

Abteilung IIC, 3. Semester

Elektrotechnik

Prof. Dr. R. Vahldieck

Prof. Dr. Ch. Hafner

0. Einführung

Elektrotechnik beinhaltet die Nutzbarmachung aller bekannten elektrischen Phänomene. Die Vielfalt dieser Phänomene ergibt einen riesigen Spielraum. Gleichzeitig ist eine physikalisch präzise Beschreibung sehr aufwendig und erfordert profunde Kenntnisse der Quantentheorie und der Feldtheorie. Um effizient mit dem Entwurf, der Analyse und der Betreuung von komplexen elektrischen Systemen umgehen zu können, wurden verschiedene, oft stark vereinfachende Betrachtungsweisen und Methoden eingeführt, welche sich in der Praxis gut bewährt haben. Besonders wichtig ist dabei die makroskopische Betrachtung der Phänomene, welche bereits in der Maxwell'schen Theorie der elektromagnetischen Felder eingeführt wurde. Die Maxwell Theorie bildet den auch die physikalische Grundlage der Elektrotechnik. Allerdings ist diese Theorie für die Beschreibung elektrischer Systeme zu aufwendig. Sie wird deshalb meist nur für Detailabklärungen und zur Herleitung einfacherer Theorien verwendet. Die in der Elektrotechnik bedeutendste, vereinfachte Theorie ist die Netzwerktheorie. Diese ist sehr praxisbezogen und effizient. Sie bildet deshalb die wichtigste Grundlage für diese Vorlesung.

Die grosse Anzahl verschiedener elektromagnetischer Systeme macht deren genaue Beschreibung unmöglich. Allerdings kann man diese Systeme grob in vier verschiedene Gruppen einteilen: 1) Kommunikationssysteme, 2) Computersysteme, 3) Automatisierungssysteme und 4) Energiesysteme. Die ersten beiden Gruppen beinhalten Informationsübertragung und Informationsverarbeitung und haben damit einen offensichtlichen Bezug zur Informatik. Ohne die elektrotechnisch entwickelten Kommunikations- und Computersysteme hätte die Informatik sicher nicht die Bedeutung erlangt, welche ihr in den modernen Gesellschaften zukommt. Dass dies auch für Automatik und Energiesysteme zutrifft ist etwas weniger offensichtlich macht sich aber spätestens dann bemerkbar, wenn ein Computer abstürzt, weil die Temperaturregelung versagt oder weil die Energieversorgung zusammenbricht. Informatiker sind also zweifellos Nutzniesser der Elektrotechnik.

Umgekehrt bedienen sich die Elektroingenieure zunehmend der Informatik, um neue, komplexe Systeme zu entwickeln. Dadurch ergibt sich automatisch eine enge, fruchtbare Zusammenarbeit von Elektroingenieuren und Automatikern. Damit diese möglich wird, müssen Informatiker mit den wichtigsten Grundlagen der Elektrotechnik ebenso vertraut sein wie die Elektroingenieure mit den wichtigsten Grundlagen der Informatik.

Es ist das Ziel dieser Vorlesung, die dafür notwendigen Grundlagen zu schaffen. Wir werden uns dabei auf die Netzwerktheorie konzentrieren, aber auch physikalische Grundlagen einstreuen, welche für das Verständnis hilfreich sind.

1. Grundlagen

Da die physikalischen Grundlagen der Elektrotechnik aus der Physik bereits bekannt sind und eine ausführliche Darstellung den Rahmen dieser Vorlesung sprengen würde, beschränken wir uns hier auf eine kurze Übersicht über die wichtigsten Phänomene und Grundgrößen.

1.1. Elektrische Ladung, Feld und Kraft

Die elektrische Ladung Q (Masseinheit Coulomb, $1\text{C} = 1\text{As}$) ist zunächst die Quelle des elektrischen Feldes. Eine am Ort \vec{r}_Q ruhende Punktladung der Stärke Q erzeugt im gesamten Raum ein elektrisches Feld. Am Ort \vec{r} ergibt sich das Feld

$$\vec{E}(\vec{r}) = \frac{Q(\vec{r} - \vec{r}_Q)}{4\pi\epsilon_0 |\vec{r} - \vec{r}_Q|^3}.$$

Dabei ist ϵ_0 eine Proportionalitätskonstante, die sogenannte Dielektrizitätskonstante des freien Raumes. Das elektrische Feld ist eine vektorielle Größe und zeigt von der erzeugenden Punktladung radial weg. Zudem ist die Feldstärke $E = |\vec{E}|$ (Masseinheit V/m) proportional zur Ladung Q und nimmt mit dem Quadrat des Abstandes von der Punktladung ab.

Für das elektrische Feld gilt das *Superpositionsprinzip*, d.h. das von mehreren Punktladungen erzeugte Gesamtfeld ergibt sich aus der (vektoriellen) Addition des von jeder einzelnen Ladung erzeugten Feldes. Es genügt daher, zu wissen, wie das Feld einer einzigen Punktladung berechnet wird, um das Feld für jede beliebige Ladungsanordnung zu berechnen. Dem Superpositionsprinzip in etwas anderer Form (Quellenüberlagerung) werden wir auch in der Netzwerktheorie wieder begegnen.

Einerseits erzeugt eine Ladung ein elektrisches Feld, andererseits erfährt eine Ladung Q in einem elektrischen Feld (das von anderen Ladungen erzeugt wird) eine Kraft

$$\vec{F} = Q\vec{E},$$

welche messbar ist.

Eines der tiefgründigsten physikalischen Gesetze besagt dass Ladungen weder aus nichts entstehen noch verschwinden. Die Gesamtladung eines geschlossenen Systems bleibt deshalb konstant. Auch diesem Ladungserhaltungssatz werden wir in abgewandelter Form in der Netzwerktheorie begegnen.

1.2. Arbeit, Energie, Spannung und Potential

Bewegt man eine Ladung in einem elektrischen Feld längs eines Weges s von einem Punkt A zu einem Punkt B , so muss Arbeit geleistet werden. Damit wird es möglich, elektrische Energie in mechanische Energie bzw. Arbeit umzuformen. Es gilt

$$W = \int_A^B \vec{F}(s) d\vec{s} = \int_A^B Q \vec{E}(s) d\vec{s} = Q \int_A^B \vec{E}(s) d\vec{s} = QU .$$

Die Masseinheit für Arbeit bzw. Energie ist Joule ($1\text{J} = 1\text{Nm}$). Folglich gilt für die Masseinheit Volt mit der die elektrische Spannung U gemessen wird: $1\text{V} = 1\text{J/C}$.

Die Berechnung der Spannung beinhaltet ein Wegintegral über das elektrische Feld:

$$U = \int_A^B \vec{E}(s) d\vec{s} .$$

Es ist eine spezielle Eigenschaft des elektrostatischen Feldes, dass dieses Integral von der geometrischen Form des Weges unabhängig ist und nur von den beiden Endpunkten A und B abhängt. Für den Spezialfall eines geschlossenen Weges (mit $A = B$) gilt dann:

$$\oint \vec{E}(s) d\vec{s} = 0 .$$

Felder mit dieser Eigenschaft nennt man übrigens wirbelfrei.

Diese Eigenschaft ermöglicht es, jedem Punkt des Raumes ein elektrisches Potential Φ so zuzuordnen, dass die elektrische Spannung zwischen den Punkten sich aus der Potentialdifferenz ergibt:

$$U_{AB} = \Phi_B - \Phi_A .$$

Das elektrische Potential hat natürlich dieselbe Masseinheit (Volt), wie die elektrische Spannung. Letztere ist übrigens eine zentrale Grösse der Netzwerkanalyse. Wir werden sehen, dass die eben erwähnte Eigenschaft des elektrostatischen Feldes Konsequenzen für die Netzwerktheorie hat.

Es ist wichtig zu beachten, dass die elektrische Spannung eine gerichtete Grösse ist. Es gilt $U_{BA} = -U_{AB}$. In der Praxis werden die Indices AB meist weggelassen, was bei Anfängern gerne zu Vorzeichenfehlern führt.

1.3. Bewegte Ladungen, Magnetfeld, elektrischer Strom

Wird eine elektrische Ladung bewegt, so wird neben dem elektrischen Feld ein magnetisches Feld erzeugt, welches nicht nur proportional zur Ladung, sondern auch zu deren Ge-

schwindigkeit ist. Eine am Ort \vec{r}_Q mit der Geschwindigkeit \vec{v} bewegte Punktladung der Stärke Q erzeugt im gesamten Raum ein magnetisches Feld. Am Ort \vec{r} ergibt sich die magnetische Induktion, d.h. das B Feld

$$\vec{B}(\vec{r}) = \frac{\mu_0 Q \vec{v} \times (\vec{r} - \vec{r}_Q)}{4\pi |\vec{r} - \vec{r}_Q|^3}.$$

Dabei ist μ_0 eine Proportionalitätskonstante, die sogenannte Permeabilitätskonstante des freien Raumes. Die Richtung der magnetischen Induktion ist wegen des Vektorproduktes in obiger Formel etwas weniger leicht auszumachen als die Richtung des elektrischen Feldes. Das B Feld steht senkrecht zur Bewegungsrichtung und zum Abstandsvektor. Die Stärke des B Feldes nimmt ebenso wie die Stärke des E Feldes mit dem Quadrat des Abstandes ab. Wichtig ist die Proportionalität zur Geschwindigkeit. Unbewegte Ladungen erzeugen demzufolge kein B Feld.

Für das B Feld gilt das Superpositionsprinzip.

Eine ruhende Ladung Q erfährt in einem B Feld (welches von bewegten Ladungen erzeugt wird) keine Kraft. Hingegen erfährt eine mit der Geschwindigkeit \vec{v}_q bewegte Ladung die messbare Kraft

$$\vec{F} = Q \vec{v}_q \times \vec{B}.$$

Bei makroskopischen Beschreibungen kümmert man sich nicht um die Bewegungen einzelner Ladungen, sondern um Bewegungen von Ladungsströmen. Der elektrische Strom I bezeichnet den Gesamtfluss aller elektrischen Ladungen durch eine bestimmte Fläche pro Zeiteinheit. Praktisch werden elektrische Ladungen häufig in metallischen Drähten geführt, d.h. sie bewegen sich hauptsächlich längs des Drahtes. Die Fläche für die Definition des Stromes ist dann sinnvollerweise eine Querschnittsfläche senkrecht zur Drahtachse. Häufig wird die Querabmessung des Drahtes vernachlässigt und man spricht von einem Fadenstrom.

Dünne Drähte vermögen in der Praxis kaum elektrische Ladungen zu speichern. Wegen des Ladungserhaltungssatzes folgt daraus, dass der Strom, welcher am Drahtanfang A in den Draht hinein fließt, den Draht in gleicher Stärke und zur gleichen Zeit am Ende B verlassen muss. Dies ist eine Idealisierung, welche in der elektrischen Schaltungstechnik häufig gemacht wird. Wir werden diese Idealisierung zunächst ebenfalls verwenden, später jedoch genauere Modelle betrachten.

Da elektrische Ströme die Summe bewegter Ladungen darstellen, erzeugen sie B Felder und erfahren durch diese Felder eine messbare Kraft. Beispielsweise erfährt ein Fadenstrom I der Länge d in einem gegebenen B Feld die Kraft

$$\vec{F} = I \vec{d} \times \vec{B},$$

wobei \vec{d} in Richtung des Fadenstromes zeigt.

Wie die elektrische Spannung, ist auch der elektrische Strom eine gerichtete Grösse. Diese zwei Grössen bilden den Kern der Netzwerkanalyse.

1.4. Ladungserhaltung

Für die elektrischen Ladungen gilt der Ladungserhaltungssatz, d.h. Ladungen können weder aus dem Nichts entstehen noch verschwinden. Betrachtet man die Gesamtladung Q in einem Volumen V , so muss deren zeitliche Änderung den Strömen durch die Oberfläche ∂V des Volumens gleich sein. Wichtig ist dabei die Festlegung für die *Zählrichtung* der Ströme. Zählt man die Ströme, welche aus dem Volumen *heraus* fliessen *positiv*, so nimmt die Ladung im Volumen ab, wenn der Gesamtstrom aus dem Volumen positiv ist. Also gilt:

$$\frac{\partial Q}{\partial t} = -\sum_k I_k .$$

Diese Form des Ladungserhaltungssatzes ist für die Netzwerktheorie sehr wichtig.

1.5. Vernachlässigungen und Vereinfachungen

In den bisherigen Abschnitten haben wir elektrische Ströme und Spannungen mit einigen offensichtlichen Vereinfachungen eingeführt. Für die Definition der Spannung sind wir von unbewegten Ladungen ausgegangen, haben dann aber für die Definition des Stromes bewegte Ladungen betrachtet. Bewegte Ladungen erzeugen nicht nur zusätzlich zum elektrischen Feld ein Magnetfeld, sie erzeugen auch eine Verwirbelung des elektrischen Feldes, welche zur Folge hat, dass die Definition der elektrischen Spannung nicht mehr wegunabhängig ist. Ausserdem breitet sich das elektromagnetische Feld wellenartig, mit endlicher Geschwindigkeit aus, was verschiedene Konsequenzen hat. So erscheint ein Strompuls am Ende eines Drahtes nicht zur gleichen Zeit wie am Drahtanfang. Die Pulsausbreitung benötigt Zeit. Als Folge davon kann der Draht wiederum — wegen des Ladungserhaltungssatzes — nicht ladungsfrei sein.

Glücklicherweise ist die Ausbreitungsgeschwindigkeit elektromagnetischer Wellen für unser Empfinden enorm hoch. Im freien Raum breiten sich diese Wellen mit $c \approx 3 \cdot 10^8$ m/s aus. Eine ähnliche Ausbreitungsgeschwindigkeit erreichen auch elektrische Impulse auf gut leitenden Drähten. Die Laufzeiten können deshalb oft vernachlässigt werden. Trotzdem muss man sich dieser Vernachlässigungen bewusst sein und allenfalls nach besseren Modellen suchen. Solche Modelle sind beispielsweise bei langen Datenleitungen und hohen Datenraten erforderlich. Beispielsweise ergibt sich für eine 3m lange Datenleitung eine Laufzeit von ca. 10^{-8} s. Bei schnellen Pulsen von beispielsweise $1\text{ns} = 10^{-9}$ s ist diese Laufzeit natürlich erheblich. Laufzeitunterschiede an Eingängen logischer Schaltungen führen

offensichtlich zu Fehlverhalten. Bei einer Taktrate von 1GHz können Laufzeitunterschiede von 1ns nicht toleriert werden. Solche Laufzeitunterschiede ergeben sich, wenn die Zuleitungen sich um 30cm unterscheiden.

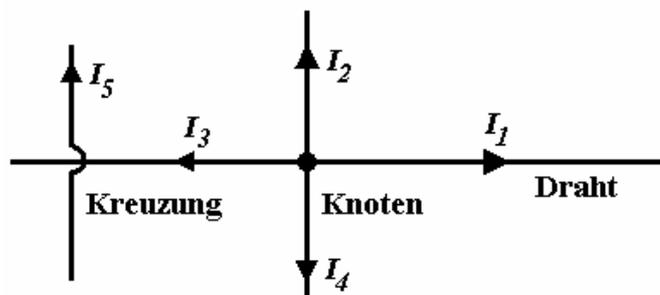
Wir werden im Folgenden mit den grösstmöglichen Vereinfachungen beginnen und diese nur aufgeben, wenn es erforderlich ist.

2. Elektrische Netzwerke bei Gleichstrom

In der Netzwerktheorie wird die Methode der grösstmöglichen Vereinfachung und der makroskopischen Betrachtung, welche wir bei der Einführung des Fadenstromes verwendet haben, konsequent angewendet. Darüber hinaus werden graphische Symbole für die Darstellung sowie Methoden zur Berechnung der Ströme und Spannungen in elektrischen Schaltungen eingeführt.

Das graphische Symbol für einen idealisierten Fadenstrom ist eine einfache Linie, die der tatsächlichen Form des verwendeten Drahtes i.a. nicht entspricht, sondern lediglich angibt, welche Elemente der Draht verbindet. Für die Schaltungsanalyse wird dann angenommen, dass das Verhalten dieses Drahtes nahezu ideal sei, dass also der Strom am Eingang des Drahtes identisch mit dem Strom am Ausgang sei, dass sich auf dem Draht keine Ladungen ansammeln usw. Es reicht daher, jeden Draht mit einem einzigen Strom zu charakterisieren. Da der Strom eine gerichtete Grösse ist, wird der Strom mit einem Pfeil auf dem Draht markiert, welcher die positive Stromrichtung angibt. Ist die Stromrichtung unbekannt, so wird eine willkürliche Richtung eingesetzt. Ergibt die Rechnung schliesslich einen negativen Stromwert, so heisst dies, dass der positive Strom in die umgekehrte Richtung fliesst.

Mehrere Drähte können miteinander in einem Knoten verbunden werden. Knoten werden graphisch durch Punkte markiert. Praktisch haben Drahtverbindungen eine gewisse Ausdehnung. Sie können also beispielsweise Ladungen speichern. Idealisierte Knoten der Netzwerktheorie speichern keine Ladungen.



Figur 1: Drähte mit Strompfeilen, Knoten und Kreuzung. Die Ströme I_1 , I_2 , I_3 , I_4 sind miteinander in einem Knoten elektrisch verbunden. Zwischen I_3 und I_5 besteht keine elektrische Verbindung. Der kleine „Umweg“ von I_5 verdeutlicht die Kreuzung, wird aber bequemlichkeitshalber in Europa weggelassen. Umgekehrt wird der Punkt, welcher den Knoten verdeutlicht, in Amerika weggelassen. Wir verwenden im Folgenden die europäische Konvention.

Elektrische Schaltungen, welche nur aus idealisierten Drähten und Knoten bestehen sind ohne praktisches Interesse. Um zu interessanten Schaltungen zu kommen, müssen etwas weniger ideale Bauteile eingeführt werden. Die einfachsten derartigen Bauteile haben dieselbe Form wie ein Draht, aber ein interessantes Innenleben. Das heisst, diese Bauteile ha-

ben zwei drahtförmige Anschlüsse und werden deshalb Zweipole genannt. Das Innenleben der Zweipole wird durch physikalische Effekte charakterisiert. Wir werden verschiedene Zweipole im folgenden Abschnitt genauer betrachten.

Selbstverständlich kann man Bauteile mit mehr als zwei Anschlüssen herstellen. Das gilt insbesondere für Transistoren mit drei Anschlüssen und für integrierte Schaltungen mit nahezu beliebig vielen Anschlüssen. Die Charakterisierung und Berechnung von derartigen Mehrpolen ist natürlich viel schwieriger als die Behandlung von Zweipolen. Es gelingt meist, für Mehrpole Ersatzschaltungen aus mehreren Zweipolen anzugeben, welche die Mehrpole in guter Näherung beschreiben. Die Behandlung von Zweipolen ist deshalb zentral.

Sind die Bauteile charakterisiert, so muss untersucht werden, wie sich verschiedene Bauteile zu elektrischen Schaltungen verknüpfen lassen, welche ein gewünschtes Verhalten aufweisen. Um derartige Schaltungen bzw. Netzwerke analysieren zu können, werden wir in diesem Kapitel die wichtigsten Werkzeuge vorstellen.

Da wir nicht nur im Raum, sondern auch in der Zeit leben, ist natürlich das Zeitverhalten elektrischer Schaltungen von besonderem Interesse. Trotzdem beschränken wir und in diesem Kapitel zunächst auf den scheinbar uninteressanten Gleichstromfall, d.h., wir nehmen an dass sich Ströme und Spannungen in der betrachteten Schaltung nicht ändern. Dies führt zu einer enormen Vereinfachung, weil wir damit jeden Strom und jede Spannung mit einer einzigen, reellen Zahl vollständig beschreiben können, während im zeitveränderlichen Fall eine viel aufwendigere Beschreibung mit Zeitfunktionen erforderlich wäre.

