstellung mit 16 Zellen ist in Abb. 15.9 dargestellt.

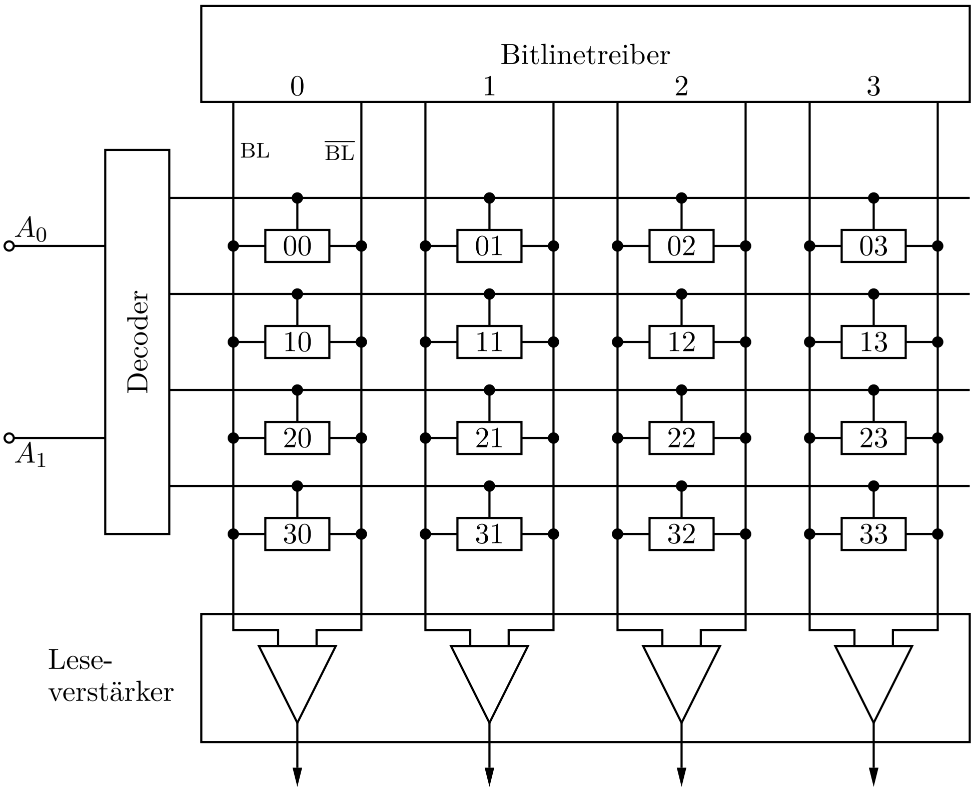


Abb. 15.9 Schematischer Aufbau einer SRAM-Speicherzellenmatrix.

Hausaufgabe: Schauen Sie sich das Datenblatt des in Abb. 15.10 dargestellten SRAM an.

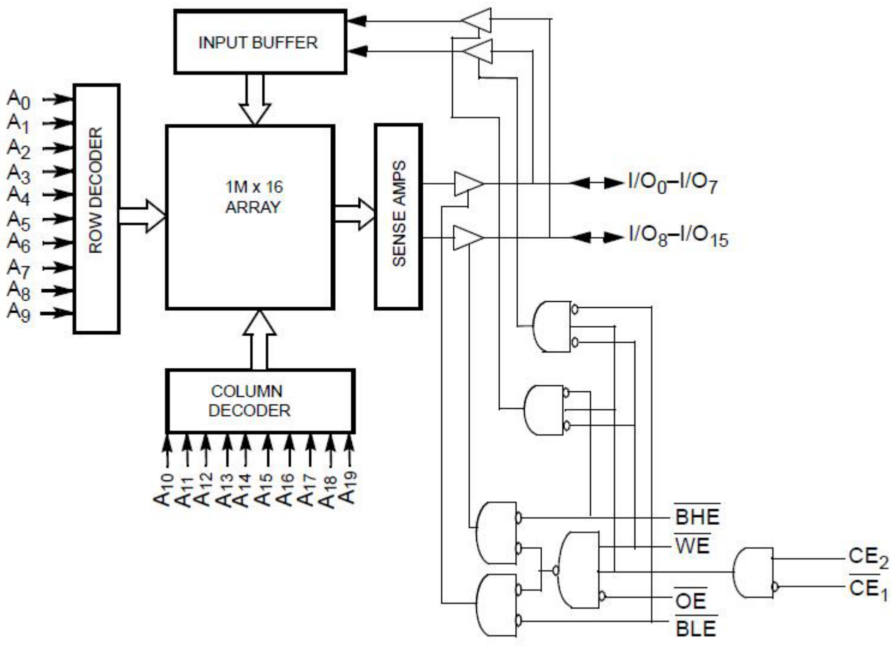


Abb. 15.10 Prinzipaufbau eines SRAMS (Quelle: Datenblatt CY7C1061AV33).

15.3 Flash-Speicher

Ein Flash-Speicher entspricht der oben vorgestellten Version des EEPROM. Er ist somit ein nichtflüchtiger Speicher, der die Information in Floating Gates speichert. Flash-Speicher kommen z. B. in USB-Sticks und Solid-State-Drives[3] (SSD) vor. Durch den Aufbau ohne bewegliche Teile ist flash memory im Vergleich zu Festplatten unempfindlich gegen mechanische Erschütterungen und leise. SSDs sind schnell im Vergleich zu Magnetspeichern und ermöglichen Datentransferraten von rund 500 MB/s. Mit rund 0,7 € pro GB (Stand 2014) sind SSDs deutlicher teuerer. Die Speichergröße erreicht 0,5 bis 1 TB.

Nachteilig ist die begrenzte Anzahl von Schreibzugriffen pro Zelle. Für NAND-Flash werden abhängig vom Typ mehrere Tausend bis Millionen Schreib- und Löschzyklen erreicht (program/erase cycles). Durch viele Schreibzugriffe degeneriert die Isolationsschicht um das Floating Gate, und die Ladung kann nicht mehr dauerhaft gespeichert werden. Um diesen Nachteil auszugleichen, werden die Firmware und das Dateisystem so eingerichtet, dass die Schreibzugriffe möglichst gleichmäßig auf den ganzen Speicher verteilt werden. Defekte Zellen können in der Regel erkannt und ausgeblendet werden.

15.4 Magnetspeicher

Magnetspeicher gehören zu den leistungsfähigsten und billigsten Datenspeichern. In Abbildung 15.11 ist ein typischer Festplattenspeicher (Hard Disk Drive HDD), wie er in den meisten PCs und Laptops zu finden ist, dargestellt. Die Kosten für 4 TB Speicherplatz betragen nur rund 150 € (Stand 2014), also weniger als 5 Cent/GB. Die Schreib- und Lesezeiten sind dafür begrenzt (in der Regel unter 150 MB/s). Die Mechanik der Festplatten ist eindrucksvoll. Die typische Rotationsgeschwindigkeit einer Platte beträgt 5400 oder 7200 Umdrehungen pro Minute. Das Speicherprinzip beruht auf dem Riesenmagnetowiderstand ferromagnetischer Filme, für dessen Entdeckung Peter Grünberg und Albert Fert 2007 den Nobelpreis für Physik erhielten. Durch ein magnetisches Feld am Schreibkopf wird die Magnetisierung des Films wie gewünscht eingestellt.

[3] deutsch: *Festkörperlaufwerk*

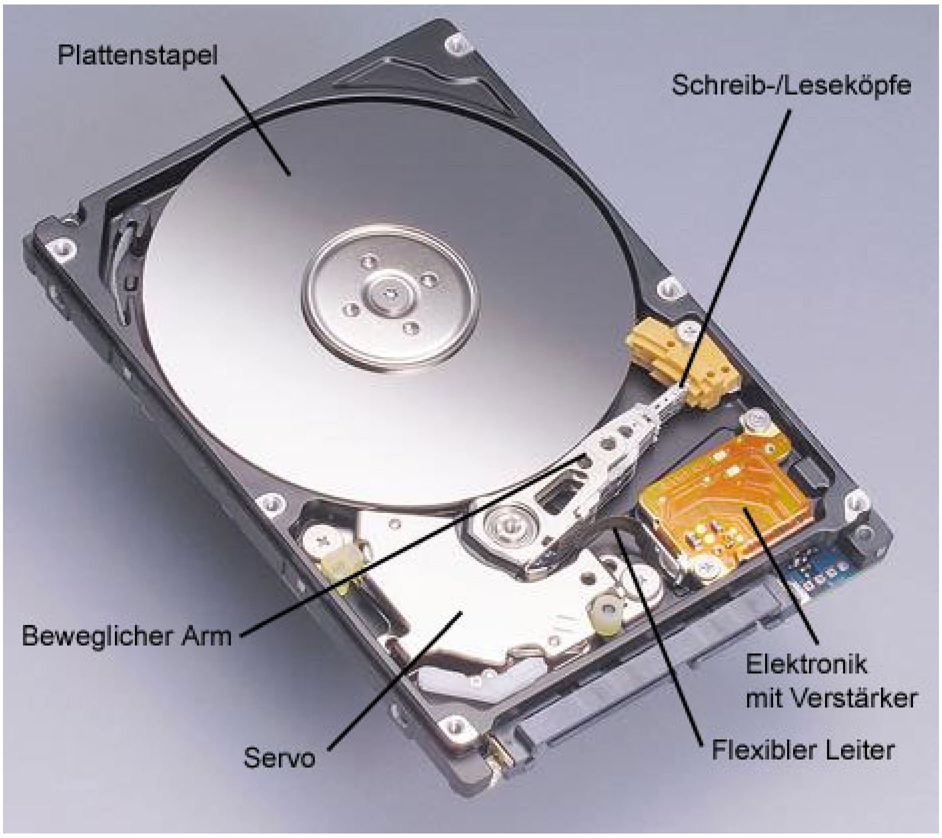


Abb. 15.11 Photo eines Festplatte (HDD) [16].

Für Langzeitspeicherung und als Back-up werden Magnetbänder (Tape) eingesetzt. Die Bandlänge beträgt für ein Band der Generation LTO-6 rund 850 m. Die Haltbarkeit der Bänder liegt zwischen 15 bis 30 Jahren. Hersteller geben die Speicherkapazität und Datentransferraten gerne für komprimierte Daten an. Für einen Komprimierungsfaktor von 2,5 erreichen diese Bänder dann 6 TB Speicherkapazität und Datenraten von 400 MB/s. Hier liegen die Kosten für ein (komprimiertes) GB bei rund 0,01 €.

15.5 Optische Datenspeicher*

Bei optischen Datenspeichern werden die Bits berührungslos mit einem Laser ausgelesen (Laser-Interferenzdetektor). Typische optische Speicher sind die Compact Disc (CD), Digital Versatile Disc (DVD) und Blu-ray Disc. Dabei handelt es sich um Wechselmedien, die mit entsprechenden Laufwerken ausgelesen werden können. Normalerweise handelt es sich um ROM, das einmal beschrieben werden kann. Selbst bei sachgemäßer Behandlung (insbesondere Vermeidung von Kratzern, Feuchtigkeit und UV-Einstrahlung) ist die Haltbarkeit von optischen Medien leider auf einige Jahrzehnte beschränkt. Medien mit höherer Informationsdichte (DVD, Blu-ray Disc) benötigen eine präzisere Fokussierung des Laserstrahls. Bei der Blu-Ray Disc wird violettes Licht (405 nm Wellenlänge) verwendet. Die im Vergleich zu DVD und CD kleine Wellenlänge ermöglicht eine höhere Informationsdichte.

15.6 Assoziativspeicher*

Ein Standardspeicher hat den Nachteil, dass man die Speicheradresse zu einem gegeben Inhalt nicht unbedingt kennt. Das macht elektronische Suchen langsam. Ein einfaches Beispiel wäre ein Speicher von Telefonnummern und dazugehörige Personen. Ein Wort hat die Struktur [Name,Telefonnummer]. Weiß man nicht, ob und wo eine gesuchte Person samt Telefonnummer im Speicher abgelegt ist, muss man eventuel lange suchen. Die Suche entspricht dem sukzessiven Auslesen von Werten und Vergleich mit dem gewünschten Personennamen.

In dieser Fragestellung sind die selteneren Assoziativspeicher[4] interessant. Der Aufbau einen Assoziativspeichers ist in Abb. 15.12 illustriert.

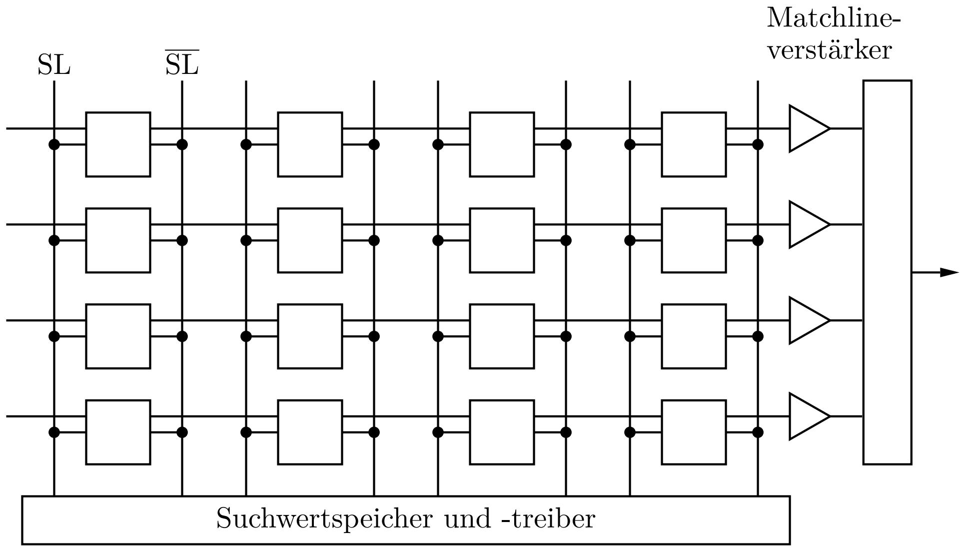


Abb. 15.12 Aufbau eines Assoziativspeichers mit CAM-Zellen, Matchlines, Matchlineverstärkern und Encoder.

Hier wird über die sogenannte Matchline (ML) der Inhalt des Suchwertes mit dem Speicherinhalt aller Zellen gleichzeitig verglichen. Die Matchline ist auf „1“ vorgespannt. CAM-Zellen, die nicht mit dem entsprechenden Bit des Suchwortes übereinstimmen, ziehen die Matchline auf „0“. Die CAM-Zellen, die mit den entsprechenden Bit des Suchwortes übereinstimmen, halten die Matchline im Zustand „1“. Der Zustand der Matchlines wird von den Matchlineverstärkern detektiert und die Adresse des gespeicherten Wertes an den Ausgang des CAM weitergegeben.

Das CAM liefert also nicht die Antwort auf die Frage „Welche Telefonnummer hat die Person X“, sondern auf die Frage „Steht Person X im Telefonbuch“. Diese Frage wird sehr schnell (innerhalb weniger Taktzyklen) beantwortet. Ein möglicher Aufbau einer Zelle ist in Abb. 15.13 dargestellt.

[4] Engl. Content-addressable memory (CAM)

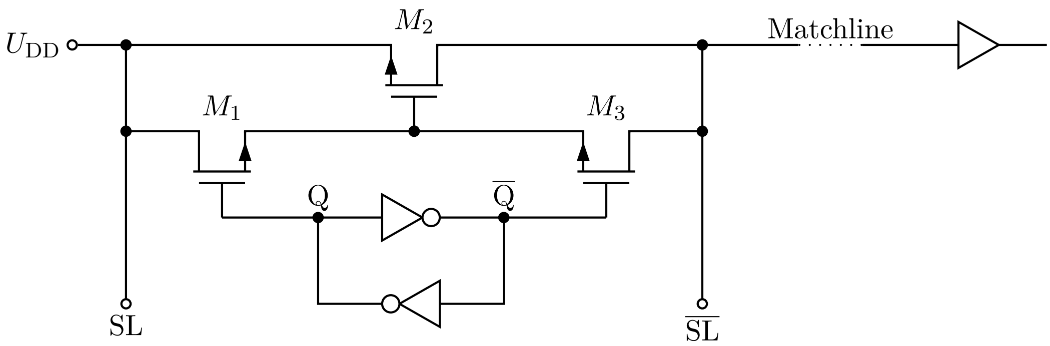


Abb. 15.13 Aufbau einer CAM-Zelle.

Beispiel: o.B.d.A. SL= “1“, $\overline{\mathrm{SL}}$ = “0“

a)	Q= “1“, $\overline{\mathrm{Q}}$ = “0“	$\Rightarrow$	M_1 ist an und M_3 ist aus	$\Rightarrow$	M_2 ist an.
b)	Q= “0“, $\overline{\mathrm{Q}}$ = “1“	$\Rightarrow$	M_1 ist aus und M_3 ist an	$\Rightarrow$	M_2 ist aus, damit ist ML unterbrochen.

15.7 Tabellarische Zusammenfassung

Eine Übersicht der Charakteristika von Speicherbausteinen ist in Tabelle 15.2 zusammengestellt. Die angegebenen Zahlen können nun Richtwerte sein und die Leistungsfähigkeit von Speichern nimmt stetig zu. Einige Werte für ROM und SRAM fehlen in der Tabelle, da es zum einen viele und zu verschiedene Arten von ROM gibt. Zum anderen werden SRAMs im wesentlichen als interne Blöcke in Prozessoren, FPGAs, etc. eingesetzt, so dass der Vergleich mit externen Speichern wenig sinnvoll ist.

Typ	**R**	**W**	**Größe**	**Datentransfer-raten** [MB/s]	**Kosten** [€/GB]
ROM	X	(X)	8 - 128 kB		
SRAM	X	X	kB - MB		
DRAM	X	X	8 GB	$\leq$ 6 400	10,00
Solid-State-Drive	X	X	1 TB	500	0,70
Magnetspeicher	X	X	4 TB	150	0,05
Optisch: CD	X	(X)	700 MB	10	0,45
Optisch: DVD	X	(X)	4,7 GB	10	0,10

Tab. 15.2 Die wichtigsten Eckdaten von flüchtigen (grau) und nichtflüchtigen (farblos) Speicherbausteinen.

16 Analog-Digital-Wandler und Digital-Analog-Wandler

Analog-Digital-Wandler[1] und Digital-Analog-Wandler[2] verbinden die digitale mit der analogen Welt. Praktisch jedes analoge Signal wird früher oder später in ein digitales Signal umgewandelt und dann digital weiterverarbeitet und gespeichert. ADCs sind komplexe, aufwendige und oft auch teure Bausteine, und es ist wichtig, ihre Eigenschaften und Grenzen zu verstehen.

Ein Analog-Digital-Wandler wandelt ein analoges Spannungssignal in eine Folge diskreter digitaler Werte um (Abb. 16.1). Dabei ist die Genauigkeit der digitalen Darstellung durch den Wert der digitalen „1“, des Least Significant Bit (LSB) und durch den Abstand der digitalen Werte auf der Zeitachse festgelegt. Zum Beispiel kann die digitale „1“ dem analogen Wert von 0,1 V entsprechen und jede Mikrosekunde ein digitaler Wert bestimmt werden. Es tritt also ein sogenannter Quantisierungsfehler auf, da im obigen Beispiel alle Analogwerte im Intervall $0,1 \pm 0,005$, also zwischen 0,05 V und 0,15 V, der digitalen „1“ zugeordnet werden. Der Quantisierungsfehler ist allerdings nicht relevant, wenn er (deutlich) unter dem Rauschen des Analogsignals liegt.

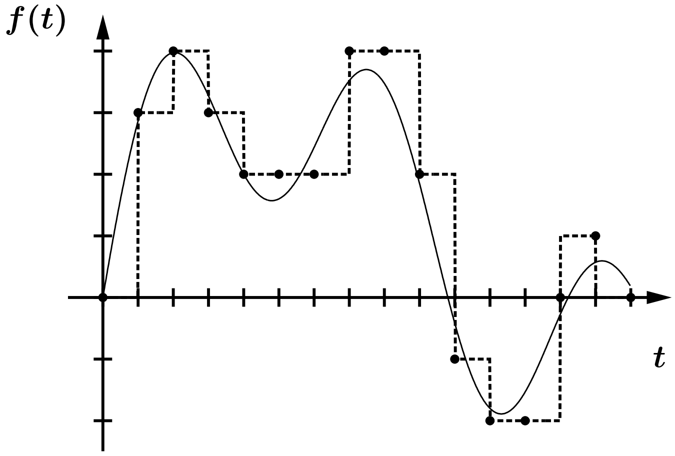


Abb. 16.1 Illustration eines analogen (durchgezogene Linie) und des entsprechenden diskretisierten Signals (Punkte und gestrichelte Linie).

Auch der digitale Maximalwert des ADC und die maximal zulässige analoge Eingangsspannung ist beschränkt. Da man den digitalen Wert des ADC meistens in Binärdarstellung angibt, redet man von z. B. einem 8-Bit-ADC. Ein 8-Bit-ADC kann 256 unterschiedliche Werte dar-

[1] Engl. *analog-to-digital converter (ADC)*
[2] Engl. *digital-to-analog converter (DAC)*

stellen, ein 12-Bit-ADC 4096, ein 16-Bit-ADC 65 536 und ein 20-Bit-ADC rund 1M Werte. Sechzehn Bit entspricht 96 db und 20 Bit 120 db, da

$$20 \cdot \log(\frac{65536}{1}) = 96\,\mathrm{db} \quad \text{und} \quad 20 \cdot \log(\frac{10^6}{1}) = 120\,\mathrm{db}$$

Diese Zahlen darf man sich merken.

Falls das Eingangssignal zeitlich variiert, entsteht auch auf der Zeitachse ein Fehler, da der ADC eine endliche Umwandlungszeit benötigt. Oft variieren Signale aber langsam, z. B. aufgrund von Temperaturschwankungen.

Die ideale Übertragungsfunktion eines ADC, d.h. die Zuordnung der analogen Eingangsspannungen zu dem entsprechenden Digitalwert ist in Abb. 16.2 dargestellt. Wichtig ist, dass „die Höhe und Breite (eigentliche Tiefe) aller Stufen“ immer gleich ist. D.h. alle Stufenkanten liegen auf einer Verbindungsgraden. Abweichungen von der regelmässigen Stufenfolge werden wir am Ende des Kapitels betrachten.

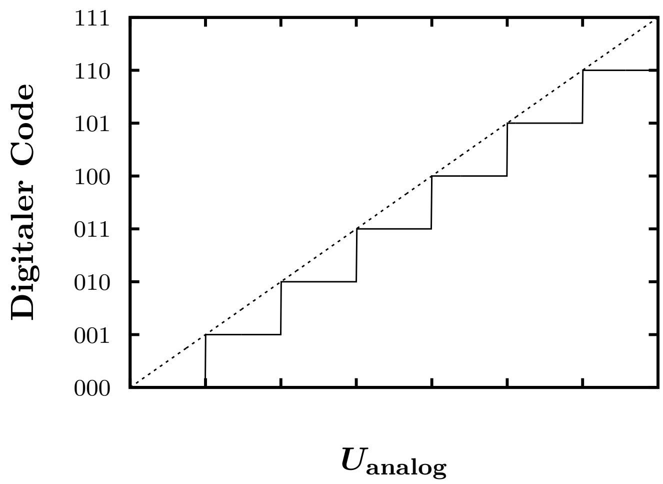


Abb. 16.2 Übertragungsfunktion eines idealen ADC. Die gestrichelte Gerade dient nur zur Veranschaulichung der Treppenfunktion.

Zum Verständnis des ADCs ist es hilfreich, zuerst den Digital-Analog-Wandler kennenzulernen.

16.1 Digital-Analog-Wandler

Ein Digital-Analog-Wandler (DAC) erzeugt zu einem digitalen Eingangswert eine entsprechende analoge Ausgangsspannung. Es hier gibt drei grundsätzliche Konversionsverfahren – das Parallel- oder Flashverfahren, das Wägeverfahren und das Zählverfahren. Für die ersten beiden Verfahren wird ein Addierer eingesetzt (siehe Abb. 16.3).

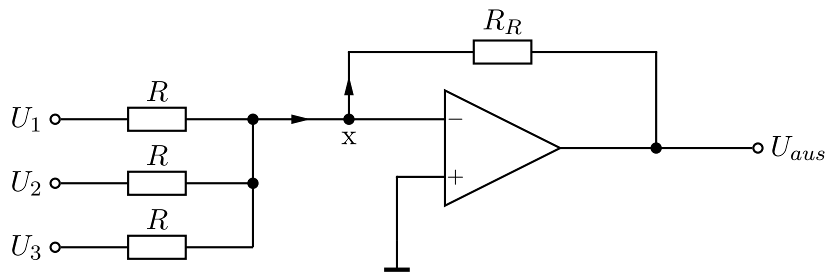


Abb. 16.3 Schaltdiagramm eines Addierers. Das Symbol x bezeichnet hier die virtuelle Masse.

Wegen der negativen Rückkopplung gelten die goldenen Regeln. Also ist der Strom, der von links in den Knoten fließt, gleich dem Strom, der rechts wieder herausfließt. Die Ein- und die Ausgangsspannung werden jeweils auf den virtuellen Massepunkt x bezogen. Für den Strom gilt:

$$I_{\text{ein}} = \frac{U_1}{R} + \frac{U_2}{R} + \frac{U_3}{R} = -\frac{U_{\text{aus}}}{R_R} \quad . \tag{16.1}$$

Aus dieser Schaltung bildet man durch gezieltes Zuschalten von Widerständen einen DAC (siehe Abb. 16.4).

Parallel-Verfahren

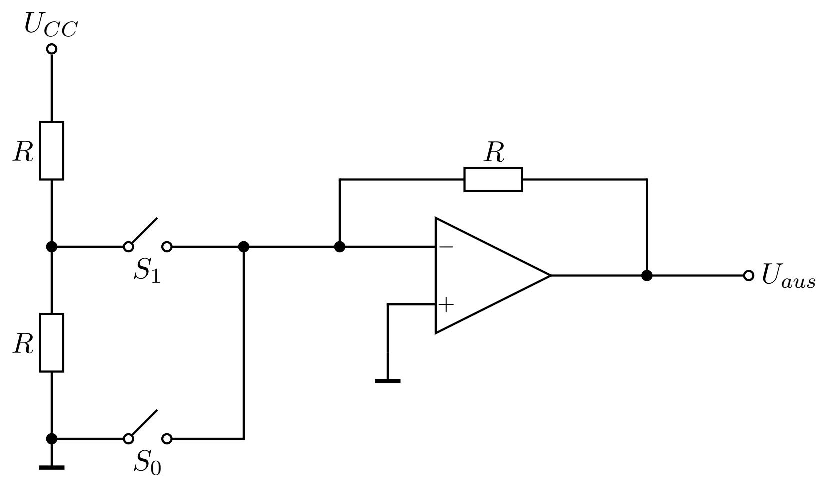


Abb. 16.4 DAC basierend auf dem Parallelverfahren. Der Dekoder zur Schaltersteuerung ist nicht gezeigt.

Das Prinzip ist einfach. Die Versorgungsspannung U_{CC} wird durch n gleich große Widerstände in n gleich große Teilspannungen unterteilt. Genau ein Schalter S ist jeweils geschlossen. Dazu braucht man einen Dekoder. Die Schaltung ist ein DAC, da abhängig von dem digitalen Eingangswert der entsprechende Schalter geöffnet und eine Analogspannung erzeugt wird. Hier benötigt man pro möglicher Ausgangsspannung einen Schalter und einen Widerstand (also für einen 8-bit-DAC schon 256 Schalter). Das ist aufwendig, daher sind Parallel-DACs selten.

Wägeverfahren

Hier verwendet man (binär) skalierte Widerstände. Es gibt zwei plausible Arrangements (Abb. 16.5 und 16.6).

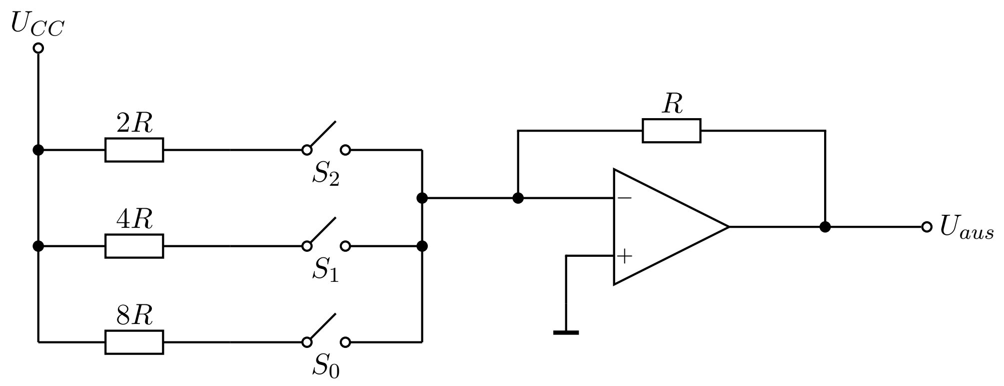


Abb. 16.5 DAC basierend auf dem Wägeverfahren.

In beiden Fällen gilt $U_{LSB} = -\frac{1}{8}\,U_{aus}$. In diesem Fall ist S_0 geschlossen, und S_1 und S_2 sind offen. Jede mögliche Kombination offener und geschlossener Schalter ist zulässig und produziert entsprechende Analogausgangsspannungen.

Die Nachteile der obigen Schaltung sind:

- Die Potentiale am Schalter schwanken zwischen den Extremwerten U_{CC} und $0\,\mathrm{V}$. Entsprechend ist der Strom nicht konstant, und die Spannungsquelle wird ungleichmäßig belastet. Das ist bei einem endlichen Innenwiderstand $R_i \neq 0$ der Spannungsquelle U_{CC} schlecht.

- Die Widerstände müssen genau den Sollwerten $8\,R$, $4\,R$ und $2\,R$ entsprechen.

Eine bessere Schaltung ist in Abbildung 16.6 dargestellt. Dabei handelt es sich um einen R-$2R$-Leiter. Den Trick mit parallelen und seriell geschalteten Widerständen in einem Netzwerk kennen wir schon aus dem π- und T-Glied in Hohlleitern. Die Spannungsquelle wird gleichmäßig belastet. Hier fließt immer Strom. Es gibt keine Schwankungen und damit kein Umladen der Schalterkapazität.

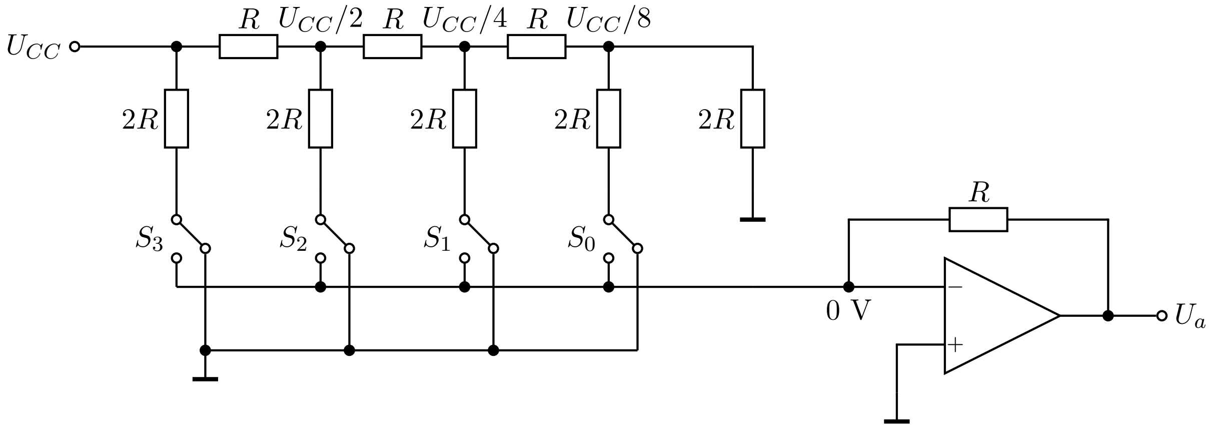


Abb. 16.6 Aufbau eines R-2R-DAC.

Das MSB liegt an, wenn der Schalter S_3 am OPV liegt und alle anderen Schalter auf Masse. Die Kombination von parallelem Widerständen ($2\,R \,\|\, 2\,R \;=\; R$) und seriellen Widerstand ($R + R = 2R$) ergibt immer $2\,R$. Dies gilt unabhängig von der Schalterstellung, da auch der Eingang des OPV auf virtuellem 0 V liegt. Generell gilt

$$U_{\text{aus}} \;=\; -R\,I\;.$$

Die Spannungen an den Knoten der Leiter stellen sich unabhängig von der Schalterstellung zu U_{CC}, $\text{U}_{\text{cc}}/2$, $\text{U}_{\text{cc}}/4$, $\text{U}_{\text{cc}}/8$ ein. Damit gilt für das **MSB**

$$I \;=\; \frac{U_{\text{CC}}}{2\,R} \quad\Rightarrow\quad U_{\text{aus}} = -\,\frac{U_{\text{CC}}}{2}\quad. \tag{16.2}$$

für das **NMSB**[3]

$$I \;=\; \frac{\frac{U_{\text{CC}}}{2}}{2\,R} \quad\Rightarrow\quad U_{\text{aus}} = -\,\frac{U_{\text{CC}}}{4}\quad. \tag{16.3}$$

für das **LSB**

$$I \;=\; \frac{\frac{U_{\text{CC}}}{8}}{2\,R} \quad\Rightarrow\quad U_{\text{aus}} = -\,\frac{U_{\text{CC}}}{16}\quad. \tag{16.4}$$

Da alle Widerstände gleich groß sind, kann man diesen DAC leichter in einem IC integrieren. Nur das „Matching“ der Widerstände ist entscheidend, nicht der Absolutwert. Natürlich sollten hier wie auch in der vorangehenden Schaltung die Leitungs- und Schalterwiderstände klein gegenüber R sein.