Die Betrachtung des Gleichstromverhaltens einer Schaltung bringt nicht nur eine grosse Vereinfachung, sondern in vielen Fällen auch eine nützliche Übersicht. Bei langsamen Veränderungen mag eine Betrachtung des Gleichstromverhaltens für das Verständnis einer Schaltung völlig ausreichen, aber auch bei rasch veränderlichen Strömen wird den interessierenden Strompulsen oder Signalen oft ein Gleichstromanteil überlagert, welcher dazu dient, den Arbeitspunkt einzelner Bauelemente einzustellen. Die meisten elektronischen Schaltungen weisen eine Gleichspannungsquelle auf, welche die Schaltung mit elektrischer Energie versorgen. Der erste Schritt bei der Analyse solcher Schaltungen ist also immer die Betrachtung des Gleichstromverhaltens, selbst wenn die interessierenden Signale sehr rasch veränderlich sind. Diese Signale werden erst in einem zweiten Schritt betrachtet, wobei man vom Überlagerungsprinzip profitiert, von dem bereits im vorhergehenden Kapitel die Rede war.

2.1. Zweipole

Die makroskopische Betrachtungsweise der Netzwerktheorie kümmert sich nicht um das Innenleben, bzw. die inneren physikalischen Effekte, sondern nur um das, was von aussen sichtbar bzw. messbar ist. Da wir uns zunächst auf die Ströme und Spannungen konzentrieren, charakterisieren wir einen Zweipol durch die Spannung zwischen den beiden An-

schlüssen und die Ströme welche beim einen Anschluss in den Zweipol hinein und beim andern Anschluss herausfließen. Wie beim idealisierten Draht, nehmen wir auch hier an, der Zweipol könne keine Ladungen speichern und die Ströme an den beiden Anschlüssen seien deshalb identisch. Damit lässt sich ein Zweipol durch die Abhängigkeit des Stromes von der angelegten Spannung oder umgekehrt durch die Abhängigkeit der Spannung vom Strom vollständig charakterisieren. Wichtig ist, dass zu jedem Zweipol ein Strompfeil und ein Spannungspfeil angegeben werden, welche die entsprechenden Zählrichtungen angeben.

2.1.1 Pfeilsysteme und Verlustleistung

Meist wird das sogenannte Verbraucherpfeilsystem verwendet, bei welcher Strom- und Spannungspfeil in dieselbe Richtung zeigen. Der Sinn dieser Konvention wird klar, wenn wir die im Zweipol umgesetzte Leistung

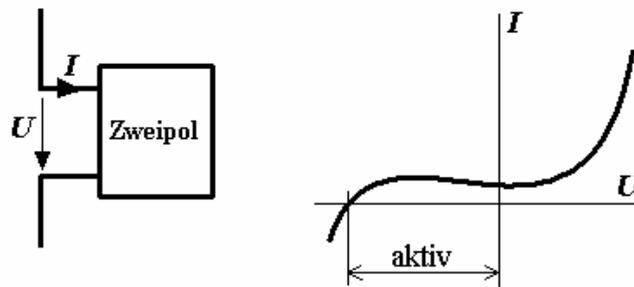
$$P = UI$$

betrachten. Haben U und I — unter Verwendung des Verbraucherpfeilsystems — dasselbe Vorzeichen, so wird die Verlustleistung P positiv, d.h. der Zweipol entzieht der elektrischen Schaltung Energie (welche von mindestens einer Quelle oder einem Generator eingespiessen wird). Man sagt, der Zweipol verbrauche Energie.

Wegen der Energieerhaltung wird diese Energie übrigens nicht vernichtet, sondern in eine andere Form umgewandelt. Meist ist dies thermische Energie, welche zu einer Erhitzung des Bauelementes führt. Dieser „Energieverbrauch“ von Zweipolen ist nicht immer unerwünscht. Er wird insbesondere bei Wandlern angestrebt. Das gilt bei elektrischen Heizelementen, aber auch bei Glühlampen, Elektromotoren (elektro-mechanische Wandler), etc.

Natürlich gibt es auch Wandler, welche andere Formen von Energie in elektrische Energie umwandeln, also elektrische Energie erzeugen. Beispiele sind Generatoren (mechanisch-elektrische Wandler) und Batterien (chemisch-elektrische Wandler). Bei derartigen Wandlern wird die Verlustleistung P negativ. Man verwendet dann manchmal das Erzeugerpfeilsystem mit entgegengesetztem Strom- und Spannungspfeil. In diesem Pfeilsystem wird $P = UI$ für Erzeuger positiv und für Verbraucher negativ, bezeichnet dann also die erzeugte, d.h. dem elektrischen System zugeführte Leistung.

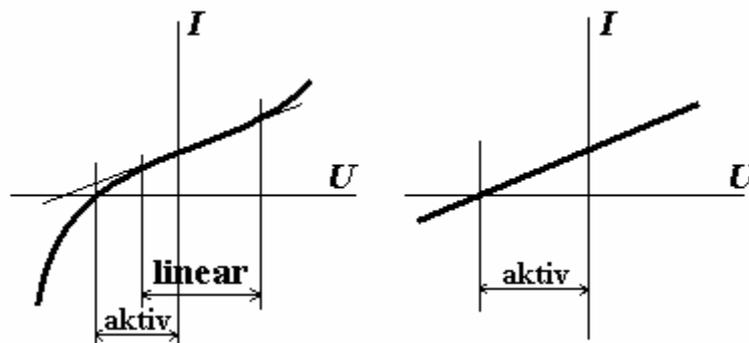
Im Allgemeinen können Zweipole aktive und passive Bereiche aufweisen, d.h. es existieren Bereiche auf der Kennlinie wo elektrische Energie erzeugt wird und Bereiche wo elektrische Energie verbraucht wird.



Figur 2: Zweipol mit Strom und Spannungspfeilen (Verbraucherpeilsystem) und Kennlinie. Die Kennlinie weist einen aktiven Bereich auf, in dem Strom- und Spannungsrichtung entgegengesetzt sind. In diesem Bereich gibt der Zweipol Energie ab. Ausserhalb des aktiven Bereichs ist der Zweipol passiv, d.h. er nimmt Energie auf.

2.1.2. Kennlinien

Die Funktion $I(U)$ bzw. $U(I)$ lässt sich einfach graphisch darstellen und wird Kennlinie genannt. Je nach den physikalischen Effekten, welche in den Zweipolen eine Rolle spielen, können sehr komplizierte Kennlinien entstehen, welche ausserdem von andern physikalischen Grössen, z.B. der Temperatur, beeinflusst werden. Eine Beeinflussung durch andere physikalische Grössen ist in den elektrischen Schaltungen meist unerwünscht. Man sucht deshalb meist nach Bauteilen, bei welchen äussere Einflüsse vernachlässigbar sind. Eine wichtige Ausnahme sind Sensoren, mit denen andere physikalische Grössen als Strom und Spannung elektronisch gemessen werden können. Im Folgenden vernachlässigen wir nicht elektrische Einflüsse und nehmen an, dass die Spannungs-Strom-Kennlinie den Zweipol vollständig beschreibt.



Figur 3: Links: Nichtlineare Kennlinie mit aktiven und passiven Bereichen und einem nahezu linearen Bereich. Rechts: Idealisierte, lineare Kennlinie mit aktiven und passiven Bereichen.

Bei allgemeinen Kennlinien besteht ein nichtlinearer Zusammenhang zwischen Strom und Spannung. Dieser führt bei der Netzwerkbeschreibung auf nichtlineare Gleichungssysteme.

Für Informatiker ist sicher klar, dass die Lösung nichtlinearer Gleichungssysteme viel schwieriger ist, als die Lösung linearer Gleichungssysteme. Um schliesslich zu solchen Gleichungssystemen zu gelangen, werden häufig Bauteile verwendet, welche nahezu lineare Kennlinien aufweisen. Sind nichtlineare Bauteile vorhanden, so versucht man in einem ersten Schritt, den Arbeitspunkt, d.h. den Punkt auf der Kennlinie in dem der Zweipol betrieben wird, zu erraten. Dann wird die Kennlinie durch eine linearisierte Kennlinie, d.h. eine Tangente im Arbeitspunkt, ersetzt und mit der linearisierten Kennlinie weitergerechnet. Ergibt die Rechnung einen vom geschätzten Arbeitspunkt deutlich abweichenden Punkt, so wird der Arbeitspunkt modifiziert, die Kennlinie angepasst und die Prozedur so lange wiederholt, bis die Berechnung zum tatsächlichen Arbeitspunkt konvergiert. Neben iterativen, numerischen Verfahren zur Berechnung nichtlinearer Bauelemente gibt es auch graphische Verfahren und analytische Verfahren, welche aber nur in einfachen Schaltungen zum Erfolg führen.

2.1.3. Linearisierung

Eine linearisierte Kennlinie lässt sich — wie jede Gerade — durch zwei Punkte definieren. Betrachten wir eine beliebige, lineare Funktion $U(I)$, so schneidet diese normalerweise sowohl die U als auch die I Achse. Diese Schnittpunkte sind offenbar prominente Punkte auf der Kennlinie. Da im Spezialfall einer der Schnittpunkte im Unendlichen liegen kann, ist es vielleicht sinnvoller, die Kennlinie durch einen Schnittpunkt und die Steigung zu charakterisieren. Man schreibt dann

$$U(I) = U_0 + RI \text{ oder } I(U) = I_0 + GU.$$

Dabei bezeichnet R den Ohm'schen Widerstand und G den Leitwert. Es gilt offensichtlich

$$G = 1/R.$$

Für die Masseinheiten von R bzw. G erhält man 1Ω (Ohm) = 1 V/A , bzw. 1 S (Siemens) = 1 A/V .

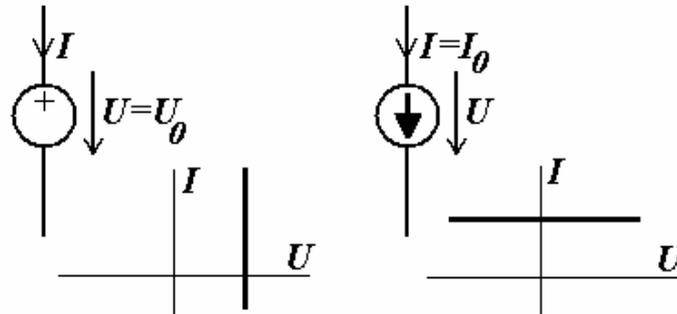
Sind R und G ungleich Null, so können beide Darstellungen verwendet und nach Belieben ausgetauscht werden. Für die Spezialfälle $R = 0$ und $G = 0$ kommt hingegen nur eine Darstellung in Frage, die aber besonders einfach ist.

2.1.4. Ideale Strom- und Spannungsquellen

Für $R = 0$ wird offenbar $U(I) = U_0 = \text{konstant}$. Man nennt einen derartigen Zweipol ideale Spannungsquelle. Analog ergibt sich eine ideale Stromquelle mit $I(U) = I_0$, wenn $G = 0$ ist. Offenbar genügt zur Charakterisierung einer idealen Quelle eine einzige, reelle Zahl.

Ideale Quellen existieren in der Praxis natürlich nicht, werden aber mit elektronischen Schaltungen innerhalb eines begrenzten Arbeitsbereiches sehr gut nachgeahmt.

Für ideale Strom- und Spannungsquellen werden spezielle Symbole verwendet.



Figur 4: Links: Ideale Spannungsquelle mit Kennlinie, rechts: Ideale Stromquelle mit Kennlinie. Strom- und Spannungspfeile gemäss Verbraucherpfeilsystem. Die Spannungsquelle ist deshalb unterhalb der U Achse (negative Werte von I) aktiv; die Stromquelle ist links der I Achse (negative Werte von U) aktiv.

Zu beachten ist, dass auch für ideale Strom- und Spannungsquellen passive Bereiche existieren, wo Energie verbraucht wird. Quellen müssen also nicht immer Energie in ein elektrisches Netzwerk einspeisen.

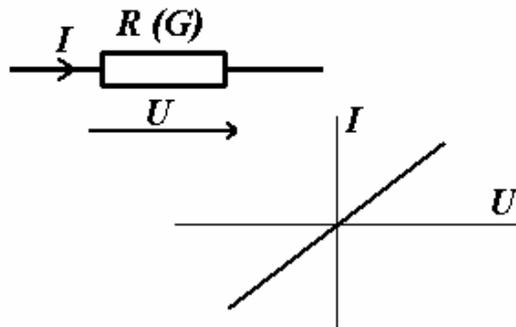
Einen Spezialfall des Spezialfalls erhält man übrigens mit $R = 0$ und $U_0 = 0$. Dies ist nichts anderes als ein idealer Draht. Was für ein Bauteil ist durch $G = 0$ und $I_0 = 0$ charakterisiert?

2.1.5. Der Ohm'sche Widerstand

Ein anders gearteter Spezialfall eines linearen Zweipols findet man, wenn die Quellterme U_0 und I_0 verschwinden. Dann gilt

$$U = RI = I/G \text{ und } I = GU = U/R.$$

Für den damit beschriebenen Ohm'schen Widerstand verwendet man ein einfaches Rechteck als Symbol.



Figur 5: Widerstand mit Kennlinie, Verbraucherpfilsystem. Der Widerstand ist überall passiv; es existiert kein aktiver Bereich.

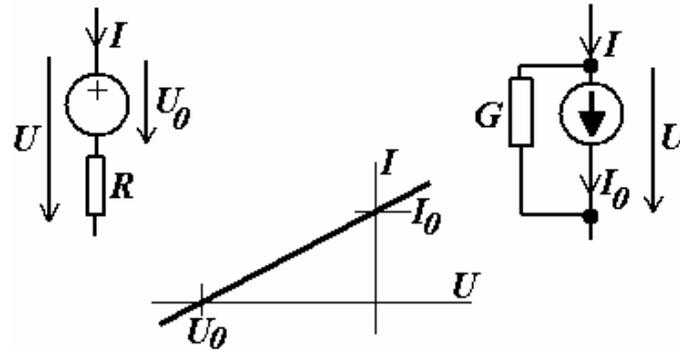
Ohm'sche Widerstände haben offensichtlich keinen aktiven Bereich, d.h. sie nehmen immer elektrische Energie auf, sind also ideale Verbraucher.

Im Spezialfall $R = 0$ erhält man aus einem Ohm'schen Widerstand einen idealen Draht.

Wie für die idealen Quellen genügt zur Charakterisierung eines Ohm'schen Widerstandes eine einzige, reelle Zahl. Damit kennen wir bereits die elementarsten Zweipole für Gleichstromnetzwerke.

2.1.6. Verlustbehaftete Quellen

Ein allgemeiner, linearer Zweipol kann als Kombination einer idealen Quelle mit einem Ohm'schen Widerstand bzw. Leitwert aufgefasst werden. Offensichtlich gibt es die folgenden zwei Möglichkeiten: 1) Die Formel $U(I) = U_0 + RI$ wird aufgefasst als Serieschaltung einer idealen Spannungsquelle mit einem Widerstand R . Da bei realen Quellen R von außen nicht sichtbar ist, nennt man R auch Innenwiderstand. 2) Die Formel $I(U) = I_0 + GU$ wird aufgefasst als Parallelschaltung einer idealen Stromquelle mit einem Widerstand R . Serie- und Parallelschaltungen von Zweipolen werden später genauer untersucht.



Figur 6: Verlustbehaftete Spannungs- und Stromquelle mit Verbraucherpfleilsystem und Kennlinie. Die Kennlinien beider Quellen sind formal identisch. Gilt $R = 1/G = U_0/I_0$, so sind die beiden Quellen von aussen ununterscheidbar. Im Bereich $U_0 < U < 0$ sind die Quellen aktiv. Im verwendeten Pfeilsystem ist die Leerlaufspannung U_0 der Spannungsquelle negativ. Üblicherweise würde man in diesem Falle die Pfeilrichtungen umdrehen, so dass U_0 positiv wird.

2.1.7. Quellenumwandlung

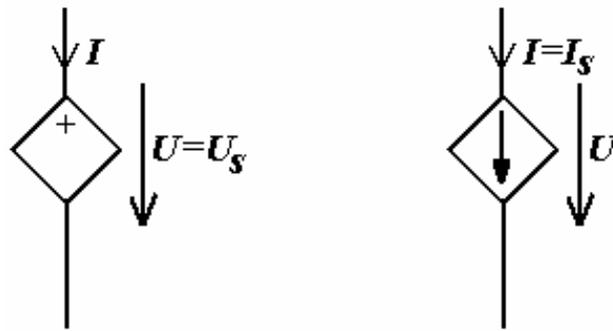
Da $U(I) = U_0 + RI$ und $I(U) = I_0 + GU$ ein und dieselbe lineare Kennlinie beschreiben, ist klar, dass die verlustbehaftete Spannungsquelle und die verlustbehaftete Stromquelle denselben Zweipol beschreiben, dass also die obigen Serie- und Parallelschaltungen ineinander umgewandelt werden können, bzw. von aussen ununterscheidbar sind. Dabei gelten folgende Zusammenhänge:

$$R = 1/G \text{ und } U_0 = RI_0.$$

Ob für eine reale Quelle eine Strom- oder Spannungsquellenersatzschaltung bevorzugt wird, hängt von der konkreten Schaltung ab, in der sich die Quelle befindet. Mit Hilfe der Quellenumwandlung gelingt es oft, bestehende Schaltungen formal zu vereinfachen. Die Quellenumwandlung ist deshalb ein nützliches Instrument der Schaltungsanalyse.

2.1.8. Gesteuerte Quellen

Bei gesteuerten Quellen sind Quellenspannung bzw. Quellenstrom von einer äusseren Stellgrösse abhängig. Für gesteuerte Quellen werden folgende Symbole verwendet:



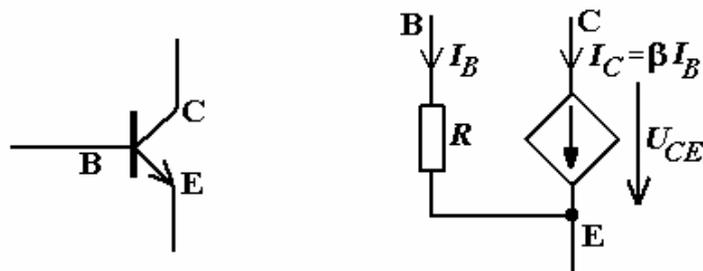
Figur 7: Ideale, gesteuerte Quellen mit Verbraucherpfilsystem. Links: Spannungsquelle, rechts: Stromquelle.

Die Steuergrößen können Ströme oder Spannungen sein, welche an irgendeiner Stelle der elektronischen Schaltung abgenommen werden. Es lassen sich damit folgende Fälle unterscheiden:

- Stromgesteuerte Stromquelle ($I_0 = \beta I_x$)
- Spannungsgesteuerte Stromquelle ($I_0 = \alpha U_x$)
- Stromgesteuerte Spannungsquelle ($U_0 = \rho I_x$)
- Spannungsgesteuerte Spannungsquelle ($U_0 = \mu U_x$)

Die Annahme, dass die Quellengröße (U_0 bzw. I_0) der Stellgröße (U_x bzw. I_x) proportional sei, ist natürlich i.a. eine Vereinfachung. Diese Linearisierung ist aber in vielen Fällen ausreichend genau.