Zählverfahren (Pulsweitenmodulation)

Das Zählverfahren beruht auf dem Aufladen und Entladen eines Kondensators über einen Widerstand.

[3]Next-to-most-sigificant bit

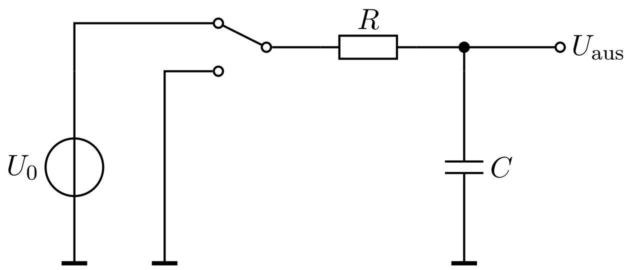


Abb. 16.7 DAC basierend auf der Pulsweitenmodulation.

Hier braucht man eine intelligente Logik für den Schalter. Das Tastverhältnis bestimmt U_{aus}. Man kann mit hohen Frequenzen arbeiten und damit das Tastverhältnis und U_{aus} mit $U_{aus} = U_{CC} \cdot T_{ein}/(T_{ein} + T_{aus})$ genau einstellen. Deshalb wird das Verfahren auch DAC mit Pulsweitenmodulation genannt. Steuert man den Schalter mit einem digitalen Rechtecksignal an, so treten abhängig von den Anstiegszeiten und Abklingzeiten hohe Frequenzen auf, die die Taktfrequenz des Schalters eventuell weit überschreiten. Durch den RC-Tiefpass werden die entsprechenden Spannungsvariationen aber abgeschwächt. Auch die in Abbildung 16.8 dargestellte Schaltung (mit einer Stromquelle) funktioniert. Durch Hinzufügen eines Operationsverstärkers als Puffer könnte man die Qualität der Ausgangsspannung noch verbessern.

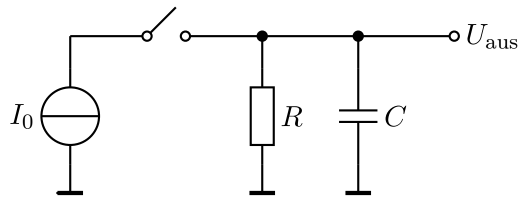


Abb. 16.8 DAC basierend auf Pulsweitenmodulation mit einer Stromquelle.

16.2 ADC-Typen

Wir betrachten drei einfache Wandlertypen, nämlich

- Parallelwandler oder Flash-ADC
- Wägeverfahren[4]
- Zählverfahren

und den $\sum\Delta$-Wandler (Sigma-Delta-Wandler). Wichtig ist für alle ADCs, dass die analoge Eingangsspannung sich nicht ändert bevor die Digitalisierung abgeschlossen ist. Diesem Zweck dient eine sogenannte „Sample-and-Hold"-Schaltung (S/H)[5] oder eine „Track-and-Hold"-Schaltung (T/H)[6]:

[4] Engl. *successive approximation*
[5] Deutsch *Abtast-Halte-Glied*
[6] Deutsch *Folge-Halte-Glied*

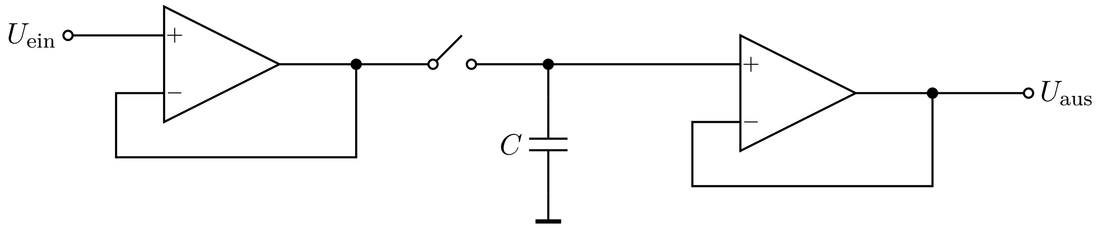


Abb. 16.9 Sample-and-Hold-Schaltung.

Der Spannungsfolger links im Bild dient dazu, den Kondensator schnell auf die Spannung U_{ein} aufzuladen, falls der Schalter S geschlossen ist. Der Spannungsfolger rechts reduziert die Last, die der Kondensator sieht, und verhindert, dass der Kondensator sich aufgrund dieser Last entlädt. Der Schalter wird durch einen Taktimpuls gesteuert. Ein S/H tastet das Eingangssignal nur zu der fallenden Flanke ab („sample“) und hält die Kondensatorspannung sonst konstant („hold“). Ein T/H folgt dem Eingangssignal kontinuierlich („track“), solange der Taktpuls den Wert „0“ annimmt und hält das Signal, wenn der Taktpuls den Wert „1“ annimmt. In den folgenden Diagrammen ist die S/H-Stufe der größeren Übersicht halber nicht eingezeichnet.

Flash-ADC

Der Aufbau des Flash-ADC (FADC) ist in Abb. 16.10 dargestellt. Die Eingangsspannung eines n-Bit-FADC wird gleichzeitig auf die positiven Eingänge von 2^n-Komparatoren gelegt. Die negativen Eingänge werden über eine Spannungsteilerkette an 2^n gleich große Teilspannungen angeschlossen.

FADCs werden zur Digitalisierung von schnellen Signalen und für hohe Zeitauflösung eingesetzt und mit entsprechend hohen Tastraten betrieben. Man kann Umwandlungsraten im Bereich von GSPS (Giga samples per second) erzielen, da nach der Reaktionszeit der Komparatoren das digitale Signal im wesentlichen schon vorliegt (in „Thermometercodierung“). Damit können Signale mit einer Zeitauflösung weit unter einer Nanosekunde digitalisiert werden. Nachteile sind der hohe Leistungs- und Platzverbauch. Daher ist in der Praxis die Auflösung von FADCs oft auf 8 Bits begrenzt oder die Zahl der Kanäle beschränkt. Außerdem sind FADCs teuer.

Der Thermometercode wird durch etwas Logik leicht in Binärode übersetzt (siehe Tab. 16.1). So entsprechen die Komparatorausgänge (0,1,1,1) der Binärzahl 11.

FADCs werden z.B. in Digitaloszilloskopen eingesetzt. Hier liefern die teuren Exemplare bis zu 80 GSPS bei 8 Bits für vier Kanäle.

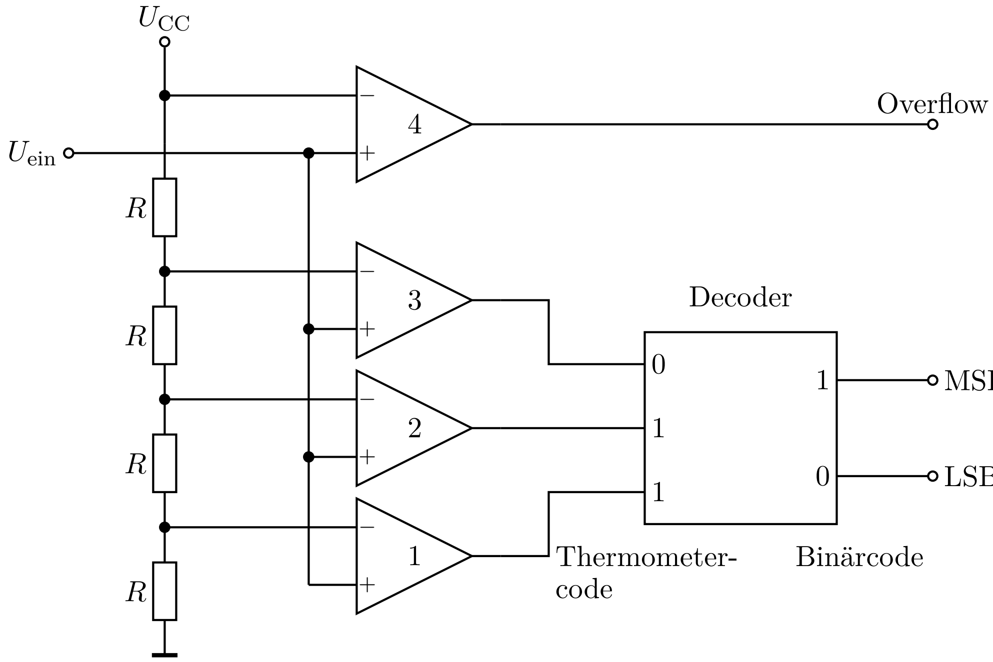


Abb. 16.10 Aufbau eines Flash-ADC (2-Bit) mit Spannungsteilerkette, Komparatoren und Decoder.

Komparator	Digitaler Code				
4	0	0	0	0	1
3	0	0	0	1	1
2	0	0	1	1	1
1	0	1	1	1	1
Overflow	0	0	0	0	1
MSB	0	0	1	1	0
LSB	0	1	0	1	0

Tab. 16.1 Komparatorausgänge und entsprechende Ausgänge des Decoders im Binärcode.

Wägeverfahren[7]

Der prinzipielle Aufbau eines ADC nach dem Wägeverfahren ist in Abb. 16.11 dargestellt. Beim Wägeverfahren wird das Analogsignal nicht gleichzeitig mit vielen Referenzsignalen verglichen wie beim Parallelverfahren (FADC), sondern nacheinander. Die Referenzsignale werden von der Steuerlogik so ausgewählt, dass möglich wenig Vergleichsschritte nötig sind. ADCs, die auf dem Wägeverfahren beruhen, benötigen nur einen Komparator, einen DAC, einen Speicher (SAR[8]) und etwas Logik.

Der Ablauf der Digitalisierung ist in Abb. 16.12 illustiert. Im ersten Taktzyklus wird U_{ein} mit der Referenzspannung $U_{\text{MSB}}/2$ verglichen. Falls $U_{\text{ein}} > U_{\text{MSB}}/2$, so ist der Ausgang des

[7]Engl. *successive approximation (SAR)*
[8]Engl. *successive approximation register*

Komparators positiv, sonst negativ. Damit ist der Wert des MSB schon bestimmt. Für n Bits werden n Taktzyklen und n Referenzspannungen benötigt.

Das Verfahren ermöglicht mittlere Umwandlungsgeschwindigkeiten ($\leq$ 5 Msps) und ist relativ genau (8 bis 18 Bits). Vorzüge sind die geringe Leistungsaufnahme und der moderate Platzverbrauch.

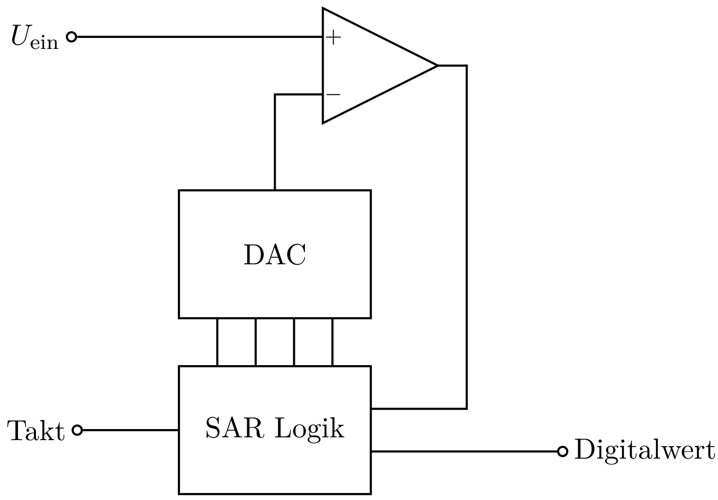


Abb. 16.11 Aufbau eines ADC nach dem Wägeverfahren.

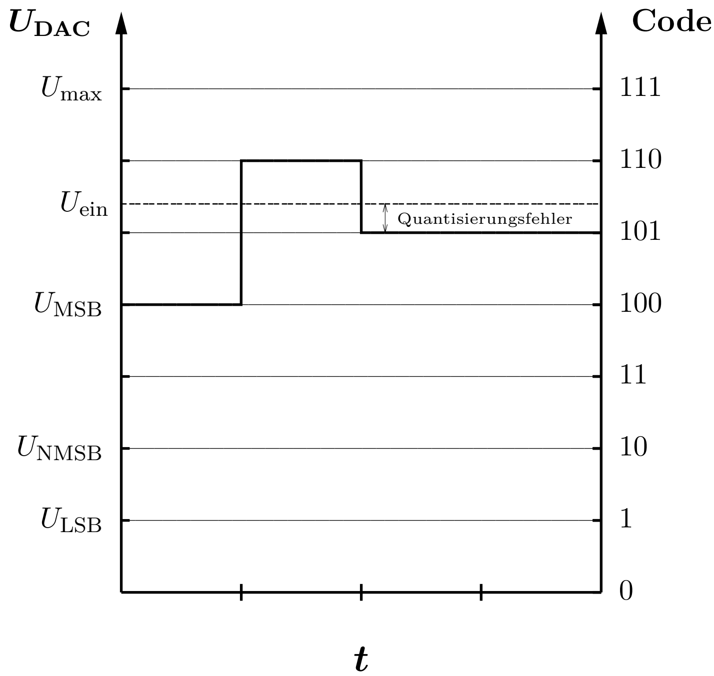


Abb. 16.12 Ilustration der Analog-Digital-Wandlung nach dem Wägeverfahren: Wert der Referenzspannung als Funktion der Zeit.

Zählverfahren[9]

Im Zählverfahren wird die analoge Eingangsspannung nicht mit diskreten Vergleichsspannungen verglichen wie beim Parallel- oder Wägeverfahren. Stattdessen wird eine Spannungsrampe erzeugt, die kontinuierlich ansteigende Rampenspannung mit der Eingangsspannung verglichen und die Zeit bzw. Anzahl der Taktzyklen gemessen bis die Rampenspannung die Eingangsspannung erreicht hat. Das Blockdiagramm eines ADC nach dem Zählverfahren ist in Abb. 16.13 dargestellt.

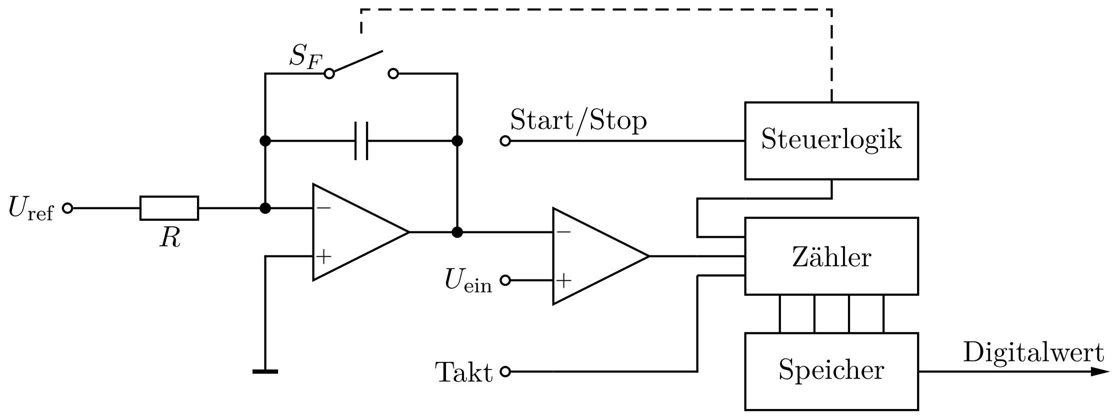


Abb. 16.13 Schematischer Aufbau eines ADC nach dem Zählverfahren mit Rampe, Zähler, Steuerlogik und Speicher.

Die Rampe wird durch einen Integrator realisiert, der von einer Strom- oder Spannungsquelle aufgeladen wird (Abb. 16.14, siehe auch Kapitel Operationsverstärker Abb. 4.19 (a)). Nach Ablauf einer Analog-Digital-Wandlung wird der Rückkopplungskondensator über den Schalter S wieder entladen. Ein realistischer Ablauf ist etwas komplexer und ist in Abb. 16.15 dargestellt. Hier wird der Offset (pedestal) berücksichtigt. Um Nichtlinearitäten zu vermeiden, wird der Zähler außerdem nicht bei der Rampenspannung Null gestartet, sonder einige Zyklen später.

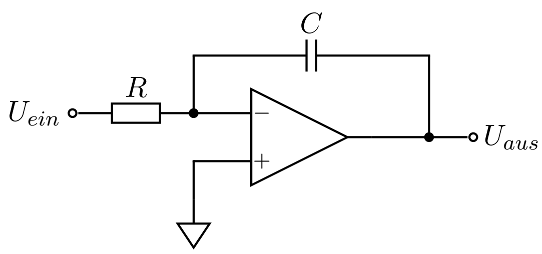


Abb. 16.14 Ein Integrator, realisiert mithilfe eines OPVs. Die integrierte Spannung ist invertiert, z. B. bei positiver, konstanter U_{ref} läuft die Ausgangsrampe ins Negative.

[9]Engl. *ramp, single slope*

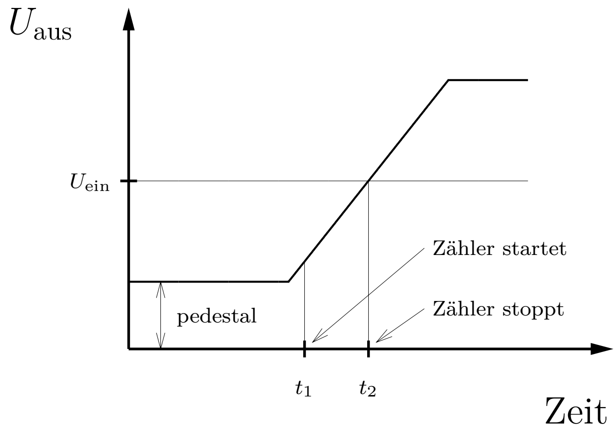


Abb. 16.15 Schematischer Aufbau der Digitalisierung eines ADC basierend auf dem Zählverfahren. Das Diagramm zeigt die Eingangsspannung und die Rampenspannung als Funktion der Zeit.

Für einen mit Operationsverstärker (OPV) realisierten Integrator gilt

$$U_{\text{aus}} \quad = \quad -\frac{1}{R\,C}\int U_{\text{ein}}\,\mathrm{d}t \quad ,\text{da} \quad \frac{U_{\text{ein}}}{R} = -\,C\,\frac{\mathrm{d}U_{\text{aus}}}{\mathrm{d}t} = I_C \quad . \tag{16.5}$$

Somit ist der Zählerstand nach Erreichen einer Schwellenspannung proportional zur Eingangsspannung U_{ein}: $Z_{\text{Zähler}} \sim U_{\text{ein}}$.

Das Zählverfahren bietet sich auch für ADCs mit vielen Eingangskanälen an. Dann wird die Rampenspannung mit allen Eingangsspannungen parallel verglichen und muss nur einmal erzeugt werden.

Für einen idealen ADC bestimmt die Steigung der Rampe und die Zählrate die mögliche Auflösung. Eine langsame Anstiegszeit erhöht die Spannungsauflösung auf Kosten der Zeitauflösung des Signals.

Zweirampenverfahren[10]

Eine Verfeinerung des Aufbaus aus Abb. 16.13 ist in Abb. 16.16 dargestellt. Hier wird das Zweirampenverfahren („dual slope“) angewendet. Zunächst wird die Eingangsspannung für eine feste Zeit t_1 integriert, dann wird die Integration mit der negativen Referenzspannung fortgesetzt. Die Zeit wird mit einem Taktgeber plus Zähler gemessen.

[10] Engl. *dual slope*

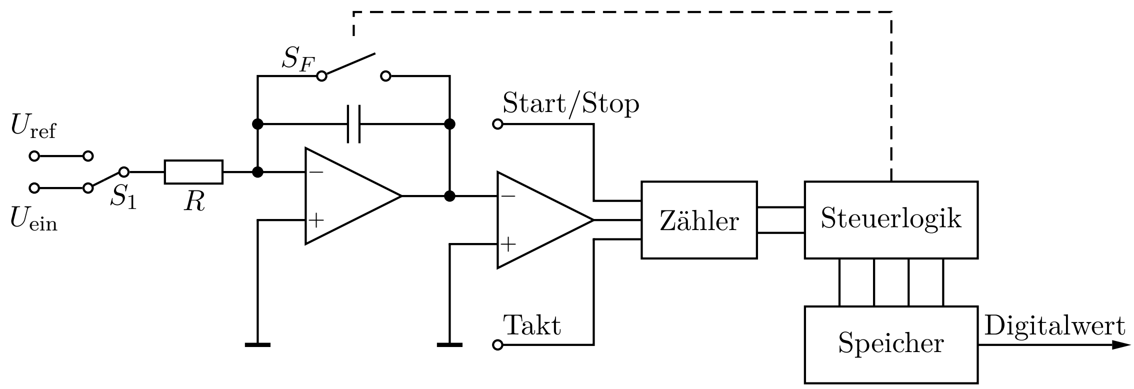


Abb. 16.16 Prinzipieller Aufbau eines Dual-slope ADC: Schalter, Integrator, Komparator (Schwellschalter), Zähler und Steuerung.

Gemäß dem Prinzipaufbau eines Dual-Slope-ADC erfolgt die Konvertierung der Eingangsspannung in eine Zahl in folgenden Schritten:

1. Vorbereitung: Der Kondensator ist entladen und die Ausgangsspannung ist 0 . Der Zähler steht auf „0“. Der Schalter S_F ist offen.

2. Aufladen: Der Kondensator wird aufgeladen, der Zähler zählt bis Z_{max}. Der Eingang des Integrators ist mit der zu digitalisierenden Spannung U_{ein} verbunden.

3. Entladen: Der Kondensator wird entladen, der Zähler zählt bis zum Erreichen der Ausgangsspannung Null. Der Eingang des Integrators ist mit der konstanten Referenzspannung verbunden. Das Vorzeichen von U_{ref} wird so gewählt, dass es entgegengesetzt zu U_{ein} ist.

Der Ausgangswert des ADC wird aus dem Zählerstand gewonnen, der bis zum Umkippen des Schwellschalters erreicht wird. Es gilt:

$$\begin{aligned} U_{\mathrm{aus}} = 0\,\mathrm{V} &= -\frac{1}{RC}\int_0^{t_1} U_{\mathrm{ein}}\,\mathrm{d}t - \frac{1}{RC}\int_{t_1}^{t_2} U_{\mathrm{ref}}\,\mathrm{d}t \\ &= -\frac{1}{RC} U_{\mathrm{ein}}\, t_{\mathrm{Laden}} - \frac{1}{RC} U_{\mathrm{ref}}\, t_{\mathrm{Entladen}} \quad . \end{aligned} \tag{16.6}$$

Daraus folgt

$$U_{\mathrm{ein}} = -U_{\mathrm{ref}} \cdot \frac{t_{\mathrm{Entladen}}}{t_{\mathrm{Laden}}} \quad . \tag{16.7}$$

Der Dual-slope-ADC ist somit unabhängig von den Werten von R und C sowie der Taktperiode des Zählers T bzw. der Taktfrequenz $f = \frac{1}{T}$. Die Taktfrequenz f braucht nur für kurze Zeit stabil zu sein, nämlich bis zum Erreichen von t_2. Ein Dual-slope-ADC benötigt wenig elektrische Leistung, relativ wenig Platz und ist sehr genau ($\approx$ 100 ppm), aber nicht sehr schnell.

Der Ablauf der Digitalisierung ist in Abb. 16.17 für zwei unterschiedliche Eingangsspannungen U_{ein} dargestellt. Die Strecken links haben unterschiedliche Steigungen abhängig von U_{ein}. Die Strecken rechts haben identische Steigungen, da U_{ref} konstant ist.

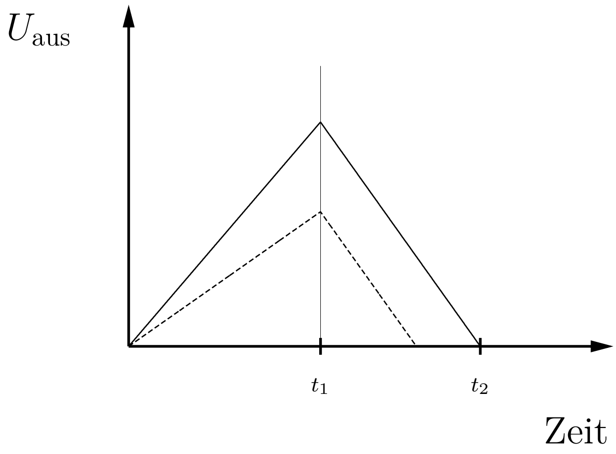


Abb. 16.17 Ausgangsspannung als Funktion der Zeit für zwei unterschiedliche Eingangsspannungen.

Der Sigma-Delta-Wandler*

Der Sigma-Delta-Wandler ($\Sigma\Delta$-Wandler) ist ein wichtiger Wandlertyp, der für langsame Signale sehr hohe Auflösung (22 Bit) bei geringen Kosten ermöglicht.

Der schematische Aufbau des $\Sigma\Delta$-Wandlers ist in Abb. 16.18 dargestellt. Wesentliche Blöcke sind ein Differenzverstärker, Integrator, Komparator und DAC, sowie ein digitaler Filter und Logik.

Die Analogbausteine sind einfach gehalten, DAC und Komparator haben nur eine Auflösung von einem Bit. Die hohe Auflösung des $\Sigma\Delta$-Wandlers wird durch Methoden der digitalen Signalverarbeitung erzielt, nämlich durch „Oversampling“, Rauschformung und digitale Filter.

Wir wollen zuerst die grundlegende Funktionsweise betrachten. Im Schaltkreis ist eine Rückkopplungsschleife enthalten, da der Ausgang des Komparators über den 1-Bit-DAC die Spannung am invertierenden Eingang des Differenzverstärkers steuert.

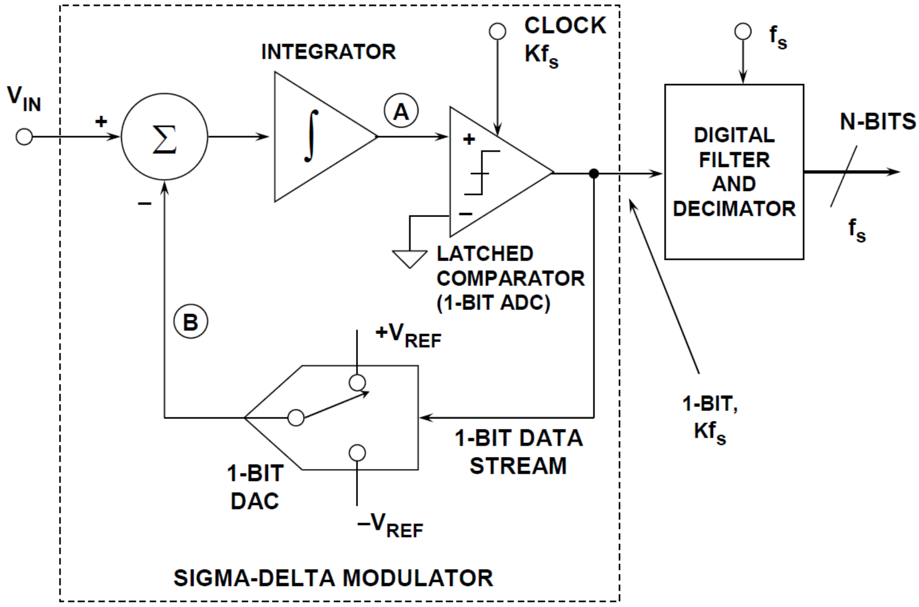


Abb. 16.18 Schematischer Aufbau eines ΣΔ-Wandlers mit Differenzverstärker (hier als Σ symbolisiert), Integrator, neuem Komparator, 1-Bit-DAC und digitalen Filtern (Quelle: MT-022).

Wir wollen uns den zeitlichen Ablauf der Digitalisierung des ΣΔ-Wandlers veranschaulichen und wählen dazu besonders anschauliche Parameter.

Der Differenzverstärker habe die Verstärkung 1 V/V. Seine Eingangsspannungen sind E_1 und E_2. Die Differenzspannung $x = E_1 - E_2$. Der Integrator summiert einfach seine Eingangsspannung $x(n)$ zu der bisherigen Ausgangsspannung $y(n) = y(n+1) = x(n) + y(n)$, die Komparatorschwelle liegt bei 0,5 V und der DAC-Ausgang sei 0 V für den Eingangszustand „0“ und 1 V für „1“. Bei einer Eingangsspannung von 0,5 V (bzw. 0,75 V) ergibt sich die Bitfolge z am Ausgang des Komparators wie folgt

n	E_1	E_2	$x = E_1 - E_2$	$y(n) = x(n) + y(n-1)$	z
1	$0,5$	0	$0,5$	$0,5$	1
2	$0,5$	1	$-0,5$	0	0
3	$0,5$	0	$0,5$	$0,5$	1
4	$0,5$	1	$-0,5$	0	0
1	$0,75$	0	$0,75$	$0,75$	1
2	$0,75$	1	$-0,25$	$0,5$	1
3	$0,75$	1	$-0,25$	$0,25$	0
4	$0,75$	0	$0,75$	1	1
5	$0,75$	1	$-0,25$	$0,75$	1
6	$0,75$	1	$-0,25$	$0,5$	1
7	$0,75$	1	$-0,25$	$0,25$	0
8	$0,75$	0	$+0,75$	1	1

Tab. 16.2 Illustration des Funktionsprinzips eines ΣΔ-Wandlers

Man sieht an diesem einfachen Beispielen, dass sich abhängig von der Eingangsspannung eine charakteristische Bitfolge einstellt. Bestimmt man die mittlere Zahl der „1"-sen z.B. $\frac{2}{4} = 0,5$ oder $\frac{3}{4} = 0,75$ und multipliziert man sie mit der DAC-Spannung ergibt sich das Eingangssignal. Je höher die Taktfrequenz ist, desto genauer kann man die Eingangsspannung abbilden. Man kann also mit einem 1-Bit-ADC und dem Komparator eine hohe Auflösung erzielen, wenn die Taktfrequenz, also die Abtastfrequenz fs, deutlich höher ist als die Nyquist-Frequenz f_N (Die Nyquist-Frequenz wiederum entspricht der doppelten Bandbreite des Signals $f_\mathrm{N} = 2\,f_B$, damit kein Aliaseffekt auftritt (siehe Abschnitt 16.3). Abb. 16.19 veranschaulicht die Wirkung des Oversampling und der Rauschformung.

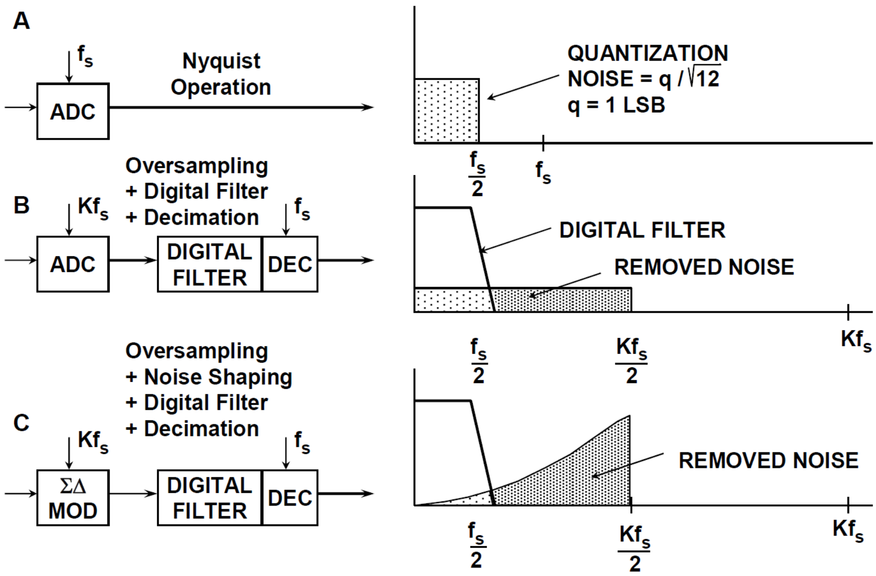


Abb. 16.19 Darstellung der Frequenzantwort des ΣΔ-Wandlers im Frequenzraum (Quelle: MT-022). Ohne Oversampling (oben), mit Oversampling und digitalen Filtern (Mitte) und zusätzlich mit Rauschformung (unten).

Schneidet man nun im Frequenzraum mithilfe eines digitalen Tiefpassfilters die Frequenzen oberhalb der Signalteile heraus, so nimmt das Signal-zu-Rausch-Verhältnis SNR deutlich zu. Ein sehr einfacher Digitalfilter wäre z.B. der sogenannte „Moving average"-Filter. Hier würde der Datenstrom $z[n]$ durch die Werte $z'[n] = \frac{1}{N}\sum_{0}^{N-1} z[n-k]$ ersetzt, also durch den Mittelwert von $z(n)$ und der $(N-1)$ vorangegangenen Werte z.

Für die oben genannten Beispiele wäre ein Mittel über 4 Werte ausreichend. Aus der Bitfolge 1010 1010 würde dann $\frac{1}{4}\frac{1}{4}\frac{2}{4}\frac{2}{4}\;\frac{2}{4}\frac{2}{4}\frac{2}{4}\frac{2}{4}$ und aus 1101 1101 1101 würde $\frac{1}{4}\frac{2}{4}\frac{2}{4}\frac{3}{4}\;\frac{3}{4}\frac{3}{4}\frac{3}{4}\frac{3}{4}\;\frac{3}{4}\frac{3}{4}\frac{3}{4}\frac{3}{4}$. (Hier wurde für den ersten Mittelwert angenommen, dass die unbekannten vorangegangenen drei Werte alle gleich 0 sind.) Durch das Oversampling ist die Datenrate jetzt sehr hoch und enthält redundante Information. Es reicht, wenn man nur den m-ten Wert berücksichtigt. Dieser Schritt heißt Dezimierung oder Downsampling. D.h. im obigen Beispiel wird statt der Bitfolge $\frac{1}{4}\frac{2}{4}\frac{2}{4}\frac{3}{4}\ldots\frac{3}{4}$ bei der Wahl von $m = 4$ nur die Folge $\frac{3}{4}\frac{3}{4}\frac{3}{4}$ übertragen.

Der Wert von m muss so gewählt werden, dass gilt $f_\mathrm{s}/m \geq 2\,f_\mathrm{N}$, damit kein Aliaseffekt auftritt.

Der $\Sigma\Delta$-Wandler hat noch eine weitere Besonderheit, die die Auflösung wesentlich verbessert und zwar die Rauschformung. Das Rauschen wird nicht gleichförmig auf alle Frequenz-Intervalle verteilt, sondern bevorzugt auf hohe Frequenzen (siehe Abb. 16.19 unten).

Warum das so ist, sieht man an Abb. 16.20. Hier wird das Verhalten des $\Sigma\Delta$-Wandlers im Laplace-Raum betrachtet. Der Differenzverstärker ist als Summierer dargestellt, der Integrator hat die Übertragungsfunktion $\frac{1}{f}$. Der Komparator als die Quelle des Quantisierungsrauschen wird auch als Summierer dargestellt.

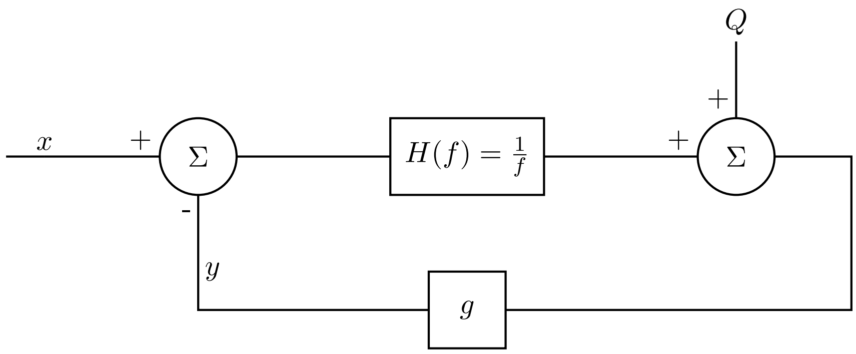


Abb. 16.20 Illustration des Effekts der Rauschformung.

Dieser Schaltblock des $\Sigma\Delta$-Wandlers ist hier vereinfacht, indem der DAC durch einen Verstärkungsfaktor g ersetzt wurde.

Mit

$$y = g \cdot \left[\frac{1}{f}(x-y) + Q\right] \tag{16.8}$$

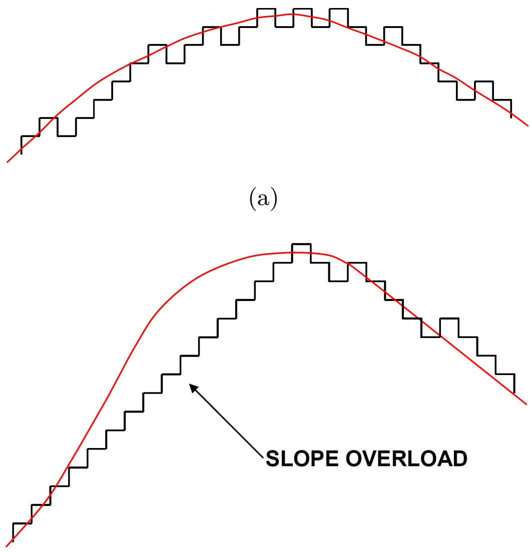
folgt

$$y = \frac{x}{\frac{f}{g}+1} + \frac{Q \cdot \frac{f}{g}}{\frac{f}{g}+1} \tag{16.9}$$

Für $f \approx 0$ gilt $y \approx x$, nur dass Signal x wird an den Ausgang weitergegeben. Für große f hingegen dominiert das Rauschen Q. Der Integrator wirkt effektiv als Tiefpass für das Signal und als Hochpass für das Rauschen. Diesen Effekt kann man noch verstärken, indem man mehr als einen Integrator und Summierer verwendet.

Wir kommt dieser Effekt zustande? Das liegt daran, dass das Quantisierungsrauschen nach dem Integrator auftritt und dadurch eine andere Übertragungsfunktion erfährt.

Wichtig ist bei der $\Sigma\Delta$-Wandlung, dass sich das Eingangssignal nicht zu schnell ändert (siehe Abb. 16.21)



(a)

(b)

Abb. 16.21 Illustration des Verhalten eines ΣΔ-Wandlers für Signale unterschiedlicher Steilheit. (Quelle: MT-022)

In Abb. 16.21(a) folgt der Wandler in guter Näherung dem Kurvenverlauf. In Abb. 16.21(b) ist der Wandler zu langsam.

ΣΔ-Wandler werden in der Audioelektronik fast aussschliesslich verwendet. Die analoge Bandbreite liegt bei rund 20 kHz, der menschlichen Hörschwelle. Die auf der CD gespeicherte Datanrate ist 44,1 kHz. Die Abtastrate des analogen Audiosignals mit 64-fachen Oversamplingfaktor (bezogen auf die Nyquistfrequenz) ist rund 3 MHz.

16.3 Abtasttheorem

Will man eine Pulsform mit hoher Präzision digital erfassen, so muß die Abtastrate groß genug sein. Wie groß? Das kommt auf die Bandbreite B, d. h. auf das im Eingangssignal enthaltene Frequenzspektrum an. Grundsätzlich wird das Signal mit mindestens der doppelten Frequenz $f > 2\,B$ abgetastet. Dieser Zusammenhang wurde von Nyquist und Shannon untersucht [17]. Die Frequenz $f_N = 2B$ wird auch als Nyquist-Frequenz bezeichnet. Im Zeitraum heißt das, dass der minimale Zeitabstand zwischen zwei Werten (Abtastungen) weniger als $\frac{1}{2\,B}$ betragen muss. Dann lässt sich das Eingangssignal ohne Informationsverlust wieder reproduzieren.

In der Praxis gilt das Theorem nur in Näherung. Hohe Frequenzen sind nie exakt abgeschnitten. Außerdem wird der Quantisierungsfehler ignoriert.

Falls die Samplingfrequenz kleiner als das Doppelte der im Signal enthaltenen hohen Frequenzen ist, tritt Aliasing auf. Dies kann man mit einem vorgeschalteten Tiefpass reduzieren. Dadurch werden Frequenzen im Signal, die höher als die halbe Samplingfrequenz sind, stark gedämpft.

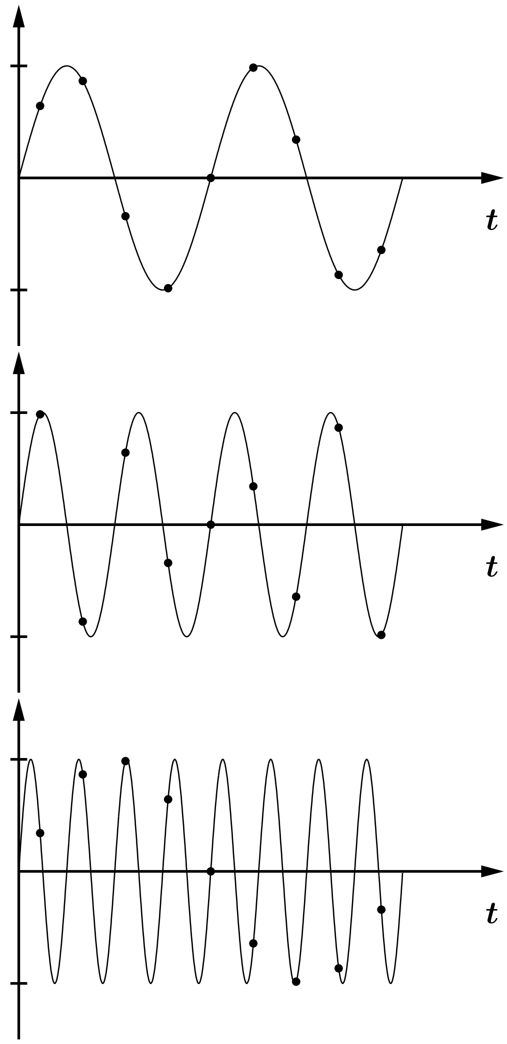


Abb. 16.22 Illustration des Aliaseffekts (Aliasing) bei unterschiedlichen Abtastraten eines harmonischen Signals. Ist die Abtastrate zu gering (unten), wird eine harmonische Funktion geringerer Frequenz suggeriert.

16.4 ADC-Fehler

Eine Analog-Digital-Wandlung ist nie exakt, sondern fehlerbehaftet. Fehler sind insbesondere

1. Qantisierungsfehler
2. Nullpunkts- bzw. Offsetfehler
3. Verstärkungs- bzw. Gainfehler [11]
4. Differentieller Nichtlinearitätsfehler [12]

[11] Engl. *gain error*
[12] Engl. *differential nonlinearity*

5. Integraler Nichtlinearitätsfehler [13]

6. Clock jitter

7. Fehlende Digitalwerte

Quantisierungsfehler

Da der ADC digitale Ausgangswerte mit beschränkter Auflösung liefert, tritt ein sogenannter Diskretisierungs- oder Quantisierungsfehler auf. Der angezeigte Digitalwert ist also meist etwas zu groß oder zu klein im Vergleich zu der dem Analogwert exakt entsprechenden Digitalzahl. Im Idealfall wäre die Übertragungskurve die gestrichelte Gerade aus Abb. 16.2. Nur die Spannungen 0 V, U_{LSB} und alle ganzzahligen Vielfachen von U_{LSB} entsprechen den idealen Digitalwerten. Die maximale Abweichung von der Idealwert tritt für Analogspannungen knapp unter dem Wert $n\,U_{\mathrm{LSB}}$ auf und beträgt fast ein U_{LSB}. Man kann die maximale Diskrepanz halbieren, wenn man die Stufenfunktion um $U_{\mathrm{LSB}}/2$ nach links verschiebt. Dies ist in Abb. 16.23 dargestellt. Die entsprechenden Abweichung von der idealen Übertragungskurve in Abb. 16.24.

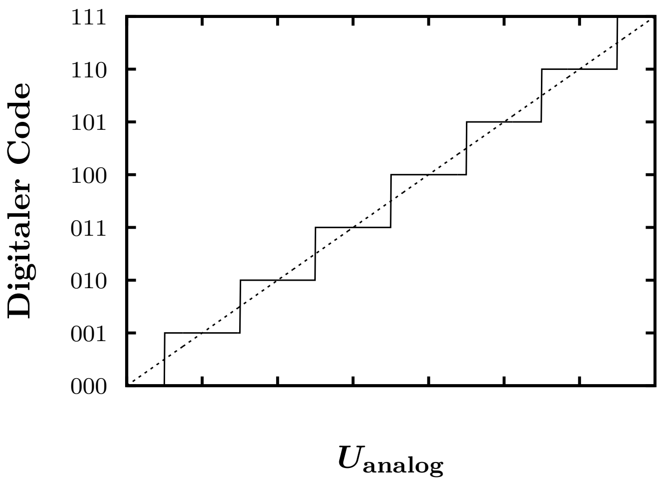


Abb. 16.23 Übertragungsfunktion eines idealen ADC. Die gestrichelte Gerade dient nur zur Veranschaulichung der Treppenfunktion.

[13] Engl. *integral nonlinearity*

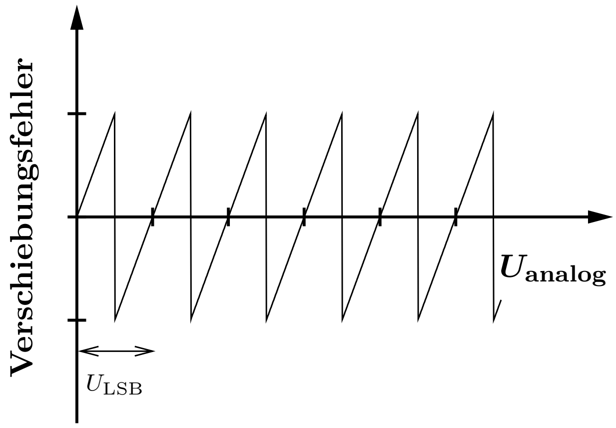


Abb. 16.24 Darstellung des Quantisierungsfehlers: Zwischen den Eingangsspannungen eines Abschnittes wird nicht binär entschieden.

Zur Analyse des Quantisierungsfehlers QE bestimmen wir die mittlere quadratische Abweichung der Eingangsspannung U_{analog} vom Digitalwert, ausgedrückt in Digitalcode U_{LSB}, im Intervall $[1/2\,U_{\text{LSB}},\,3/2\,U_{\text{LSB}}]$. Dies entspricht der einfachsten Rechnung. Natürlich hätte man auch andere geeignete Intervalle oder den gesamten Wertebereich des ADC analysieren können. Daraus ergibt sich dann

$$\begin{aligned}
\overline{QE^2} &= \frac{1}{U_{\text{LSB}}}\int_{\frac{1}{2}\,U_{\text{LSB}}}^{\frac{3}{2}\,U_{\text{LSB}}} (U_{\text{analog}} - U_{\text{LSB}})^2 \; dU_{\text{analog}} \\
&= \frac{1}{U_{\text{LSB}}}\int_{-\frac{1}{2}\,U_{\text{LSB}}}^{+\frac{1}{2}\,U_{\text{LSB}}} U_{\text{analog}}^2 \; dU_{\text{analog}} \\
&= \frac{1}{U_{\text{LSB}}}\left[\frac{U_{\text{analog}}^3}{3}\right]_{-\frac{U_{\text{LSB}}}{2}}^{+\frac{U_{\text{LSB}}}{2}} = \frac{U_{\text{LSB}}^2}{12}\,.
\end{aligned}$$

Das RMS des Quantisierungsfehlers ist also $U_{\text{LSB}}/\sqrt{12}$.

Mit Hilfe des so bestimmten Rauschens kann das Signal-zu-Rauschverhältnis SNR[14] für ein Sinussignal bei Vollausschlag berechnet werden:

Das Rauschen entspricht dem Quantisierungsfehler QE (siehe oben). Das mittlere Signal (RMS) ergibt sich aus

$$\frac{U_{\text{LSB}}}{2}\,2^n \sin 2\,\pi\,f\,t \quad\Rightarrow\quad S \;=\; \frac{U_{\text{LSB}}\,2^n}{2\,\sqrt{2}}\,.$$

Und somit der db-Wert des SNR

$$SNR \;=\; 20\,\log_{10}\frac{\frac{2^n\,U_{\text{LSB}}}{2\sqrt{2}}}{\frac{U_{\text{LSB}}}{\sqrt{12}}} = 20\,\log 2^n + 20\,\log\sqrt{\frac{3}{2}} \approx 6\,n + 1,76\,\text{db}\,.$$

[14] signal-to-noise ratio

Offsetfehler

Der Offsetfehler (Gleichanteilverschiebung) verschiebt die gesamte Übertragungskennlinie um einen festen Wert. „Die Treppe beginnt zu früh oder zu spät.“

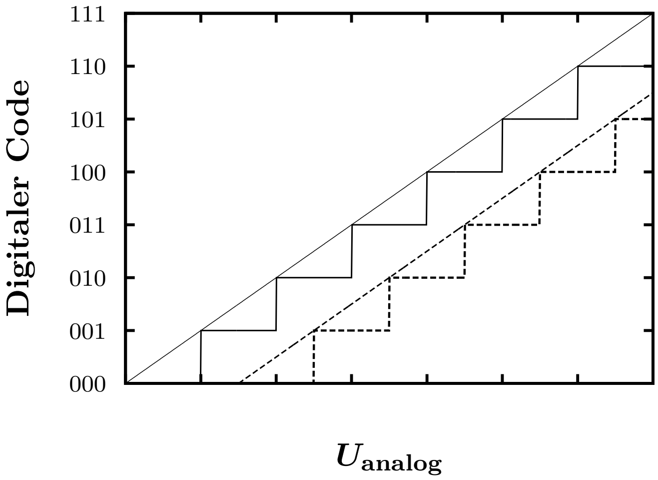


Abb. 16.25 Darstellung des Offsetfehlers. Die gepunktete Übertragungskennlinie enthält einen Offsetfehler und ist gegenüber der idealen Lennlinie (durchgezogen) verschoben.

Gainfehler

Beim Gain-Fehler oder Skalenfaktor-Fehler ist die Skalierung der Übertragungskennlinie größer oder kleiner als 1. „Die Treppe hat zu breite oder zu schmale Stufen, aber alle Stufen sind gleich hoch.“

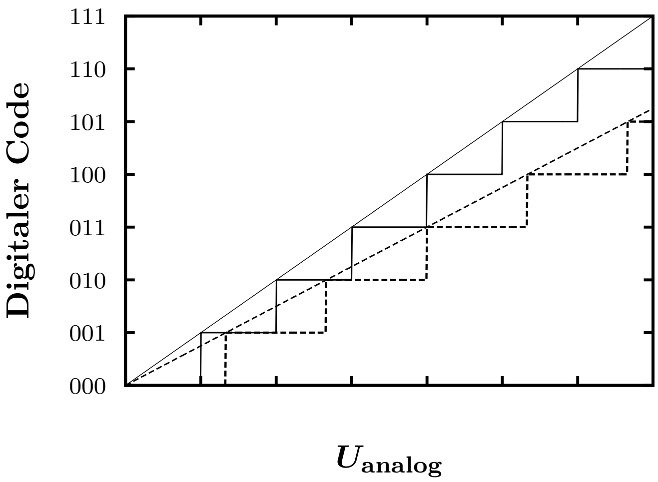


Abb. 16.26 Darstellung des Gainfehlers: Der Skalenfaktor der gepunkteten Kennlinie ist fehlerhaft.

Gain- und Offsetfehler werden in Summe als „Full Scale Error“ bezeichnet.

Differentieller Nichtlinearitätsfehler (DNL)

$$\mathrm{DNL}_i = \frac{(U_{i+1} - U_i) - U_{\mathrm{LSB}}}{U_{\mathrm{LSB}}} = \frac{U_{i+1} - U_i}{U_{\mathrm{LSB}}} - 1 \tag{16.10}$$

U_{i+1} entspricht der Analogspannung, an der der i-te Digitalwert zum $(i+1)$-ten Digitalwert springt. Der DNL ist ein Maß für die Abweichung der Stufenbreite vom Sollwert (U_{LSB}). Jede Stufe kann individuell vom Soll abweichen. Bei einem idealen ADC gilt DNL$= 0$ für alle Stufen.

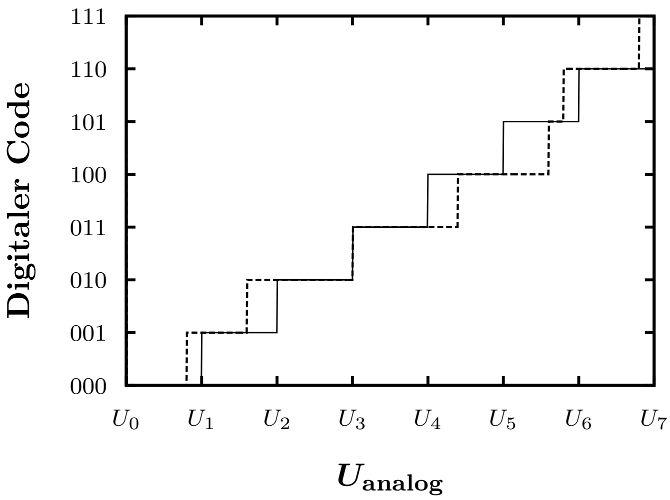


Abb. 16.27 Darstellung des differentiellen Nichtlinearitätsfehlers.

Fehlende Digitalwerte

Hier variert die Stufenhöhe, bzw. es fehlt eine Stufe oder die Stufe hat die Breite 0. Falls $\mathrm{DNL}_i > 1$, wird ein Digitalwert nie erreicht.

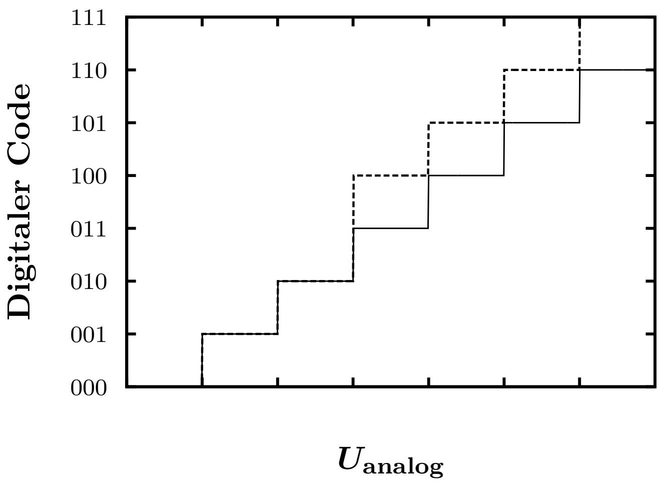


Abb. 16.28 Ein fehlender Digitalwert.

Integraler Nichtlinearitätsfehler (INL)

$$\mathrm{INL}_i = \frac{(U_i - U_0) - i\,U_{\mathrm{LSB}}}{U_{\mathrm{LSB}}} = \frac{U_i - U_0}{U_{\mathrm{LSB}}} - i \tag{16.11}$$

Der INL-Fehler ist das Maß für die Abweichung der Höhe (nicht der Breite) der i-ten Stufe vom Sollwert bzw. für die Abweichung der Position der i-ten Stufe zu der nominellen Stufenkante. Der INL der i-ten Stufe entspricht der Summe aller vergangenen DNLs, z.B.

$$\begin{aligned}\mathrm{INL}_2 &= \frac{U_2 - U_0}{U_{\mathrm{LSB}}} - 2 \\ &= \mathrm{DNL}_1 + \mathrm{DINL}_2 = \frac{U_1 - U_0}{U_{\mathrm{LSB}}} - 1 + \frac{U_2 - U_1}{U_{\mathrm{LSB}}} - 1\ .\end{aligned}$$

Die DNLs können sich kompensieren, wenn sowohl zu schmale und zu breite Stufen auftreten oder verstärken, z.B. wenn alle Stufen zu breit sind. Falls $\mathrm{INL} > 1\,\mathrm{LSB}$, wird ein Digitalwert des ADC übersprungen.

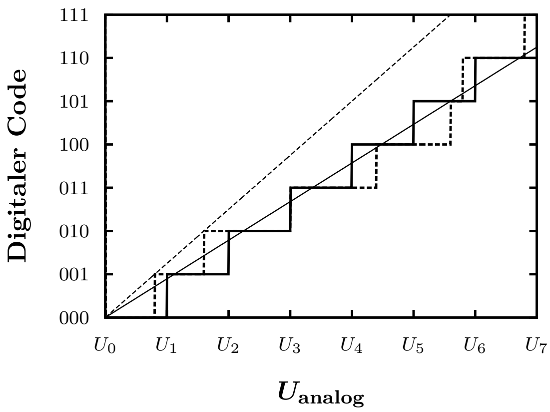


Abb. 16.29 Darstellung des integralen Nichtlinearitätsfehlers.

Im Gegensatz zu Offset und Verstärkungsfehlern kann man DNL und INL nicht kompensieren.

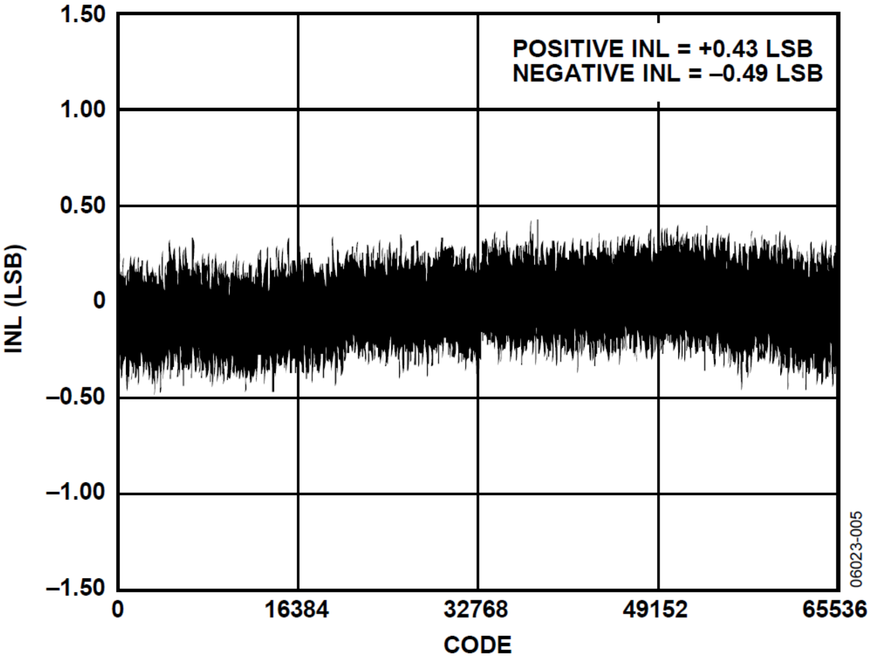


Abb. 16.30 Verlauf des INL-Fehlers über dem Digitalen Code für den AD7622 von Analog Devices (siehe Datenblatt).

Clock-Jitter

Ein Analogsignal wird in der Regel nicht direkt auf den ADC-Eingang gelegt,sondern erst auf eine vorangehende „Sample and Hold"-Schaltung (siehe Abb. 16.9). Hier wird das Signal zu festen Zeiten abgetastet und auf einen Kondensator gespeichert. Der Abtastzeitpunkt wird durch ein Tastsignal vergeben. Eine Abweichung des Tastsignals vom Soll nennt man „Clock-Jitter". Es ist äußerst wichtig, den Clock-Jitter sehr klein zu halten, am besten im Bereich von ps oder hunderten von Femtosekunden. Dies sieht man in Abb. 16.32.

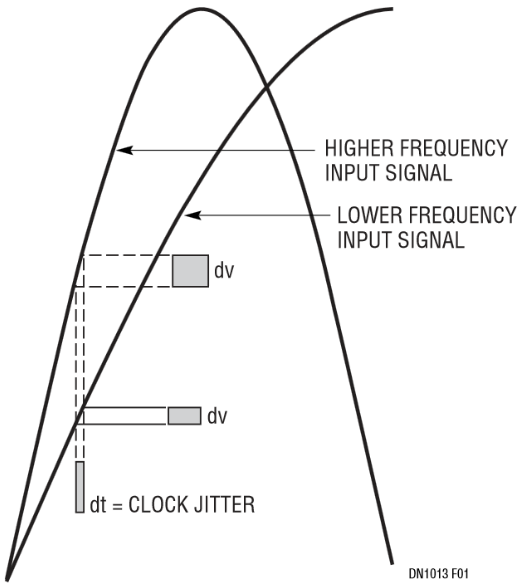


Abb. 16.31 Illustration des Effekts von Clock-Jitters ΔT auf die gespeicherte Signalamplitude für zwei mögliche Eingangssignale (Quelle: Design Note 1013).

Ein schnelles Signal, dessen Amplitude stark mit der Zeit variiert (hohe „ slew rate") führt bei einer Variation Δt des Taktpulses zu einen größeren Fehler als ein langsames Signal. Dies ist in anderer Form auch in Abb. 16.32 dargestellt.

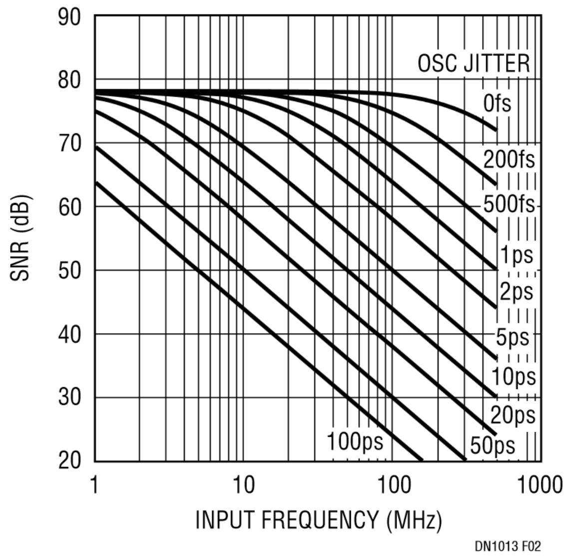


Abb. 16.32 Signal-zu-Rausch-Verhältnis (SNR) in db für ein harmonisches Eingangssignal fester Frequenz für verschiedene Jitterwerte (Quelle: Design Note 1013).

Man sieht, dass die Anforderungen an die Genauigkeit der Clock recht hoch sind. Die Kurven der Abb. 16.32 entsprechen der Gleichung

$$\mathrm{SNR(db)} \quad = \quad -20 \log (2\pi\, f_{\mathrm{ein}} \cdot \delta) \tag{16.12}$$

Hier ist f_{ein} die Frequenz des Eingangssignals (nicht der Clock!) und δ der mittlere Clock-Jitter.

Dynamisches Verhalten eines ADC

Ein ADC hat keine ideale lineare Übertragungsfunktion. Dadurch können nach der Digitalisierung im Ausgangsignal Frequenzkomponenten enthalten sein, die im Eingangssignal nicht vorkommen. Es gibt verschiedene Möglichkeiten diese Verzerrung zu quantifizieren. Eine davon ist die sogenannte „Total Harmonic Distortion“ (THD). Hier vergleicht man die Amplitude A_i der ersten fünf Oberwellen mit der Amplitude A_0 der Grundfrequenz.

$$THD = \sqrt{\frac{A_1^2 + A_2^2 + A_3^2 + A_4^2 + A_5^2}{A_0^2}} \quad . \tag{16.13}$$

Ein Beispiel ist in Abb. 16.33 gezeigt. Hier sind nur die Harmonischen 1. und 2. Ordnung bei 200 kHz und 300 kHz von Belang. Die höheren Frequezkomponenten sind stark abgeschwächt. Da die Amplitude der 2. Oberwelle bei 300 kHz mit rund -100 db dominiert, ergibt sich der THD dieses ADC zu rund -100 db.

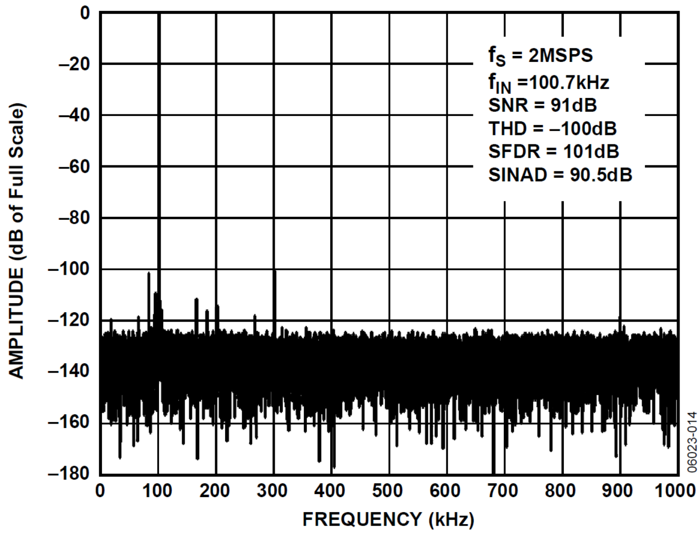


Abb. 16.33 Dynamischen Verhaltens und Rauschen des AD7622 (siehe Datenblatt) für eine harmonische Eingangsschwingung von 100 kHz.

17 Programmierbare Digitallogik

Programmierbare Digitallogik besteht aus den uns bekannten Logikgattern und Speicherbausteinen. Im Gegensatz zu diesen kann der Benutzer die Funktion dieser Elemente aber selbst einstellen und dadurch äußerst flexibel maßgeschneiderte komplexe Schaltungen aufbauen.