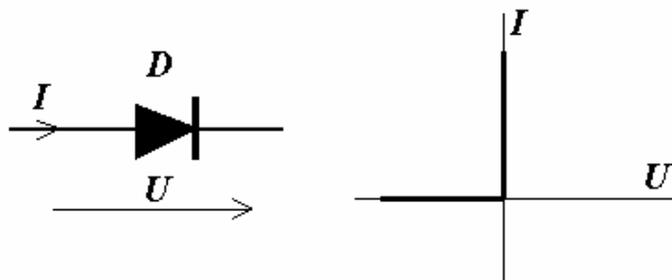
Genau genommen sind gesteuerte Quellen keine Zweipole. Praktische Realisierungen von gesteuerten Quellen brauchen mindestens einen zusätzlichen Pol, an welchem die Stellgröße abgenommen werden kann. Prominentes Beispiel ist der Transistor, für den es verschiedene Ersatzschaltungen mit gesteuerten Quellen gibt. Dadurch gelingt eine erhebliche Vereinfachung der Schaltungsanalyse.



Figur 8: Symbol (links) und einfaches, stark idealisiertes Ersatzschaltbild mit stromgesteuerter Stromquelle (rechts) für NPN Transistor mit den drei Anschlüssen Basis (B), Collector (C) und Emittter (E). Dieses Ersatzschaltbild gilt nur sehr begrenzt, für das sogenannte Kleinsignalverhalten. Für genauere Rechnungen werden kompliziertere Ersatzschaltungen verwendet, welche reale Transistoren genauer modellieren.

2.1.9. Ideale Dioden

Dioden sind nichtlineare Halbleiterbauteile, welche in erster Näherung Strom in eine Richtung gut leiten, in die umgekehrte Richtung hingegen nicht. Idealisiert man dieses Verhalten, so erhält man eine ideale Diode, welche Strom in einer Richtung widerstandslos leitet und in der umgekehrten Richtung vollständig sperrt. Die Kennlinie einer idealen Diode besteht also aus zwei linearen Bereichen. Im einen leitenden Bereich wird die Diode durch einen Widerstand $R = 0$ und im Sperrbereich durch $G = 0$ beschrieben. Dioden können deshalb als eine spezielle Art von Schaltern, sogenannten Gleichrichtern betrachtet werden. Auch für Dioden wird ein spezielles Symbol verwendet.



Figur 9: Symbol mit Verbraucherpeilsystem und Kennlinie einer idealen Diode.

Zur Behandlung von Dioden versucht man meist zu erraten, ob die Diode im Sperr- oder Durchgangsbereich arbeitet und ersetzt dann die Diode durch $G = 0$ (Strom durch die Diode = 0) bzw. $R = 0$ (Spannung über der Diode = 0).

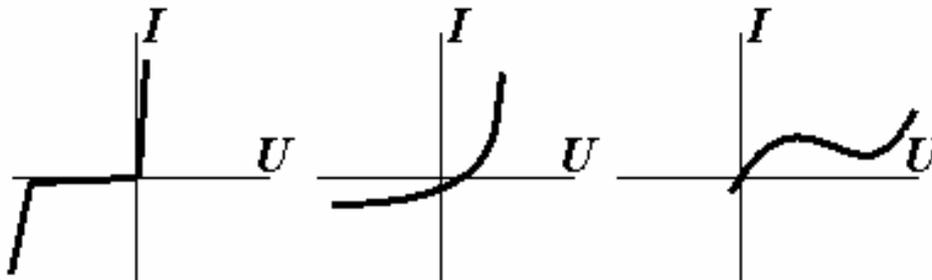
2.1.10. Reale Dioden

Je nach Bauart und Dotierung der Halbleiter erhält man viele unterschiedliche Dioden mit recht unterschiedlichen Kennlinien, welche von den idealen Dioden zum Teil stark abweichen.

Zunächst ist zu beachten, dass reale Dioden in Durchlassrichtung nicht ideal leiten. Insbesondere bei kleinen Spannungen unterhalb einem bestimmten Wert U_D (bei Si Dioden ca. 0.6V) ist der Durchlassstrom nahezu Null. Eine verbesserte Ersatzschaltung einer Halbleiterdiode im Durchlassbereich besteht deshalb aus einer Spannungsquelle mit Innenwiderstand, wobei der Innenwiderstand bei vielen Anwendungen vernachlässigbar klein ist.

Im Sperrbereich sind viele Dioden bei nicht zu hohen Spannungen nahezu ideal, d.h. der Sperrstrom ist meist vernachlässigbar klein. Wird eine Spannung $-U_Z$ unterschritten, so beginnen Halbleiterdioden zu leiten, d.h. der Sperrstrom wächst rasant. Dieser Durchbrucheffekt wird Zener effekt genannt und bei Zenerdioden bewusst ausgenutzt. U_Z ist die Zenerspannung. Unterhalb der Zenerspannung kann die Diode — ähnlich wie im Durchlassbereich — gut durch eine Spannungsquellenersatzschaltung beschrieben werden. Der Wert der Zenerspannung hängt von der Bauart der Diode ab. Da die Verlustleistung in diesem Bereich mit der angelegten Spannung $-U$ stark wächst, besteht natürlich die Gefahr, dass die Diode irreversibel geschädigt wird.

Weitere Abarten sind Photodioden und Tunnelioden.



Figur 10: Kennlinien von Zener-, Photo- und Tunnelioden im Verbraucherpeilsystem. Die Kennlinie der Photodiode ist abhängig vom einfallenden Licht und weist einen aktiven Bereich (im 4. Quadranten) auf der mit der Lichtstärke anwächst. Die Tunneliode weist einen rückläufigen Bereich auf. Wird dieser Bereich linearisiert, so erhält man eine Ersatzschaltung mit negativem Innenwiderstand. Trotzdem ist die Tunneliode auch in diesem Bereich passiv.

Wie Figur 10 zeigt, kann eine Zenerdiode vereinfacht durch drei lineare Bereiche charakterisiert werden. In der Praxis werden Zenerdioden vor allem im Bereich unterhalb der Zenerspannung betrieben und zur Spannungsstabilisierung bzw. zur Erzeugung von Referenzspannungen verwendet.

Tunnelioden sind wie die meisten Dioden im gesamten Kennlinienbereich passiv, weisen aber einen „rückläufigen“ Abschnitt auf. Linearisiert man diesen Bereich, so erhält man

eine Quellenersatzschaltung mit negativem Innenwiderstand. Dadurch wird die Tunneldiode für die Herstellung von Kompensationsschaltungen und von Oszillatoren interessant.

Photodioden sind lichtempfindliche Dioden. Im Prinzip sind alle Dioden lichtempfindlich. Die Lichtempfindlichkeit wird aber meist technisch durch Einpacken in lichtundurchlässige Gehäuse unterdrückt. Je nach Stärke des Lichteinfalls wird die Diodenkennlinie verschoben. Dadurch entsteht ein aktiver Bereich. Wird die Diode in diesem Bereich betrieben, so speist sie Energie in die elektronische Schaltung ein. Photodioden können deshalb als Detektoren für Licht verwendet werden.

2.2. Mehrpole

Die Behandlung von Mehrpolen ist um einiges schwieriger als die Behandlung von Zweipolen. Es wird daher gerne versucht, Mehrpole durch Ersatzschaltungen bestehend aus mehreren Zweipolen anzunähern und lediglich die Ersatzschaltungen zu analysieren. Ein typisches und besonders wichtiges Beispiel ist der Transistor (siehe Figur 8). Wir verzichten an dieser Stelle auf eine explizite Behandlung von Mehrpolen und nehmen im Folgenden an, dass Mehrpole durch Ersatzschaltungen bestehend aus Zweipolen angenähert beschrieben werden können.

2.3. Knotengleichungen

Bisher haben wir einzelne Bauelemente isoliert betrachtet. Nun wollen wir verschiedene Bauelemente miteinander verknüpfen, d.h. die entsprechenden Anschlussdrähte miteinander verbinden. Dabei idealisieren wir die Anschlussdrähte und beschreiben sie durch Fadenströme. Durch die Verknüpfung von K Fadenströmen I_k entsteht ein *Knoten*. Praktisch nimmt ein Knoten ein gewisses Volumen V ein, in welchem elektrische Ladungen enthalten sein können. Für die Gesamtladung des Knotens gilt natürlich der Ladungserhaltungssatz aus Abschnitt 1.4. Nun idealisieren wir den Knoten ebenso wie die Anschlussdrähte der Bauelemente indem wir annehmen, das Volumen des Knotens sei so klein, dass die Knotenladung Q stets vernachlässigbar klein sei. Damit folgt aus dem Ladungserhaltungssatz sofort die sogenannte Knotengleichung:

$$\sum_k I_k = 0.$$

D.h. die Summe der Fadenströme, welche aus einem Knoten heraus fließen, verschwindet. Ebenso verschwindet natürlich auch die Summe der Fadenströme, welche in einen Knoten hinein fließen. Wichtig ist einzig die einheitliche Orientierung der Ströme. Praktisch werden immer einzelne Ströme in den Knoten hinein und andere aus dem Knoten heraus fließen. Erstere erhalten üblicherweise ein positives und letztere ein negatives Vorzeichen. Figur 1 zeigt einen Knoten mit vier Strömen und nach aussen gerichteten Strompfeilen. Für diesen Knoten gilt die Knotengleichung $I_1 + I_2 + I_3 + I_4 = 0$

Theoretisch ist die Anzahl K der Fadenströme für den einfachsten Knoten 1. Es folgt dann sofort, dass der Fadenstrom I_1 verschwindet. Fließt I_1 durch einen Zweipol, so verhält sich die entsprechende Schaltung genau gleich, wie wenn der entsprechende Zweipol weggelassen würde. Praktisch ist deshalb $K > 1$.

Für $K = 2$ erhält man sofort $I_1 = -I_2$. In diesem Falle wird die Knotengleichung häufig nicht explizit aufgeschrieben, sondern lediglich ein einziger Fadenstromwert in die Schaltung eingezeichnet.

Da elektrische Ströme nur in geschlossenen Kreisen fließen können, bestehen alle interessanten elektrischen Schaltungen aus mindestens zwei Knoten.

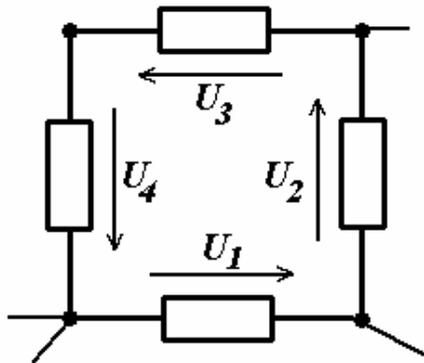
2.4. Maschengleichungen

Es stellt sich die Frage, ob die Aufstellung aller Knotengleichungen und aller Gleichungen, welche die Bauelemente beschreiben, ausreichen um die Ströme in einem Netzwerk zu berechnen. Betrachten wir das einfachste Netzwerk aus zwei miteinander verbundenen Zweipolen, so notieren wir für dieses zwei Zweipolgleichungen $U_1 = f_1(I_1)$, $U_2 = f_2(I_2)$ und zwei identische Knotengleichungen $I_1 = -I_2$. Dabei kann das Vorzeichen in den Knotengleichungen je nach Definition der Stromzählrichtung unterschiedlich sein. Offenbar gelingt mit diesen Gleichungen die Berechnung der Fadenströme nicht, weil wir für die Berechnung der vier Unbekannten U_1 , I_1 , U_2 , I_2 nur drei unterschiedliche Gleichungen haben. Wir suchen deshalb nach einem physikalischen Gesetz, welches eine weitere Einschränkung für elektrische Schaltungen liefert. Einem solchen Gesetz sind wir bereits in Abschnitt 1.2 begegnet: Das elektrische Feld ist wirbelfrei.

Um die Wirbelfreiheit ausnutzen zu können, suchen wir in unseren elektrischen Schaltungen nach geschlossenen Wegen. Diese werden Maschen genannt. Eine Masche besteht aus einer (geschlossenen) Aneinanderreihung von K Bauelementen. Das Umlaufintegral über das elektrische Feld gemäss Abschnitt 1.2 wird dadurch in K Abschnitte zerlegt. Das Integral über das elektrische Feld über jeden einzelnen Abschnitt, d.h. über jedes einzelne Bauelement ist offenbar gleich der Spannung über dem betreffenden Bauelement. Aus der Wirbelfreiheit des elektrischen Feldes folgt dann die Maschengleichung:

$$\sum_{k=1}^K U_k = 0.$$

Ähnlich wie bei der Knotengleichung ist hier die *Orientierung* der Maschenspannungen wichtig: Man kann entweder die Spannungen im Uhrzeigersinn oder im Gegenuhrzeigersinn positiv zählen.



Figur 11: Masche in einem Netzwerk mit Maschenspannungen im Gegenurzeigersinn. Maschengleichung für diese Masche: $U_1 + U_2 + U_3 + U_4 = 0$

Der einfachste Fall einer Masche mit $K = 1$ Maschenspannungen entspricht einem „kurzgeschlossenen“ Bauelement mit $U_1 = 0$ und hat praktisch dieselbe (geringe) Bedeutung wie ein Knoten mit $K = 1$. Für $K = 2$ erhält man die einfachste sinnvolle Masche: $U_1 = -U_2$. Auch für diesen Fall lässt sich sinngemäss dasselbe sagen wie für einen Knoten mit $K = 2$.

Kehren wir zu unserm einfachen Netzwerk mit zwei Zweipolen zurück, so liefert die Maschengleichung $U_1 = -U_2$ die fehlende vierte Gleichung.

2.5. Vollständigkeit, Linearität und Abhängigkeiten

Zusammen mit den Gleichungen, welche die Bauelemente beschreiben, erhält man aus den Knoten- und Maschengleichungen immer ausreichend viele Gleichungen um sämtliche Ströme und Spannungen in einem Netzwerk berechnen zu können. Das System bestehend aus Bauelementgleichungen, Knotengleichungen und Maschengleichungen ist also vollständig.

Die Knoten- und Maschengleichungen sind offensichtlich lineare Gleichungen für Ströme bzw. Spannungen. Besteht ein Netzwerk aus linearen Bauelementen (mit linearen Bauelementgleichungen), so wird es durch ein lineares Gleichungssystem beschrieben, welches bequemerweise in Matrizenform notiert und mit passenden Algorithmen der linearen Algebra gelöst werden kann.

Wie bei den Knotengleichungen können auch bei den Maschengleichungen (in Netzwerken mit mehreren Maschen) Abhängigkeiten auftreten, d.h. die entsprechenden Gleichungen können voneinander linear abhängig sein. Die Kunst der Netzwerkberechnung besteht also darin, sämtliche voneinander unabhängigen Knoten- und Maschengleichungen zu notieren, bzw. abhängige Gleichungen wegzulassen. Wir werden in den Abschnitten 2.11 und 2.12 zwei Methoden für Netzwerke bestehend aus linearen Zweipolen skizzieren, welche voll-

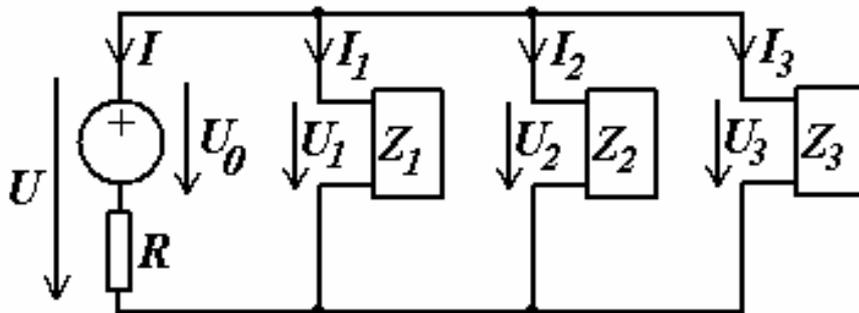
ständige Gleichungssysteme liefern. Diese Methoden werden gerne für Computerprogramme verwendet.

Wird ein Netzwerk von Hand berechnet, so ist es vorteilhaft, mit den einfachsten unabhängigen Knoten- und Maschengleichungen zu arbeiten. Diese in komplizierten Netzwerken aufzufinden, ist eine Sache der Erfahrung und Intuition.

Da der Schwierigkeitsgrad der manuellen Netzwerkberechnungen mit der Anzahl der Bauelemente stark wächst, gibt es verschiedene Methoden um Netzwerke zu vereinfachen. Die wichtigsten davon werden in den folgenden Abschnitten skizziert. Dabei beschränken wir uns der Einfachheit halber auf Netzwerke von Zweipolen.

2.6. Parallelschaltungen von Zweipolen

In komplizierteren Netzwerken findet man oft Teile, welche aus einer Parallelschaltung von zwei oder mehreren Zweipolen bestehen. Die folgende Figur zeigt eine einfache Schaltung mit einer Parallelschaltung von drei Zweipolen:



Figur 12: Parallelschaltung von drei Zweipolen Z_1, Z_2, Z_3 an einer Spannungsquelle U_0 mit Innenwiderstand R . Beim verwendeten Pfeilsystem gilt: $U_1 = U_2 = U_3 = U = U_0 + RI$, $I = -(I_1 + I_2 + I_3)$.

Offenbar ist die Spannung über allen parallel geschalteten Zweipolen gleich, d.h. es gilt $U_1 = U_2 = \dots = U_K$, während die Ströme i.a. unterschiedlich sind.

Durch die Parallelschaltung entsteht ein neuer Zweipol mit $U_{parallel} = U_1 = U_2 = \dots = U_K$ und

$$I_{parallel} = \sum_{k=1}^K I_k .$$

Dabei ist zu beachten, dass die Zählrichtungen der Ströme in dieselbe Richtung zeigen.

Sind alle parallel geschalteten Zweipole linear, so ist auch der resultierende Zweipol linear. Gilt insbesondere $I_k = I_{k0} + G_k U_k$, so folgt

$$I_{parallel} = \sum_k I_k = \sum_k I_{k0} + \sum_k G_k U_{parallel} = I_{parallel0} + G_{parallel} U_{parallel}$$

Als Spezialfälle hat man dabei die Parallelschaltung von idealen Stromquellen mit

$$I_{parallel0} = \sum_k I_{k0}$$

und die Parallelschaltung von Leitwerten mit

$$G_{parallel} = \sum_k G_k \cdot$$

Diese Parallelschaltung wird auch *Stromteiler* genannt, wobei sich der Gesamtstrom $I_{parallel}$ in die „Teilströme“

$$I_k = G_k U_{parallel} = G_k I_{parallel} / G_{parallel} = I_{parallel} \frac{G_k}{\sum_{i=1}^K G_i}$$

aufteilt.

Eine Parallelschaltung idealer Spannungsquellen mit unterschiedlichen Quellenspannungen ist wegen $U_{parallel} = U_1 = U_2 = \dots = U_K$ verboten.