Programmierbare Bausteine treten in den unterschiedlichsten Varianten auf. Eine Übersicht ist in Abbildung 17.1 dargestellt. Verwandt mit den aufgeführten Bausteinen ist der FPGA, der aber weit leistungsfähiger als selbst ein CPLD ist.

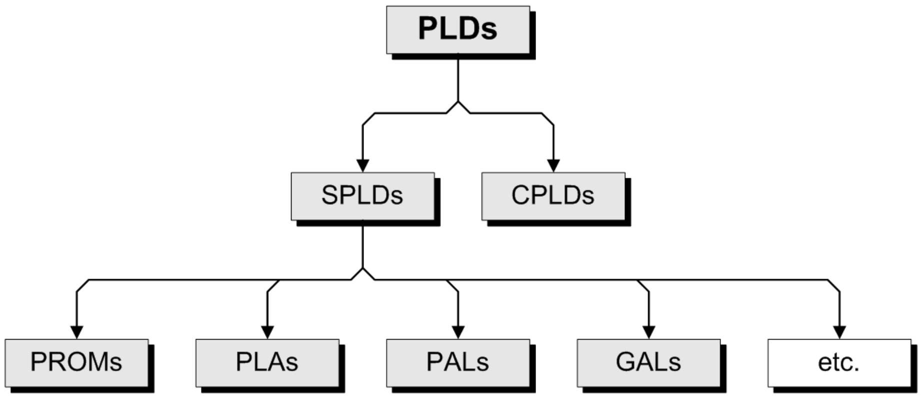


Abb. 17.1 Programmierbare Logikbausteine [18].

Die benutzten gängigen Akronyme sind in Tabelle 17.1 zusammengefasst.

PLD	Programmable Logic Device
SPLD	Simple Programmable Logic Device
CPLD	Complex Programmable Logic Device
PROM	Programmable Read-only-Memory
PLA	Programmable Logic Array
PAL	Programmable Array Logic
GAL	Gate Array Logic
FPGA	Field Programmable Gate Array

Tab. 17.1 Gängige Akronyme programmierbarer Digitallogikbausteine.

FPGA sind in der Detektorinstrumentierung, der in Datenauslese und für schnelle Triggersystemen unersetzlich. FPGAs sind auch in industriellen Anwendungen als sogenannte „embedded systems“ weit verbreitet und kommen z. B. im Auto zum Einsatz.

Wir behandeln eine moderne FPGA-Familie am Ende dieses Kapitels.

17.1 Rückblick: Schaltalgebra

Leider gibt es eine Vielzahl von Symbolen für die grundlegenden Verknüpfungen. Nicht alle sind gleich bequem. Für ODER, UND und NICHT werden z. B. die in der Tabelle 17.2 eingetragenen Symbole gerne verwendet. Ich empfehle die erste Zeile, die angelsächsische Notation, da hier die Rechenregeln für logische Operationen denen der Multiplikation und Addition reeller Zahlen sehr ähnelt.

ODER	**UND**	**NICHT**
$a+b$	$a \cdot b$	$\overline{a}$
a | b	a & b	!a
$a \vee b$	$a \wedge b$	$\neg a$

Tab. 17.2 Symbole für die logischen Verknüpfungen ODER, UND und NICHT.

17.2 SPLDs

SPLDs sind die Klasse der sogenannten **einfachen** programmierbaren Logikbausteine. Dazu gehören die PROMs, PLAs und PALs. Wie bei Speichern gibt es auch hier sowohl einmalig wie mehrfach programmierbare Varianten dieser Bausteine.

Ein **PROM** (Kap. 15) ist ein Logikbaustein, der logische Ausdrücke als *Summe von Produkten* darstellt. Dabei ist die UND-Matrix vorgegeben und die ODER-Matrix kann programmiert werden (Abb. 17.2 und 17.3).

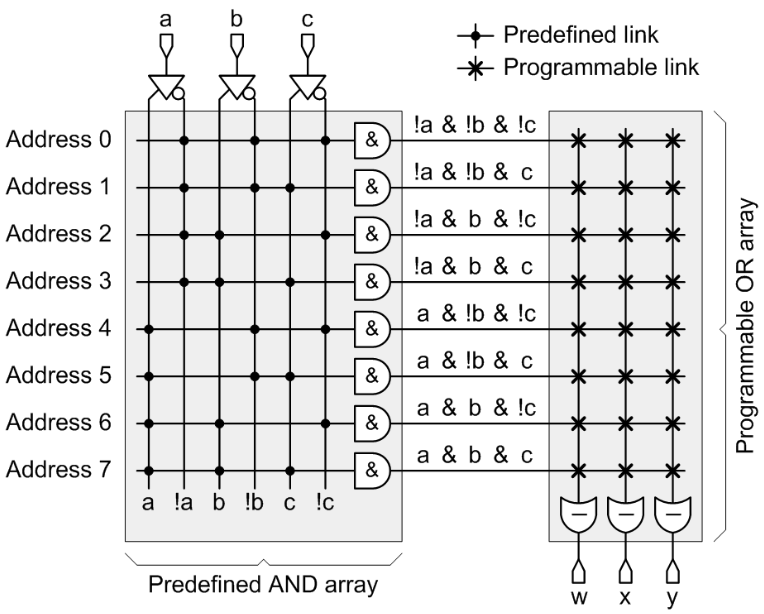


Abb. 17.2 PROM mit den Adresseingängen a, b, c und den Datenausgängen w, x, y vor der Programmierung. Links sind die UND- oder Produktterme, rechts die ODER- oder Summenterme [18].

Das PROM in Abbildung 17.2 und 17.3 hat acht Adressen und dementsprechend drei Eingangsvariablen. Alle kombinatorisch möglichen UND-Terme (siehe Tab. 17.3) werden in die ODER-Matrix geleitet und dort wie benötigt verknüpft. Im Gegensatz zu den früher vorgestellten Beispielen von Normalformen treten in Abbildung 17.3 nur Summenterme mit genau drei Variablen auf. Erst durch Anwendung der Regeln der Schaltalgebra erhält man die vereinfachten Ausdrücke aus Abbildung 17.3 (siehe auch Tab. 17.4).

Adresse	**UND – Ausgang**
0	$\overline{a}\overline{b}\overline{c}$
1	$\overline{a}\overline{b}c$
2	$\overline{a}b\overline{c}$
3	$\overline{a}bc$
4	$a\overline{b}\overline{c}$
5	$a\overline{b}c$
6	$ab\overline{c}$
7	abc

Tab. 17.3 Logische Verknüpfung am Ausgang der UND-Matrix.

Ausgang	**Logische Verknüpfung am Ausgang des PROM**
w	$ab\overline{c} + abc = ab$
x	$\overline{a}\overline{b}\overline{c} + \overline{a}\overline{b}c + \overline{a}b\overline{c} + \overline{a}bc + a\overline{b}\overline{c} + a\overline{b}c = \overline{a}\overline{b} + \overline{a}b + a\overline{b} = \overline{a} + a\overline{b} = \overline{ab}$
y	$\overline{a}\overline{b}c + \overline{a}bc + a\overline{b}c + ab\overline{c} = \overline{a}c + a\overline{b}c + ab\overline{c} = \overline{a}c + \overline{b}c + ab\overline{c} =$ $\left(\overline{a} + \overline{b}\right) c + ab\overline{c} = \overline{ab}c + ab\overline{c} = ab \oplus c$

Tab. 17.4 Logische Verknüpfung am Ausgang des PROM.

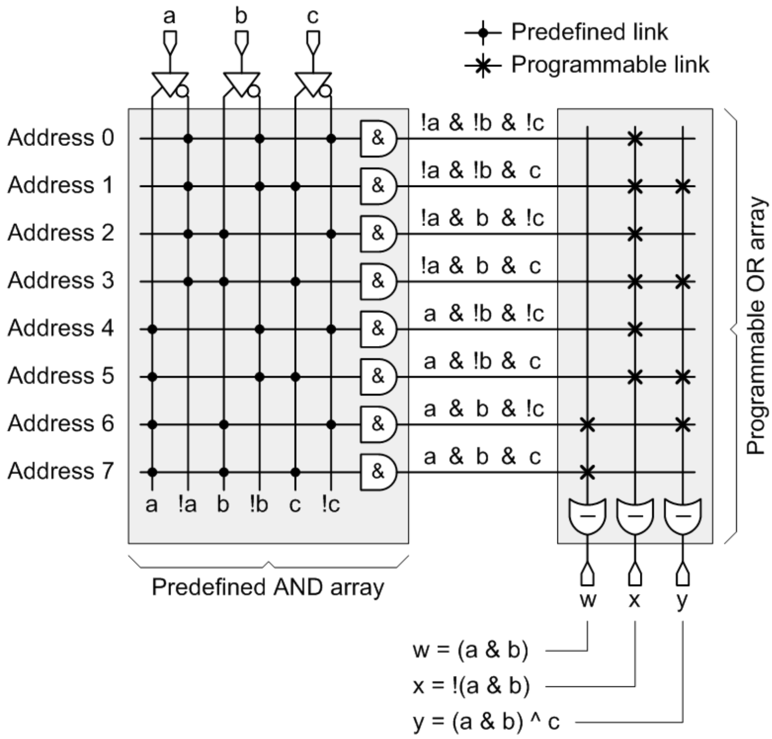


Abb. 17.3 PROM nach der Programmierung [18]. In der Definition der Variable y steht ∧ für XOR.

Natürlich könnte man sich mehr als drei Datenausgänge vorstellen. Dafür ist das dargestellte PROM aber zu klein.

Ein PLA ist ähnlich wie ein PROM aufgebaut (Abb. 17.4 und 17.5). Hier können aber sowohl die UND-Matrix wie die ODER-Matrix programmiert werden. PLAs sind flexibel, aber langsamer als PROMs und daher weniger gebräuchlich als PALs.

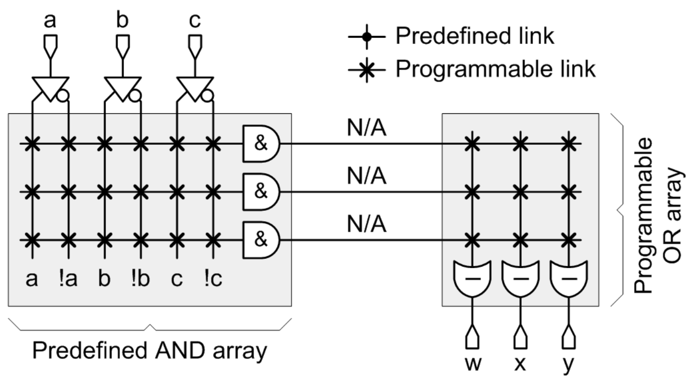


Abb. 17.4 PLA vor der Programmierung [18].

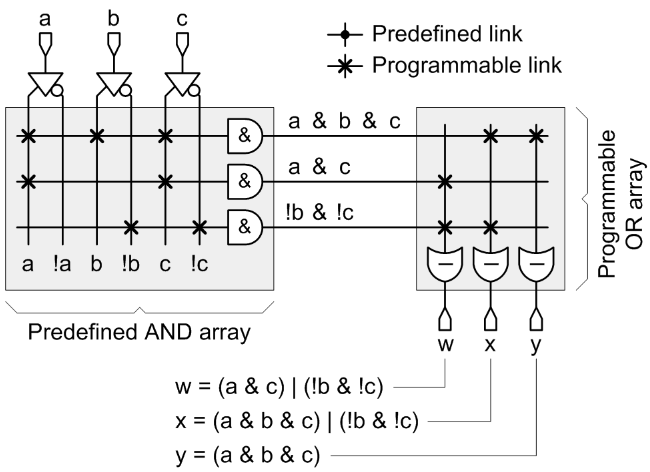


Abb. 17.5 PLA nach der Programmierung [18].

Bei einem PAL (Abb. 17.6) ist im Gegensatz zum PROM nicht die ODER-, sondern die UND-Matrix programmierbar.

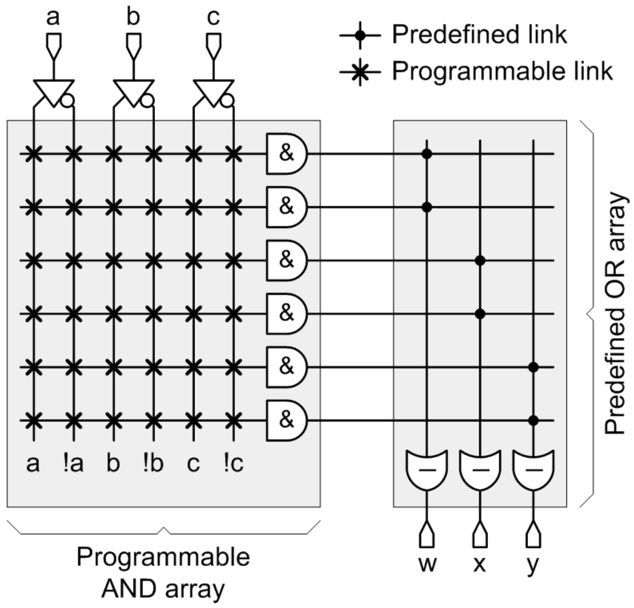


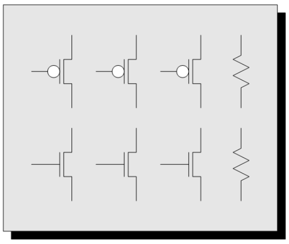
Abb. 17.6 PAL vor der Programmierung [18].

Zwei beliebte PALs sind das 16R8 und das 22V10 von AMD. Hier gibt es 8 Eingangspins und weitere 8 Pins, die entweder als Eingang oder Ausgang konfiguriert werden können.

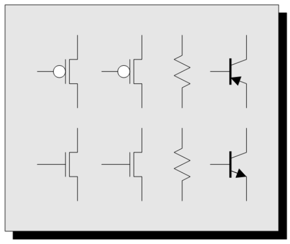
Oft gibt es eine Reihe weiterer Optionen, die festlegen, ob das PAL mit oder ohne Register betrieben werden soll, welche Polarität die Ein- und Ausgänge haben sollen, etc.

17.3 Gate Arrays

Ein „Gate Array“ ist eine Ansammlung von standardisierten Logikelementen und Transistoren auf einem Wafer, die noch nicht durch die Metalllagen verbunden sind. Die Metalllagen werden je nach Nutzer unterschiedlich auf die vorgefertigten Wafer aufgebracht. Dadurch reduzieren sich für den Nutzer die Kosten im Vergleich zu der Entwicklung eines ASIC[1] erheblich. Da der Nutzer vermutlich nur einen Bruchteil der bereitgestellten Bauteile einsetzt, ist die effektive Integrationsdichte eines Gate Arrays eher gering.



(a)

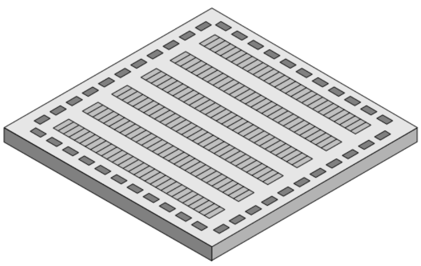


(b)

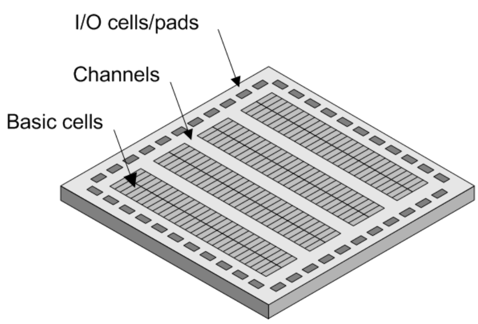
Abb. 17.7 Grundstrukturen eines Gate Array [18]: (a) Gate Array CMOS basic cell. (b) Gate Array BiCMOS basic cell.

Der Gate Array in Abbildung 17.7(a) stellt verschiedene MOSFETs und Widerstände zur Verfügung, der Gate Array in Abbildung 17.7(b) zusätzlich noch npn- und pnp-Bipolartransistoren.

Die in Abbildung 17.7 vorgestellten Komponenten werden in Blöcken bzw. Säulen auf einem Gate Array plaziert. So können sich z. B. Arrays mit einfachen Säulen oder Doppelsäulen, wie in Abbildung 17.8, ergeben.



(a)



(b)

Abb. 17.8 Platzierung der Elementarzellen auf dem Gate Array [18]: (a) Single-column Array. (b) Dual-column Array.

[1] Akronym für *Application-specific Integrated Circuit*

17.4 CPLDs

CPLDs sind aus mehreren SPLDs, meistens PALs, zusammengesetzt. Die einzelnen PALs sind durch eine Vielzahl von programmierbaren Verbindungen verbunden. In einem CPLD werden also die Verbindungen innerhalb eines PALs und zwischen den PALs programmiert. Nach dem Programmieren bleibt die Verbindung bis zum Abschalten der Versorgungsspannung oder einem Reset erhalten. Als Programmierungstechnologie verwendet man EPROMs, EEPROMs oder SRAMs. Es war historisch wichtig zu verstehen, dass es von Vorteil ist, nicht jede mögliche Verbindung zur Verfügung zu stellen. Die Struktur eines CPLDs ist in Abbildung 17.9 dargestellt. Man erkennt die globalen breiten Verbindungsstraßen und die lokalen schmaleren Verbindungen.

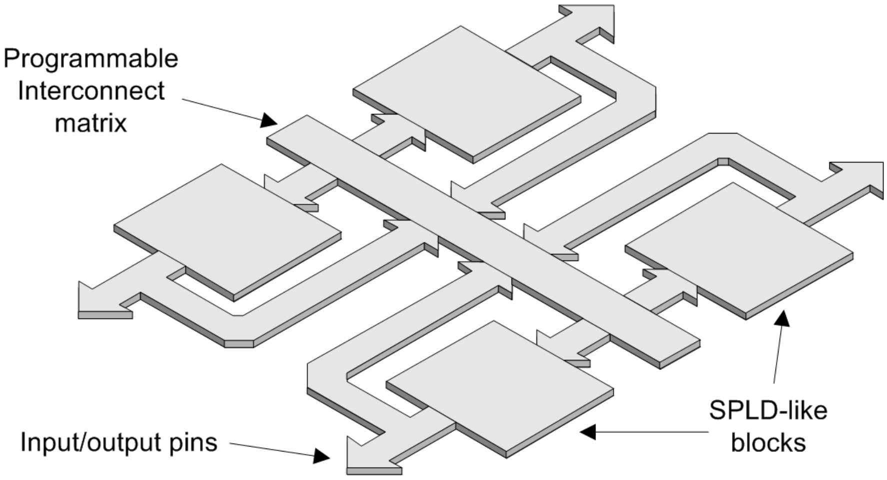


Abb. 17.9 Struktur eines CPLDs [18].

Beispiele für CPLDs sind z. B. die MAX-Familie von Altera [19, 20, 21] oder die Mach-Familie [22]. Die Durchlaufzeit eines CPLDs beträgt rund 5 ns bis 25 ns.

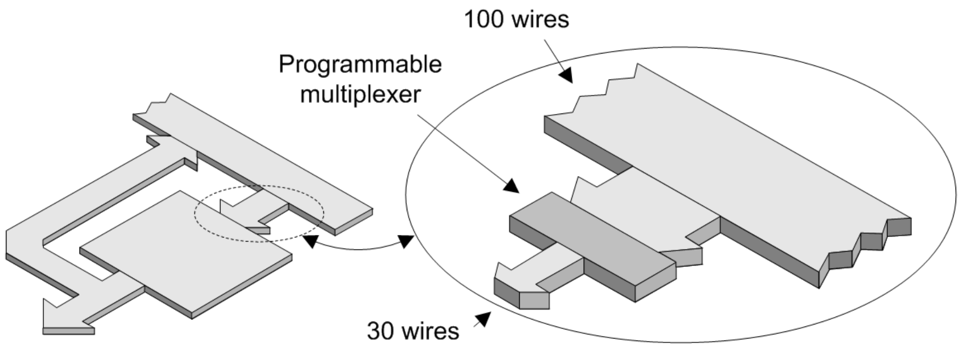


Abb. 17.10 Detailansicht eines CPLD [18].

17.5 FPGA

Ein FPGA ist ein „Field Programmable Gate Array“.

Die Bezeichnung „Field Programmable" soll ausdrucken, dass der Nutzer und nicht die Foundry die Programmierung durchführt.

Der Name ist allerdings nicht sehr passend, denn ein FPGA ist kein Gate Array, sondern ein sehr komplexer Baustein mit einer Vielzahl von Logikblöcken und Verbindungsnetzen, die beide vom Nutzer programmiert und konfiguriert werden können. Diese grundlegende Struktur ist in Abbildung 17.11 dargestellt.

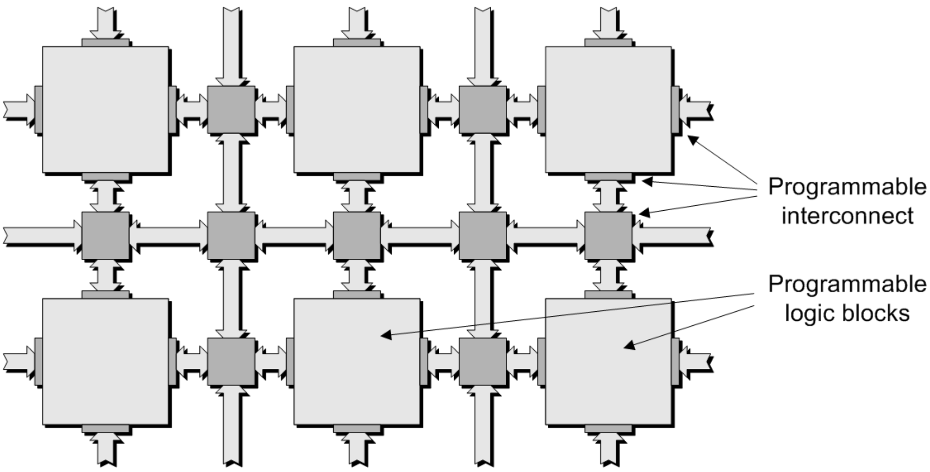


Abb. 17.11 Grundstruktur eines FPGA [18].

Im Kern besteht ein FPGA aus *logisch programmierbaren Inseln in einem Meer von programmierbaren Verbindungen.* Wie diese Analogie bereits impliziert, nehmen die Inseln sehr viel weniger Chipfläche ein als die Verbindungen.

FPGAs sind äußerst leistungsfähige Bausteine und wichtig für viele Lebensbereiche. Der Weltmarkt für FPGAs hat ein Volumen von rund 4 Milliarden $ (in 2010). Die zwei großen Firmen Xilinx und Altera dominieren mit einem Marktanteil von über 80%. Oft bemerkt man den FPGA als Teil eines Embedded System von außen nicht.

Bevor wir die Bauelemente eines FPGAs genauer betrachten, wollen wir uns das Datenblatt eines Virtex-6 der Firma Xilinx anschauen. Der Virtex-6 ist eine FPGA-Familie mit mehreren verwandten Bausteinen.

17.6 Charakteristika des Virtex-6

Abbildung 17.12 zeigt ein Entwicklungsboard mit einem Virtex-6. Diese Boards bieten viele Angenehmlichkeiten wie Stecker, Speicherbausteine, Schnittstellen etc. Sie werden zu Beginn einer Firmware-Entwicklung eingesetzt. Erst später wird der FPGA dann auf einer selbst entwickelten, maßgeschneiderten Platine eingesetzt.

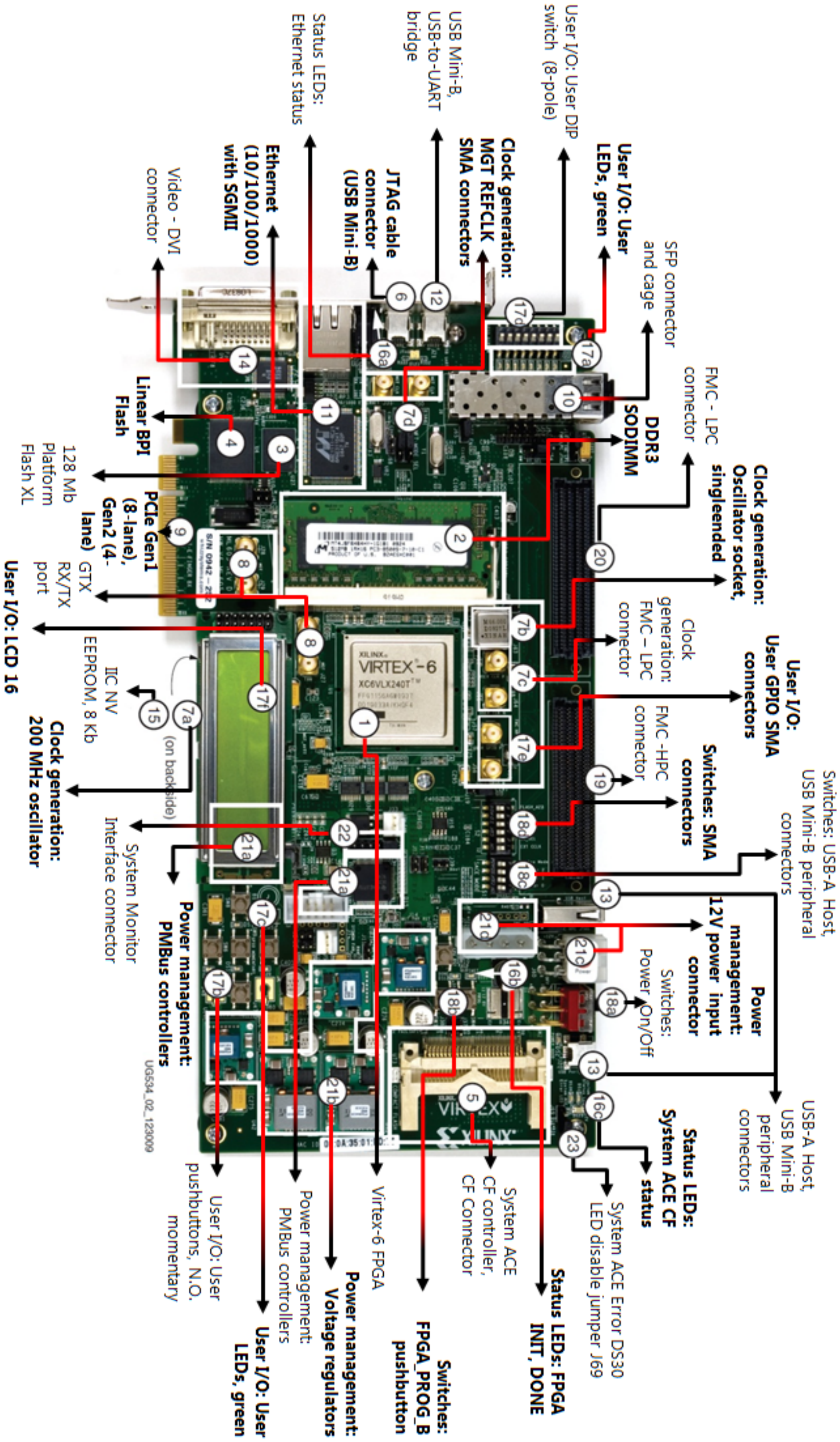


Abb. 17.12 Entwicklungsboard ML605 mit einem Virtex-6 [23].

Die Fläche eines Virtex-6 beträgt 23 mm × 23 mm bis 42,5 mm × 4, 25 mm. Dies entspricht einigen Milliarden Transistoren. Die exakte Zahl ist ein Herstellergeheimnis. Der Virtex-6 ist in einem 40 nm CMOS-Prozess der Firma UMC hergestellt. Die nächste Generation, der Virtex-7 oder der Stratix-5 von Altera sind in 28 nm CMOS-Technologie gebaut. Weitere Generationen könnten in 14 nm gebaut werden.

Betriebsspannung: Nominell 1,0 V. Im Energieeinsparmodus 0,9 V. Für die I/O-Spannungsversorgung 1,2 - 2,5 V.

Taktfrequenz: 600 MHz bis 1,6 GHz. Für Digital Signal Processing (DSP) 600 MHz. Diese hohen Taktfrequenzen werden nicht für jedes Design erreicht.

Zahl der Pins: Es gibt 240 bis 1200 I/O Pins. Die Pins sind in I/O-Bänke mit je 40 Pins pro Bank zusammengefasst. Die I/O-Elemente sind sehr flexibel gestaltet. So werden verschiedene Logikstandards unterstützt, und es ist eine optionale Terminierung mit 100 Ω vorgesehen. Außerdem können alle Pins als Eingang oder Ausgang konfiguriert werden. Darüber hinaus gibt es zahlreiche andere Pins für Power, high-speed serial links, Konfiguration, etc.

Preis: Das Entwicklungsboard ML605 mit einem *XC6VLX240T-1FFG1156* kostet ca. 1700 €, ein *XC6VLX240T-1FFG1156* rund 1400 €, weniger bei zunehmender Stückzahl. *1FFG1156* ist der Name des Verpackungsgehäuse.

Es gibt noch eine Vielzahl Pins für andere Aufgaben (Abb. 17.13).

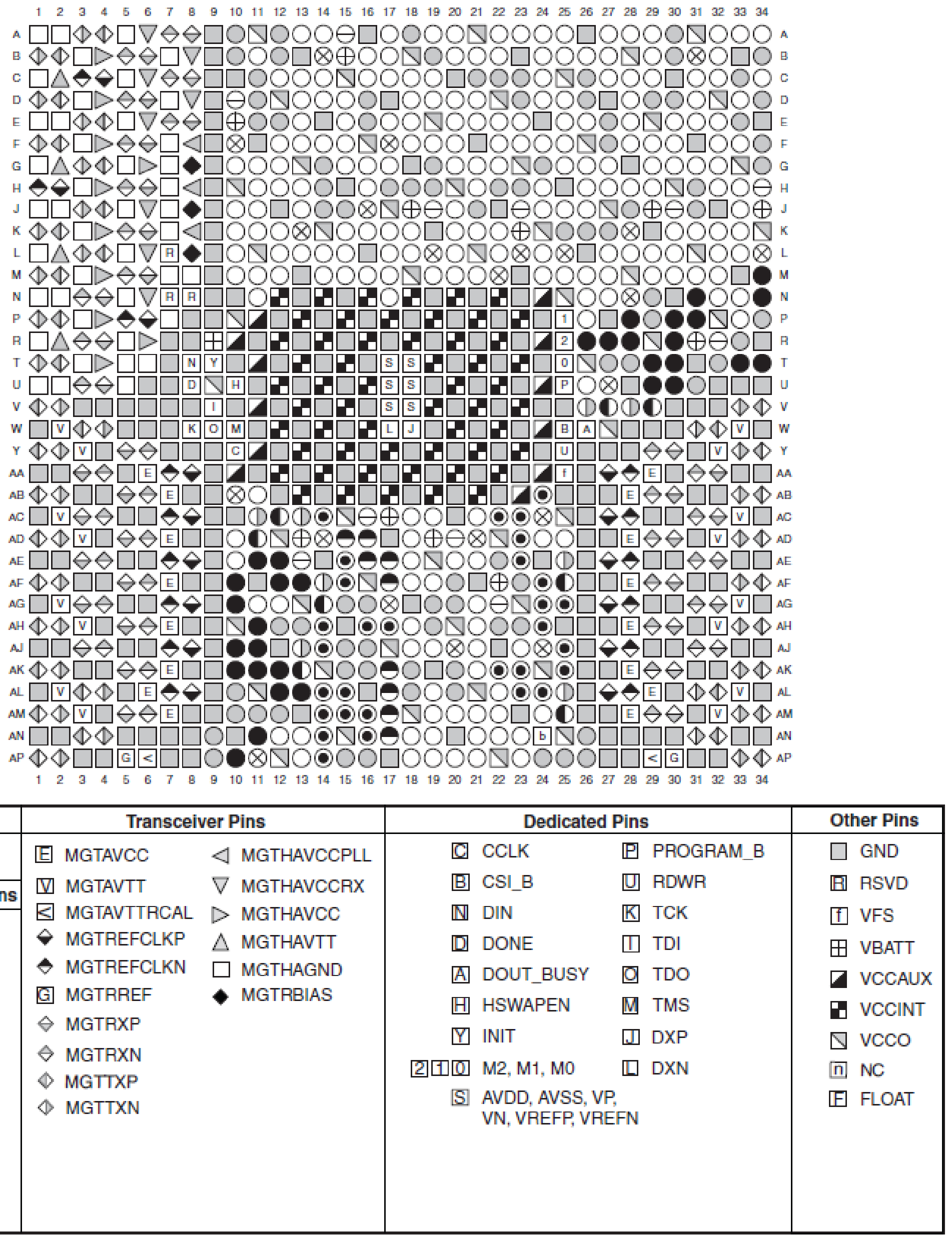


Abb. 17.13 FF1155 Package - **HX255T** and HX380T Pinout Diagram [24].

17.7 Woraus bestehen die Programmierinseln?

Die Inseln heißen bei Xilinx „Configurable Logic Blocks“ (CLB). Sie haben eine Unterstruktur, die aus den „Logic Cells“ besteht. [2]

Die Logic Cell besteht aus Look-up Tables (LUT), Multiplexern (MUX) und Flipflops.

[2]Bei Altera FPGAs werden ähnliche Unterstrukturen „Logic Elements“ genannt.

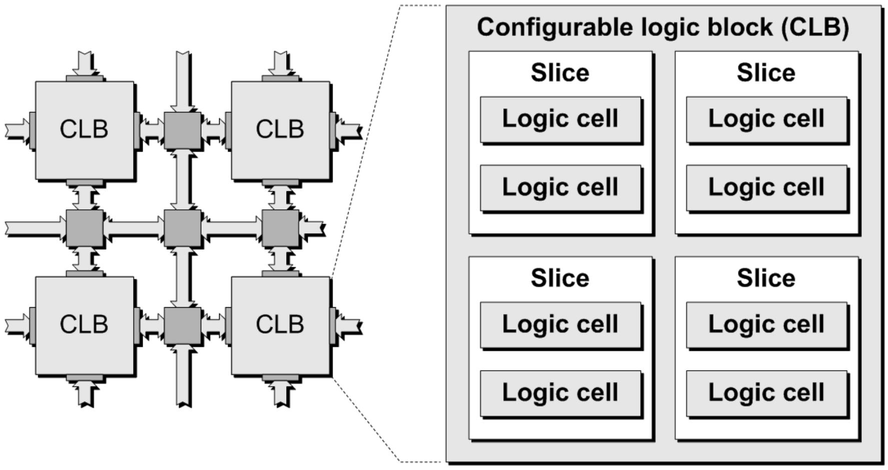


Abb. 17.14 Struktur eines Configurable Logic blocks [18].

Die Unterstruktur der in Abbildung 17.14 dargestellten Programmierinsel variiert von Hersteller zu Hersteller und eventuell auch von einer FPGA-Generation eines Herstellers zur nächsten.

In allen Logic Cells werden aber Multiplexer, LUTs und Flipflops eingesetzt.

Multiplexer

Ein Multiplexer (siehe 13.2) macht den Logikblock flexibler, wirkt als Schalter bzw. zur Realisierung von if-else Bedingungen. Durch Kombination der 2-nach-1-Multiplexer verschiedener Zellen kann man Multiplexer höherer Ordnungen aufbauen.

LUT

Jede logische Funktion kann als Wahrheitstabelle dargestellt werden. Ein Beispiel für eine Funktion mit drei Eingängen a, b und c ist in Abbildung 17.15 dargestellt. Es gibt 2^n unterschiedliche Werte der Eingangsvariablen und jeder Kombination wird eine Ausgangsvariable zugeordnet. Entsprechend den 2^n Ausgangswerten gibt es 2^n unterschiedliche Funktionen mit n Eingangsvariablen.

In LUTs werden diese Funktionen nicht durch ihre Logikgatter dargestellt, sondern durch ein 2^n-SRAM und einen 2^n-nach-1-Multiplexer mit n Adresseingängen.

Eine gute Eigenschaft der LUTs ist, dass alle Funktionen in der gleichen Zeit ausgewertet werden können. Das ist bei der Implementierung mit Logikgattern nicht unbedingt der Fall. Die Werte des SRAM werden während der Konfiguration des FPGA geladen. Im Gegensatz zu den SPLDs werden also keine Werte oder Verbindungen permanent eingebaut, sondern können beliebig oft geändert werden.

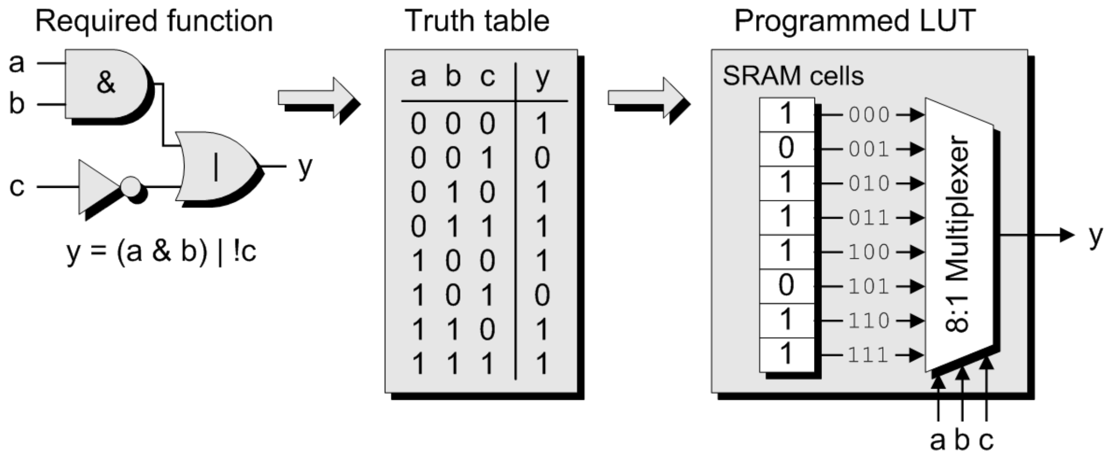


Abb. 17.15 Äquivalenz von Logikgattern und LUTs [18].

Ein CLB des Virtex-6 besteht aus zwei Slices. Jedes Slice besteht aus vier Logic Cells. Jede Logic Cell beinhaltet vier LUTs, acht Speicherelemente und acht breite Multiplexer. Es gibt Carryeingänge und -ausgänge zum Aufbau von Logik mit mehreren Logic Cells. Dies ist in Abb. 17.16 bis 17.17 vereinfacht illustiert.

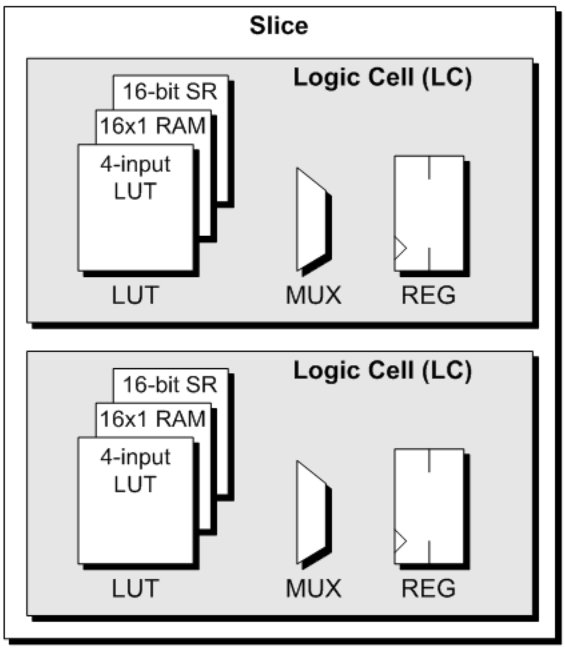


Abb. 17.16 Illustration einer Slice aus mehreren Logic Cells (hier nur zwei!) [18].

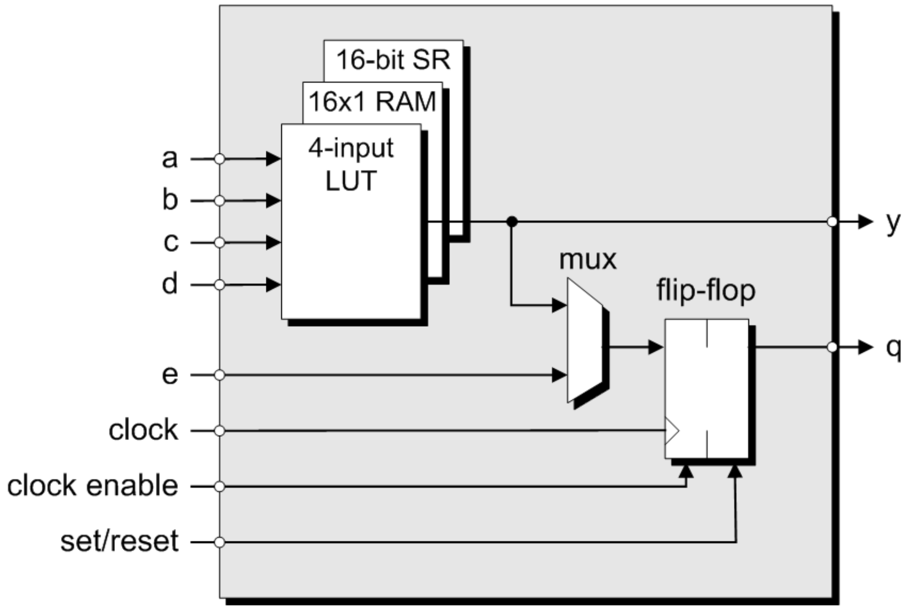


Abb. 17.17 Illustration einer Verknüpfung der Elemente einer Logic Cell [18].

Tatsächlich ist die Struktur eines CLB etwas komplexer. In Abbildung 17.16 wird ein vollständiges Diagramm eines Slice gezeigt.

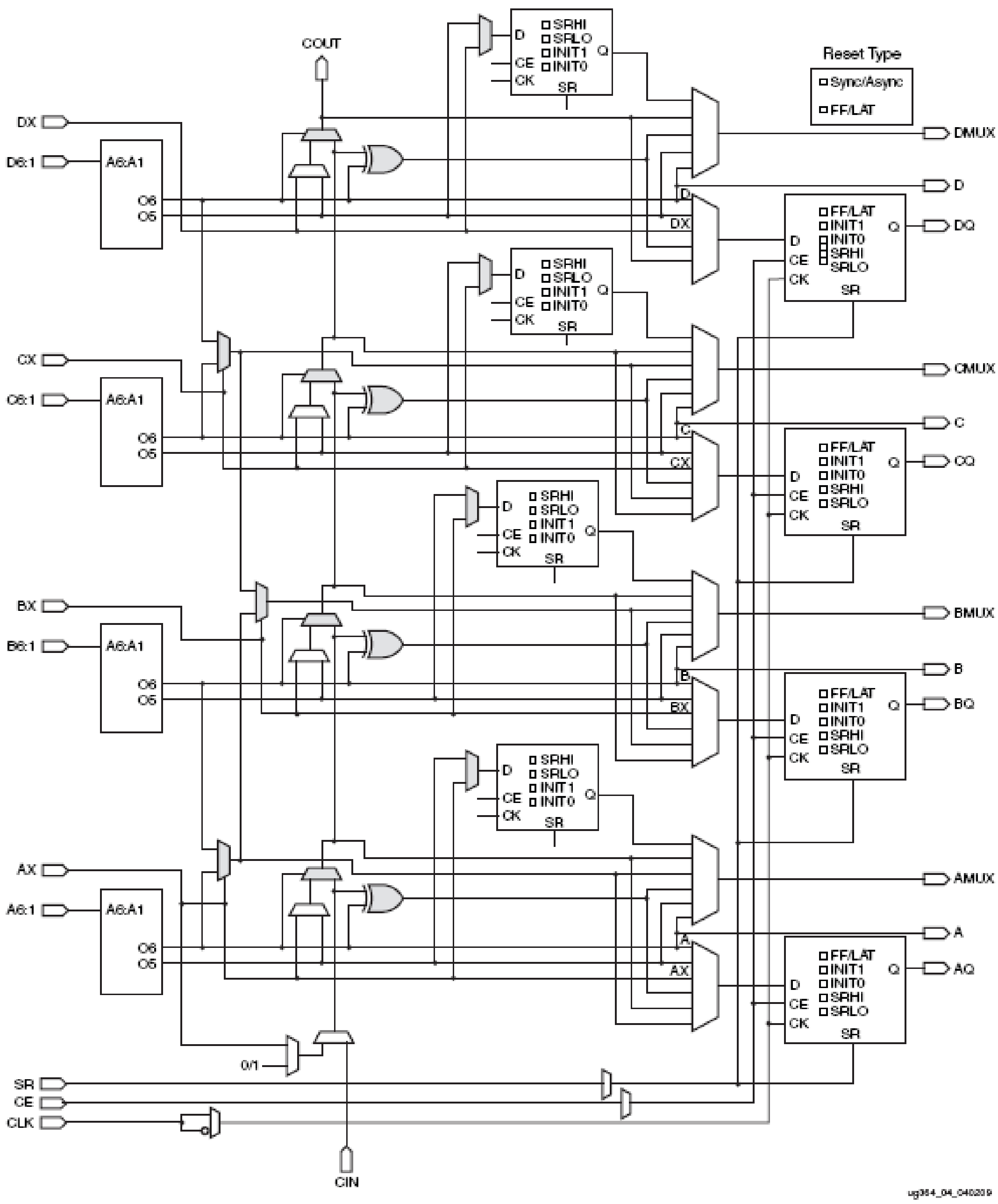


Abb. 17.18 Aufbau des sogenannten SLICEL [24].

Hier sieht man, dass es Verbindungen zwischen den LUTs gibt, erkennt die Carry-Struktur und einiges mehr. Die LUTs im obigen Schaltplan haben 6 Eingänge, aber auch LUTs mit 4 und 5 Eingängen sind häufig. Das ist technologie-, generations- und herstellerabhängig. Außerdem hängt die optimale Wahl davon ab, welche Kenngröße optimiert werden soll, z. B. die Chipfläche, der Energieverbrauch oder die Rechenleistung.

17.8 Woraus bestehen die Verbindungen?

Die Verbindungsmatrix (*sea of interconnects*) zwischen den Blöcken des FPGAs ist von größter Bedeutung. Wir betrachten hier nur von SRAMs gesteuerte Verbindungen, wie sie in den wichtigsten FPGAs von Altera und Xilinx verwendet werden. Es gibt aber auch moderne FPGAs, die nicht reversible Verbindungen (*anti-fuses*) verwenden, z. B. von Actel und auch Xilinx. Der Vorteil dieser Verbindungen ist, dass sie beim Abschalten der Stromversorgung erhalten bleiben, weniger Platz kosten und viel „sicherer“ sind.

Die Verbindungen und programmierbaren Schalter an den Kreuzungspunkten nehmen bei weitem den größten Raum in dem FPGA ein. Bis zu 90% der FPGA-Fläche ist damit belegt. Die von den Logikblöcken besetzte Fläche liegt damit in der Größenordnung von nur 10%. In den Verbindungsstrukturen entstehen die größten Signalverzögerungen.

Es gibt unterschiedliche Verbindungstypen und Verzweigungsschalter, die in Abbildung 17.19 dargestellt sind:

a) Programmable Routing Switch

b) Input Connection Block

Auch die Länge der Verbindungen variiert je nach dem, ob benachbarte oder weiter entfernte Logikblöcke verbunden werden müssen.

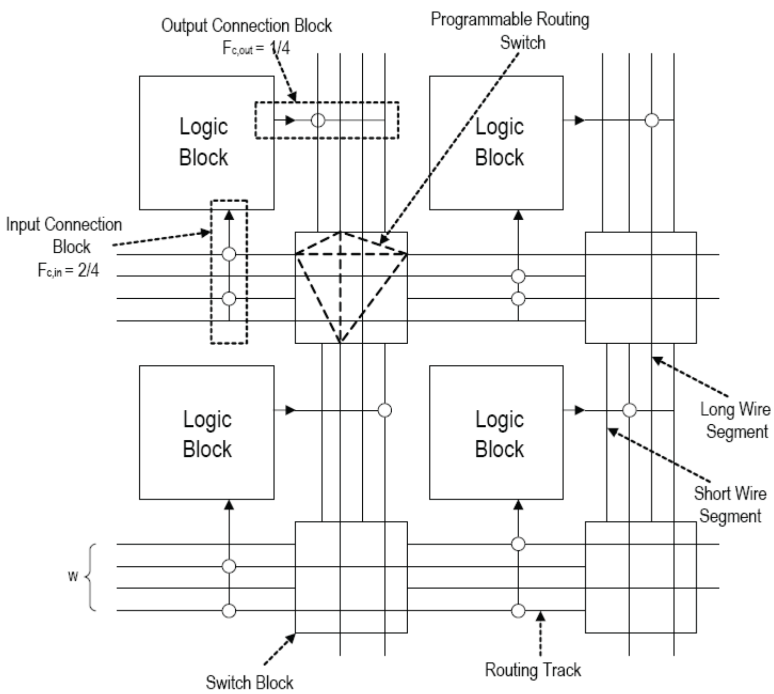


Abb. 17.19 Verbindungstypen eines FPGA [25].

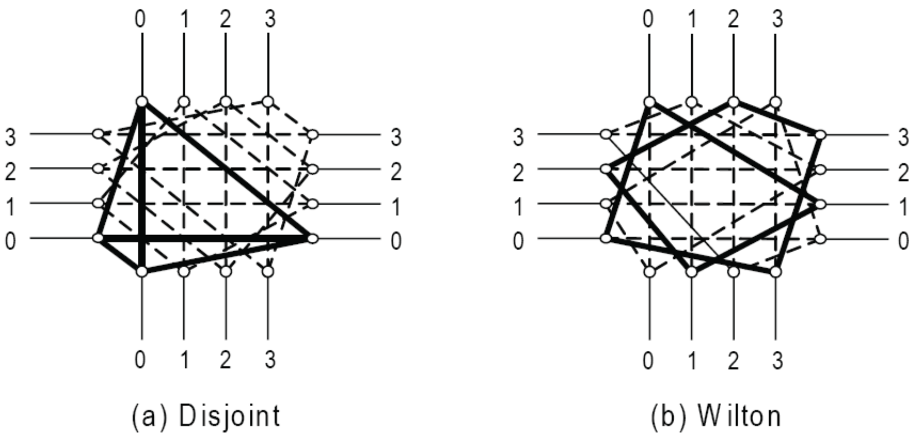


Abb. 17.20 Zwei Schaltblocktypen [25].

Abbildung 17.20 soll illustrieren, dass dieses Thema komplex ist. Im linken Bild gibt es die Möglichkeit, von der waagerechten Spur „n“ auf dieser Spur zu bleiben oder auf die senkrechte Spur „n“ umzusatteln. Hier sind beide Richtungen, nach oben oder unten möglich und entsprechen einem unterschiedlichen Schaltzustand.

Analoges gilt bei Betrachten eines Eingangssignals auf der senkrechten Spur.

Im rechten Bild sind gleich viele Schalter verwendet, aber man kann von der waagerechten Spur „n“ sowohl nach oben auf die Spur „n“ wie nach unten auf eine Spur „j“ mit „j“ $\neq$ „n“ wechseln. Man kann wie im linken Bild auch auf der waagerechten Spur bleiben. Beide Schalter haben die „Switch box connectivity“ 3. Es gibt auch Schaltarchitekturen mit noch mehr Verzweigungsmöglichkeiten.

Gleichzeitig ist es interessant, wie die Schalter realisiert werden. Auch hier gibt es viele unterschiedliche Architekturen und natürlich auch Technologien: einfache transmission gates, tri-state buffers, Multiplexer, uni- oder bi-direktionale Schalter, Schalter mit SRAM oder Flipflops, etc.

17.9 Clockverteilung

Die Taktverteilung auf einem großen FPGA ist äußerst komplex. Sogenannte „Clock Manager“ stellen unterschiedlich einstellbare Taktfrequenzen zur Verfügung, hier von 600 MHz bis 1,6 GHz. Gleichzeitig gibt es Jitterfilter und Phasenverschieber (z. B. um 30 ps).

Im Virtex-6 gibt es 32 sogenannte globale Taktleitungen, die die gesamte Fläche des FPGA abdecken. Es gibt auch Regionen, in denen eine externe Clock zugeschaltet werden kann.

Wichtig sind die „Phase Locked Loops“ (PLLs), mit denen man aus einem langsamen äußeren Takt einen phasengekoppelten schnellen internen Takt erzeugt. Der schnelle Oszillator ist ein VCO (voltage controlled oscillator), bei dem die Ausgangsfrequenz von der Steuerspannung abhängt.

Abb. 17.21 Steuerung des Taktes auf einem FPGA.

17.10 Boundary Scan

Der boundary scan wurde von der Joint Test Action Group (JTAG) entwickelt und wurde dann IEEE Standard 1149.1

Problem: wie testet man einen komplexen Baustein in einem möglicherweise fehlerhaften Board mit sehr vielen möglichen Zuständen?

Lösung: Boundary scan. Alle SRAMs für die Konfiguration bilden ein großes Schieberegister. Auch die Ausgänge des FPGA sind damit verbunden und können seriell ausgelesen werden.

Der JTAG port hat vier Eingänge:

- Data in
- Data out
- Clock
- Mode select (und einen optionalen Reset).

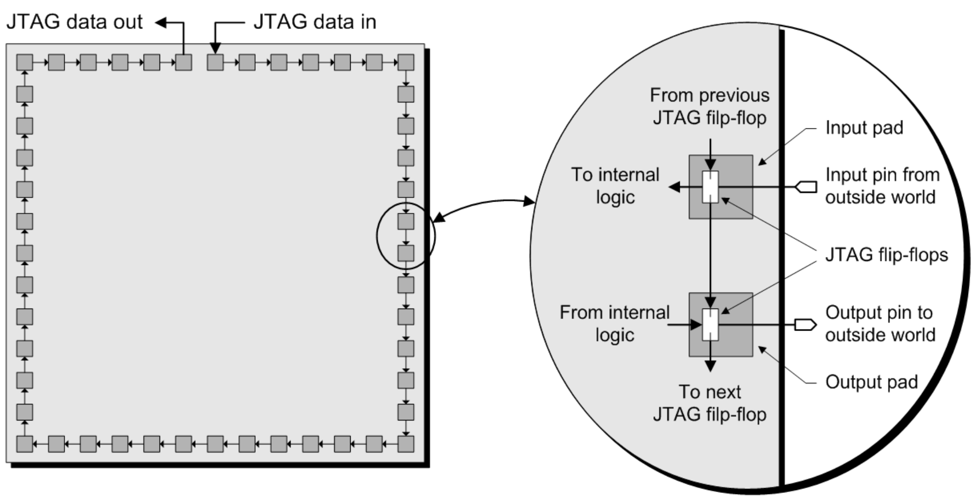


Abb. 17.22 JTAG data [18].

17.11 Datenleitungen

Um die Parallelität eines FPGA optimal zu nutzen, müßen für viele Anwendungen hohe Datenraten zur Verarbeitung in den FPGA übertragen werden. Je nach Aufgabe werden dann weniger oder aber sogar mehr Daten an den nächsten elektronischen Baustein weitergegeben. Deshalb werden eine Vielzahl von schnellen Datenübertragungsblöcken zur seriellen bi-direktionalen Datenübertragung, sogenannte Transceiver, bereitgestellt. Ein Transceiver kann, abhhängig von seiner Konfiguration, Daten senden oder Daten empfangen.

Die Transceiver sind an differentielle serielle Leitungen angeschlossen. Eine Vielzahl von Logikstandards wird unterstützt. Siehe Abschnitt über SerDes. Es gibt Transceiver, die auf geringe Leistung optimiert sind und Links für maximale Übertragungsraten.

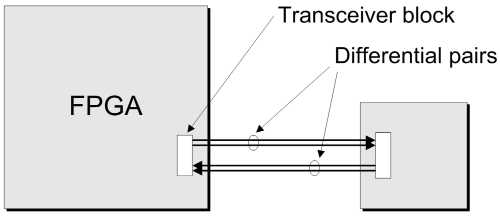


Abb. 17.23 FPGA Transceiver [18].

Der XCV6LX240T verfügt über 24 GTX-Transceiver mit insgesamt 24 × 480 Mb/s bis zu 24 × 6,6 Gb/s Bandbreite. Dadurch ist ein schneller und flexibler Datenübertrag möglich.

17.12 Zusätzliche Bauelemente eines FPGAs

Neben den vorgestellten Logikzellen und -blöcken werden in den meisten FPGAs zusätzliche optionale Bauelemente bereitgestellt.

Distributed RAM: Nur Xilinx. Manche Slices enthalten extra memory (SLICEM).

Block RAM: Es gibt dezidierte RAM-Blöcke, die über den FPGA verteilt sind. Allerdings ist ihre Größe mit 5 bis 32 Mb (Megabit, nicht Megabyte!) recht begrenzt. Hier braucht man in der Praxis immer einen externen großen Speicher.

Das BRAM hat zwei unabhängige Zugänge (Ports). Es ist getaktet.

Mikroprozessoren: Man unterscheidet zwischen „soft“ und „hard“ CPUs. „Hard“ sind fest im FPGA eingebaut. „Soft“ baut man aus den FPGA Logikzellen auf. Das wird vom Hersteller mit der entsprechenden Software unterstützt.

„Hard“ ist weniger häufig als früher. Dies gab es bei V-4 und V-5, aber nicht mehr beim V-6. Stattdessen gibt es einen „soft core“ MicroBlaze Prozessor. Der MicroBlaze hat eine 32-bit RISC architecture und wurde von Xilinx für die Implementierung in dem FPGA optimiert. Dasselbe gilt für den Nios II von Altera.

DSP: Dedizierte Prozessoren für die Digitale Signalverarbeitung (DSP). Ein einfaches Beispiel ist der Block: Multiply, Add, Accumulate (MAC) in Abbildung 17.24.

(Virtex-6: 25 × 18 Zweierkomplement Multiplizierer mit 48-bit accumulator.)

ADC: 10-bit 200 KSPS [3] als Monitor für Temperatur und Spannungen. Auch für externe Anschlüsse zugänglich.

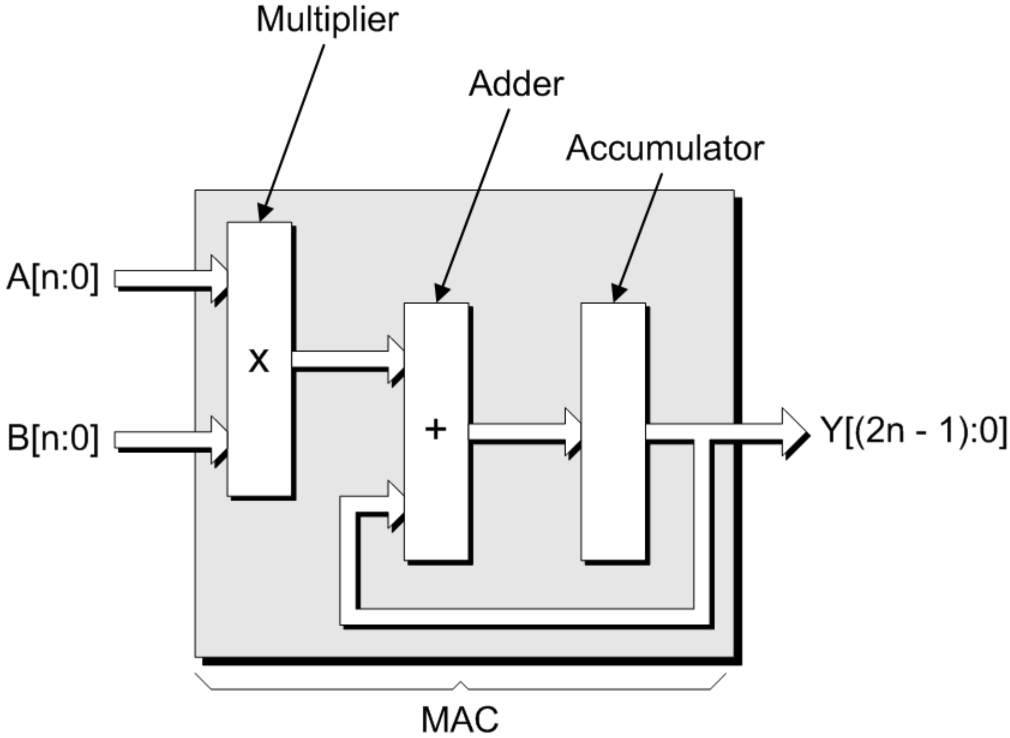


Abb. 17.24 Eine MAC-Einheit [18].

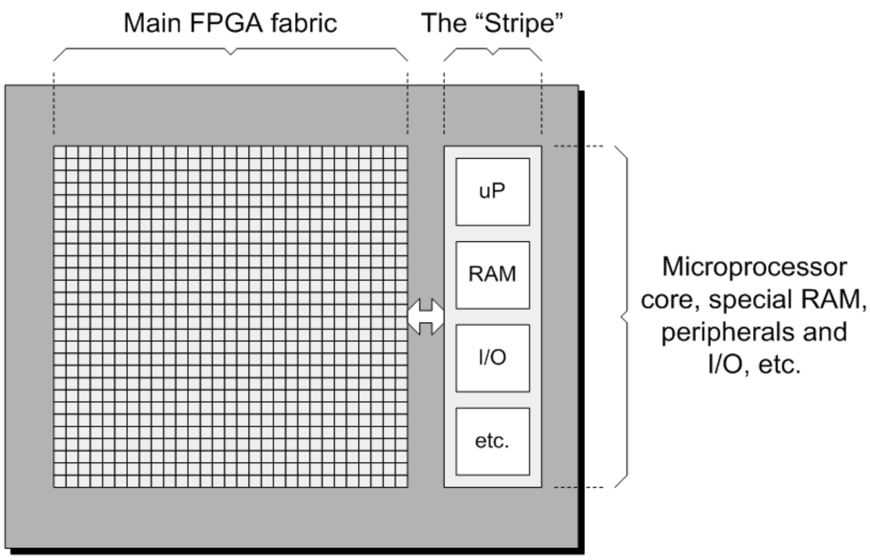


Abb. 17.25 Main FPGA fabric [18].

[3]kilosamples per second

17.13 Übersicht: Virtex-6

Die Elemente der FPGAs der Virtex-6-Familie sind in Abbildung 17.26 dargestelllt.

Device	Logic Cells	Configurable Logic Blocks (CLBs)		DSP48E1 Slices(2)	Block RAM Blocks			MMCMs(4)	Interface Blocks for PCI Express	Ethernet MACs(5)	Maximum Transceivers		Total I/O Banks(6)	Max User I/O(7)
		Slices(1)	Max Distributed RAM (Kb)		18 Kb(3)	36 Kb	Max (Kb)				GTX	GTH		
XC6VLX75T	74,496	11,640	1,045	288	312	156	5,616	6	1	4	12	0	9	360
XC6VLX130T	128,000	20,000	1,740	480	528	264	9,504	10	2	4	20	0	15	600
XC6VLX195T	199,680	31,200	3,040	640	688	344	12,384	10	2	4	20	0	15	600
XC6VLX240T	241,152	37,680	3,650	768	832	416	14,976	12	2	4	24	0	18	720
XC6VLX365T	364,032	56,880	4,130	576	832	416	14,976	12	2	4	24	0	18	720
XC6VLX550T	549,888	85,920	6,200	864	1,264	632	22,752	18	2	4	36	0	30	1200
XC6VLX760	758,784	118,560	8,280	864	1,440	720	25,920	18	0	0	0	0	30	1200
XC6VSX315T	314,880	49,200	5,090	1,344	1,408	704	25,344	12	2	4	24	0	18	720
XC6VSX475T	476,160	74,400	7,640	2,016	2,128	1,064	38,304	18	2	4	36	0	21	840
XC6VHX250T	251,904	39,360	3,040	576	1,008	504	18,144	12	4	4	48	0	8	320
XC6VHX255T	253,440	39,600	3,050	576	1,032	516	18,576	12	2	2	24	24	12	480
XC6VHX380T	382,464	59,760	4,570	864	1,536	768	27,648	18	4	4	48	24	18	720
XC6VHX565T	566,784	88,560	6,370	864	1,824	912	32,832	18	4	4	48	24	18	720

Abb. 17.26 Übersichtstabelle: Virtex-6 FPGA [26].

Die Zahl der Logic Cells ist sowohl bei Xilinx wie Altera im Namen enthalten.

Beispiele: Der XC6V (für Xilinx Virtex-6) LX (Typ LX) 75T enthält rund 75k Logic Cells. Der XC6VHV565T enthält 566k Logic Cells.

17.14 Firmwareentwicklung

Die Programmierung eines FPGAs, also die Erstellung der *Firmware* erfolgt in den folgenden Schritten:

1. Funktionsbeschreibung des FPGAs mit einer Hardware Description Language (HDL) z. B. VHDL oder Verilog. Mittels einer funktionalen Simulation (RTL-Simulation) wird abschliessend überprüft, ob das in VHDL ausgedrückte Modell der Schaltung sich wie gewünscht verhält. Hierbei wird der Datenfluss durch die kombinatorische Logik und die Speicherbausteine simuliert, ohne das Zeitverhalten abzubilden.

2. Danach Synthese, das heißt Interpretation des VHDL-Codes und Abbildung der Firmware auf Logikbausteine (LUTs, MUX, Register, etc.) und erstellen einer „Netlist“ auf Register-Transfer-Level (RTL).

3. Implementierung, d.h. Abbildung der abstrakten Logikbausteine auf den spezifischen FPGA mit den Einzelschritten: „Translate“, „Map“, „Place and Route". Jetzt ist genau festgelegt, welche Bausteine und Verbindungen auf dem FPGA benutzt werden. Damit ist auch der Platzverbrauch des Designs auf dem FPGA bekannt. Danach Simulation des Zeitverhaltens. Iteration von 1. bis 3.

4. Das Ergebnis ist eine Bitfolge (bit stream), die man in den FPGA laden kann. Das Konfigurieren dauert nur Sekunden. Die Konfigurationsregister eines V-6 sind − abhängig

vom Untertyp – 16 bis 160 Mb groß. Jetzt kann man das Verhalten der implementierten Hartware mit den obigen Simulationen vergleichen. Eventuell sind weitere Iterationen nötig.

Sehr interessant ist auch die partielle Rekonfiguration des FPGA, bzw. die aktive Rekonfiguration während der Rechenzeit.