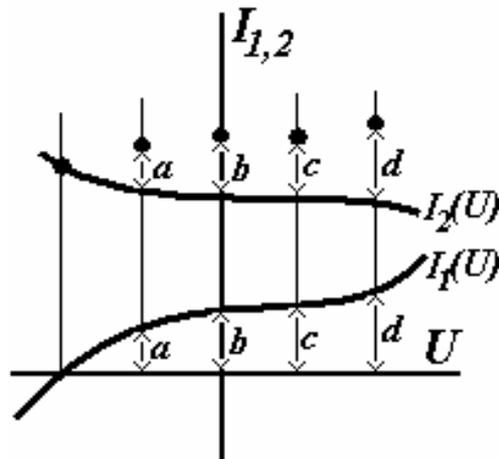
Für parallel geschaltete Widerstände gilt wegen $R = 1/G$:

$$1/R_{parallel} = \sum_k 1/R_k, \quad R_{parallel} = 1 / \left(\sum_k 1/R_k \right).$$

Werden nichtlineare Zweipole parallel geschaltet, so entsteht ein nichtlinearer Zweipol, für den ebenso wie bei linearen Zweipolen $U_{parallel} = U_1 = U_2 = \dots = U_K$ und $I_{parallel} = \sum_k I_k$ gilt. Sind die einzelnen Kennlinien analytisch durch Formeln der Form $I_k = g_k(U_k)$ gegeben, so ergibt sich einfach

$$I_{parallel} = \sum_k I_k = \sum_k g_k(U_{parallel}).$$

Sind die Kennlinien graphisch gegeben, so zeichnet man sie in dieselbe Figur ein und addiert die Ströme der einzelnen Zweipole graphisch.

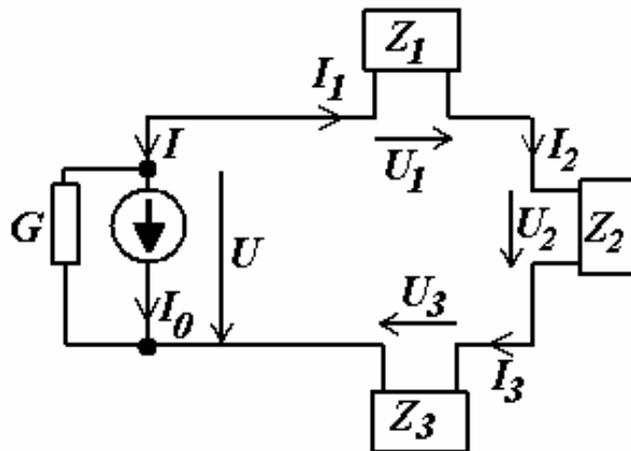


Figur 13: Graphische Ermittlung der Kennlinie einer Parallelschaltung von zwei nichtlinearen Zweipolen mit den Kennlinien $I_1(U_1)$ und $I_2(U_2)$. Die Kennlinie $I(U)$ der Parallelschaltung ist durch Punkte markiert. Sie ergibt sich aus $U = U_1 = U_2$ und $I = I_1 + I_2$. Prinzip: Addition der Ströme (vertikale Richtung in der verwendeten Darstellung).

In einem Netzwerk, welches parallel geschaltete Zweipole enthält, ersetzt man die parallel geschalteten Zweipole und erhält damit ein vereinfachtes Netzwerk mit weniger Zweipolen. Sind die Spannungen im vereinfachten Netzwerk berechnet, so lassen sich die einzelnen Teilströme der parallel geschalteten Zweipole aus den entsprechenden Kennlinien ablesen oder aus den entsprechenden Zweipolgleichungen bestimmen.

2.7. Serieschaltungen von Zweipolen

Serieschaltungen von zwei oder mehreren Zweipolen sind dual zu den Parallelschaltungen. Man kann deshalb die im vorhergehenden Abschnitt gemachten Aussagen übernehmen und Ströme mit Spannungen, Leitwerte mit Widerständen vertauschen. Die folgende Figur zeigt eine einfache Schaltung mit einer Serieschaltung von drei Zweipolen:



Figur 14: Serieschaltung von drei Zweipolen Z_1, Z_2, Z_3 an Stromquelle I_0 mit Innenleitwert G . Im verwendeten Pfeilsystem gilt: $I_1 = I_2 = I_3 = -I = -(I_0 + GU)$, $U_1 + U_2 + U_3 = U$.

Offenbar ist der Strom, welcher durch die in Serie geschalteten Zweipole fließt gleich, d.h. es gilt $I_1 = I_2 = \dots = I_K$, während die Spannungen i.a. unterschiedlich sind.

Durch die Serieschaltung entsteht ein neuer Zweipol mit $I_{parallel} = I_1 = I_2 = \dots = I_K$ und

$$U_{parallel} = \sum_k U_k .$$

Dabei ist zu beachten, dass die Zählrichtungen der Spannungen in dieselbe Richtung zeigen.

Sind alle in Serie geschalteten Zweipole linear, so ist auch der resultierende Zweipol linear. Gilt insbesondere $U_k = U_{k0} + R_k I_k$, so folgt

$$U_{serie} = \sum_k U_k = \sum_k U_{k0} + \sum_k R_k I_{serie} = I_{serie0} + R_{serie} I_{serie} .$$

Als Spezialfälle hat man dabei die Serieschaltung von idealen Spannungsquellen mit

$$U_{serie0} = \sum_k U_{k0}$$

und die Serieschaltung von Widerständen mit

$$R_{serie} = \sum_k R_k .$$

Diese Serieschaltung wird auch *Spannungsteiler* genannt, wobei sich die Gesamtspannung U_{pserie} in die „Teilspannungen“

$$U_k = R_k I_{serie} = R_k U_{serie} / R_{serie} = U_{serie} \frac{R_k}{\sum_{i=1}^K R_i}$$

aufteilt.

Eine Serieschaltung idealer Stromquellen mit unterschiedlichen Quellenströmen ist wegen $I_{pserie} = I_1 = I_2 = \dots = I_K$ verboten.

Für in Serie geschaltete Leitwerte gilt wegen $G = 1/R$:

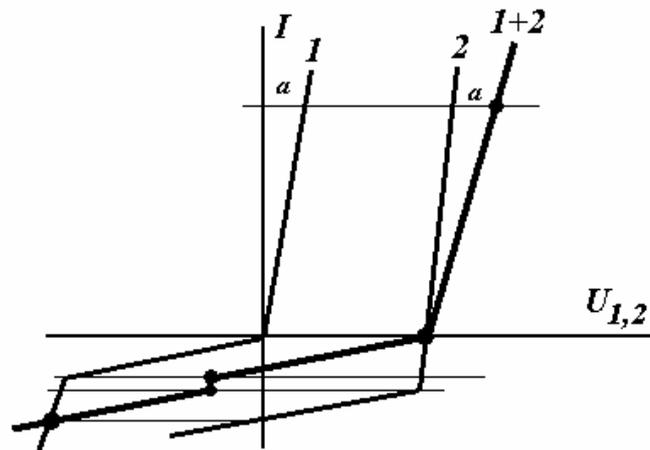
$$1/G_{serie} = \sum_k 1/G_k, \quad G_{serie} = 1/\left(\sum_k 1/G_k\right).$$

Werden nichtlineare Zweipole in Serie geschaltet, so entsteht ein nichtlinearer Zweipol, für den ebenso wie bei linearen Zweipolen $I_{pserie} = I_1 = I_2 = \dots = I_K$ und $U_{serie} = \sum_k U_k$ gilt. Sind

die einzelnen Kennlinien analytisch durch Formeln der Form $U_k = f_k(I_k)$ gegeben, so ergibt sich einfach

$$U_{serie} = \sum_k U_k = \sum_k f_k(I_{serie}).$$

Sind die Kennlinien graphisch gegeben, so zeichnet man sie in dieselbe Figur ein und addiert die Spannungen der einzelnen Zweipole graphisch.



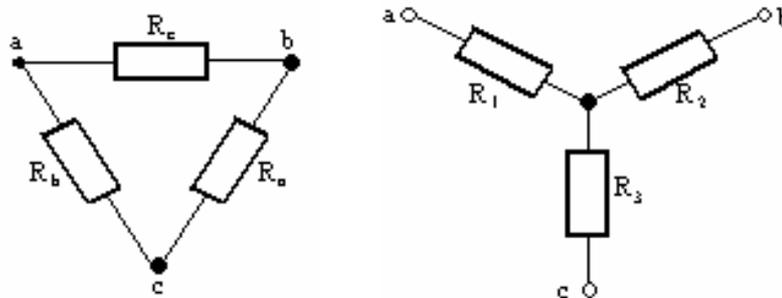
Figur 15: Graphische Ermittlung der Kennlinie einer Serieschaltung (1+2) von zwei nichtlinearen (stückweise linearen) Zweipolen (1 bzw. 2). Prinzip: Addition der Spannungen (horizontale Richtung in der verwendeten Darstellung). Bemerkung: Bei stückweise linearen Kennlinien ist es sinnvoll, mit den Knickpunkten zu arbeiten.

In einem Netzwerk, welches in Serie geschaltete Zweipole enthält, ersetzt man die in Serie geschalteten Zweipole und erhält damit ein vereinfachtes Netzwerk mit weniger Zweipolen. Sind die Ströme im vereinfachten Netzwerk berechnet, so lassen sich die einzelnen Teilspannungen der in Serie geschalteten Zweipole aus den entsprechenden Kennlinien ablesen oder aus den entsprechenden Zweipolgleichungen bestimmen.

Bemerkung: Hat man Serie- und Parallelschaltungen eines Netzwerkes zu dessen Vereinfachung berücksichtigt, so beobachtet man oft, dass das vereinfachte Netzwerk neue Serie- oder Parallelschaltungen enthält und somit nochmals vereinfacht werden kann usw. Durch mehrere, geschachtelte Vereinfachungen erhält man oft sehr einfache Netzwerke, welche sich mit geringem Aufwand manuell berechnen lassen. Beispiele findet man auf S. A10-A13.

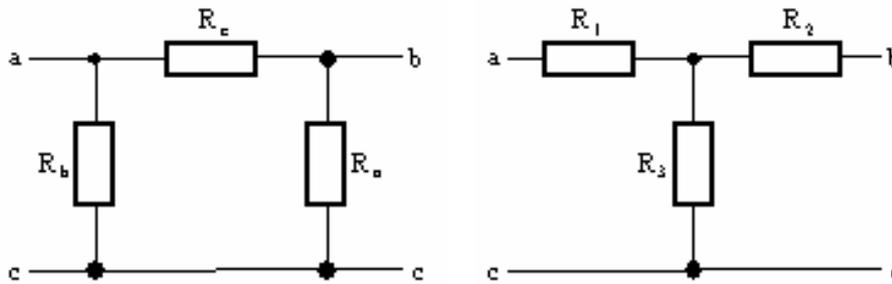
2.8. Stern-Dreieck Umwandlung

Die einfachsten Parallel- und Serieschaltungen von nur zwei Zweipolen bestehen im Grunde genommen aus einer einzigen Masche oder einem einzigen Knoten mit jeweils zwei Bauelementen. Erhöht man die Anzahl der Bauelemente um eins, so erhält man etwas kompliziertere Strukturen, die sogenannten Dreiecks- und Sternschaltungen, d.h. Maschen und Knoten mit jeweils drei Zweipolen. Der Einfachheit halber beschränken wir uns hier auf Dreiecks- und Sternschaltungen bestehend aus Widerständen.



Figur 16: Links: Dreieckschaltung, rechts: Sternschaltung

In der Praxis treten Dreiecks- und Sternschaltungen oft in der Energietechnik und in der Zweitorteknik auf. Unter einem Zweitort versteht man einen Vierpol, wobei zwei der vier Pole als Eingangsklemmen und die anderen zwei Pole als Ausgangsklemmen betrachtet werden. Wie die folgende Figur zeigt, ist eine sogenannte π -Schaltung ein Vierpol bestehend aus einer Dreieckschaltung und eine sogenannte T-Schaltung ein Vierpol bestehend aus einer Sternschaltung.



Figur 17: Links π -Schaltung, rechts: T-Schaltung

Dreiecks- und Sternschaltungen sind von aussen gesehen identisch, wenn für die von aussen beobachtbaren Ströme und Spannungen dieselben Beziehungen gelten. Für äquivalente Dreiecks- und Sternschaltungen bestehend aus Widerständen gilt:

Stern (bzw. T) - Dreieck (bzw. π) Umwandlung:

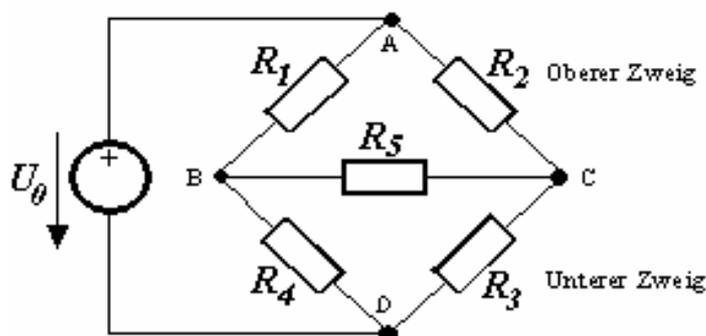
$$R_a = (R_1 R_2 + R_2 R_3 + R_3 R_1) / R_1, \quad R_b = (R_1 R_2 + R_2 R_3 + R_3 R_1) / R_2, \quad R_c = (R_1 R_2 + R_2 R_3 + R_3 R_1) / R_3$$

Dreieck (bzw. π) - Stern (bzw. T) Umwandlung:

$$R_1 = (R_b R_c) / (R_a + R_b + R_c), \quad R_2 = (R_c R_a) / (R_a + R_b + R_c), \quad R_3 = (R_a R_b) / (R_a + R_b + R_c).$$

Die Herleitung dieser Gleichungen ist etwas mühsam, weshalb wir hier darauf ebenso verzichten, wie auf die Verallgemeinerung für beliebige Zweipole.

Stern-Dreieck und Dreieck-Stern Umwandlungen lassen sich manchmal in komplizierteren Schaltungen zur Vereinfachung verwenden, wenn Vereinfachungen mit Serie- bzw. Parallelschaltungen nicht mehr weiter führen. Ein typisches Beispiel ist die Berechnung der Wheatstone'schen Brücke welche in der Messtechnik oft verwendet wird.

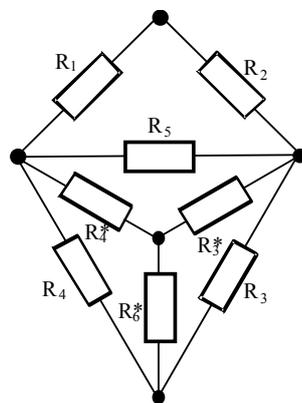


Figur 18: Wheatstone'sche Brücke bestehend aus Widerständen, an einer idealen Spannungsquelle. Die Punkte ABC bzw. BCD markieren zwei Dreiecke. Die Punkte B und C können auch als Sternknoten aufgefasst werden.

Die Wheatstone'sche Brücke besteht gewissermassen aus der Parallelschaltung zweier Dreiecksschaltungen (oder ebenso gut aus der Serieschaltung zweier Sternschaltungen). Wandelt man ein Dreieck in einen Stern um, so ergibt sich eine vereinfachte Schaltung, welche zwei Serieschaltungen enthält, die eine weitere Vereinfachung ermöglichen.

Beispiel: Berechne mit Hilfe der Dreieck-Stern Umwandlung den Gesamtwiderstand der Wheatstone Brücke gemäss Figur 18 an den Klemmen A und D.

Lösung: Man wählt entweder das obere oder untere Dreieck zur Umwandlung aus. Nach der Umwandlung (hier des unteren Dreiecks) ersetzt die Sternschaltung mit den Widerständen $R_{3,4,5}^*$ die ursprüngliche Dreieckschaltung. Die Werte dieser neuen Widerstände ergeben sich direkt aus den oben hergeleiteten Umformungen

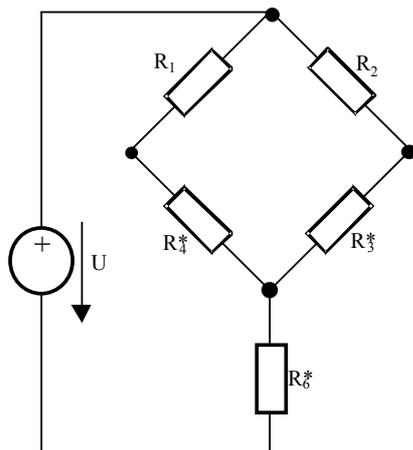


$$R_3^* = \frac{R_5 R_3}{R_3 + R_4 + R_5}$$

$$R_4^* = \frac{R_4 R_5}{R_3 + R_4 + R_5}$$

$$R_5^* = \frac{R_3 R_4}{R_3 + R_4 + R_5}$$

Somit kann die Schaltung wesentlich einfacher dargestellt werden und die Berechnung des Gesamtwiderstands gestaltet sich einfach:



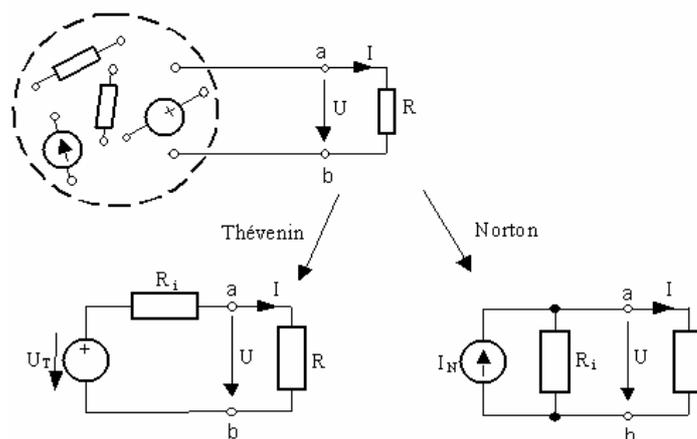
$$R_{\text{ges}} = R_6^* + \frac{(R_1 + R_4^*)(R_2 + R_3^*)}{R_1 + R_4^* + R_2 + R_3^*}$$

Selbstverständlich liesse sich die Methode der Stern-Dreiecks Umwandlung auf Knoten bzw. Maschen mit mehr als zwei Zweipolen erweitern. Die entsprechenden Umwandlungen sind aber recht kompliziert und werden daher in der Praxis nicht angewendet.

Weitere Beispiele auf S. A14

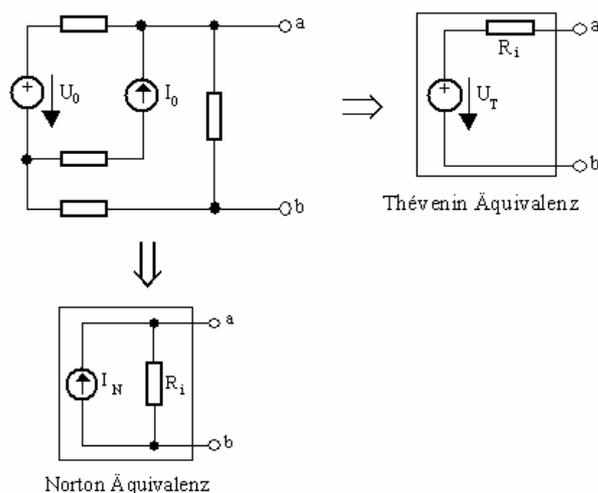
2.9. Thévenin, Norton und Quellenumwandlung

Bereits im Unterabschnitt 2.1.3, bei der Linearisierung von Zweipolen haben wir gesehen, dass jeder verlustbehaftete, lineare Zweipol entweder als Parallelschaltung einer idealen Stromquelle mit einem Leitwert oder als Serieschaltung einer idealen Spannungsquelle mit einem Widerstand betrachtet werden kann. Jeder lineare (bzw. linearisierte), verlustbehaftete Zweipol kann also durch eine Stromquellenersatzschaltung oder eine Spannungsquellenersatzschaltung dargestellt werden.



Figur 19: Veranschaulichung der Sätze von Thévenin und Norton

Ebenso sieht jedes lineare Netzwerk mit nur zwei Anschlussklemmen von aussen gesehen wie ein linearer Zweipol aus und kann demzufolge durch eine verlustbehaftete Strom- (Satz von Norton) oder Spannungsquelle (Satz von Thévenin) ersetzt werden.



Figur 20: Prinzipbeispiel für die Umwandlung einer Schaltung in eine Thévenin Äquivalenz und in eine Norton Äquivalenz.

Da die Stromquellenersatzschaltung aus einer Parallelschaltung besteht, wird sie vorteilhaft dann verwendet, wenn der betreffende Zweipol mit andern Zweipolen parallel geschaltet ist. Umgekehrt ist die Spannungsquellenersatzschaltung vorteilhaft, wenn der betreffende Zweipol mit andern Zweipolen in Serie geschaltet ist. Die Parallelschaltung mehrerer linearer Zweipole kann dann leicht in eine einzige Stromquelle (mit Innenleitwert) umgewandelt werden. Analog lässt sich eine Serieschaltung mehrerer linearer Zweipole leicht in eine einzige Spannungsquelle (mit Innenwiderstand) umwandeln.

Weil Stromquelle (mit Innenleitwert) und Spannungsquelle (mit Innenwiderstand) von außen ununterscheidbar sind, lassen sie sich auch ineinander umformen. Es gilt offenbar:

$$I_N = U_T/R_i \text{ bzw. } U_T = I_N/G_i, \text{ wobei } R_i = 1/G_i = U_T/I_N.$$

Diese Quellenumwandlung ist oft hilfreich zur weiteren Vereinfachung einer Schaltung. Die manuelle Schaltungsberechnung besteht oft aus einer Kaskade von Vereinfachungen mit Serie- und Parallelschaltungen, Stern-Dreiecks Umformungen, sowie Quellenumwandlungen.

Beispiele sind auf S. A17 und A20 ff.

2.10. Quellenüberlagerung, Superpositionsprinzip

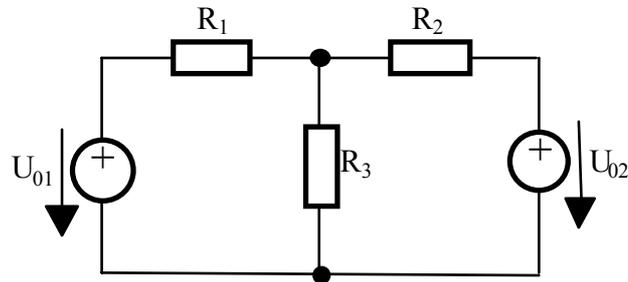
Hat man ein Netzwerk so stark wie möglich vereinfacht, so kann es sein, dass immer noch mehrere Quellen vorhanden sind. Sind alle Bauelemente linear, so sind alle Gleichungen, welche das Netzwerk beschreiben linear. Die Linearität kann nun ausgenutzt werden. Wegen der Linearität gilt das Superpositionsprinzip, nach dem jede Linearkombination von Ursachen eine Linearkombination von Wirkungen ergibt. In einem linearen elektrischen Netzwerk sind die Quellen die Ursachen und die Ströme und Spannungen die Wirkungen.

Sind N Quellen vorhanden, so kann man zunächst die Wirkungen jeder Quelle separat berechnen und zum Schluss die N Resultate aufsummieren. Bei der Berechnung der Wirkung einer Quelle, setzt man alle andern Quellenspannungen bzw. -ströme gleich null. Das heisst, eine ideale Spannungsquelle wird zu einem „Kurzschluss“, bzw. einem idealen Draht, eine ideale Stromquelle zu einem „Leerlauf“, d.h. sie wird einfach weggelassen und ihre Klemmen bleiben offen (unverbunden). Spannungsquellen mit Innenwiderständen bzw. Stromquellen mit Innenleitwerten werden zu einfachen Widerständen bzw. Leitwerten.

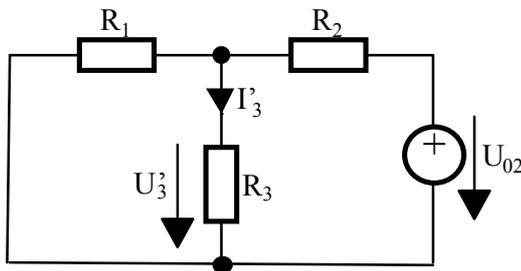
Die vereinfachten Netzwerke mit einer einzigen Quelle lassen sich oft mit den bereits beschriebenen Verfahren weiter vereinfachen und somit bequem rechnen.

Zur Veranschaulichung betrachten wir ein einfaches Beispiel mit zwei Spannungsquellen.

Beispiel: Bestimme den Strom durch den Widerstand R_3

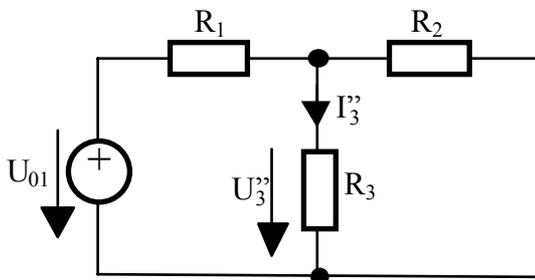


Lösung: Wir setzen zunächst die Quelle $U_{01} = 0$ und berechnen den durch die Quelle U_{02} verursachten Strom I_3'



$$I_3' = \frac{U_{R_3}'}{R_3} = \frac{U_{02}R_1}{R_1R_2 + R_2R_3 + R_3R_1}$$

Im 2. Schritt wird die Quelle $U_{02} = 0$ gesetzt und der Strom I_3'' verursacht durch die Quelle U_{01} berechnet



$$I_3'' = \frac{U_{R_3}''}{R_3} = \frac{U_{01}R_2}{R_1R_2 + R_2R_3 + R_3R_1}$$

Schliesslich werden beide Ergebnisse addiert

$$I_3 = I_3' + I_3'' = \frac{U_{01}R_2}{R_1R_2 + R_2R_3 + R_3R_1} + \frac{U_{02}R_1}{R_1R_2 + R_2R_3 + R_3R_1}$$

Hinweis: Das betrachtete Beispiel liesse sich auch mit Hilfe der Quellenumwandlung lösen.

Weitere Beispiele zur Superposition auf S. A16.

2.11. Systematische Methoden

Die bisher betrachteten Methoden zur Analyse von Netzwerken werden für die manuellen Berechnungen häufig verwendet, werden aber bei grossen Netzwerken problematisch und sind ausserdem für die Implementierung von Computerprogrammen nicht gut geeignet. Wir suchen deshalb nach systematischen Methoden, beschränken uns hier aber auf die Berechnung von Netzwerken bestehend aus N linearen Zweipolen.

Zunächst ist es sinnvoll, die Zweipole fortlaufend zu nummerieren, wobei die Reihenfolge willkürlich sein kann. Für jeden Zweipol Z_k werden nun die Spannung U_k und der Strom I_k festgelegt. Dazu gehört nicht nur die Angabe der Werte, sondern auch die Angabe der Richtungen von Spannung und Strom mit Pfeilen in der Schaltung. Die Festlegung der Richtungen ist an sich willkürlich, es ist jedoch empfehlenswert für jeden Zweipol Spannung und Strom in dieselbe Richtung einzutragen, d.h., das sogenannte *Verbraucherpfeilsystem* zu verwenden. Wir werden dies im Folgenden voraussetzen.

Offenbar gibt es N zu berechnende Spannungen U_k über den Zweipolen und N zu berechnende Ströme I_k , welche durch die Zweipole fließen, also $2N$ Unbekannte. Um diese berechnen zu können, suchen wir $2N$ linear unabhängige, lineare Gleichungen.

Ein erstes Set von N linear unabhängigen, linearen Gleichungen erhalten wir sofort aus den Gleichungen, welche die N Zweipole beschreiben und die Form $U_k = U_{k0} + R_k I_k$ oder $I_k = I_{k0} + G_k U_k$ haben. Diese Gleichungen sind besonders einfach und erlauben es uns sofort, entweder U_k oder I_k aus dem Gleichungssystem zu eliminieren. Damit bleiben nur noch N Unbekannte, für die wir insgesamt N linear unabhängige Gleichungen aus den Maschen- und Knotenbedingungen gewinnen müssen.

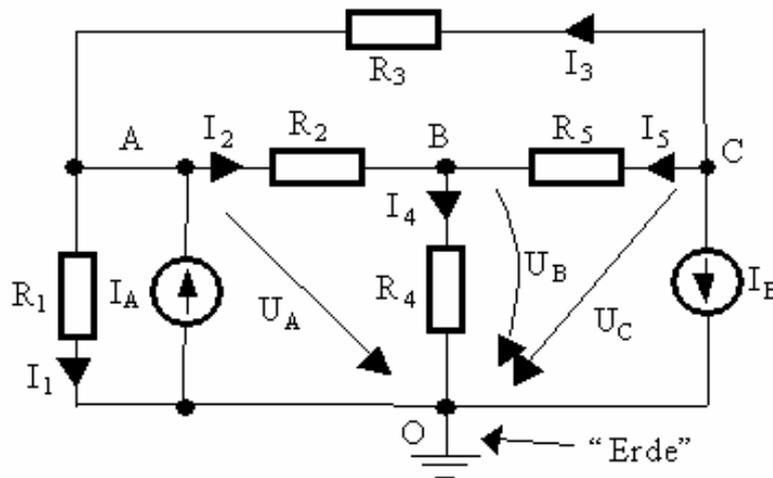
Es existieren verschiedene, teils recht anspruchsvolle Methoden zur Definition der gesuchten N linear unabhängige Gleichungen. Wir werden im Folgenden zwei gängige Methoden skizzieren. Dabei ersetzen wir jede nichtideale Spannungsquelle durch eine ideale Spannungsquelle mit in Serie geschaltetem Innenwiderstand und jede nichtideale Stromquelle durch eine ideale Stromquelle mit parallel geschaltetem Innenleitwert. Das Netzwerk besteht damit aus den besonders einfachen Bauelementen 1) Widerständen (bzw. Leitwerten), 2) idealen Spannungsquellen, 3) idealen Stromquellen.

2.11.1. Methode der Knotengleichungen, Knotenanalyse

Wir wählen zunächst einen beliebigen Knoten als Bezugsknoten. In der Praxis ist dieser Knoten meist der Masseknoten, welcher mit Erde verbunden wird. Wenn kein Masseknoten ausgezeichnet ist, ist es empfehlenswert, einen möglichst grossen Knoten auszuwählen.

Im zweiten Schritt werden für jeden Knoten die sogenannten Knotenspannungen zugeordnet. Die entsprechenden Spannungspfeile beginnen beim betreffenden Knoten und enden beim Bezugsknoten. Sind M Knoten vorhanden, so definiert man offenbar $M-1$ Knotenspannungen. Die Knotenspannungen sind offenbar Potentiale, welche sich als einfache

Summen von Spannungen über den Zweipolen ergeben. Das Potential des Bezugsknotens („Erde“) ist Null.



Figur 21: Knotenanalyse einer Schaltung mit vier Knoten A, B, C, Bezugsknoten O und den zugehörigen Knotenspannungen.

Wenn wir die Knotengleichungen für die $M-1$ Knoten aufschreiben, so erhalten wir offenbar ebensoviele Gleichungen wie Knotenspannungen. Allerdings enthalten diese Gleichungen zunächst Ströme durch die angeschlossenen Zweipole. Diese Ströme sind entweder bekannt (wenn der betreffende Zweipol eine ideale Stromquelle ist) oder zu berechnen (wenn der betreffende Zweipol ein Widerstand bzw. Leitwert ist). Unbekannte Ströme in den Knotengleichungen lassen sich mit den Zweipolgleichungen durch (unbekannte) Spannungen über den Zweipolen ersetzen. Diese Zweipolspannungen lassen sich wiederum sofort durch Knotenspannungen ausdrücken. Damit erhalten wir tatsächlich $M-1$ lineare Gleichungen, welche nur die $M-1$ unbekanntes Knotenspannungen enthalten. In Matrixschreibweise:

$$\mathbf{GU} = \mathbf{I}.$$

Dabei ist \mathbf{G} die Leitwertmatrix. Diese Matrix ist symmetrisch und enthält Summen der Leitwerte der verschiedenen Zweipole. \mathbf{U} ist ein Vektor, welcher die Knotenspannungen enthält und \mathbf{I} ist ein Vektor, welcher die Stromquellen enthält.

Sind die Knotenspannungen durch Auflösen des eben erhaltenen Gleichungssystems berechnet, so erhält man die Spannungen über den Zweipolen sofort aus der Differenz der Knotenspannungen an den beiden Klemmen des Zweipols. Schliesslich erhält man die Ströme der Zweipole aus den Zweipolspannungen und den (linearen) Zweipolgleichungen.

Beispiel: Zur Veranschaulichung betrachten wir nun ein relativ einfaches Beispiel mit zwei Stromquellen gemäss Figur 21. Dabei notieren wir die Knotengleichungen und setzen die Gleichungen der angeschlossenen Zweipole ein.

Knoten A: Hier ergibt die Summe der Ströme

$$I_1 - I_A + I_2 - I_3 = 0$$

oder mit den Spannungen und Widerständen geschrieben

$$\frac{U_A}{R_1} - I_A + \frac{U_A - U_B}{R_2} - \frac{U_C - U_A}{R_3} = 0$$

Aufgelöst nach I_A folgt

$$U_A \left(\frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} \right) - U_B \frac{1}{R_2} - U_C \frac{1}{R_3} = I_A$$

Knoten B:

$$-I_2 + I_4 - I_5 = 0$$

bzw. Spannungen und Widerstände berücksichtigt

$$-\frac{U_A - U_B}{R_2} + \frac{U_B}{R_4} - \frac{U_C - U_B}{R_5} = 0$$

daraus folgt

$$-U_A \frac{1}{R_2} + U_B \left(\frac{1}{R_2} + \frac{1}{R_4} + \frac{1}{R_5} \right) - U_C \frac{1}{R_5} = 0$$

Knoten C: Die Beschreibung dieses Knotens erfolgt auch nach dem gleichen Schema. Als Ergebnis erhält man

$$-U_A \frac{1}{R_3} - U_B \frac{1}{R_5} + U_C \left(\frac{1}{R_3} + \frac{1}{R_5} \right) = -I_B$$

Die Knotenanalyse hat also drei simultane lineare Gleichungen mit drei unbekanntem Knotenspannungen geliefert. Diese Gleichungen können auch in Matrixform folgendermassen geschrieben werden:

$$\begin{bmatrix} \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} & -\frac{1}{R_2} & -\frac{1}{R_3} \\ -\frac{1}{R_2} & \frac{1}{R_2} + \frac{1}{R_4} + \frac{1}{R_5} & -\frac{1}{R_5} \\ -\frac{1}{R_3} & -\frac{1}{R_5} & \frac{1}{R_3} + \frac{1}{R_5} \end{bmatrix} \begin{bmatrix} U_A \\ U_B \\ U_C \end{bmatrix} = \begin{bmatrix} I_A \\ 0 \\ -I_B \end{bmatrix}$$

Dieses lineare Gleichungssystem lässt sich mit verschiedenen bekannten Algorithmen auflösen.

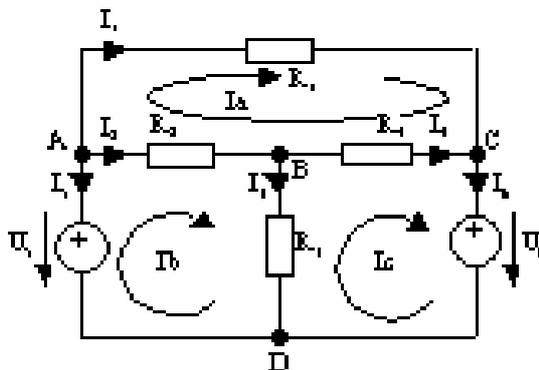
Es stellt sich vielleicht die Frage, wo denn die Maschengleichungen geblieben sind, haben wir diese doch gar nicht explizite aufgestellt. Es ist zu beachten, dass die Maschengleichungen implizit bei der Einführung der Knotenspannungen berücksichtigt wurden.

Offenbar ergeben sich Probleme bei der eben skizzierten Methode, wenn das Netzwerk ideale Spannungsquellen enthält. Diese Probleme können meist durch Quellenumformung gelöst werden. Ist dies unmöglich, so kann die Methode der Maschengleichungen verwendet werden. Diese Methode wird im folgenden Unterabschnitt skizziert und ist im Wesentlichen dual zur Methode der Knotengleichungen. Das heisst auch, dass die Methode der Maschengleichungen Probleme mit idealen Stromquellen hat. Gelingt es nicht, entweder alle idealen Stromquellen aus einem Netzwerk zu eliminieren und die Methode der Knotengleichungen zu verwenden oder alle idealen Spannungsquellen aus einem Netzwerk zu eliminieren und die Methode der Maschengleichungen zu verwenden, so kann man das Superpositionsprinzip verwenden und gelangt so auf zwei Netzwerke, eines mit Stromquellen und eines mit Spannungsquellen.

Weitere Beispiele zur Knotenanalyse auf S. A27,

2.11.2. Methode der Maschengleichungen, Umlaufanalyse

Bei der Methode der Maschengleichungen definieren wir Maschenströme (anstelle der Knotenspannungen bei der Knotenanalyse) und notieren für diese die Maschengleichungen (anstelle der Knotengleichungen). Maschenströme sind in einer Masche zirkulierende Ströme. Dabei kann die Zirkulationsrichtung im Uhrzeigersinn oder im Gegenuhrzeigersinn definiert werden. Sinnvoll ist natürlich eine einheitliche Wahl. Die Definition der Maschenströme garantiert übrigens die implizite Erfüllung der Knotengleichungen.



Figur 22: Einfache Schaltung mit Maschenströmen für die Maschenanalyse.

Das Auffinden von unabhängigen Maschen ist etwas schwieriger als das Auffinden unabhängiger Knoten bei der Knotenanalyse. Man kann systematisch Maschen solange erzeugen, bis jeder Zweipol zu mindestens einer Masche gehört und muss dabei lediglich beachten, dass jede neue Masche einen zuvor noch nicht berücksichtigten Zweipol enthält. Soll das erhaltene Gleichungssystem manuell ausgewertet werden, so versucht man am besten, möglichst kleine Maschen zu wählen.

Sind M unabhängige Maschen gewählt und die entsprechenden M Maschenströme definiert, so geht man ganz analog zur Knotenanalyse vor: 1) Man notiert die Maschengleichungen, in denen die Spannungen über idealen Spannungsquellen bekannt sind und die Spannungen über Widerständen R_k durch $I_k = U_k/R_k$ ersetzt werden. 2) Man ersetzt die Zweipolströme durch die Maschenströme und erhält die Matrixgleichung

$$\mathbf{RI} = \mathbf{U}.$$

Dabei ist \mathbf{R} die (symmetrische) Widerstandsmatrix. Die Vektoren \mathbf{I} und \mathbf{U} enthalten die Maschenströme und die Quellenspannungen. 3) Das Gleichungssystem wird aufgelöst. 4) Die unbekannt Ströme in den Zweipolen werden aus den Maschenströmen bestimmt. Dabei werden alle Maschenströme aufsummiert, welche durch den betreffenden Zweipol fließen (negative Werte, falls die Stromrichtung mit der Maschenstromrichtung nicht übereinstimmt). 5) Die unbekannt Spannungen über den Widerständen werden mit $U_k = R_k I_k$ berechnet.