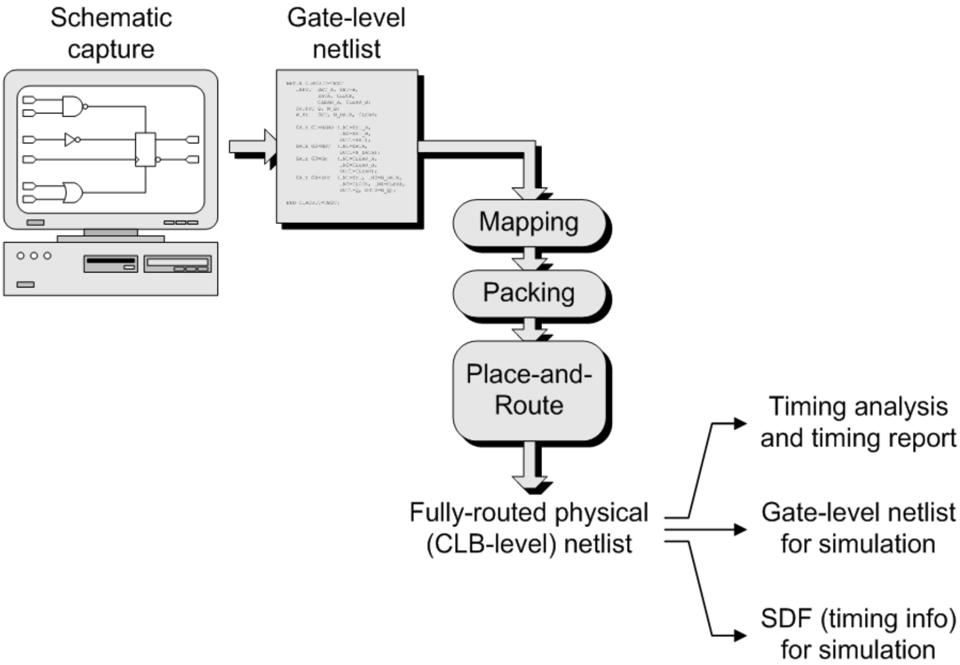


Abb. 17.27 Schematischer Ablauf einer Firmwareentwicklung [18].

17.15 Warum FPGAs?

Das Problem der ASICs sind die hohen Entwicklungs- und Maskenkosten, bei einer möglicherweise begrenzten Nutzerzahl.

Natürlich gibt es commodity ASICs, z. B. CPUs und GPUs, in hohen Stückzahlen, die aber fest verdrahtet und nicht billig sind.

Was macht den FPGA so leistungsstark?

- Er ist zwar nicht so hoch getaktet wie eine CPU arbeitet dafür aber parallel!
- Dadurch ist er im Vergleich zu CPUs und GPUs energieeffizient.
- Im Vergleich zum ASIC sind FPGAs schon bei kleinen Stückzahlen billig.
- Die Entwicklungszeit ist viel kürzer als beim ASIC-Design (nur Firmware).

Für Detektorbauer in der Teilchenphysik, Schwerionenphysik oder an Neutronenquellen, Synchrotonstrahlungsquellen, etc. haben FPGAs verschiedene Nachteile:

- Im Wesentlichen gibt es keine Analogfunktion.
- Die Abwärme bzw. der Stromverbrauch ist im Vergleich zum maßgeschneiderten integrierten Schaltkreisen zu hoch.
- FPGAs sind nicht notwendigerweise sehr strahlenhart.

Dadurch werden FPGAs in der Regel nicht in der unmittelbaren Detektornähe eingesetzt (Frontend), sondern erst nach Übertragung der Daten in eine bequeme Umgebung zur Weiterverarbeitung (Backend).

	ASIC	**FPGA**	**GPU**	**CPU**
Rechenleistung	„optimal“	hoch	hoch	mittel
Speicher (on-chip)	„optimal“	mittel	mittel	gering
Bandbreite/Datenrate (off-chip)	“optimal“	hoch	hoch	mittel
Taktrate	“optimal“	gering	hoch	hoch
Leistungsaufnahme	“optimal“	gering	sehr hoch	hoch
Energieeffizienz	“optimal“	gut	schlecht	sehr schlecht
Entwicklungsaufwand	sehr hoch	hoch	mittel	gering
Kosten	sehr hoch	mittel	gering	gering

Tab. 17.5 Qualitativer Vergleich der Eigenschaften von ASICs, FPGAs, GPUs und Multicore bzw. Mancore CPUs.

ASICs sind per definitionem immer auf die jeweilige Anwendung optimiert. Doch sie sind, bis auf sehr hohe Stückzahlen, für den Nutzer mit höchstem Aufwand bzw. höchsten Kosten verbunden. Falls der Nutzer das Design bezahlen muss, kosten sowohl das Design wie die Produktion der Chipmasken leicht mehere Millionen EUR.

	FPGA	**GPU**	**CPU**
Beispiel	Xilinx Virtex-6 LX240T	Nvidia GTX 580	Intel Core i7 980 Extreme
Halbleiterprozess [nm]	40	40	32
Anzahl Transistoren [10^9]	3	3	1,17
Chipfläche [mm^2]	250 - 400	520	248
Taktrate [GHz]	0,1 - 0,4	1,5	3,2
Anzahl Rechenkerne	$\sim$ 150k [a]	512	6 [b]
Peak Rechenleistung [GFlops]	$\sim$ 200	1581	15,8
Speicher (on-chip) [MB]	1,8	1	12,75
Bandbreite/Datenrate (off-chip) [GB/s]	20	192	25,6
Peak Leistungsaufnahme [W]	< 5 [c]	244	130
Energieeffizienz [GFlop/J]	$\sim$ 40	6,5	0,12

Tab. 17.6 Technologische Daten ausgewählter FPGAs, GPUs und Multicore CPUs.

[1] Logikgatter

[2] je 4-fach SIMD und Hyperthreading-fähig

[3] applikationsspezifisch

18 Aufbau- und Verbindungstechnik*

Aufgabe der Aufbau- und Verbindungstechnik (AVT) ist, ausgehend von einem Schaltungsentwurf und mikroelektronischen Komponenten, ein funktionierendes elektronisches Modul oder Gerät herzustellen. Dabei gilt es zahlreiche Randbedingungen zu berücksichtigen und eventuell Kompromisse zu schließen. Wichtig sind:

- die maximal zulässige Größe des Moduls und eventuell auch der Zuleitungen.
- die erzeugte Abwärme.
- Umgebungsparameter wie Außentemperatur, Feuchtigkeit, Umgebungsmedium (Vakuum, Luft, Stickstoff, Wasser, usw.), Erschütterungen beziehungsweise maximale Beschleunigungen.
- Elektromagnetische Verträglichkeit und Abschirmung.
- Wartungsfreundlichkeit und die Mindestlebensdauer des Moduls.
- Stückzahlen, Lieferraten, automatische Produktions- und Qualitätsicherungsverfahren.
- Preis einer Einheit.

Für spezielle Anwendungen in physikalischen Experimenten z.B. der Teilchenphysik treten weitere Besonderheiten auf, wie das Funktionieren eines Moduls in hohen Magnetfeldern und in externen Strahlungsfeldern (meistens γ, n, p, e, π-Strahlen und Ionen). Außerdem sollen die Module möglichst „leicht“ sein, das heißt ihre Strahlungslänge groß, um die Vielfachstreuung zu minimieren. Leichte Bauteile sind aber auch im Fahrzeugbau von Vorteil (geringerer Benzinverbrauch).

Ausgehend von dem Schaltkreis und den genannten Randbedingungen müssen vor allem die Substrattechnologie (z.B. Leiterplatten), die Bauformen der aktiven und passiven Komponenten (z.B. „Surface mount devices“) und die Verbindungstechnik (z.B. Drahtverbindungen) festgelegt werden. Es ist wichtig, schon beim Schaltungsentwurf eine klare Vorstellung davon zu haben, wie das Modul gebaut werden soll.

In den Bildern 18.1, 18.2 und 18.3 sind drei sehr unterschiedliche Elektronikmodule gezeigt. Wir wollen ausgehend von diesen Beispielen unterschiedliche Aufbau- und Verbindungstechnologien besprechen.

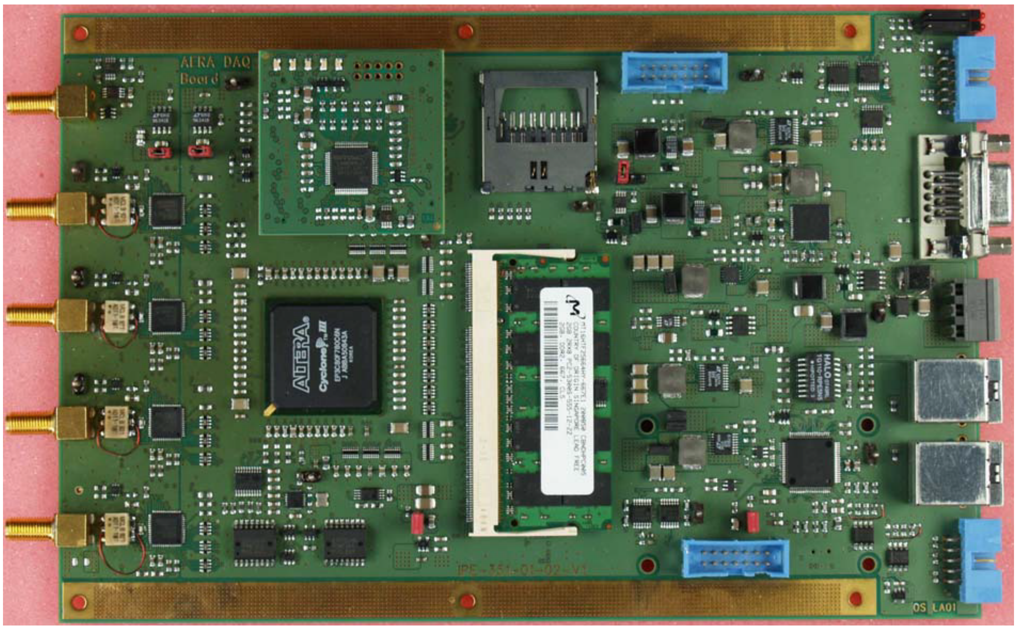


Abb. 18.1 Karte zur Auslese von Radiosignalen aus kosmischen Schauern.

Die Karte aus Abb. 18.1 ist im Einsatz im Pierre Auger Observatory in Argentinien. Sie besteht aus einer klassischen Leiterplatte mit einer Vielzahl elektronischer Bausteine wie Speicher, FPGAs, Vorverstärker und ADCs. Ein- und Ausgangssignale werden mit Steckern übertragen.

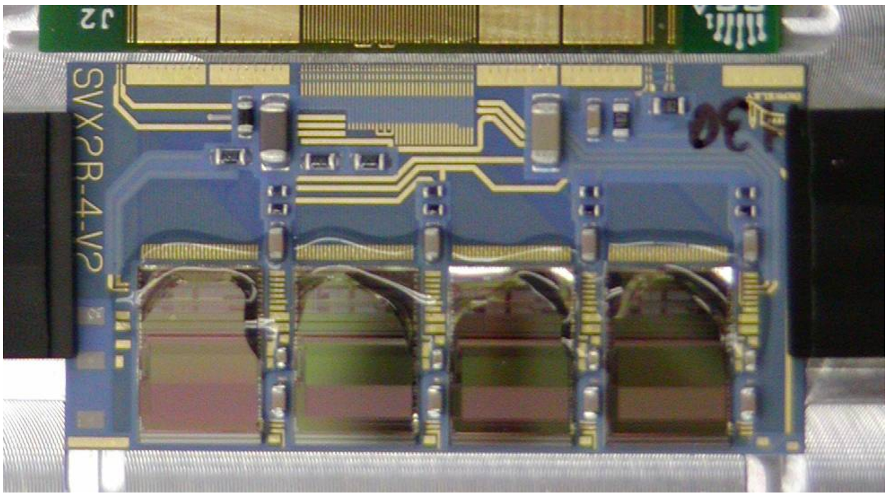


Abb. 18.2 Multichipmodul zur Auslese des CDF Run2b Siliziumdetektors[1].

Im Gegensatz dazu ist das Multichipmodul (MCM) in Abb. 18.2 auf einer BeO-Keramik aufgebaut. Es hat nur wenige Komponenten, insbesondere vier ungehäuste Chips und einige Widerstände und Kondensatoren in SMD-Bauweise. Es gibt keine Stecker sondern Drahtverbindungen.

[1]Leider wurde dieser Detektor nur entworfen, aber nie gebaut.

Das MCM in Abb. 18.3 ist auf einer flexiblen Leiterplatte aufgebaut und deckt auch die Rückseite des Siliziumsensors ab.

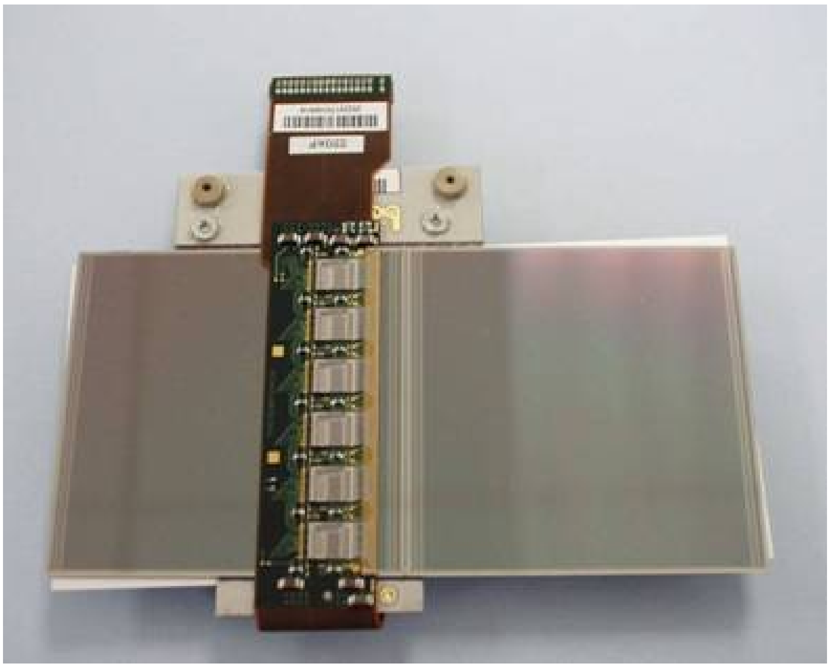


Abb. 18.3 Siliziumsensor und Multichipmodul des ATLAS Semiconductortrackers (SCT).

18.1 Aufbautechnologien

Die aktiven und passiven Komponenten sind auf einem rein passiven Substrat mechanisch befestigt und durch die Leiterbahnen elektrisch verbunden. Es gibt vier wesentliche Substrattypen.

- Feste Leiterplatten aus Glasfaserepoxidharzgemischen.
- Flexible Leiterbahnen aus Polyimidfolien (Kapton).
- Mehrschichtige Keramikträger in Dickschichttechnologie.
- Dünnschichtmodule auf Siliziumträgern.

Substrattyp	Zahl der Metall-lagen	Minimaler Abstand/Breite der Leiterbahnen	Größe	Kosten	Wärme-ablei-tung
Starre Leiterplatte	2-24	100 μm	$1\,\mathrm{cm}^2 - 1\,\mathrm{m}^2$	günstig	schlecht
Flexible Leiterplatte	2-8	75-100 μm	$1\,\mathrm{cm}^2 - 100\,\mathrm{cm}^2$	günstig	schlecht
Keramik / Dickschicht	2-6	75-100 μm	$1\,\mathrm{cm}^2 - 100\,\mathrm{cm}^2$	teuer	sehr gut
Silizium / Dünnschicht	2-4	20 μm	$1\,\mathrm{cm}^2 - 100\,\mathrm{cm}^2$	sehr teuer	gut

Tab. 18.1 Die wichtigsten Eckdaten von Substrattypen.

18.1.1 Leiterplatten

Die klassische Leiterplatte besteht aus einem mit Epoxidharz vergossenen Glasfasergewebe (G10 bzw. FR4) als Träger- und Isolatormaterial. Darauf werden im einfachsten Fall beidseitig dünne Kupferschichten aufgebracht und dann die gewünschten Leiterbahnen durch photolitographische Prozesse hergestellt[1].

Komplexere Leiterplatte besteht meistens aus mehreren Lagen (2/4/6/8-24), die durch elektrische Durchführungen (Vias) verbunden sind. Die Lagen werden aufeinander geklebt und gepresst. Oft ist wichtig möglichst große Funktionalität und Integrationsdichte zu realisieren.

Leiterplatten sind robust, billig und auch geeignet für große Boards mit vielen Leiterlagen. In der Teilchenphysik werden oft viele identische Boards benutzt, um Trigger- und Datenauslesesysteme mit vielen Millionen Kanälen gleichzeitig und dadurch schnell auszulesen.

18.1.2 Flexible Leiterplatten

Flexible Leiterplatten werden entweder als Kabel oder für Spezialanwendungen wie zum Beispiel Multichipmodule (MCMs) verwendet, oft als billige Alternative zu Keramiksubstraten. Der Träger ist eine Polyimidifolie, die sehr robust ist und niedrige wie hohe Temperaturen verkraftet. Meistens wird Kupfer als Leiter verwendet, in seltenen Fällen auch Aluminium. Die Zahl der Leiterbahnen ist in der Praxis auf zwei (Kabel) bis sechs (MCM) begrenzt. Die Strahlenlänge ist groß, und entsprechend gering ist die Vielfachstreuung (und das Gewicht) beim Einsatz in Teilchendetektoren.

[1] Arbeitsschritte: 1. Aufbringen des Photolacks. 2. Belichtung des Photolacks durch eine Maske, die dem Negativ der gewünschten Leiterbahnen entspricht. 3. Enwickeln und Entfernen des entwickelten Photolacks. 4. Wegätzen des nicht duch Photolack geschützten Kupfers.

Die Foliendicke beträgt typischerweise 25 μm. Als Kupferdicke wählt man meistens 18 μm[2] oder 36 μm. Für Aufbauten mit mehreren Lagen werden Kleberschichten zwischen die Polyimid/Kupferfolien eingefügt und das Paket unter Hitze gepresst. Typische Kleberdicken sind 25 μm. Es werden Acryl- und Epoxidkleber benutzt. Die Wärmeableitung durch die vielen Kleber- und Polyimidschichten ist mäßig. Hier sind Keramikträger deutlich überlegen.

Man kann zuverlässig Leiterbahnbreiten- und abstände von 100 μm herstellen, auch über größere Flächen. Auch 75 μm ist noch sicher und bezahlbar. Für kleine Flächen und anspruchsvolle hoche Integrationsdichten wurden auch schon 50 μm schmale Leiterbahnen produziert. Kritisch für den Platzverbrauch und die Kosten sind die Durchführungen (Vias) zwischen unterschiedlichen Lagen. Ein Viadurchmesser von 150 μm und eine Vialichtung von 375 μm Durchmesser sind gut machbar (allerdings nicht für jeden Anbieter).

18.1.3 Keramikträger

Als Keramikträger benutzt man Aluminiumoxid[3] (Al_2O_3), Aluminiumnitrid[4] (AlN) und für Spezialanwendungen Berylliumoxid[5] (BeO). Diese Matrialen haben hervorragende elektrische Eigenschaften. Sie sind sehr gute Isolatoren und die Dielektrizitätskonstante ist hoch ($\epsilon_{\mathrm{Al_2O_3}}= 9,5$, $\epsilon_{\mathrm{AlN}} = 8,8$, $\epsilon_{\mathrm{BeO}} = 6,5$). Außerdem sind Keramiken gute Wärmeleiter, unempfindlich gegenüber Feuchtigkeit und allgemein robust.

Auf die Keramikträger werden im Siebdruckverfahren Isolierpasten und in einem seperaten Arbeitsschritt Leiterpasten aufgebracht und bei hohen Temperaturen von 500 bis 800 °C in Sinteröfen (siehe Abb. 18.4) eingebrannt. Die Leiterbahnen werden durch das Siebmuster (Stencil) definiert.

Nachteile der Keramikträger ist ihr Preis, manchmal die relativ große Dicke von typischen 500 bis 700 μm und die relativ begrenzte Zahl der Leiterlagen. Abhängig von dem Dickschichtverfahren können minimale Leiterbahndicken und Leiterbahnabstände von bis zu 80 μm erreicht werden.

[2] 18 μm entsprechen einem Gewicht von „$\frac{1}{2}$ounce per square foot“.
[3] Engl. alumina
[4] Engl. aluminium nitride
[5] Engl. beryllia

Abb. 18.4 Sinterofen für den Dickschichtdruck auf Keramiksubstraten.

18.1.4 Dünnschichttechnologie

Dünnschichttechnologien auf Siliziumträgern sind für sehr integrierte Aufbauten mit hoher Leiterbahndichte sinnvoll, werden aber eher selten eingesetzt. Dünnschichttechnologie wird auch beim Bau der Metalllagen in integrierten Schaltkreisen verwendet. Für Multichipmodule (MCMs) auf Siliziumträgern kann man Leiterbreiten und Leiterabstände von 10 bis 20 μm erreichen.

18.2 Verbindungstechnologien

Wie werden elektrische Bauteile miteinander verbunden? Wir unterscheiden:

- Löten
- Kleben
- Drahtverbindungen
- Dickdrahtbonden
- Flipchiptechnologie

In der Regel sind integrierte Schaltkreise oder andere aktive Komponenten so verpackt (gehäust), dass man sie leicht auf die Platine löten kann. Manchmal verwendet man aber auch unverpackte Chips (bare die). In der Teilchenphysik meistens um Platz zu sparen oder um die „Masse“ und damit die Vielfachstreuung geladener Teilchen zu minimieren. Die unverpackten Chips kann man nicht löten, sie haben keine „Beinchen“. Stattdessen werden sie auf die Platine oder Keramik aufgeklebt. Abhängig vom Design und der Komplexität des Chips mit elektrisch leitendem oder nichtleitendem Kleber. Häufig verwendet werden Epoxidkleber mit Zusatz von Silberkörnern.

18.2.1 Löten

Dies ist die wichtigste und billigste Mehode, die erst bei sehr kleinen Strukturen nicht mehr funktioniert. Fürs Basteln, Reparaturen und für Schaltungen mit wenigen Komponenten kann man das Löteisen zum Erwärmen und Aufbringen des Lötzinns benutzen. Es ist empfehlenswert unter einem Mikroskop zu arbeiten, wogegen manche Elektroniker und Physiker allerdings eine ausgeprägte Abneigung haben. Alternativ kann man mit Lötpaste und einer „Herdplatte“ arbeiten.

Für komplexe Leiterplatten mit vielen Bausteinen oder in der Massenproduktion benutzt man Maschinen. Dazu wird die Lötpaste mithilfe einer dünnen Metallschablone („stencil“) auf die Lötpads der Platine gebracht. Dann werden die passiven Bauteile von eine automatischen Bestückungsmaschine („pick-and-place machine“) auf ihre Sollposition gesetzt. Dies funktioniert nur für geeignete planare Bauformen, die sogenannten „surface mount devices“ (SMD). Eine gute Bestückungsmaschine setzt knapp ein Bauteil pro Sekunde. Die Positioniergenauigkeit der Komponenten ist selbst bei großen Platinenflächen deutlich besser als 0,1 mm. Abhängig von der Zahl, Art und Verpackung der aktiven Komponenten werden auch die aktiven Komponenten, zum beispiel ICs in diesem Schritt aufgesetzt, eventuell manuell.

Danach wird die Platine in einen Reflowofen geschoben und durchläuft auf einen Transportband ein genau festgelegtes Temperaturprofil (siehe Abb. 18.5). Eine Platine durchläuft den Ofen in rund 20 Minuten. Bei der Massenproduktion von Leiterplatinen würden immer mehrer Leiterplatten den Ofen gleichzeitig druchlaufen. Es gibt verschiedene Lötzinnvarianten, die sich in ihrem Schmelzpunkt etc. unterscheiden. Das klassische Pb(50%)/Sn(50%) ist für kommerzielle Produkte in Europa seit einigen Jahren verboten.

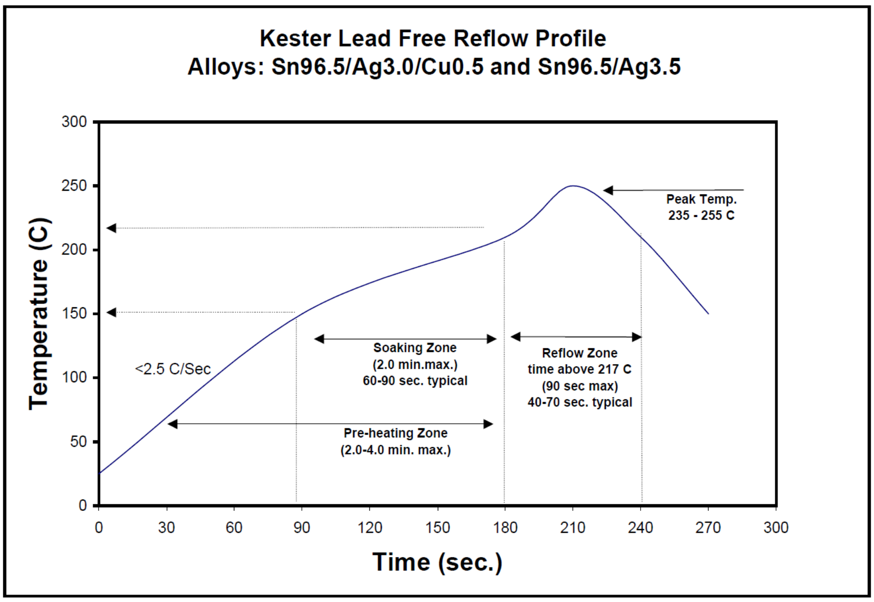


Abb. 18.5 Exemplarischer Temperaturverlauf beim Reflow-Löten (Quelle: Kester Inc.).

18.2.2 Kleben

ICs haben auf ihrer Rückseite, das heißt ihrer der Platine zugewandten Seite, meist eine dünne Metallisierungsschicht, falls es sich nicht um Flipchips handelt. Diese Schicht wird durch den leitenden Kleber mit Masse verbunden. (Eine beliebte und immer kontrovers diskutierte Frage bei „mixed-signal“ ASICs ist, welchen Teil des Chips man mit welcher Masse verbindet.) Unverpackte Chips haben keinen Kühlkörper, die Abwärme kann durch die Oberfläche an die Luft oder durch den Boden an die Platine abgegeben werden. Dies ist fast immer ein Problem, das besondere Aufmerksamkeit verlangt. Hier ist es wichtig, alle Klebeschichten sehr dünn zu halten, da selbst die besten Kleber schlechte Wärmeleiter sind. Eine große Kupferfläche in der Nähe des Chips ist hilfreich. Oft werden sogenannte „thermal vias“ unter dem Chip plaziert, die keine elektrische Funktion haben, sondern nur den themrischen Kontakt mit der Platine verbessen. Auch die Platine mit ihren Kunststoff- und Kleberlagen ist natürlich ein schlechter Wärmeleiter. Wieder sollte man auf die Dicke der Lagen achten. Beim Einsatz von Elektronik in der Nähe von gekühlten Halbleiterdetektoren ist die Abwärme von den Auslesechips besonders störend. Hier wird oft die Rückseite der Platine oder des MCM gekühlt. Die beste Wärmeleitung erreicht man wie oben ausgeführt mit Keramiksubstraten.

18.2.3 Drahtverbindungen

Drahtverbindungen sind maschinell hergestellte elektrische Verbindungen zwischen den Pads der ICs und der Platine oder dem MCM. Der Draht hat meistens einen Durchmesser von 20 μm bis 75 μm und besteht aus Aluminium oder auch Gold. Die Länge der Drahtverbindung ist recht unterschiedlich und variiert typischer Weise von 1 mm bis 10 mm, aber es sind auch Drahtverbindungen >10 cm möglich. Beim klassischen Aluminiumdrahtbonden wird der Draht durch Ultraschall mit den jeweiligen Pads verschmolzen. Dabei ist die Oberflächenbeschaffenheit und die Festigkeit des Pads wichtig. Auf dem Chip haben die Pads eine Aluminiummetallisierung. Die Pads auf der Platine haben meistens eine Nickel-Goldbeschichtung mit rund 1 μm Ni und 0,3 μm Gold.

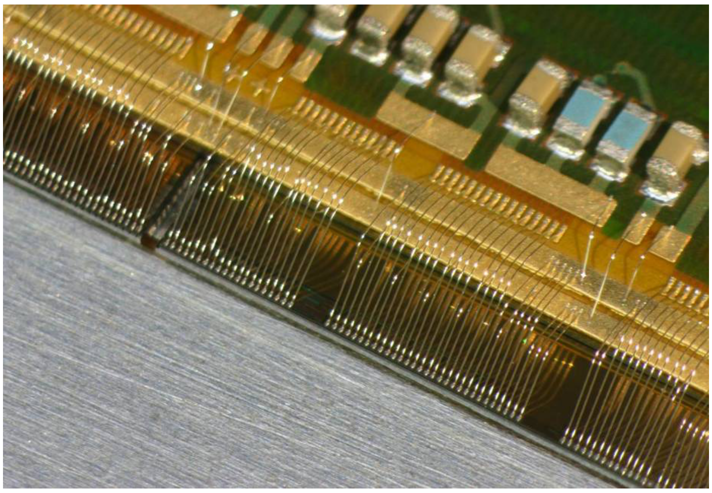


Abb. 18.6 Drahtverbindungen des Pixeldetektors des Experiments ALICE am Large Hadron Collider (LHC).

Eine Nahaufnahme einer Reihe von Drahtverbindungen ist in Abb. 18.6 gezeigt. Der Abstand der Drähte beträgt ≈200 μm. Es sind Drahtabstände von nur 50 μm realisierbar.

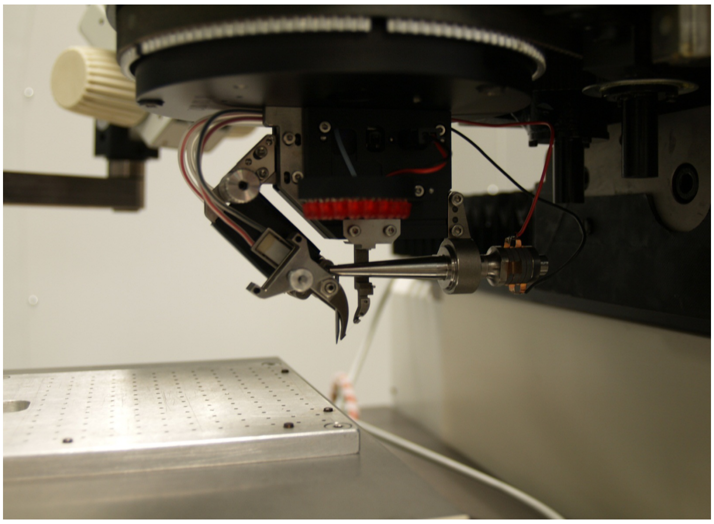


Abb. 18.7 Bondkopf mit Ultraschallhorn, Wedge- und Drahthalter.

Eine Aufnahme eines Bondkopfes ist in Abb.18.7 zu sehen. Moderne Maschinen sind sehr schnell, so dass mehere Verbindungen pro Sekunde gesetzt werden können. Eine zuverlässige Drahtverbindung hält eine Zugkraft von >15 g aus. Fehlerhafte Verbindungen kann man mit etwas Erfahrung durch optische Inspektion unter dem Mikroskop erkennen. So treten gerne Fehler an dem „Fuß“ eines Bondes auf (siehe Abb. 18.8). Eine hohe Integrationsdichte kann man mit überlappenenden Bonds erreichen (siehe Abb. 18.9).

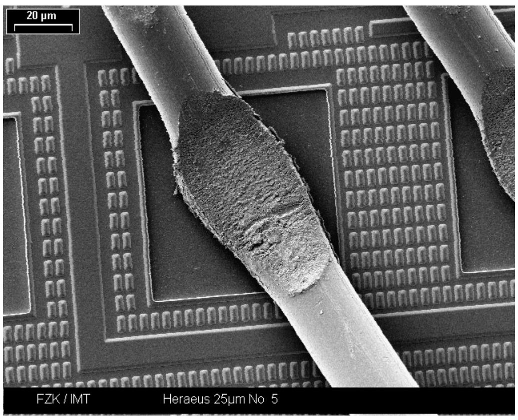


Abb. 18.8 „Fuß“ und „Ferse“ einer Drahtverbindung

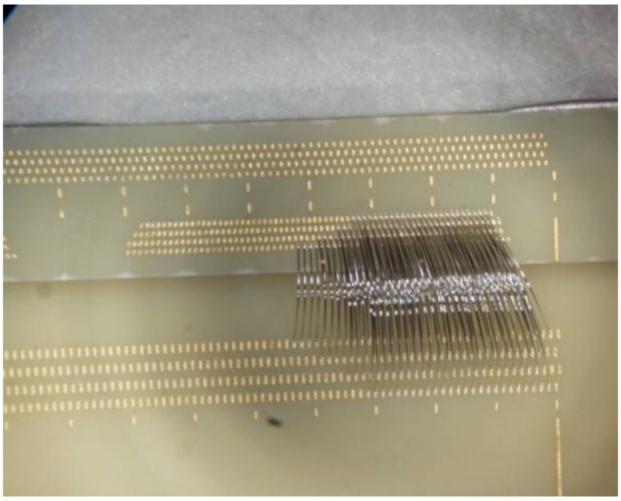


Abb. 18.9 Vierfache versetzte Reihe von Bonds.