Beispiel: Zur Veranschaulichung betrachten wir ein einfaches Beispiel mit zwei Spannungsquellen gemäss Figur 22. Dabei beschreiben die Maschenströme I_a , I_b , I_c die gesamte Stromverteilung. Die Maschengleichungen lauten (unter Berücksichtigung der Gleichungen der involvierten Zweipole):

$$\text{Umlauf a: } R_1 I_1 - R_4 I_4 - R_2 I_2 = 0$$

$$\text{Umlauf b: } R_2 I_2 + R_3 I_3 - U_5 = 0$$

$$\text{Umlauf c: } U_6 + R_4 I_4 - R_3 I_3 = 0$$

Die einzelnen Ströme lassen sich als Überlagerung der Maschenströme I_a , I_b und I_c darstellen. Zum Beispiel für Umlauf b

$$-R_2(I_a - I_b) + R_3(I_b - I_c) - U_5 = 0$$

Bzw. nach U_5 aufgelöst

$$-R_2 I_a + (R_2 + R_3) I_b - R_3 I_c = U_5$$

Auch hier kann man das Bildungsgesetz bereits erkennen: In der Maschengleichung für Umlauf b tritt als Koeffizient für den Umlaufstrom I_b die Summe der in diesem Umlauf befindlichen Widerstände auf, der sogenannte Umlaufwiderstand.

Als Koeffizient für den Maschenstrom I_a tritt der Widerstand R_2 auf. Dieser tritt in beiden Maschen auf und wird deshalb Kopplungswiderstand genannt. Die Quellenspannung U_5 erscheint auf der rechten Seite der Gleichung mit positivem Vorzeichen wenn sie dem Umlauf entgegen gerichtet ist, ansonsten mit negativem Vorzeichen. Das aus dem Umlauf b bereits erkennbare Bildungsgesetz kann auf die anderen Maschen übertragen werden und man erhält in Matrizenform

$$\begin{bmatrix} R_1 + R_2 + R_4 & -R_2 & -R_4 \\ -R_2 & R_2 + R_3 & -R_3 \\ -R_4 & -R_3 & R_3 + R_4 \end{bmatrix} \begin{bmatrix} I_a \\ I_b \\ I_c \end{bmatrix} = \begin{bmatrix} 0 \\ U_5 \\ -U_6 \end{bmatrix}$$

Weitere Beispiele zur Maschenmethode auf S. A31.

3. Schaltvorgänge

Bisher haben wir statische Netzwerke betrachtet, in denen die Spannungen und Ströme zeitlich konstant sind. Solche Netzwerke sind besonders einfach zu berechnen, aber auch nicht sehr interessant. Bei den allen praktischen Anwendungen sind die Spannungen und Ströme zeitabhängige Grössen. Dies gilt selbst für einfachste Anwendungen, wie beispielsweise Taschenlampen. Diese bestehen lediglich aus einer Spannungsquelle (der Batterie), einem Verbraucher (der Lampe) und einem Schalter. Mit Hilfe des Schalters bedient der Benutzer seine Taschenlampe und erzwingt damit zeitliche Veränderungen im elektrischen System. Fast alle elektrischen Geräte sind mit Schaltern ausgerüstet, welche dem Benutzer die gewünschte Kontrolle über das System ermöglichen. Dies gilt insbesondere auch für Computer, welche neben dem Netzschalter über zahlreiche weitere Schalter verfügen. So ist beispielsweise jede Taste der Tastatur ein Schalter. Neben diesen benutzergesteuerten Schaltern enthält ein Computer Millionen von elektronisch gesteuerten Schalter, welche Milliarden von Schaltvorgängen pro Sekunde auslösen. In diesem Kapitel wollen wir die wichtigsten Effekte betrachten, welche sich in elektrischen Systemen bei Schaltvorgängen ergeben. Dabei werden wir uns mit einfachen Schaltungen mit meist nur einem einzigen Schalter begnügen. Eine Erweiterung auf komplexere Systeme mit vielen Schaltern liegt auf der Hand. Die numerische Analyse solcher Systeme kann aber nur mit geeigneten Computerprogrammen erfolgen.

Zunächst sind Schalter das wichtigste Bauelement, welches wir in unser Arsenal von elektrischen Bauelementen aufnehmen müssen. Dies wird im folgenden Abschnitt gemacht. Zudem müssen wir berücksichtigen, dass zeitliche Veränderungen der Ströme und Spannungen auch zeitliche Veränderungen der elektromagnetischen Felder in den Bauelementen erzwingen und diese führen wiederum zu physikalischen Effekten, welche wir bisher ausser Acht gelassen hatten, nun aber berücksichtigen müssen. In der Schaltungstechnik geschieht dies durch Einführung neuer Bauelemente. Die wichtigsten und zugleich einfachsten „dynamischen“ Bauelemente sind Kondensatoren und Spulen. Die zugehörigen Idealisierungen sind Kapazitäten und Induktivitäten. Wir werden diese einfachen Zweipole und deren Einbettung in Netzwerke im Folgenden genauer betrachten.

Formal haben wir konstante Spannungen und Ströme mit den Grossbuchstaben U bzw. I bezeichnet. Für zeitveränderliche Grössen werden wir Kleinbuchstaben verwenden und allenfalls zur Verdeutlichung das Argument (t) hinzufügen. Die Bezeichnungen u und $u(t)$ sind also im Folgenden gleichbedeutend.

3.1. Bauelemente

Im vorhergehenden Kapitel haben wir vorwiegend einfache Zweipole betrachtet und diese meist so idealisiert, dass möglichst einfache Rechnungen durchgeführt werden konnten. Insbesondere haben wir die Kennlinien der Zweipole meist linearisiert, um lineare Glei-

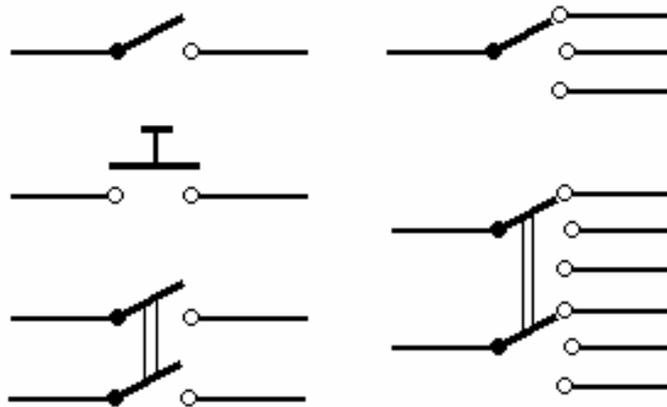
chungssysteme für die Netzwerke zu erhalten. Die Linearisierung ermöglichte es, mit lediglich drei idealen Bauelementen, der idealen Spannungsquelle, der idealen Stromquelle und dem Ohm'schen Widerstand auszukommen. Wir werden in diesem Abschnitt drei neue, ideale Zweipole hinzufügen: Schalter, Kapazität und Induktivität.

Neben zweipoligen Schaltern werden oft auch mehrpolige Schalter verwendet. Deren Behandlung bietet keine besondere Schwierigkeit. Hingegen basieren Induktivitäten auf dem Induktionsgesetz, welches auch mehrpolige Bauelemente, die Transformatoren, ermöglicht. Deren Idealisierung führt auf die Gegeninduktivität. Da Transformatoren in der Energietechnik, aber auch in der Nachrichtentechnik und in den Netzteilen von Computern eine bedeutende Rolle spielen, widmen wir der Gegeninduktivität einen separaten Unterabschnitt.

3.1.1. Schalter

In der Praxis existieren Dutzende von Schalterformen, welche im Wesentlichen in zwei Gruppen eingeteilt werden können: 1) Von aussen (durch den Benutzer) bediente Schalter und 2) gesteuerte Schalter welche durch Stellgrössen im Netzwerk selber ausgelöst werden. Ob die Schalter vom Benutzer oder durch Stellgrössen gestellt werden, ist für die Schaltungsanalyse sekundär.

Es gibt zahlreiche Bauformen von durch den Benutzer bedienten Schaltern. Diese unterscheiden sich durch Stromstärken, welche geschaltet werden können, erlaubte Spannungen über den Schaltkontakten, Schaltgeschwindigkeit, erzeugte Störungen usw. Beispielsweise können mechanische Schalter beim Schliessen „prellen“, d.h. sie öffnen und schliessen sich abwechselnd mehrmals in kurzer Folge, bevor sie endgültig geschlossen bleiben. Wird ein solcher Schalter beispielsweise für eine Computertastatur verwendet, so löst ein einzelner Tastendruck mehrere Signale aus. Da die Herstellung nicht prellender mechanischer Schalter schwierig ist, werden meist prellende Schalter mit nachfolgender Elektronik, welche das unerwünschte Prellen unterdrückt, eingesetzt. Andere Probleme ergeben sich bei Umschaltern, welche mehrere Zustände haben und beispielsweise bei Verstärkern zur Wahl der Eingangsquelle verwendet werden. Derartige Umschalter haben beispielsweise mehrere Eingänge und einen einzigen Ausgang. Es wird erwartet, dass — je nach Schalterstellung — genau ein Eingang mit dem Ausgang verbunden ist. Praktisch muss es aber bei jedem Umschalten einen Zwischenzustand geben, wo normalerweise kein Eingang mit dem Ausgang verbunden ist. Dies führt beispielsweise bei Verstärkern zu unangenehmen Knackgeräuschen. Auch hier schaffen elektronische Zusatzschaltungen Abhilfe. Wir werden im Folgenden unerwünschte Nebeneffekte von realen Schaltern ausser Acht lassen und uns auf idealisierte Schalter konzentrieren.



Figur 23: Symbole für verschiedene, idealisierte Schalter. Links: Ein/Aus Schalter (oben: Kippschalter, Mitte: Tastenschalter, unten: gekoppelte Kippschalter), rechts: Umschalter mit drei Ausgängen (oben: einzelner Umschalter, unten: zwei gekoppelte Umschalter).

Idealisierte „Ein/Aus“ Schalter haben zwei Zustände: Offen oder geschlossen. Im offenen Zustand besteht keine Verbindung zwischen den Klemmen (Leerlauf, bzw. unendlicher Ohm'scher Widerstand, bzw. Leitwert Null). Im geschlossenen Zustand werden sind die Klemmen ideal miteinander verbunden (Kurzschluss, bzw. Ohm'scher Widerstand Null, bzw. unendlicher Leitwert). Der ideale „Ein/Aus“ Schalter ist damit ein ideales Bauelement für binäre Logik und er wird auch durch binäre Logik beschrieben. Wird beispielsweise ein solcher Schalter zur Zeit t_1 eingeschaltet, so gilt für Schalterstrom i und Schalterspannung u :

$$i = 0 \text{ falls } t < t_1 \text{ und } u = 0 \text{ falls } t > t_1.$$

Zur Analyse einer Schaltung, welche einen derartigen Schalter enthält, muss man im Prinzip zwei verschiedene Schaltungen (vor und nach dem Schaltmoment) aufzeichnen und analysieren. Dabei ergeben sich die Ströme und Spannungen vor dem Schaltmoment aus einer statischen Berechnung — sofern zuvor kein anderer Schalter betätigt wurde und lediglich statische Spannungs- und Stromquellen vorhanden sind. Nach dem Schaltmoment wird die Analyse meist komplizierter, da verschiedene dynamische Effekte auftreten können. Dies gilt insbesondere dann, wenn Kapazitäten und Induktivitäten vorhanden sind, welche in den folgenden Unterabschnitten betrachtet werden. Nach dem Schaltmoment werden dann Ströme und Spannungen mit Differentialgleichungen beschrieben und die statischen Werte vor dem Schaltmoment ergeben die Anfangsbedingungen für diese Differentialgleichungen.

In logischen Schaltungen kann man das dynamische in erster Näherung vernachlässigen und auch die Spannungen und Ströme nach dem Schaltvorgang statisch behandeln. Praktisch ist dies allerdings insbesondere bei schnellen Schaltvorgängen problematisch und kann dazu führen, dass die so entworfenen Schaltungen nicht zuverlässig oder gar nicht richtig funktionieren.

Eine erhebliche Komplizierung der Schaltungsanalyse ergibt sich in der Praxis dadurch dass oft viele Schalter vorhanden sind, welche nicht nur einmal und ausserdem nicht gleichzeitig betätigt werden. Insbesondere bei komplizierten logischen Schaltungen führt mangelhafte Synchronisierung verschiedener Schalter zu Fehlverhalten. Exakt gleichzeitiges Betätigen von Schaltern an unterschiedlichen Stellen eines Netzwerkes ist praktisch unmöglich. Um Abhilfe zu schaffen, werden deshalb Computer und andere elektronische Schaltungen „getaktet“. Für die Schaltungsanalyse ergibt sich daraus, dass nur diskrete Zeitwerte $t_k = k \cdot dt$ beobachtet werden müssen, wobei dt das Taktintervall ist. Ausserdem kann angenommen werden, dass alle Schalter, welche innerhalb eines Taktintervalls betätigt werden, gleichzeitig, d.h. synchron arbeiten.

Von manuell betätigten Schaltern kann man kaum erwarten, dass sie synchron gestellt werden, selbst wenn das Taktintervall relativ gross ist — ausser wenn mechanisch oder elektronisch gekoppelte Schalter (siehe Figur 23) verwendet werden. Computer bestehen deshalb jeweils nicht nur aus synchron arbeitenden, sondern auch aus asynchronen Schaltungsteilen. Letztere sind insbesondere typisch für den Eingabebereich. Für die synchron arbeitenden Bereiche kommen fast ausschliesslich elektronisch gesteuerte Schalter in Frage, welchen wir uns nun kurz zuwenden möchten.

Gesteuerte Schalter weisen offenbar eine klare Verwandtschaft mit den gesteuerten Quellen auf, welche wir im Unterabschnitt 2.1.8 skizziert hatten. Berücksichtigt man den Steuerungsteil, so sind gesteuerte Schalter ebenso wie gesteuerte Quellen im Prinzip oft Drei- oder Vierpole. Es gibt aber auch *zweipolige*, gesteuerte Schalter, bei denen die Steuergrösse die Spannung über dem Schalter selber oder der Strom, welcher durch den Schalter fliesst, ist. Typische Beispiele dafür sind *Sicherungen*, welche „durchbrennen“, wenn der Strom zu gross wird, *Überspannungsschalter*, aber auch einfache *Dioden*, welche bei positiver Spannung als geschlossene Schalter und bei negativer als offene Schalter betrachtet werden können.

In der Praxis werden tatsächlich *Transistoren* nicht nur als gesteuerte Quellen, sondern auch als gesteuerte Schalter eingesetzt. In der Leistungselektronik werden Transistoren meist durch *Thyristoren* ersetzt. Sowohl Transistoren als auch Thyristoren sind dreipolige Bauelemente.

Gesteuerte Schalter können in der Praxis ein recht kompliziertes, dynamisches Verhalten aufweisen. Dies wird schon beim stark vereinfachten Ersatzschaltbild eines Transistors gemäss Figur 8 klar. Dabei wird der Pfad Kollektor-Emitter durch den Strom durch die Basis gesteuert. Ist der Basisstrom Null, so ist der Schalter Kollektor-Emitter offen, d.h. es fliesst kein Strom von Kollektor zu Emitter. Nimmt der Basisstrom geringfügig zu, so fliesst nur ein geringer Strom von Kollektor zu Emitter und es fällt schwer zu sagen, der Schalter sei geschlossen. Damit der Kollektor-Emitter Strom ausreichen gross wird, muss also der Basisstrom ausreichend gross sein. Ein anderes Problem ergibt sich dadurch, dass die Kollektor-Emitter bei einem NPN Transistor positiv sein muss und ausserdem — je nach Bauart — einen gewissen Wert übersteigen muss (wie bei den Dioden liegt dieser Wert für Silizium Halbleiter bei ca. 0.6V). Schaltungstechnisch kann man diese und weitere Probleme vermindern, indem man kompliziertere gesteuerte Schalter entwirft, welche nicht

nur aus einem einzigen Transistor bestehen. Ähnlich wie bei den Thévenin und Norton Äquivalenzen kann man dann für diese Schaltungen einfache Ersatzschaltungen angeben. Da man mit komplizierten Schaltungen versucht, idealen Schaltern möglichst nahe zu kommen, sind die Ersatzschaltungen für komplizierte Schaltungen oft sogar besonders einfach. Idealisierte, gesteuerte Schalter lassen sich dann durch einfache logische Sätze beschreiben, z.B. für einen gesteuerten Ein/Aus Schalter: „Der Schalter ist geschlossen, wenn die Stellgröße einen vorgegebenen Wert übersteigt. Andernfalls ist er offen.“ Als Folge davon wird die Schaltungsanalyse relativ einfach — zumindest, wenn das dynamische Schaltverhalten ausser Acht gelassen werden kann.

Da es nicht zum vorneherein klar ist, wann das dynamische Schaltverhalten berücksichtigt werden muss, müssen wir einfache Schaltungen analysieren, welche zeigen worin das dynamische Schaltverhalten überhaupt besteht und was dabei wichtig ist. Um dies tun zu können, müssen wir nun zwei einfache Zweipole definieren und analysieren, welche die dynamischen Grundelemente beinhalten.

Die einfachste Form eines elektrischen Schalters besteht aus zwei (metallischen) Kontakten, welche über ein bewegliches Metallteil verbunden werden (wenn der Schalter geschlossen wird. Dabei wird angenommen, zwischen den beiden Kontakten bestehe eine elektrische Spannung solange der Schalter offen ist, ohne dass ein Strom fliesse. Da die Kontakte eine gewisse räumliche Ausdehnung haben müssen, sammeln sich auf ihnen elektrische Ladungen und zwar negative Ladungen auf dem einen und positive Ladungen auf dem andern Kontakt. Wird der Schalter geschlossen, so fließt zunächst ein hoher Strom, welcher die auf den Kontakten angesammelten Ladungen ausgleicht. Die statischen Betrachtungen des vorhergehenden Kapitels haben Ladungen an Kontakten (und in Knoten bzw. auf Drähten eines Netzwerks) ignoriert. Im nächsten Unterabschnitt werden wir diese berücksichtigen. Nach einiger Zeit fließt in einem statischen Netzwerk ein konstanter Strom durch unsern Schalter. Wenn wir den Schalter öffnen, so müsste der Strom abrupt geändert werden. Dies ist praktisch unmöglich, weil die zeitliche Änderung eines Stromes ein Magnetfeld erzeugt, dessen zeitliche Änderung wiederum Ströme induziert. Dies werden wir im übernächsten Unterabschnitt genauer betrachten. Die zeitliche Änderung eines Stromes induziert nicht nur Ströme im betreffenden Draht, sondern auch in andern Drähten einer Schaltung. Diesen Effekt der Gegeninduktion werden wir ebenfalls kurz skizzieren.