18.2.4 Dickdrahtbonden

Eine Alternative zu dem vorgestellten Verfahren ist das „Thermosonic-ball-wedge-Bonding“. Hier wird der Draht durch eine Kapillare geführt und an der ersten Verbindung eine kleine Kugel geformt, die durch Druck, Ultraschall und hier auch Wärme mit dem Pad verbunden wird. Dies setzt vorraus, dass das pad und damit das Modul erwärmt werden kann und darf. Abb.18.10 zeigt einige Goldbälle (gold studs) ohne Drahtverbindung auf einer Siliziumoberfläche, Abb. 18.11 einen Kugelbonder. Will man nur die Kugeln setzen, z.B. für Flipchipprozesse, so kann man sehr hohe Geschwindigkeiten von fast 20 Kugeln pro Sekunde erreichen.

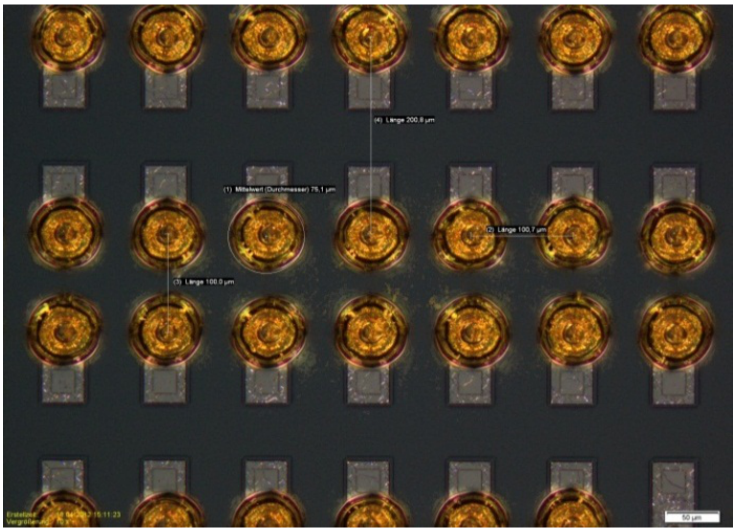


Abb. 18.10 Goldstuds auf Siliziumsubstrat.

Für Leistungselektronik oder Stromversorgungen will man oft sehr solide Verbindungen schafen. Dazu ist das Dickdrahtbonden oder das Bänderbonden geeignet. Die Drähte sind 100 μm bis 500 μm dick und können einen Zug von 150 bis 2200 g aushalten. Man kann Aluminium- oder Kupferdraht verwenden.

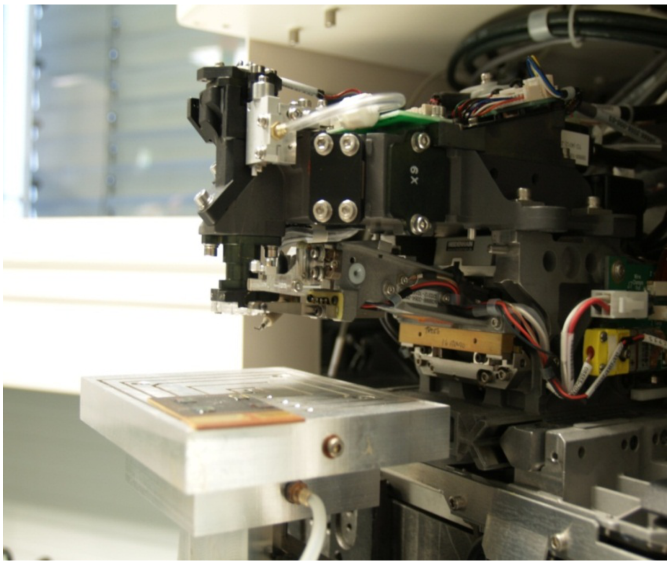


Abb. 18.11 Photo eines „Thermosonic gold stud bonder“.

18.2.5 Flipchiptechnologie

Drahtverbindungen ist immer noch die einfachste und sehr zuverlässige Technologie für hohe Packungsdichten, aber in der Praxis werden nur Pads an den Rändern von Chips verbunden. Die höhere verbindungsdichte erzieFlt man mittels Flipchiptechnologie (siehe Abb. 18.12). Hier kann man die gesamte Fläche eines Chips für Verbindungen genutzt und je nach Bedarf entsprechend viele Pads angebracht. Nach Aufbringen von Verbindungskugeln wird der Chip von einem Flipchipbonder kopfüber auf dem Substrat positioniert und die Verbindung zu den entsprechenden Substratpads hergestellt. Die Kugeln können aus Gold, Lot (Pb/Sn oder Sn/Ag/Cu) oder Indium bestehen. Je nach Verfahren werden dauerhafte Kontakte durch Wärme, Druck oder eine Kombination von beidem hergestellt. Wichtig ist es, dass die Metallisierung, die „Under bump metalization“ (UBM), der Pads zu den Kugeln passt. Nach dem Verbindungsvorgang wird zwischen Chip und Substrat meistens ein Kleber eingebracht, der die mechanische Stabilität erhöht.

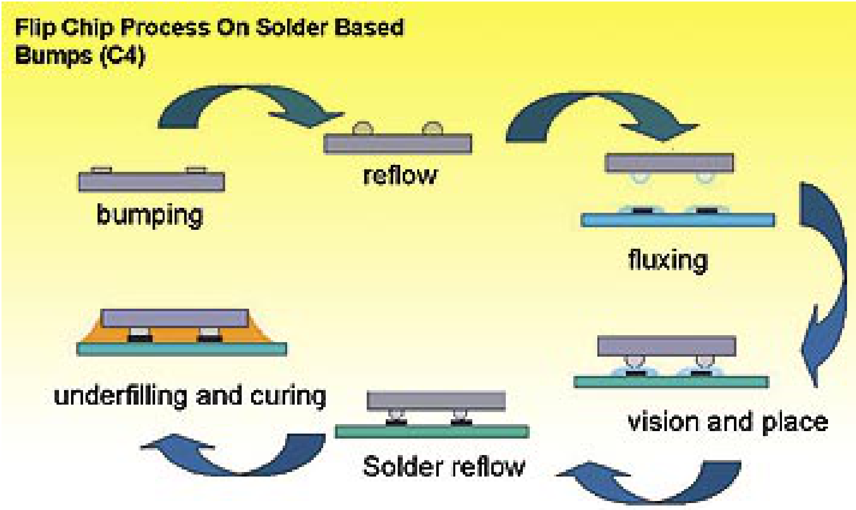
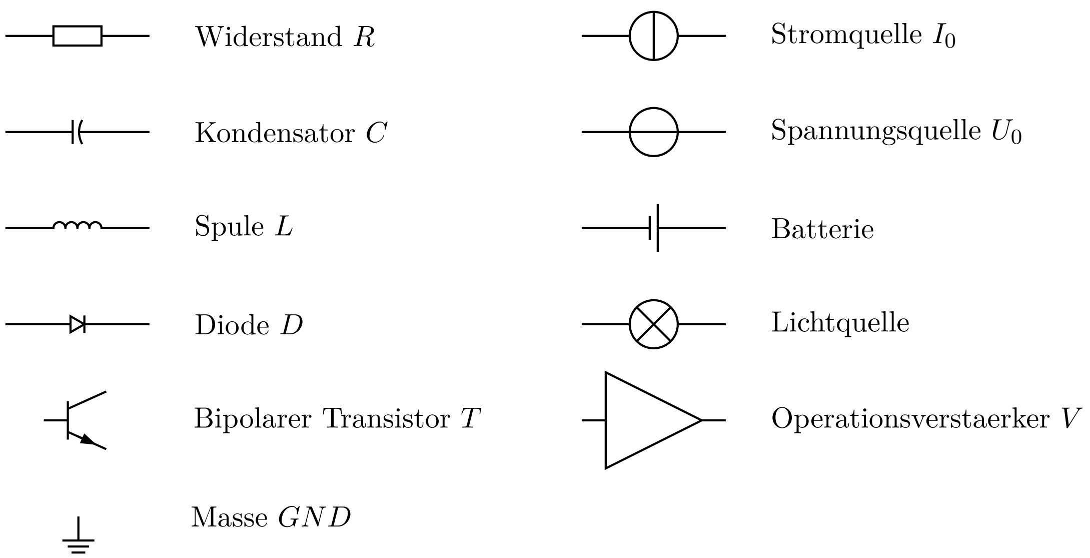


Abb. 18.12 Prozessschritte der Flipchiptechnologie.

19 Appendix

19.1 Elekrische Symbole



19.2 Einheitenpräfixe

Symbol	Präfix	Zehnerpotenz	Beispiel	Wert
Y	Yotta	10^{24}		Quadrillion
Z	Zetta	10^{21}		Trilliarde
E	Exa	10^{18}	EeV	Trillion
P	Peta	10^{15}	PByte	Billiarde
T	Tera	10^{12}	THz	Billion
G	Giga	10^{9}	GΩ	Milliarde
M	Mega	10^{6}		Million
k	Kilo	10^{3}		Tausend
h	Hekto	10^{2}	Hl	Hundert
da	Deka	10^{1}	10	Zehn
		10^{0}	1	Eins
d	Dezi	10^{-1}	dB	Zehntel
c	Zenti	10^{-2}		Hundertstel
m	Milli	10^{-3}		Tausendstel
μ	Mikro	10^{-6}	μm	Millionstel
n	Nano	10^{-9}	nV	Milliardstel
p	Piko	10^{-12}	pA	Billionstel
f	Femto	10^{-15}	fF	Billiardstel
a	Atto	10^{-18}	As	Trillionstel
z	Zepto	10^{-21}		Trilliardstel
y	Yokto	10^{-24}		Quadrillionstel

Tab. 19.1 Einheitenpräfixe.

19.3 Hilfreiche Zahlenwerte

$e = 2.7182818284\ldots$

$\frac{1}{e} = 0.3678\ldots$

$\sqrt{2} = 1.4142\ldots$

$\frac{1}{\sqrt{2}} = 0.7071\ldots$

$\sqrt{3} = 1.7320\ldots$

$\frac{1}{\sqrt{3}} = 0.5773\ldots$

$\log_{10}(2) = 0.30102\ldots$

19.4 Elementarladung

$e = 1.602 \times 10^{-19}\,\mathrm{C} \quad \Leftrightarrow \quad 1\,\mathrm{fC} \approx 6000\,e$

19.5 Grundlagen

Ohmsches Gesetz: $U = R\,I$

Spulengleichung: $U = L\,\dfrac{dI}{dt}$

Kondensatorgleichung: $Q = C\,U$ bzw. $I = C\,\dfrac{dU}{dt}$

Kirchhoff's Current Law (KCL): $\sum_i I_i = 0$

Der Gesamtstrom in jedem Knoten ist 0.

Annahme: Zufließende Ströme sind positiv, abfließende negativ.

Kirchoff's Voltage Law (KVL): $\sum_i V_i = 0$

Die Summe aller Spannungsabfälle in jeder geschlossenen Masche ist 0.

19.6 Parallele Widerstände

2 Widerstände: $R_1 \parallel R_2 \quad = \dfrac{R_1\,R_2}{R_1 + R_2}$, da $\dfrac{1}{R_1} + \dfrac{1}{R_2} = \dfrac{R_1 + R_2}{R_1\,R_2}$

3 Widerstände: $R_1 \parallel R_2 \parallel R_3 = \dfrac{R_1\,R_2 + R_1\,R_3 + R_2\,R_3}{R_1\,R_2\,R_3}$

19.7 Schaltungen mit R, C und L

(siehe Kapitel 2.12)

$$\underline{Z} = \frac{dU_{\mathrm{aus}}}{dU_{\mathrm{ein}}} = \frac{Z}{Z+R} - \frac{R_1}{R_1 + R_1} \tag{19.1}$$

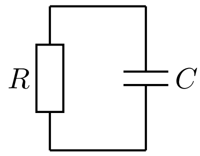
$$= \frac{1}{2}\left(\frac{1 - 2j\omega RC - \omega^2 R^2 C^2}{1 + \omega^2 R^2 C^2}\right) \tag{19.2}$$

$$= \frac{1}{2}\frac{\sqrt{[(1-\omega^2 C^2 R^2) - 2j\omega CR] \cdot [(1-\omega^2 C^2 R^2) + 2j\omega CR]}}{1 + \omega^2 C^2 R^2} \tag{19.3}$$

$$= \frac{1}{2}\frac{\sqrt{(1-\omega^2 C^2 R^2)^2 + 4\omega^2 C^2 R^2}}{1 + \omega^2 C^2 R^2} \tag{19.4}$$

$$= \frac{1}{2}\frac{\sqrt{1 + 2\omega^2 C^2 R^2 + \omega^4 C^4 R^4}}{1 + \omega^2 C^2 R^2} = \frac{1}{2} \quad . \tag{19.5}$$

19.8 Parallelschaltung von R und C



$$Z = R \,\|\, C = \frac{Z_1\,Z_2}{Z_1 + Z_2} = \frac{R\,\frac{1}{\mathrm{j}\,\omega\,C}}{R + \frac{1}{\mathrm{j}\,\omega\,C}} = \frac{R}{\mathrm{j}\,\omega\,R\,C + 1}$$

19.9 Superpositionsverfahren (zu Kap. 1.21)

Mit

$$R_B = \frac{R_L \cdot R_2}{R_L + R_2} \qquad \text{und} \qquad R_A = \frac{R_L \cdot R_1}{R_L + R_1} \tag{19.6}$$

folgt

$$R_1 + R_B = \frac{R_1 \cdot (R_L + R_2) \cdot R_L \cdot R_2}{R_L + R_2} \tag{19.7}$$

$$R_2 + R_A = \frac{R_2 \cdot (R_L + R_1) \cdot R_L \cdot R_1}{R_L + R_1} \tag{19.8}$$

$$I_L = I_{L_1} + I_{L_2} = \frac{U_1 \cdot R_B}{(R_1 + R_B) \cdot R_L} + \frac{U_2 \cdot R_A}{(R_2 + R_A) \cdot R_L} \tag{19.9}$$

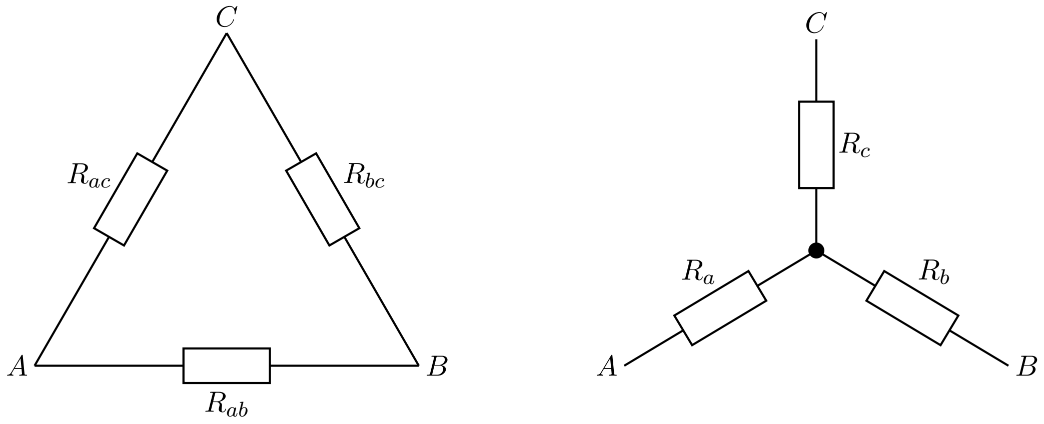
$$= \frac{U_1 \cdot \cancel{R_L} \cdot R_2}{(R_L + R_2) \cdot (R_1 + R_B) \cdot \cancel{R_L}} + \frac{U_2 \cdot \cancel{R_L} \cdot R_1}{(R_L + R_1) \cdot (R_2 + R_A) \cdot \cancel{R_L}} \tag{19.10}$$

$$= \frac{U_1 \cdot R_2}{\cancel{(R_L + R_2)} \cdot \frac{R_1 \cdot (R_L + R_2) + R_L \cdot R_2}{\cancel{R_L + R_2}}} + \frac{U_2 \cdot R_1}{R_2 \cdot (R_L + R_1) + R_L \cdot R_1} \,. \tag{19.11}$$

Mit Umsortieren des Nenners folgt

$$I_L = \frac{U_1 \cdot R_2 + U_2 \cdot R_1}{R_L \cdot (R_1 + R_2) + R_1 \cdot R_2} \,. \tag{19.12}$$

19.10 Dreieck-Stern-Transformationen



Dreieck → Stern

$$R_a = \frac{R_{ac}\,R_{ab}}{R_{ac}+R_{ab}+R_{bc}}$$
$$R_b = \frac{R_{ab}\,R_{bc}}{R_{ac}+R_{ab}+R_{bc}}$$
$$R_c = \frac{R_{ac}\,R_{bc}}{R_{ac}+R_{ab}+R_{bc}}$$

Stern → Dreieck

$$R_{ac} = \frac{R_a\,R_b + R_b\,R_c + R_c\,R_a}{R_b}$$
$$R_{ab} = \frac{R_a\,R_b + R_b\,R_c + R_c\,R_a}{R_c}$$
$$R_{bc} = \frac{R_a\,R_b + R_b\,R_c + R_c\,R_a}{R_a}$$

19.11 Elementare Mathematik

19.11.1 Trigonometrie

Winkel in Grad/Rad	$0° \mid 0$	$30° \mid \frac{\pi}{6}$	$45° \mid \frac{\pi}{4}$	$60° \mid \frac{\pi}{3}$	$90° \mid \frac{\pi}{2}$	$180° \mid \pi$
Sinus	$\frac{1}{2}\sqrt{0}=0$	$\frac{1}{2}\sqrt{1}=\frac{1}{2}$	$\frac{1}{2}\sqrt{2}$	$\frac{1}{2}\sqrt{3}$	$\frac{1}{2}\sqrt{4}=1$	0
Cosinus	1	$\frac{1}{2}\sqrt{3}$	$\frac{1}{2}\sqrt{2}$	$\frac{1}{2}$	0	-1
Tangens	0	$1\sqrt{3}$	1	$\sqrt{3}$	∞	0

19.11.2 Quadratische Gleichungen

$\mathrm{x}^2 + p\mathrm{x} + q = 0$ mit den Lösungen (abhängig von den Werten) p, q

$$\mathrm{x}_{1,2} = -\frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q}$$

19.11.3 Determinante einer Matrix

2×2-Matrix:

$$\det A = \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}\, a_{12} - a_{21}\, a_{22}$$

3×3-Matrix:

$$\begin{aligned} \det A &= \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \\ &= a_{11}\, a_{22}\, a_{33} + a_{12}\, a_{23}\, a_{31} + a_{13}\, a_{21}\, a_{32} - a_{13}\, a_{22}\, a_{31} - a_{11}\, a_{23}\, a_{32} - a_{12}\, a_{21}\, a_{33} \end{aligned}$$

n×n-Matrix

$$\det A = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1} A_{i,\sigma(i)}$$

Wie ist das Signum einer Permutation σ, $\operatorname{sgn}(\sigma)$, definiert? Der Übergang

$(1, 2, 3) \rightarrow (2, 3, 1)$ entspricht zwei Vertauschungen benachbarter Elemente und ist somit gerade $\Leftrightarrow \operatorname{sgn}(\sigma) = 1$.

$(1, 2, 3) \rightarrow (2, 1, 3)$ entspricht einer Vertauschung benachbarter Elemente und ist somit ungerade $\Leftrightarrow \operatorname{sgn}(\sigma) = -1$.

$(1, 2, 3) \rightarrow (3, 2, 1)$ entspricht drei Vertauschungen benachbarter Elemente und ist somit ungerade $\Leftrightarrow \operatorname{sgn}(\sigma) = -1$, etc.

Beispiel:

$$\det A = \det \begin{pmatrix} G_L & G_1 & G_2 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix}$$

$$\begin{aligned} &= G_L \cdot (-1) \cdot (-1) + G_1 \cdot 0 \cdot 1 + G_2 \cdot 1 \cdot 0 - G_2 \cdot (-1) \cdot 1 - G_1 \cdot 1 \cdot (-1) - G_L \cdot 0 \cdot 0 \\ &= G_L + G_1 + G_2 \end{aligned}$$

19.11.4 Kramersche Regel

$A \cdot \vec{x} = \vec{b}$, $\vec{x} = (x_1, x_2, \ldots, x_n)$ und $|A| \neq 0$ mit adjungierten Matrizen A_i.

Anmerkung: A_i erhält man aus A durch Ersetzten der i-ten Spalte von A mit $\vec{b}$

Dann gilt $x_i = \frac{|A_i|}{|A|}$

Beispiel (siehe Kap. 1.9):

$$\begin{pmatrix} G_L & G_1 & G_2 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix} \times \begin{pmatrix} U_{30} \\ U_{31} \\ U_{32} \end{pmatrix} = \begin{pmatrix} 0 \\ U_1 \\ U_2 \end{pmatrix} \quad .$$

Es wird hier der Wert von U_{30} gesucht.

$$\Leftrightarrow \quad U_{30} = \det \begin{pmatrix} 0 & G_1 & G_2 \\ U_1 & -1 & 0 \\ U_2 & 0 & -1 \end{pmatrix} \cdot \frac{1}{\det \begin{pmatrix} G_L & G_1 & G_2 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix}}$$

$$= \frac{0 + 0 + 0 - G_2 \cdot (-1) \cdot U_2 - G_1 \cdot U_1 \cdot (-1) - 0}{G_L + G_1 + G_2}$$

$$= \frac{G_1\, U_1 + G_2\, U_2}{G_L + G_1 + G_2}$$

19.11.5 Quotientenregel

Aus

$$f(x) \quad = \quad \frac{g(x)}{h(x)} \tag{19.13}$$

mit $g(x)$ und $h(x)$ differenzierbar sowie $h(x_0) \neq 0$ folgt

$$f'(x_0) \quad = \quad \frac{g' \cdot h - g \cdot h'}{h^2} \quad . \tag{19.14}$$

Beispiel:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}x} \frac{x}{x^2 + y^2} \quad &= \quad \frac{1 \cdot (x^2 + y^2) - x \cdot 2x}{(x^2 + y^2)^2} = \frac{x^2 + y^2 - 2x^2}{(x^2 + y^2)^2} \\ &= \quad \frac{y^2 - x^2}{(x^2 + y^2)^2} \end{aligned}$$

19.11.6 Partielle Integration

$$\int_a^b f'(x)\, g(x)\, dx + \int_a^b f(x)\, g'(x)\, dx = [\, f(x)\, g(x)\,]_a^b \tag{19.15}$$

19.11.7 Gammafunktion

Die Gammafunktion (Γ-Funktion) ist eine Fortsetzung der Fakultät.

Die Definition der Γ-Funktion ist

$$\Gamma(t) = \int_0^\infty \mathrm{e}^{-x}\, x^{t-1} dx \quad .$$

Für natürliche Zahlen n gilt

$$\Gamma(n+1) = n! \quad .$$

Ferner gilt

$$\Gamma(p+1) = p\, \Gamma(p)$$

und entsprechend

$$p\,(p+1)\,(p+2)\cdots(p+n-1) = \frac{\Gamma(p+n)}{\Gamma(p)} \quad .$$

19.12 Elementare Funktionentheorie

D sei eine offene Teilmenge von $\mathbb{C}$, d.h. D enthält nicht ihre Randpunkte[1]. Dann ist eine komplexe Funktion

$$f\colon\ D \subset \mathbb{C} \to \mathbb{C} \tag{19.16}$$

differenzierbar im Punkt z_0 der offenen Menge D, falls der Differentialquotient

$$\lim_{z\to z_0} \frac{f(z) - f(z_0)}{z - z_0} \quad = \quad f'(z_0) \tag{19.17}$$

existiert.

Ist $f(z)$ in allen Punkten $z_0 \in D$ differenzierbar, so bezeichnet man $f(z)$ als **holomorph**. Ist eine komplexe Funktion $f(z)$ mit $z = x + iy$ und

$$\mathrm{Re}\, f(z) + \mathrm{Im}\, f(z) \quad = \quad u(x,y) + v(x,y) \tag{19.18}$$

auf einem Teilgebiet $D \in \mathbb{C}$ differenzierbar, so gelten die

Cauchy-Riemanschen Differentialgleichungen:

$$\frac{\partial\, u}{\partial x} = \frac{\partial\, v}{\partial y} \quad \text{und} \quad \frac{\partial\, u}{\partial y} = -\frac{\partial, v}{\partial x} \quad . \tag{19.19}$$

[1]Zum Beispiel ist ein offener Kreis mit Radius R um z_0, also die Menge aller z mit $(z - z_0) < R$ so eine Menge.

$f(z)$ kann also auch als eine Vektorfunktion im Reellen $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ aufgefasst werden.

Beispiele:

a) $f(z) = z$

$$\begin{aligned} u = x \quad &, \quad v = y \\ \frac{\partial u}{\partial x} = \ 1 \ &= \frac{\partial v}{\partial y} \\ \frac{\partial u}{\partial y} = \ 0 \ &= -\frac{\partial v}{\partial x} \end{aligned}$$

b) $f(z) = z^2 = (x + jy)^2 = x^2 + 2jxy - y^2$

$$\begin{aligned} u = x^2 - y^2 \quad &, \quad v = 2xy \\ \frac{\partial u}{\partial x} = \ 2x \ &= \frac{\partial v}{\partial y} \\ \frac{\partial u}{\partial y} = \ -2y \ &= -\frac{\partial v}{\partial x} \end{aligned}$$

c) $f(z) = e^z = e^{x+jy} = e^x \cdot e^{jy} = e^x \left[\cos y + j \sin y\right]$

$$\begin{aligned} u = e^x \cos y \quad &, \quad v = e^x \sin y \\ \frac{\partial u}{\partial x} = \ e^x \cos y \ &= \frac{\partial v}{\partial y} \\ \frac{\partial u}{\partial y} = \ -e^x \sin y \ &= -\frac{\partial v}{\partial x} \end{aligned}$$

Die obigen drei Funktionen sind in ganz $\mathbb{C}$ differenzierbar. Es gibt natürlich auch nicht differenzierbare Funktionen, so z.B.

$$f(z) = z^* = x - jy \tag{19.20}$$

mit $u = x$ und $v = -y$, da

$$\frac{\partial u}{\partial x} = 1 \neq \frac{\partial v}{\partial y} = -1 \quad . \tag{19.21}$$

Diese Funktion ist in ganz $\mathbb{C}$ nicht differenzierbar.

Eine auf einem Gebiet[2] differenzierbare komplexe Funktion ist **beliebig oft** differenzierbar. Außerdem lässt sich so eine Funktion an jedem Punkt $z_0 \in D$ als Potenzreihe

$$f(z) = \sum_{k=0}^{\infty} c_k \left(z - z_0\right)^k \tag{19.22}$$

mit komplexen Koeffizienten c_k darstellen. Diese Beziehung gilt für alle Punkte z, die in einer von D vollständig enthaltenen offenen Kreisscheibe liegen. Es gilt

$$c_k = \frac{1}{k!} f^{(k)}(z_0) \quad , \tag{19.23}$$

[2]Ein Gebiet ist eine zusammenhängende, nichtleere offenen Menge.

wobei mit $f^{(k)}(z_0)$ die k-te Ableitung von $f(z)$ an der Stelle z_0 gemeint ist.

Anmerkung: Die Reihe enthält nur Terme c_k mit $k \geq 0$.

Interessanter und relevanter in der Elektronik sind Funktionen $f(z)$ mit einer (oder endlich vielen) Singularitäten. Außerhalb von z_0 sei $f(z)$ definiert und holomorph. Dann kann man $f(z)$ als **Laurent-Reihe** um den Punkt z_0 entwickeln

$$f(z) \quad = \quad \sum_{k=-\infty}^{\infty} c_k \, (z - z_0)^k \quad . \tag{19.24}$$

Hier treten auch Terme $k < 0$ auf.

Um die Koeffizienten c_k zu bestimmen, muss man ein **komplexes Kurvenintegral**

$$\int_C \mathrm{d}z \; f(z) \tag{19.25}$$

berechnen, wobei C eine einmal differenzierbare und zusammenhängende Kurve $C(t)\colon \mathbb{R} \to \mathbb{C}$ mit $t \in [a, b]$ ist[3]. Das Kurvenintegral ist dann folgendermaßen definiert:

$$\int_C \mathrm{d}z \; f(z) \quad = \quad \int_a^b \mathrm{d}t \; \frac{\mathrm{d}C(t)}{\mathrm{d}t} \cdot f\left[C\left(t\right)\right] \quad . \tag{19.26}$$

Zwei Beispiele mit Integralen über geschlossene Kreise $C(t) = e^{jt}$ mit $t \in [0, 2\pi]$:

a) $f(z) = z$

$$\int_C \mathrm{d}z \; z \quad = \quad \int_0^{2\pi} \mathrm{d}t \, j \, e^{jt} \cdot e^{jt} = j \int_0^{2\pi} \mathrm{d}t \; e^{2jt} = \frac{j}{2j} \left[e^{2jt}\right]_0^{2\pi} = 0 \quad . \tag{19.27}$$

b) $f(z) = \frac{1}{z}$

$$\int_C \mathrm{d}z \; \frac{1}{z} \quad = \quad \int_0^{2\pi} \mathrm{d}t \, j \, e^{jt} \cdot \frac{1}{e^{jt}} = j \int_0^{2\pi} \mathrm{d}t = 2\pi j \quad . \tag{19.28}$$

Ein Beispiel für nicht geschlossene Kurvenintegrale mit $C(t) = at$ wobei $a > 0$ und $t \in (0, 1)$:

$$\int_C \mathrm{d}z \; z^2 = \int_0^1 \mathrm{d}t \, a \cdot (at)^2 \cdot = a^3 \left[\frac{t^3}{3}\right]_0^1 = \frac{a^3}{3} \quad . \tag{19.29}$$

Allgemeiner gilt für einfach zusammenhängende Gebiete D, d.h. für Gebiete ohne Löcher bzw. Gebiete, in denen sich jede geschlossene Kurve auf einen Punkt zusammenziehen lässt, der

Integralsatz von Cauchy: Für jede geschlossene Kurve in D gilt

$$\int_C \mathrm{d}z \; f(z) = 0 \; , \tag{19.30}$$

wenn $f(z)$ holomorph und D einfach zusammenhängend ist.

[3] $C(t)$ könnte zum Beispiel ein Kreisbogen mit $t \in [0, 2\pi]$ und $C(t) = e^{jt}$ sein.