3.1.2. Kondensatoren und Kapazität

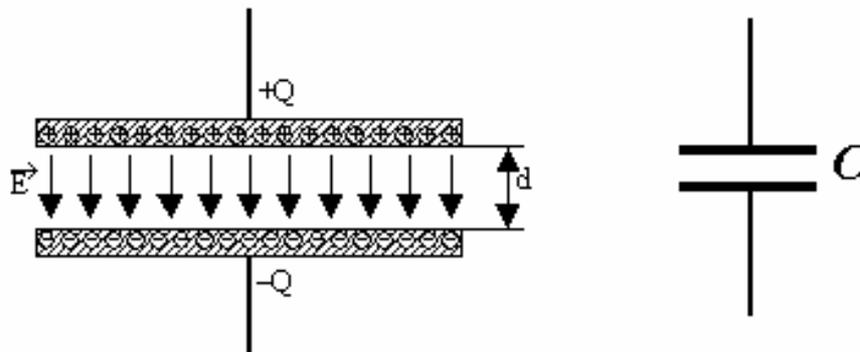
Wie ein offener Schalter besteht ein Kondensator im Prinzip aus zwei voneinander getrennten, d.h. isolierten, metallischen Kontakten, welche nun Elektroden genannt werden. Ein Kondensator ist also ein Zweipol. Dabei ist die Ladung auf den entgegengesetzten Elektroden entgegengesetzt gleich, die Gesamtladung also Null. Ausserdem ist die Ladung der positiven Elektrode proportional zum absoluten Betrag der angelegten Spannung:

$$Q = CU \text{ oder dynamisch } q(t) = Cu(t).$$

Die Kapazität C ist eine Konstante (Einheit $1\text{F} = 1\text{As}/1\text{V}$), welche von den verwendeten Materialien und der Geometrie der Anordnung abhängt. Um einen möglichst ausgeprägten Effekt, d.h. eine hohe Kapazität zu erhalten, vergrössert man dabei die Oberfläche der Elektroden so weit wie möglich und verkleinert gleichzeitig den Abstand der Elektroden. Grundsätzlich nimmt die Gesamtladung einer Elektrode bei konstanter angelegter Spannung mit der Grösse der Oberfläche zu und mit dem Abstand der Elektroden ab. Besonders beliebt sind deshalb Plattenkondensatoren mit plattenförmigen Elektroden. Sind die beiden Elektroden parallele Platten deren Abmessungen deutlich grösser als der Abstand zwischen den Platten ist, so ist C nahezu proportional zur Fläche und indirekt proportional zum Abstand der Elektroden. Weil ausgedehnte Platten praktisch unhandlich sind, wickelt man diese gerne auf und erhält damit höhere Kapazitäten auf relativ kleinem Raum.

Da bei kleinen Abständen bzw. hohen Spannungen Funken zwischen den Elektroden auftreten, werden die Elektroden meist isoliert, d.h., es wird ein isolierendes Material zwischen die Elektroden gelegt. Dabei gibt es Materialien, welche die Ladung auf den Elektroden bei fester Spannung erhöhen. Diese Materialeigenschaft wird durch die Dielektrizitätskonstante beschrieben. Dadurch gelingt es, die Abmessungen von Kondensatoren weiter zu verringern.

Wir wollen nun die wesentlichen physikalischen Effekte kurz beleuchten und betrachten ein vereinfachtes Modell eines geladenen Plattenkondensators.



Figur 24: Vereinfachtes Modell eines Plattenkondensators und Schaltsymbol.

Vernachlässigen wir Effekte in der Nähe der Plattenränder, so ist das elektrische Feld nahezu homogen und senkrecht zu den Platten und die Ladungen sind nahezu homogen auf den einander zugewandten Oberflächen der Elektroden verteilt. Es gilt dann für den Betrag der elektrischen Feldstärke die einfache Beziehung

$$E = U/d,$$

wobei U die angelegte Spannung und d der Plattenabstand ist. Weil E (gemäss dem Superpositionsprinzip für Ladungen) zur Ladung Q proportional ist wird auch U proportional zur Ladung, d.h. es gilt die bereits erwähnte Gleichung $Q = CU$. Diese Gleichung gilt übrigens

ganz allgemein, nicht nur für Plattenkondensatoren. Für die Kapazität eines Plattenkondensators ergibt sich ausserdem als Näherungsformel:

$$C = \varepsilon F/d,$$

Wobei F die Plattenfläche und ε die Dielektrizitätskonstante des Materials zwischen den Platten ist. Im Anhang sind eine Tabelle mit Dielektrizitätskonstanten für gebräuchliche Materialien und eine Tabelle mit Näherungsformeln zur Berechnung der Kapazitäten für gebräuchliche Anordnungen zu finden.

Da wir bereits wissen, dass Ströme bewegte Ladungen sind und den Ladungserhaltungssatz kennen, ergibt sich aus $q(t) = Cu(t)$ sofort

$$i(t) = dq(t)/dt = C du(t)/dt.$$

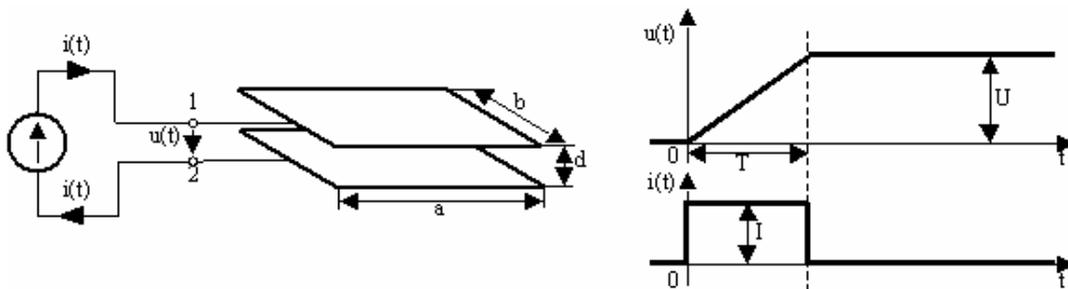
Durch Integration beider Seiten folgt

$$u(t) = u(0) + \frac{1}{C} \int_0^t i(x) dx.$$

$u(0)$ ist die an der Kapazität zur Zeit $t = 0$ liegende Spannung, die Anfangsbedingung für die obige Differentialgleichung. Die Kapazität ist also ein Zweipol, welcher durch eine lineare Differentialgleichung erster Ordnung beschrieben wird. Die Linearität ermöglicht ein einfaches Arbeiten mit Kapazitäten.

Die Verwandtschaft obiger Differentialgleichung mit der Gleichung $I = GU$ aus der Gleichstromtechnik legt die Verwandtschaft von Kapazitäten und Leitwerten nahe. Dies wird im Abschnitt über Serie- und Parallelschaltungen noch deutlicher werden.

Wir betrachten nun den Ladevorgang eines Plattenkondensators, d.h. wir nehmen an, der zunächst ungeladene Plattenkondensator werde zur Zeit $t = 0$ bis zur Zeit $t = T$ an eine ideale Stromquelle angeschlossen:



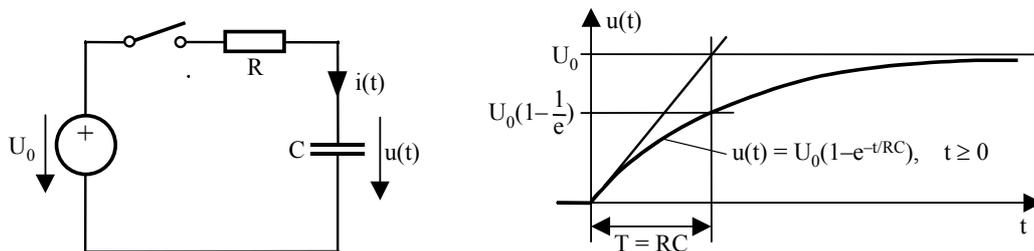
Figur 25: Ladung eines Plattenkondensators mit idealer Stromquelle während einer Zeitspanne T .

Die Anfangsbedingung lautet $u(0) = 0$. Danach ergibt sich ein linearer Spannungsanstieg für $0 < t < T$ wegen

$$u(t) = u(0) + \frac{1}{C} \int_0^t i(x) dx = \frac{1}{C} \int_0^t I dx = \frac{tI}{C}.$$

Schliesslich bleibt für $t > T$ die Spannung über der Kapazität weil nach $t = T$ kein Strom mehr fliesst. Die Spannung ist dann $u(t) = IT/C$.

Etwas komplizierter wird die Berechnung der Aufladung einer Kapazität mit einer Stromquelle mit Innenleitwert, welche äquivalent zu einer Spannungsquelle mit Innenwiderstand ist. Dies ist wiederum dasselbe wie die Aufladung einer Kapazität über einen Seriewiderstand:



Figur 26: Aufladung einer Kapazität C über einen Widerstand R .

Nehmen wir an, dass für $t < 0$ der Schalter offen und die Kapazität ungeladen sei und dass der Schalter für $t > 0$ geschlossen sei, so ist die Spannung über Kapazität und Widerstand

$$u_{RC} = 0 \text{ für } t < 0 \text{ und } u_{RC} = U_0 \text{ für } t > 0.$$

Diese Spannung ist (Maschengleichung) die Summe der Spannungen über dem Widerstand u_R und über der Kapazität u :

$$u_{RC} = u_R + u = Ri + u.$$

Dabei haben wir das Ohm'sche Gesetz für den Widerstand benützt. Nun ersetzen wir den Strom i mit Hilfe der Differentialgleichung, welche die Kapazität beschreibt und erhalten die einfache Differentialgleichung erster Ordnung:

$$RC \frac{du}{dt} + u = u_{RC}.$$

In dieser Differentialgleichung ist u_{RC} die *Inhomogenität*. Für $t < 0$ ist die Lösung trivial, d.h. $u = 0$. Für $t > 0$ ist offenbar $u = U_0$ eine einfache, *spezielle Lösung* dieser Differentialgleichung. Wie aus der Analysis bekannt ist, müssen wir dazu eine *allgemeine Lösung* der zugehörigen *homogenen Differentialgleichung*

$$RC \frac{du}{dt} + u = 0 \text{ bzw. } \frac{du}{dt} + u/RC = 0$$

überlagern. Bekanntlich sind diese Lösungen Exponentialfunktionen der Form

$$u_{\text{homogen}}(t) = A \exp(-t/RC),$$

wobei A eine aus der Anfangsbedingung $u(0) = 0$ zu bestimmende Amplitude ist. Für $t = 0$ gilt:

$$u_{inhomogen} + u_{homogen} = U_0 + A \exp(-0/RC) = U_0 + A = 0 \Rightarrow A = -U_0.$$

Damit erhalten wir für die Spannung u über der Kapazität nach dem Einschalten, d.h. für $t > 0$:

$$u(t) = U_0 - U_0 \exp(-t/RC) = U_0[1 - \exp(-t/RC)].$$

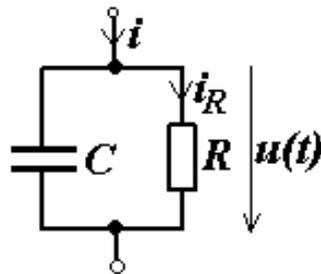
Zur Veranschaulichung dieser Lösung siehe Figur 26. Eine charakteristische Grösse, welche die Geschwindigkeit des Aufladevorgangs beschreibt, ist die *Zeitkonstante*

$$T = RC.$$

Offenbar ist T proportional zu C und R , d.h. die Aufladung braucht um so mehr Zeit, je grösser R und C sind. Nachdem die Zeit $t = T$ verstrichen ist, ist die Spannung an der Kapazität auf $U_0(1 - 1/e)$ angewachsen, sie hat also etwa 63% des asymptotischen Wertes U_0 erreicht, welchen sie nach unendlich langer Zeit erreichen würde.

In der Praxis wird der Aufladevorgang immer irgendwann abgebrochen. Wird die Kapazität vom Netzwerk getrennt, so dass kein Strom mehr fließen kann, so bleibt die Spannung konstant.

Betrachtet man aufgeladene, reale Kondensatoren, so stellt man fest, dass diese sich mit der Zeit entladen, weil die verwendeten Isolatoren zwischen den Kondensatorplatten nichtideal sind und eine gewisse Leitfähigkeit aufweisen. Um diese Nichtidealität zu berücksichtigen, kann man reale Kondensatoren mit einer Ersatzschaltung bestehend aus einer Kapazität C und einem parallelen Leitwert G bzw. Widerstand $R = 1/G$ approximieren.



Figur 27: Ersatzschaltung eines realen Kondensators (mit Innenwiderstand R , bzw. Innenleitwert $G = 1/R$) und Verbraucherpeilsystem.

Nehmen wir an, ein solcher Kondensator habe zur Zeit $t = 0$ die Spannung $u(0) = U_0$ und der Strom durch seine Klemmen sei für alle Zeiten Null, so fließt im Kondensator der Strom $i_R(t)$ von der Kapazität zum Widerstand. Mit dem Ohm'schen Gesetz und der Differentialgleichung, welche die Kapazität beschreibt, erhalten wir sofort wieder eine *einfache Differentialgleichung erster Ordnung*:

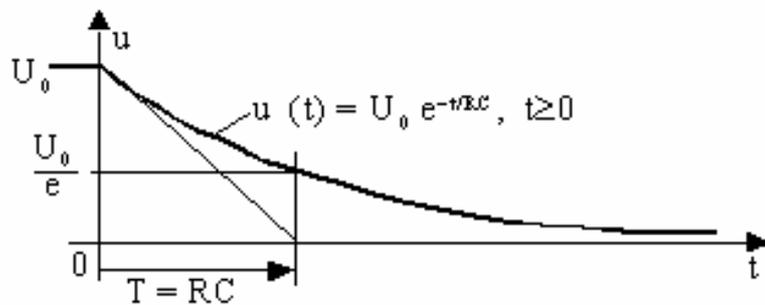
$$RC \, du/dt + u = 0.$$

Offenbar ist dies die homogene Differentialgleichung, der wir schon beim Aufladevorgang begegnet sind. Es gilt also wieder

$$u(t) = A \exp(-t/RC).$$

Unterschiedlich ist hingegen die Anfangsbedingung und damit die Amplitude A . Ausserdem fehlt die Inhomogenität. Es ist also keine spezielle Lösung zur allgemeinen Lösung zu überlagern. Für A finden wir mit der Anfangsbedingung $u(0) = U_0$ sofort $A = U_0$ und damit:

$$u(t) = U_0 \exp(-t/RC).$$



Figur 28: Entladung eines Kondensators.

Wegen des Ohm'schen Gesetzes ($i_R = u/R$) ist der Entladestrom i_R übrigens zu u proportional, nimmt also ebenso wie die Spannung u exponentiell mit der Zeitkonstante $T = RC$ ab.

3.1.3. Spulen und Induktivität

Das Induktionsgesetz besagt, dass die zeitliche Änderung eines Magnetfeldes ein (verwirbeltes) elektrisches Feld erzeugt. Da Ströme in Drähten (verwirbelte) Magnetfelder erzeugen, ist klar, dass zeitlich veränderliche Ströme zeitlich veränderliche Magnetfelder erzeugen, welche ihrerseits elektrische Felder erzeugen die schliesslich auf die Ladungen in den Drähten einwirken und so Spannungen in die Drähte „induzieren“. Dieser Effekt wird Selbstinduktion genannt. Es ist klar, dass dieser Effekt von der zeitlichen Änderung der Ströme abhängt. Freundlicherweise verhält sich die Natur so, dass die induzierte Spannung proportional zur ersten Ableitung des Stromes ist, d.h. wir schreiben (ohne auf die physikalischen Hintergründe genauer einzugehen):

$$u(t) = L \, di(t)/dt.$$

Dabei ist L (Einheit Henry, $1H = 1Vs/1A$) die Induktivität des Drahtes. Die Grösse der Induktivität hängt von der Drahtform ab. Um möglichst grosse Induktivitäten zu erhalten,

wickelt man üblicherweise Drähte auf und erhält damit Spulen. Näherungsformeln zur Berechnung der Induktivität einiger Spulenformen sind im Anhang zu finden.

Der magnetische Fluss einer Spule mit N Windungen ist N mal so gross wie der Fluss einer einzigen Windung. Damit wird in jede der N Windungen die N -fache Spannung, in die gesamte Spule jedoch die N^2 -fache Spannung induziert. Die Induktivität von Spulen ist also proportional zu N^2 .

Es ist zu beachten, dass auch gerade Drähte eine Induktivität haben. Diese ist jedoch so klein, dass sie meist vernachlässigt werden kann — ausser bei sehr raschen zeitlichen Veränderungen der Ströme und langen Drähten. In der Schaltungstechnik arbeiten wir deshalb nach wie vor mit idealisierten Drähten und setzen allenfalls Induktivitäten für die realen Drähte ein, wenn dies nötig ist.

Da Induktivitäten oft die Form von Spulen haben, wird ein Schaltsymbol verwendet, welches eine stilisierte Spule zeigt:



Figur 29: Induktivität mit Verbraucherpfeilsystem.

Offenbar hat die Differentialgleichung, welche die Induktivität beschreibt dieselbe Form, wie die Differentialgleichung, welche die Kapazität beschreibt. Um die beiden Gleichungen ineinander überzuführen, braucht man nur u mit i und L mit C zu vertauschen. Dies nennt man Dualität. Es gilt insbesondere:

$$\begin{array}{lcl}
 C & \leftrightarrow & L \\
 u(t) & \leftrightarrow & i(t) \\
 i(t) & \leftrightarrow & u(t) \\
 \\
 i(t) = C \frac{du(t)}{dt} & \leftrightarrow & u(t) = L \frac{di(t)}{dt} \\
 \\
 u(t) = \frac{1}{C} \int_0^t i(x) dx + u(0) & \leftrightarrow & i(t) = \frac{1}{L} \int_0^t u(x) dx + i(0)
 \end{array}$$

Die Dualität geht sogar noch weiter. So sind ideale Stromquellen zu idealen Spannungsquellen, Widerstände zu Leitwerten, Kurzschlüsse zu Leerläufen, Parallelschaltungen zu Serieschaltungen, Dreiecksschaltungen zu Sternschaltungen usw. dual. Selbst die einfache *Ersatzschaltung* eines realen Kondensators (Kapazität C parallel zu Leitwert G) ist dual zur einfachen Ersatzschaltung einer *realen Spule*. Diese besteht aus einer Induktivität L in Serie zu einem Widerstand R . Dabei berücksichtigt R die Ohm'schen Verluste in den Drähten der Spule.

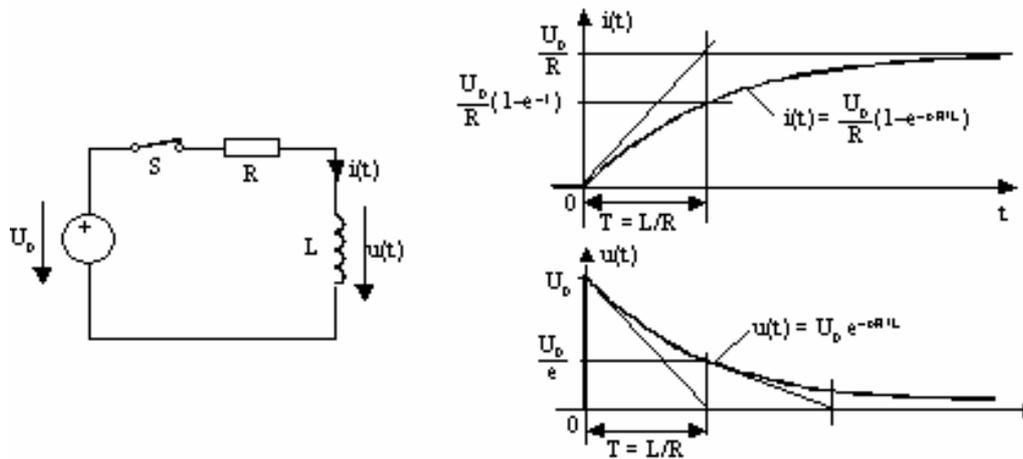
Wegen der Dualität können wir beispielsweise für die Schaltvorgänge mit einer Induktivität die dazu duale Schaltung mit einer Kapazität analysieren und die Resultate übertragen. Beispielsweise haben wir im vorhergehenden Unterabschnitt für die Entladung eines Kondensators über seinen Innenleitwert $G = 1/R$ folgende Lösung gefunden:

$$u(t) = U_0 \exp(-t/RC) = U_0 \exp(-tG/C).$$

Für die „Entladung“ einer Spule über ihren Innenwiderstand (d.h. wenn die Spulenden kurzgeschlossen werden – was zum Offenlassen der Anschlüsse des Kondensators dual ist) ergibt sich also:

$$i(t) = I_0 \exp(-t/GL) = U_0 \exp(-tR/L).$$

Dabei ist die Zeitkonstante $T = GL = L/R$. Die folgende Figur zeigt das Einschalten einer Spule.