Das Beispiel $\int \mathrm{d}z\, \frac{1}{z}$ genügt nicht dem Satz von Cauchy, da im eingeschlossenen Kreis bei $z = 0$ eine Singularität liegt. Hier gilt der

Residuensatz. Sei f holomorph im offenen Gebiet $D \setminus \{z_0\}$, dann ist das Residuum von f definiert als der Koeffizient c_{-1} der Laurentreihe von f:

$$f(z) \quad = \quad \sum_{k=-\infty}^{\infty} c_k \,(z - z_0)^k \quad . \tag{19.31}$$

Die Koeffizienten der Laurentreihe sind gegeben durch

$$c_k \quad = \quad \frac{1}{2\pi j} \int \mathrm{d}z \, \frac{f(z)}{(z - z_0)^{k+1}} \quad . \tag{19.32}$$

Beispiele für die Bestimmung von Residuen:

a) $f(z) = \dfrac{1}{z^2 + z(a+b) + ab}$

Solche Terme sind sehr typisch für Anwendungen der Laplacetransformation und treten zum Beispiel beim Entwurf von Filtern höherer Ordnung auf. Die Pole sind bei $z_0 = -a$ und $z_0 = -b$, da

$$\frac{1}{z^2 + z(a+b) + ab} \quad = \quad \frac{1}{(z+a)\,(z+b)} \quad . \tag{19.33}$$

Hier kommt man mit folgendem Satz weiter: Falls $f(z)$ einen einfachen Pol in z_0 besitzt, gilt

$$f(z) \quad = \quad \frac{g(z)}{(z - z_0)} \quad . \tag{19.34}$$

mit einem $g(z)$, welches in z_0 holomorph ist. Dann gilt

$$\operatorname{Res}(f, z_0) \quad = \quad g(z_0) = \lim_{z \to z_0} (z - z_0)\, f(z) \quad . \tag{19.35}$$

Damit ergibt sich

$$\operatorname{Res}\left(\frac{1}{(z+a)\,(z+b)}, -a\right) \quad = \quad \lim_{z \to z_0} \frac{1}{z+b} = \frac{1}{b-a} \tag{19.36}$$

und entsprechend für

$$\operatorname{Res}(f, -b) \quad = \quad \frac{1}{a-b} \quad . \tag{19.37}$$

b) $f(z) = \dfrac{1}{z^2}$

Hier tritt ein Pol zweiter Ordnung auf. Wir verwenden dass

$$\mathrm{Res}\left(\frac{1}{z^2},0\right) = 0 \quad , \tag{19.38}$$

da der Term $\frac{1}{z}$ in der Laurentreihe nicht auftaucht. Dies bestätigt auch die direkte Integration:

$$\frac{1}{2\pi j}\int_C \mathrm{d}z\,\frac{1}{z^2} = \frac{1}{2\pi j}\int_0^{2\pi}\mathrm{d}t\,\frac{1}{\left(e^{jt}\right)^2}\cdot j\,e^{jt} = \frac{1}{2\pi}\int \mathrm{d}t\,e^{-jt} \tag{19.39}$$

$$= \frac{-1}{2\pi j}\left[e^{-jt}\right]_0^{2\pi} = 0 \quad . \tag{19.40}$$

Alternativ

$$\frac{1}{2\pi j}\int_C \mathrm{d}z\,\frac{1}{z^2} = \frac{-1}{2\pi j}\left[\frac{1}{z}\right]_C = \frac{-1}{2\pi j}\left[e^{-jt}\right]_0^{2\pi} = 0 \quad . \tag{19.41}$$

c) $f(z) = \dfrac{e^z}{z^3}$

Wir erweitern den obigen Satz: Sei

$$f(z) = \frac{g(z)}{(z-z_0)^n} \tag{19.42}$$

mit $n \geq 1$ und einer holomorphen Funktion $g(z)$, dann gilt

$$\mathrm{Res}\left(f(z), z_0\right) = \frac{g^{(n-1)}(z_0)}{(n-1)!} \quad . \tag{19.43}$$

Hier ist $g^{(n-1)}$ die $(n-1)$-te Ableitung von g. Damit ergibt sich

$$\mathrm{Res}\left(\frac{e^z}{z^3},0\right) = \left(\frac{\mathrm{d}}{\mathrm{d}z}\right)^{3-1}\frac{e^z}{(3-1)!}\bigg|_{z=0} = \frac{1}{2} \quad . \tag{19.44}$$

Dies sieht man alternativ durch die Entwicklung

$$f(z) = \frac{1}{z^3}e^z = \frac{1}{z^3}\cdot\sum_{n=0}^{\infty}\frac{z^n}{n!} = \frac{1}{z^3}\left(1+\frac{z}{1!}+\frac{z^2}{2!}+\frac{z3}{3!}+\cdots\right) \tag{19.45}$$

$$= \frac{1}{z^3}+\frac{1}{z^2}+\frac{1}{2z}+\frac{1}{6}+\cdots \quad . \tag{19.46}$$

d) $f(z) = e^{\frac{1}{z^2}}$ Es gilt

$$e^{\frac{1}{z^2}} = 1+\frac{1}{z^2}+\frac{1}{2!z^4}+\frac{1}{3!z^6}+\cdots \tag{19.47}$$

Also

$$\mathrm{Res}\left(e^{\frac{1}{z^2}},0\right) = 0 \quad . \tag{19.48}$$

Residuensatz:

Sei D ein einfach zusammenhängendes Gebiet und $f(z)$ holomorph in D mit Ausnahme der Punkte $z_0, z_1, \ldots, z_n$, in denen f singulär ist. Dann gilt für das Kurvenintegral über eine geschlossene Kurve C in $D \backslash \{z_0, z_1, \ldots, z_n\}$:

$$\frac{1}{2\pi j} \int_C \mathrm{d}z \; f(z) \quad = \quad \sum_{k=0}^{n} \mathrm{Res}\,(f, z_k) \quad . \tag{19.49}$$

In dieser Darstellung wird vorrausgesetzt, dass der Umlaufsinn von C um z_k positiv ist. Falls nicht, müsste man die $\mathrm{Res}\,(f, z_k)$ mit einem negativen Vorzeichen versehen.
Anwendungen:

a) $\displaystyle\int_0^{2\pi} \mathrm{d}t \; \frac{1}{2 + \cos t}$

Wir führen dies auf ein Kurvenintegral im Komplexen zurück mit der Substitution $z = e^{jt}$. Da

$$\cos t \quad = \quad \frac{e^{jt} + e^{-jt}}{2} = \frac{z + z^{-1}}{2} \tag{19.50}$$

und

$$\frac{\mathrm{d}z}{\mathrm{d}t} \quad = \quad j e^{jt} = jz \quad , \tag{19.51}$$

ist das obiges Integral identisch zu

$$\int_C \frac{\mathrm{d}z}{jz} \, \frac{1}{2 + \frac{z+z^{-1}}{2}} \quad = \quad \int_C \mathrm{d}z \; \frac{-2j}{z^2 + 4z + 1} \tag{19.52}$$

mit Polen bei

$$z_1 = -2 + \sqrt{3} \quad \text{und} \quad z_2 = -2 - \sqrt{3} \quad .$$

Nur die Singularität $-2 + \sqrt{3}$ liegt innerhalb der geschlossenen Kurve e^{jt}, daher ignorieren wir den anderen Pol. Damit gilt

$$\mathrm{Res}\left(f, -2 + \sqrt{3}\right) \quad = \quad \frac{-2j}{-\cancel{2} + \sqrt{3} + \cancel{2} + \sqrt{3}} = \frac{-j}{\sqrt{3}} \tag{19.53}$$

und

$$\int_C \mathrm{d}z \; f(z) \quad = \quad 2\pi j \cdot \frac{-j}{\sqrt{3}} = \frac{2\pi}{\sqrt{3}} \quad . \tag{19.54}$$

b) $\displaystyle\int_{-\infty}^{\infty} \mathrm{d}x \; \frac{1}{1 + x^2}$

Dieses Integral ist reel. Wir setzen es ins Komplexe fort, indem wir $x \in \mathbb{R}$ durch $z \in \mathbb{C}$ ersetzen und das Integral längs der reellen Achse durch einen Halbkreis in der Halbebene $\mathrm{Im}\, z \geq 0$ abschließen mit $R \cdot e^{jt}$, $t \in [0, \pi]$. Dies ist ein häufig verwendeter Trick.

Für

$$\lim_{R\to\infty}\int_C \mathrm{d}z\,\frac{1}{1+z^2} \tag{19.55}$$

geht dieser Teil des Integrals gegen Null. Damit gilt

$$\int_{-\infty}^{\infty}\mathrm{d}z\,\frac{1}{1+z^2} \quad = \quad 2\pi j\,\mathrm{Res}\left(\frac{1}{1+z^2}=\frac{1}{(z+j)\,(z-j)},j\right)=\frac{2\pi j}{2j}=\pi \quad . \tag{19.56}$$

Der Pol $-j$ liegt außerhalb der betrachteten Kurve und ist damit irrelevant. Insgesamt gilt also

$$\int_{-\infty}^{\infty}\mathrm{d}x\,\frac{1}{1+x^2} \quad = \quad \pi \quad . \tag{19.57}$$

Dies hätte man auch über

$$\int_{-\infty}^{\infty}\mathrm{d}x\,\frac{1}{1+x^2} \quad = \quad [\arctan x]_{-\infty}^{\infty}=\frac{\pi}{2}-\frac{-\pi}{2}=\pi \tag{19.58}$$

berechnen können.

19.13 Schaltalgebra

19.13.1 Exklusives OR und Exklusives NOR

Die Verknüpfung XOR (siehe Tabelle 12.9)

a	**b**	**c**
0	0	0
0	1	1
1	0	1
1	1	0

XOR

a	**b**	**c**
0	0	1
0	1	0
1	0	0
1	1	1

XNOR

hat die algebraische Darstellung

$$c \quad = \quad a\overline{b}+\overline{a}b\,, \tag{19.59}$$

die sich aus der Regel „Summe der Produkte“ sofort ergibt. Die Darstellung des Exlusiven NOR ist

$$c=\overline{a}\overline{b}+ab\,. \tag{19.60}$$

19.13.2 Volladdierer

Aus der Wahrheitstabelle des Volladdierers (siehe Tabelle 13.2)

a	**b**	**c_v**	**c**	**s**
0	0	0	0	0
0	0	1	0	1
0	1	0	0	1
0	1	1	1	0
1	0	0	0	1
1	0	1	1	0
1	1	0	1	0
1	1	1	1	1

ergibt sich:

$$\begin{aligned} c &= \overline{a}bc_\mathrm{v} + a\overline{b}c_\mathrm{v} + ab\overline{c}_\mathrm{v} + abc_\mathrm{v} \\ &= c_\mathrm{v}\left(\overline{a}b + a\overline{b}\right) + ab\left(\overline{c}_\mathrm{v} + c_\mathrm{v}\right) \\ &= c_\mathrm{v}\left(a \oplus b\right) + ab \\ c &= c_\mathrm{v}\, s_1 + c_1 \end{aligned} \tag{19.61}$$

(siehe Abb. 13.2) und

$$\begin{aligned} s &= \overline{a}\overline{b}c_\mathrm{v} + \overline{a}b\overline{c}_\mathrm{v} + a\overline{b}\overline{c}_\mathrm{v} + abc_\mathrm{v} \\ &= c_\mathrm{v}\left(\overline{a}\overline{b} + ab\right) + \overline{c}_\mathrm{v}\left(\overline{a}b + a\overline{b}\right) \\ &= c_\mathrm{v}\left(\overline{a \oplus b}\right) + \overline{c}_\mathrm{v}\left(a \oplus b\right) \\ s &= c_\mathrm{v} \oplus \left(a \oplus b\right) \ . \end{aligned} \tag{19.62}$$

19.13.3 Multiplexer

Aus der Wahrheitstabelle des Multiplexer (siehe Tabelle **??**)

E_0	**E_1**	**S**	**A**
0	0	0	0
0	1	0	0
1	0	0	1
1	1	0	1
0	0	1	0
0	1	1	1
1	0	1	0
1	1	1	1

Tab. 19.2 Wahrheitstabelle zum MUX.

ergibt sich

$$A = \left(E_0 + E_1 + S\right)\left(E_0 + \overline{E}_1 + S\right)\left(E_0 + E_1 + \overline{S}\right)\left(\overline{E}_0 + E_1 + \overline{S}\right) \ . \tag{19.63}$$

Hier ist es hilfreich die Beziehungen

$$\begin{aligned} (A+B)(A+B) &= (A+B)^2 = A+B\ , & (19.64)\\ (A+B)\left(A+\overline{B}\right) &= A^2+BA+A\overline{B}+B\overline{B} = A + A\left(B+\overline{B}\right)+0 & \\ &= A+A=A & (19.65)\end{aligned}$$

und

$$(A+B+C)\left(A+B+\overline{C}\right) = A+B \tag{19.66}$$

zu verwenden, anstatt einfach loszurechnen. Darum gilt:

$$A = (E_0+S)\left(E_1+\overline{S}\right) = E_0E_1 + SE_1 + E_0\overline{S} + S\overline{S} = E_0E_1 + SE_1 + \overline{S}E_0\ . \tag{19.67}$$

Dieser Ausdruck enthält zwar den zusätzlichen Term E_0E_1, gegenüber Gleichung 13.1, entspricht aber derselben Wahrheitstabelle!

19.14 Kettenschaltungen von RC-Gliedern (zu Kap. 2.8)

$U_{aus} = U_{ein} \cdot A \cdot B \cdot C$

mit $A = \dfrac{R}{Z+R}$,

mit $B = \dfrac{B_1}{B_2} = \dfrac{R\,(Z+R)}{(Z+R)^2 + R\,Z}$

mit $B_1 = R \,\|\, (Z+R) = \dfrac{R\,(Z+R)}{Z+2R}$

und $B_2 = Z + B_1 = \dfrac{Z\,(Z+2R) + R\,(Z+R)}{Z+2R}$,

$$\text{mit } C = \frac{C_1}{C_2} = \frac{R\,(Z^2 + 3RZ + R^2)}{Z^3 + R^3 + 6R^2Z + 5RZ^2}$$

$$\begin{aligned} \text{mit } C_1 = R \parallel B_2 &= R\,B_2 \cdot \frac{1}{R + B_2} \\ &= \frac{RZ(Z+2R) + R^2(Z+R)}{Z+2R} \cdot \frac{Z+2R}{RZ + 2R^2 + Z(Z+2R) + R(Z+R)} \\ &= \frac{RZ^2 + 3R^2Z + R^3}{3R^2 + 4RZ + Z^2} \end{aligned}$$

$$\begin{aligned} \text{und } C_2 = Z \parallel C_1 &= R\,B_2 \cdot \frac{1}{R + B_2} \\ &= \frac{3R^2Z + 4RZ^2 + Z^3 + RZ^2 + 3R^2Z + R^3}{3R^2 + 4RZ + Z^2} \\ &= \frac{Z^3 + 5RZ^2 + 6R^2Z + R^3}{3R^2 + 4RZ + Z^2}, \end{aligned}$$

$$U_{aus} = U_{ein} \cdot \frac{R^3}{6R^2Z + 5RZ^2 + Z^3 + R^3}$$

Anmerkungen:

- Es ist bequemer U_{ein}/U_{aus} zu betrachten.

- Viel schneller ist die Herleitung mittels der Kramerschen Regel [1, S. 42].

19.15 Bandsperre

Wir wollen die Übertragungsfunktion bestimmen, das heißt die Ausgangsspannung als Funktion der Eingangsspannung ausdrücken. Diese Schaltung ist etwas unübersichtlich und es gibt keine direkten Vereinfachungen. Deshalb stellen wir durch Anwendung der Kirschhoffschen Regeln ein Gleichungssystem auf. Abb. 19.1 entspricht Abb. 2.23 mit einer geeigneten Bezeichnung der passiven Komponenten und der (willkürlichen) Stromrichtung.

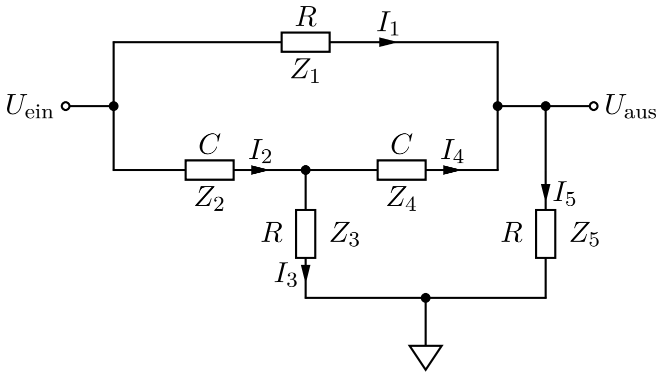


Abb. 19.1 Bandsperre.

$$U_{aus} = Z_5\, I_5 \tag{19.68}$$

$$U_{ein} - U_{aus} = Z_1\, I_1 \tag{19.69}$$

$$U_{ein} = Z_2\, I_2 + R_3\, I_3 \tag{19.70}$$

$$I_2 = I_3 + I_4 \tag{19.71}$$

$$I_1 + I_4 = I_5 \tag{19.72}$$

$$U_{ein} - U_{aus} = Z_2\, I_2 + Z_4\, I_4 \tag{19.73}$$

Die Größen U_{ein}, Z_1, Z_2, Z_3, Z_4 und Z_5 setzen wir als bekannt voraus. Es gibt sechs Unbekannte: U_{aus}, I_1, I_2, I_3, I_4 und I_5. Dabei sind die Ströme Hilfsgrößen, und uns interessiert nur die Ausgangsspannung U_{aus}. Es ergibt sich

$$U_{aus} = \frac{U_{ein}\ (Z_3\,[Z_4 + Z_2] + Z_2\, Z_4)\ Z_5 + U_2\, Z_1\, Z_3\, Z_5}{Z_5\ (Z_3\,[Z_4 + Z_2] + Z_2\, Z_4 + Z_1\,[Z_3 + Z_2]) + Z_1\ (Z_3\,[Z_4 + Z_2] + Z_2\, Z_4)} \quad . \tag{19.74}$$

Mit

$$Z_1 = Z_3 = Z_5 = R, \quad Z_2 = Z_4 = \frac{1}{j\omega C}, \quad s = j\omega, \quad Z_2\, Z_4 = \frac{1}{s^2 C^2}, \quad Z_2 + Z_4 = \frac{2}{sC}$$

und der Abkürzung

$$\tau = RC$$

ergibt sich die komplexe Übertragungsfunktion zu

$$\underline{Z} = \frac{U_{aus}}{U_{ein}} = \frac{2s\tau + 1 + s^2\tau^2}{5s\tau + 2 + s^2\tau^2} \quad . \tag{19.75}$$

Diese Gleichung kann man in die Form von Gl. 2.25 bringen

$$2 - \omega^2\tau^2 + 5j\omega\tau = 2\,(1 + \frac{j\omega\tau}{a})\,(1 + \frac{j\omega\tau}{b})$$

und damit die Koeffizienten a und b zu

$$a = \frac{5 + \sqrt{17}}{2} \approx 4{,}561 \quad \text{und} \quad a \cdot b = 2 \quad , \quad \text{also} \quad b \approx 0{,}438$$

bestimmen. Die Phasenverschiebung als Funktion von τ und ω ergibt sich zu

$$\tan\varphi = \frac{\mathrm{Im}(\underline{Z})}{\mathrm{Re}(\underline{Z})} = \frac{\omega\tau(2 - 1/a - 1/b) - \omega^3\tau^3(2/ab - 1/a - 1/b)}{1 - \omega^2\tau^2(1 + 1/ab - 2/a - 2/b) + \omega^4\tau^4/(ab)} \quad .$$

19.16 Reihenschwingkreis

zu Kapitel **??** Die Impedanz

$$Z \quad = \quad \sqrt{R^2 + \left(\omega L - \frac{1}{\omega C}\right)} \tag{19.76}$$

ist minimal für

$$\omega L - \frac{1}{\omega C} \qquad \Rightarrow \qquad Z = R \quad . \tag{19.77}$$

Z steigt auf $\sqrt{2}\,R$ für

$$\left(\omega L - \frac{1}{\omega C}\right)^2 \quad = \quad R^2 \quad . \tag{19.78}$$

Die Lösung von 19.78 ergibt sich aus

$$\pm\left(\omega L - \frac{1}{\omega C}\right) - R \quad = \quad 0 \tag{19.79}$$

bzw.

$$\pm\left(\omega^2 - \frac{1}{LC}\right) - \frac{R\omega}{L} \quad = \quad 0 \tag{19.80}$$

zu

$$\omega_{1,2} \quad = \quad \frac{R}{2L} \pm \sqrt{\left(\frac{R}{2L}^{\,2}\right) + \frac{1}{LC}} \tag{19.81}$$

und

$$\omega_{3,4} \quad = \quad \frac{-R}{2L} \pm \sqrt{\left(\frac{R}{2L}\right)^2 + \frac{1}{LC}}\,. \tag{19.82}$$

Nur ω_1 und ω_4 sind pos. Frequenzen und $\omega_1 > \omega_4$. Dami gilt

$$\Delta\omega \quad = \quad \omega_1 - \omega_4 = \frac{R}{L} \tag{19.83}$$

und

$$Q_R \quad = \quad \frac{1}{\sqrt{L}} \cdot \frac{L}{R} = \frac{1}{R}\sqrt{\frac{L}{C}} \quad . \tag{19.84}$$

19.17 Frequenzen

Die folgenden Kombinationen von R, L und C stellen jeweils eine Frequenz in [Hz] dar.

$$\left.\begin{aligned} \omega &= \tfrac{1}{RC} \\ \omega &= \tfrac{R}{L} \\ \omega &= \tfrac{1}{\sqrt{LC}} \end{aligned}\right\} \text{Hz}$$

19.18 Die Maxwell-Gleichungen

Wir werden die Maxwell-Gleichungen in dieser Vorlesung nicht oft benötigen, aber als Physiker oder Elektroingenieur sollte man sie kennen.

$$\vec{\nabla} \cdot \vec{E} = \frac{\rho}{\varepsilon_0} \quad \text{bzw.} \quad \oint_0 \vec{E} \cdot d\vec{A} = \int_V \rho\, dV = Q \tag{19.85}$$

$$\vec{\nabla} \cdot \vec{B} = 0 \tag{19.86}$$

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \quad \text{bzw.} \quad \oint_{dA} \vec{E} \cdot d\vec{s} = -\int \frac{\partial \vec{B}}{\partial t} \cdot d\vec{A} \tag{19.87}$$

$$\vec{\nabla} \times \vec{B} = \mu_0 \vec{j} + \mu_0\, \varepsilon_0\, \frac{\partial}{\partial t}\vec{E} \quad \text{bzw.} \quad \oint_{dA} \vec{B}\, d\vec{s} = \mu_0 \int_A \vec{j}\, d\vec{A} + \mu_0\, \varepsilon_0 \int_A \frac{\partial}{\partial t}\vec{E} \cdot d\vec{A} \tag{19.88}$$

Definitionen

$\vec{E}$ und $\vec{B}$ sind die Vektoren der elektrischen Feldstärke bzw. der magnetischen Flussdichte, ρ die elektrische Ladungsdichte, Q die elektrische Ladung, ε_0 die elektrische Dielektrizitätskonstante des Vakuums.

$$\operatorname{div} \vec{E} = \vec{\nabla} \cdot \vec{E} = \frac{\partial}{\partial_x} E_x + \frac{\partial}{\partial_y} E_y + \frac{\partial}{\partial_z} E_z \quad . \tag{19.89}$$

$$\operatorname{rot} \vec{E} = \vec{\nabla} \times \vec{E} = \begin{pmatrix} \frac{\partial E_z}{\partial_y} - \frac{\partial E_y}{\partial_z} \\ \frac{\partial E_x}{\partial_z} - \frac{\partial E_z}{\partial_x} \\ \frac{\partial E_y}{\partial_x} - \frac{\partial E_x}{\partial_y} \end{pmatrix} \quad . \tag{19.90}$$

Knappe Erläuterungen zu den Maxwellschen Gleichungen:

Gaußsches Gesetz (Gl. 19.85)

$$\operatorname{div} \vec{E} = \frac{\rho}{\varepsilon_0}$$

Hier wird der Gaußsche Integralsatz

$$\int \operatorname{div} \vec{E} \, dV = \int \vec{E} \, d\vec{A}$$

verwendet. Das elektrische Feld wird durch elektrische Ladungen erzeugt und der Fluß der elektrischen Feldstärke aus einer beliebig geschlossenen Oberfläche ist durch die eingeschlossene Ladung $Q = \int \rho dV$ gegeben. Die elektrischen Feldlinien gehen von positiven elektrischen Ladungen aus (Quellen) und enden an negativen Ladungen (Senken).

Gaußsches Gesetz für Magnetfelder (Gl. 19.86)

$$\operatorname{div} \vec{B} = 0$$

Die magnetische Flussdichte wird nicht durch magnetische Ladungen (Monopole) erzeugt, sondern stattdessen durch Ströme oder zeitlich variierende elektrische Felder (siehe Gl. 19.88). Magnetische Feldlinien sind geschlossen.

Faradaysches Induktionsgesetz (Gl. 19.87)

$$\vec{E} = -\frac{\partial \vec{B}}{\partial t}$$

Die Veränderung des magnetischen Flusses $\vec{B} \, d\vec{A}$ mit der Zeit führt zu einem elektrischen Wirbelfeld. Man beachte, dass $\int \vec{E} \, d\vec{s}$ nicht verschwindet. Es liegt hier also kein konservatives Feld vor! Das Induktionsgesetz ist die Grundlage für den Elektromotor.

Ampersches Gesetz mit Maxwellschem Verschiebungsstrom (Gl. 19.88)

$$\vec{B} = \mu \vec{j} + \mu_0 \, \varepsilon_0 \frac{\partial \vec{E}}{\partial t}$$

Ein (konstanter) Strom oder ein zeitlich variierendes elektrisches Feld erzeugen ein magnetisches Wirbelfeld.

Aus Gleichung 19.87 und Gleichung 19.88 kann man die Existenz elektromagnetischer Wellen im Vakuum ableiten mit $\vec{j} = 0, \rho = 0$ und den mathematischen Identitäten

$$\vec{\nabla} \cdot (\vec{\nabla} \cdot \vec{E}) = \vec{\nabla}(\vec{\nabla}\vec{E}) - \nabla^2 \vec{E}$$
$$\vec{\nabla} \cdot (\vec{\nabla} \cdot \vec{B}) = \vec{\nabla}(\vec{\nabla}\vec{B}) - \nabla^2 \vec{B}$$

mit

$$\vec{\nabla} \cdot \vec{E} = 0, \vec{\nabla} \cdot \vec{B} = 0, \vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}, \vec{\nabla} \times \vec{B} = \frac{1}{c^2}\frac{\partial \vec{E}}{\partial E}$$
$$\vec{\nabla} \cdot \frac{\partial \vec{E}}{\partial t} = \frac{\partial}{\partial t}\vec{\nabla} \cdot \vec{E}$$

und

$$\mu_0\, \varepsilon_0 = c^2$$

folgt

$$-\vec{\nabla} \cdot \frac{\partial \vec{B}}{\partial t} = 0 - \nabla^2 \vec{E} \Leftrightarrow \frac{1}{c^2}\frac{\partial^2 \vec{E}}{\partial t^2} = \nabla^2 \vec{E}$$

und

$$+\vec{\nabla} \cdot \frac{1}{c^2}\frac{\partial \vec{E}}{\partial t} = 0 - \nabla^2 \vec{B} \Leftrightarrow \frac{1}{c^2}\frac{\partial^2 \vec{B}}{\partial t^2} = \nabla^2 \vec{B} \quad .$$

Literaturverzeichnis

[1] Weddigen, C. ; Juengst, W.: *Elektronik: Eine Einfuehrung fuer Naturwissenschaftler und Ingenieure mit Beispielen zur Computer-Simulation.* Springer, 1993

[2] DIY, Electronics: *Importance of X10 oscilloscope probes.* Webseite, 2002 - 2012. – http://electronics-diy.com/electronic_schematic.php?id=970, Aufruf am Feb. 2012.

[3] Amkor: *Copper Wire Bonding.* Webseite, 2015. – http://www.amkor.com/go/packaging/copper-cu-wire-bondingl, Aufruf am Feb. 2015.

[4] Tarr, M.: *Flip Chart.* Webseite, 2015. – http://www.ami.ac.uk/courses/topics/0260_fc/index.html, Aufruf am Feb. 2015.

[5] Incorporated, Diodes: *BZX84C2V4W - BZX84C39W.* Webseite, 2009. – http://diodes.com/datasheets/ds30066.pdf, Aufruf am 1. Feb. 2012.

[6] Elektronik-Kompendium: *Spannungsstabilisierung mit Z-Diode.* Webseite, 2002 - 2012. – http://www.elektronik-kompendium.de/sites/slt/schalt/10121511.gif, Aufruf am 1. Feb. 2012.

[7] Mancini, R.: *Op Amps for Everyone.* Texas Instruments, 2001

[8] Renesas: Silicon NPN Expitaxial. (2004)

[9] Miller, J. M.: Dependence of the input impedance of a three-electrode vacuum tube upon the load in the plate circuit. In: *Scientific Papers of the Bureau of Standards* (1920), Nr. 15(351), S. 367–385

[10] Reisch, M.: *Elektronische Bauelemente: Funktion, Grundschaltungen, Modellierung mit SPICE.* Springer, 2006 http://books.google.com/books?id=8IbLAAAACAAJ. – ISBN 9783540340140

[11] Technologies AG infineon: IGBT. (2014)

[12] Bates, RL ; Bell, PJ ; Bernabeu, J. ; Bizzell, J. ; Bohm, J. ; Brenner, R. ; Renstrom, P.A.B. de ; Catinaccio, A. ; Cindro, V. ; Ciocio, A. u. a.: The ATLAS SCT grounding and shielding concept and implementation. In: *Journal of Instrumentation* 7 (2012), Nr. 03, S. P03005

[13] Ott, H.W.: *Noise reduction techniques in electronic systems.* John Wiley & Sons, 1988

[14] Hewes, J.: *Logic gates.* Webseite, 2011. – http://www.kpsec.freeuk.com/gates.htm, Aufruf am Feb. 2012.

[15] Spieler, H.: *Semiconductor detector systems.* Bd. 12. Oxford University Press, USA, 2005

[16] Microcontroller: *Festplatte*. Webseite, 2008. – `http://http://www.mikrocontroller.net/articles/Festplatte`, Aufruf am Feb. 2012.

[17] Shannon, C. E.: Communication in the presence of noise. In: *Proceedings of the IEEE* 86 (1998), Nr. 2, S. 447–457

[18] Maxfield, C.: *The Design warrior's guide to FPGAs: Devices, tools and flows*. Bd. 1. Elsevier, 2004

[19] Corporation, Altera: *MAX 5000 Programmable Logic Device Family*. Webseite, 1999. – `http://www.altera.com/literature/ds/m5000.pdf`, Aufruf am 1. Feb. 2012.

[20] Corporation, Altera: *FLEX 6000 Programmable Logic Device Family Data Sheet*. Webseite, 2001. – `http://www.altera.co.jp/literature/ds/dsf6k.pdf`, Aufruf am 1. Feb. 2012.

[21] Corporation, Altera: *Data Sheet: MAX 7000 Programmable Logic Device Family*. Webseite, 2005. – `http://www.altera.com/literature/ds/m7000.pdf`, Aufruf am 1. Feb. 2012.

[22] Ecoughli: *MACH 5 CPLD Family - Fifth Generation MACH Architecture*. Webseite, 2002. – `http://www.latticesemi.com/lit/docs/datasheets/cpld/mach5.pdf?jsessionid=f0306f3f32bc891a02cd671c531b154e4c25`, Aufruf am 1. Feb. 2012.

[23] Xilinx, Inc.: *Xilinx UG534 ML605 Hardware, User Guide*. Webseite, 2011. – `http://www.xilinx.com/support/documentation/boards_and_kits/ug534.pdf`, Aufruf am Feb. 2012.

[24] Xilinx, Inc.: *Xilinx UG364 virtex-6 FPGA configurable logic block user guide*. Webseite, 2009. – `http://www.xilinx.com/support/documentation/user_guides/ug364.pdf`, Aufruf am Feb. 2012.

[25] Kuon, I. ; Tessier, R. ; Rose, J.: Fpga architecture: Survey and challenges. In: *Foundations and Trends® in Electronic Design Automation* 2 (2008), Nr. 2, S. 135–253

[26] Xilinx, Inc.: *Xilinx DS150 virtex-6 family overview*. Webseite, 2012. – `http://www.xilinx.com/support/documentation/data_sheets/ds150.pdf`, Aufruf am Feb. 2012.

[27] Williams, T.: *The circuit designer's companion*. Newnes, 2004

[28] Sze, SM ; Ng, K.K..: *Physics of semiconductordevices*. John Wiley & Sons, 2006

[29] Bugg, D.V.: *Electronics: circuits, amplifiers and gates*. CRC Press, 2006

[30] Tietze, U. ; Schenk, C.: *Halbleiter-Schaltungstechnik*. Springer, 2002

[31] Veendrick, H.J.M.: *Deep-submicron CMOS ICs: from basics to ASICs*. Kluwer Academic Pub, 2000

[32] Horowitz, P. ; Hill, W.: *The art of electronics*. Cambridge University Press, 1989

[33] PATEL, A.: *The basics of SerDes (serializers/deserializers) for interfacing.* Webseite, 2010. – http://www.analog-eetimes.com/en/the-basics-of-serdes-serializers/deserializers-for-interfacing.html?cmp_id=71&news_id=222901121, Aufruf am Feb. 2012.

[34] STORR, W.: *Digital logic gates.* Webseite, 2011. – http://www.electronics-tutorials.ws/logic/logic_1.html, visited on February 1st, 2012.

[35] STORR, W.: *Active high pass filter.* Webseite, 2011. – http://www.electronics-tutorials.ws/filter/filter_6.html, Aufruf am Feb. 2012.

[36] ID, Maxim: *Understanding flash ADCs.* Webseite, 2010. – http://www.maxim-ic.com/app-notes/index.mvp/id/810, Aufruf am Feb. 2012.

[37] ID, Maxim: *Understanding SAR ADCs: Their architecture and comparison with other ADCs.* Webseite, 2001. – http://www.maxim-ic.com/app-notes/index.mvp/id/1080, Aufruf am Feb. 2012.

[38] KUPHALDT, T.R.: *Digital analog conversion.* Webseite, 2000 - 2002. – http://www.opamp-electronics.com/tutorials/digital_theory_ch_013.htm, Aufruf am Feb. 2012.

[39] ID, Maxim: *Pipeline ADCs come of age.* Webseite, 2001. – http://www.maxim-ic.com/app-notes/index.mvp/id/634, Aufruf am Feb. 2012.

[40] STALLER, L.: *Understanding analog to digital converter specifications.* Webseite, 2005. – http://eetimes.com/design/embedded/4025078/Understanding-analog-to-digital-converter-specifications, Aufruf am Feb. 2012.

[41] ID, Maxim: *The ABCs of ADCs: Understanding how ADC errors affect system performance.* Webseite, 2002. – http://www.maxim-ic.com/app-notes/index.mvp/id/748, Aufruf am Feb. 2012.

[42] RAZAVI, B.: *Design of analog CMOS integrated circuits.* McGraw-Hill, 2001 (McGraw-Hill series in electrical and computer engineering). http://books.google.de/books?id=X_rAQgAACAAJ. – ISBN 9780072380323