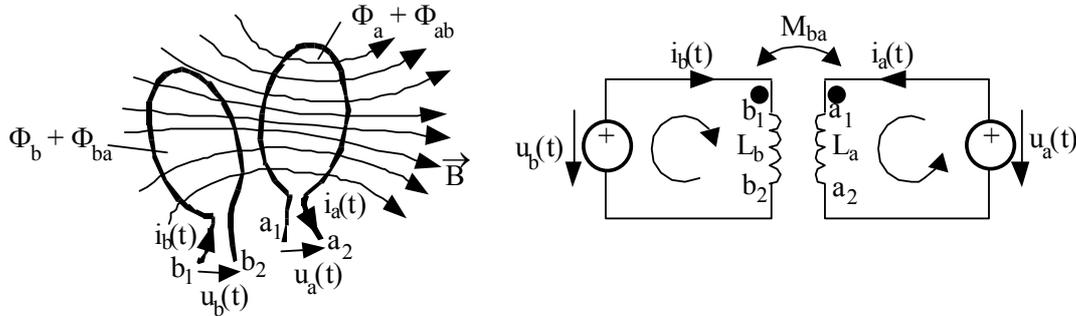


Figur 30: Einschaltvorgang einer Spule mit Innenwiderstand.

3.1.4. Gegeninduktivität

Da sich das Magnetfeld eines Stromes nicht auf die unmittelbare Umgebung des entsprechenden Drahtes beschränkt, induziert die zeitliche Änderung eines Stromes in einem Draht a nicht nur Spannungen in den Draht selbst, sondern auch Spannungen in allen andern Drähten einer elektrischen Schaltung. Diese Gegeninduktion kann zu unerwünschtem Fehlverhalten führen. Praktisch werden elektrische Schaltungen deshalb so ausgelegt, dass die entsprechenden Gegeninduktivitäten vernachlässigbar klein sind. Dies geschieht durch Vermeiden von Drähten, welche grosse Schlaufen bilden, weil die Gegeninduktivität ebenso wie die Selbstinduktivität mit der Schlaufenfläche zunimmt. In der Energietechnik und auch in der Nachrichtentechnik werden Gegeninduktivitäten aber auch bewusst ausgenutzt.

Die folgende Figur veranschaulicht den „Mechanismus“ der Gegeninduktion und zeigt das entsprechende Schaltschema, in dem die Gegeninduktivität $M_{ab} = M_{ba}$ durch einen Vierpol dargestellt wird. Hinweis: $M_{ab} = M_{ba}$ gilt für alle beliebigen Schlaufenformen, d.h. die Einwirkung von Schlaufe a auf Schlaufe b ist dieselbe wie die Einwirkung von Schlaufe b auf Schlaufe a . Die Gegeninduktion wird also durch eine einzige Grösse M beschrieben.



Figur 31: Anordnung und Ersatzschema zweier Stromschleifen mit Selbst- (L_a bzw. L_b) und Gegeninduktivität ($M_{ab} = M_{ba}$).

Die Punkte neben den Induktivitäten L_a und L_b in Figur 31 bestimmen das Vorzeichen der Glieder mit M . Für die Bestimmung des Vorzeichens von M ist die Drehrichtung der Schleifen zu beachten.

Für die Spannungen an den beiden Schleifen gelten die folgenden gekoppelten Differentialgleichungen:

$$u_a(t) = L_a di_a(t)/dt + M di_b(t)/dt$$

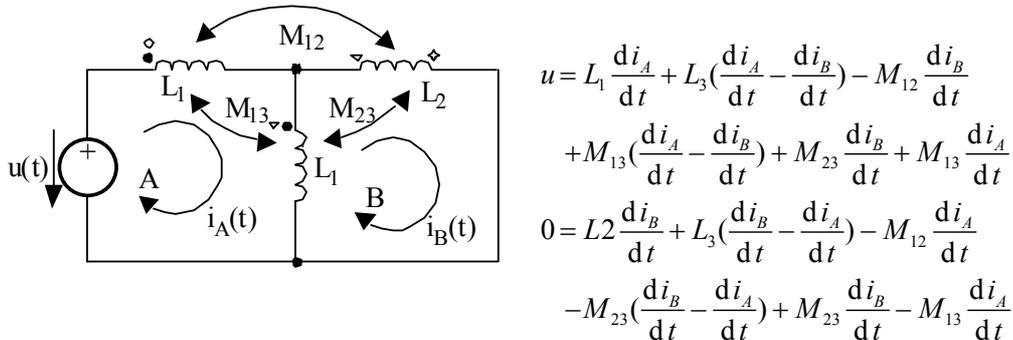
$$u_b(t) = L_b di_b(t)/dt + M di_a(t)/dt.$$

Wie schon die Induktivität ist auch die Gegeninduktivität von Spulen mit zunehmender Windungszahl grösser. Sind N_a und N_b die Windungszahlen der Schleifen a und b , so ist M proportional zu $N_a N_b$. Wichtig ist, dass die von Schlaufe a in Schlaufe b induzierte Spannung je nach Verhältnis der Windungszahlen unterschiedlich gross wird. Dies ermöglicht es, Spannungen zu „transformieren“ (siehe folgendes Kapitel).

Will man die Gegeninduktion nutzen, so arbeitet man oft mit grossen Windungszahlen. Darüber hinaus nimmt die Gegeninduktion zu, wenn ein möglichst grosser Teil des Flusses einer Spule die andere durchdringt. Beim Bau von Transformatoren werden deshalb Massnahmen ergriffen um dies zu erreichen. Dies wird insbesondere durch Materialien mit hoher Permeabilität erreicht. Werden Spulen um derartige Materialien gewickelt, so wird der Fluss des Magnetfeldes hauptsächlich im hochpermeablen Material geführt. Damit kann erreicht werden, dass nahezu der gesamte Fluss einer Spule auch durch die andere Spule fliesst. Ein bekanntes und relativ billiges, hochpermeables Material ist Eisen. Spulen und Transformatoren haben deshalb oft Eisenkerne. Allerdings hat Eisen relativ komplizierte, physikalische Eigenschaften, die insbesondere zu Verlusten und nichtlinearem Verhalten

führen, so dass Eisenkerne nicht unproblematisch sind und bei rasch veränderlichen Strömen nicht verwendet werden.

Sind mehrere Gegeninduktivitäten in einem Netzwerk enthalten, so müssen diese einzeln berücksichtigt werden. Figur 32 zeigt ein Beispiel dafür.



Figur 32: Netzwerk mit Gegeninduktivitäten und die entsprechenden Gleichungen

Die Gleichungen von Figur 32 können übrigens durch Einsetzen von

$$L_A = L_1 + L_3 + 2M_{13}, L_B = L_2 + L_3 - 2M_{23} \text{ und } L_{AB} = L_3 + M_{13} + M_{12} - M_{23}$$

vereinfacht werden (man beachte die Symmetrie):

$$u(t) = L_A di_A(t)/dt - L_{AB} di_B(t)/dt \text{ und } 0 = -L_{AB} di_A(t)/dt + L_B di_B(t)/dt.$$

3.2. Physikalische Grössen

Bisher haben wir uns nur mit zeitveränderlichen Spannungen und Strömen abgegeben. Das sind die primären Grössen, welche in einem Netzwerk berechnet werden. Im Falle von Gleichstrom haben wir bereits gesehen, dass andere physikalische Grössen für Zweipole berechnet werden können, wenn die Spannung über dem Zweipol und der Strom, welcher durch den Zweipol fliesst, bekannt sind. So wandeln verlustbehaftete Zweipole elektrische Energie in Wärme um, während Quellen Energie speichern können.

3.2.1. Leistung

Wir haben gesehen, dass für die Leistung $P = UI$ gilt. Diese Formel kann für zeitabhängige Grössen direkt übernommen werden. Es gilt für die *momentane* Leistung eines Zweipols:

$$p(t) = u(t)i(t).$$

Für einen idealen Schalter ist $p(t) = 0$, da in jedem Schaltzustand entweder $u = 0$ oder $i = 0$ gilt. Dies gilt auch für eine ideale Diode. Praktisch nehmen Schalter jedoch Energie auf und zwar vor allem während des Umschaltvorganges, welcher eine gewisse Zeit benötigt. Insbesondere bei elektronischen Schaltungen wie Computern, welche Millionen von Schaltern beinhalten, welche ihren Zustand in sehr rascher Folge ändern, ist es daher wichtig, die Schalter so zu optimieren, dass die Verlustleistung gering bleibt.

Wir haben gesehen, dass der Strom einer Kapazität proportional zur zeitlichen Ableitung der Spannung ist. Also gilt für die Momentanleistung einer Kapazität

$$p(t) = Cu(t) \, du(t)/dt.$$

Analog können wir bei der Formel für die Momentanleistung einer Induktivität die Spannung eliminieren und erhalten

$$p(t) = Li(t) \, di(t)/dt.$$

3.2.2. Arbeit und Energie

Zum Aufbau des elektrischen Feldes in einer Kapazität und des Magnetfeldes in einer Induktivität muss Arbeit geleistet werden. Da die entsprechende *Energie* weder in Wärme – wie beim Ohm'schen Widerstand – noch in eine andere Energieform (wie zum Beispiel bei Motoren) umgewandelt wird, speichern Kapazitäten und Induktivitäten die Energie, welche ihnen beim Ladevorgang zugeführt werden. Aus der Physik wissen wir, dass die Energie das Zeitintegral über die Leistung ist. Also gilt für eine Kapazität

$$w(t) = C \int_0^t u(x) \frac{du(x)}{dx} dx = C \frac{u^2(t) - u^2(0)}{2}.$$

Ist die Kapazität zur Zeit $t = 0$, d.h. beim Beginn des Ladevorgangs, ungeladen, so ist $u(0) = 0$ und wir erhalten

$$w(t) = \frac{Cu^2(t)}{2}.$$

Ganz analog erhalten wir für die gespeicherte Energie einer Induktivität

$$w(t) = \frac{Li^2(t)}{2}.$$

Kapazitäten und Induktivitäten können die gespeicherte Energie ans Netzwerk abgeben und damit ähnlich wie Strom- und Spannungsquellen wirken.

3.2.3. Kraft

Da in Kapazitäten Ladungen und elektrische Felder, in Induktivitäten Ströme und magnetische Felder vorhanden sind, treten auch Kräfte auf, welche beim Bau von entsprechenden Bauelementen berücksichtigt werden müssen. Bei zeitabhängigen Grössen werden natürlich auch die Kräfte zeitabhängig. Dabei können jedoch einfach die Formeln aus dem ersten Kapitel übernommen werden.

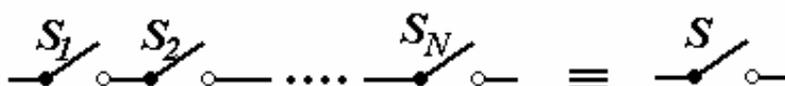
3.3. Serie- und Parallelschaltungen

3.1.1. Serieschaltungen

Bei Serieschaltungen von beliebigen Bauelementen summieren sich die Spannungen, während der Strom, welcher durch die Bauelemente fliesst, ein und derselbe ist. Daran ändert sich nichts, wenn Ströme und Spannungen zeitveränderlich sind. Es gilt also für die Serieschaltung von N beliebigen Zweipolen

$$u(t) = \sum_{n=1}^N u_n(t); \quad i(t) = i_n(t) \forall n.$$

Bei einer Serieschaltung von N idealen Schaltern S_n fliesst offenbar nur dann ein Strom, wenn alle Schalter geschlossen sind. Zur Beschreibung des „Ersatzschalters“ S , welcher dieselbe Funktion wie die Serieschaltung hat, werden also *logische Verknüpfungen* gebraucht: S ist geschlossen, wenn S_1 und S_2 und ... und S_N geschlossen ist. S ist offen, wenn S_1 oder S_2 oder ... oder S_N offen ist. Schalter sind denn auch das wichtigste Bauelement für logische Schaltungen.



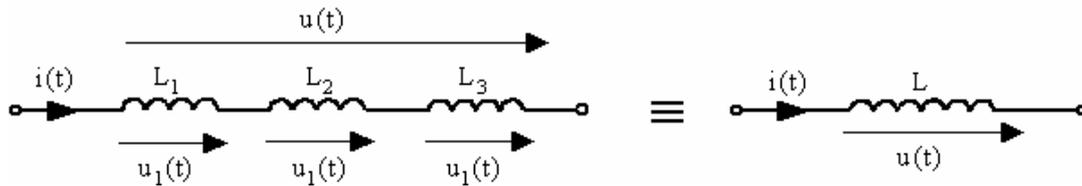
Figur 33: Serieschaltung von N Schaltern.

Auch eine Serieschaltung von N Induktivitäten L_n kann leicht durch eine einzige Induktivität L ersetzt werden. Es gilt:

$$u(t) = \sum_{n=1}^N u_n(t) = \sum_{n=1}^N L_n \frac{d}{dt} i_n(t) = \frac{di(t)}{dt} \sum_{n=1}^N L_n$$

und somit:

$$L = \sum_{n=1}^N L_n .$$

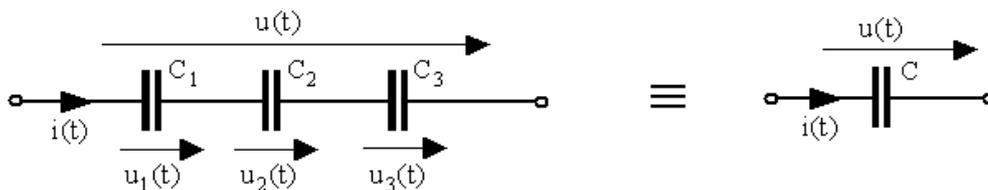


Figur 34: Serieschaltung von N Induktivitäten.

Die Serieschaltung von N Induktivitäten ist also formal identisch zur Serieschaltung von N Widerständen.

Die Serieschaltung von N Kapazitäten C_n ist etwas komplizierter. Sie ist formal identisch zur Serieschaltung von N Leitwerten. Man erhält für die Gesamtkapazität C :

$$1/C = \sum_{n=1}^N 1/C_n .$$



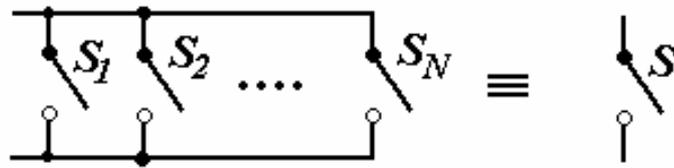
Figur 35: Serieschaltung von N Kapazitäten.

3.1.2. Parallelschaltungen

Bei Parallelschaltungen von beliebigen Bauelementen summieren sich die Ströme, während die Spannung über allen Bauelementen dieselbe ist. Daran ändert sich nichts, wenn Ströme und Spannungen zeitveränderlich sind. Es gilt also für die Parallelschaltungen von N beliebigen Zweipolen

$$i(t) = \sum_{n=1}^N i_n(t); \quad u(t) = u_n(t) \forall n .$$

Bei einer Parallelschaltung von N idealen Schaltern S_n fließt offenbar immer dann ein Strom, wenn wenigstens ein Schalter geschlossen ist. Zur Beschreibung des „Ersatzschalters“ S , welcher dieselbe Funktion wie die Serieschaltung hat, werden also *logische Verknüpfungen* gebraucht: S ist geschlossen, wenn S_1 oder S_2 oder ... oder S_N geschlossen ist. S ist offen, wenn S_1 und S_2 und ... und S_N offen ist.

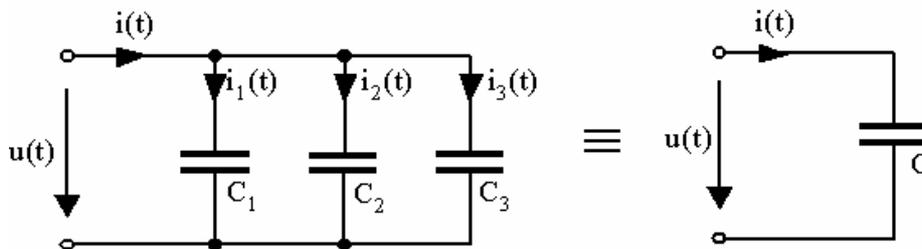


Figur 36: Parallelschaltung von N Schaltern.

Die Parallelschaltung von N Kapazitäten ist formal identisch zur Parallelschaltung von N Leitwerten und zur Serieschaltung von Induktivitäten, d.h. die Kapazitäten können einfach aufsummiert werden.

$$C = \sum_{n=1}^N C_n .$$

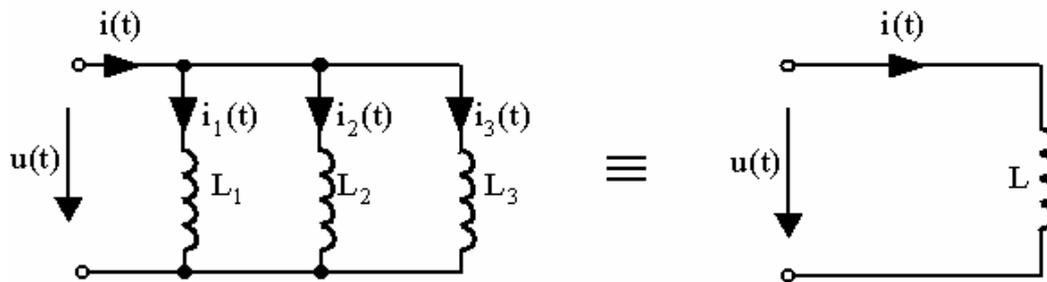
Dass diese Formel gilt, wird auch klar, wenn man sich N Plattenkondensatoren mit gleichem Plattenabstand vor Augen hält und diese zusammenfügt. Die erhaltene Gesamtfläche ist dann gleich der Summe der Elektrodenflächen der einzelnen Kondensatoren. Da die Kapazität zur Fläche proportional ist, ergibt sich obige Formel.



Figur 37: Parallelschaltung von N Kapazitäten.

Die Parallelschaltung von N Induktivitäten ist dual zur Serieschaltung von N Kapazitäten und formal identisch zur Parallelschaltung von N Widerständen. Es gilt also:

$$1/L = \sum_{n=1}^N 1/L_n .$$



Figur 38: Parallelschaltung von N Induktivitäten.

3.4. Dualität

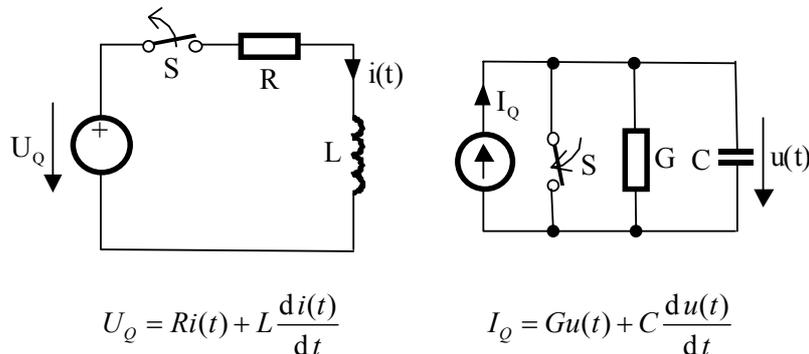
Wir haben bisher verschiedentlich von der Dualität von Schaltungen gesprochen und diese auch ausgenützt. Dabei geht es darum, dass die Analyse zweier dualer Netzwerke auf formal identische Gleichungen führt, wenn zueinander duale Grössen ausgetauscht werden. Wenn man die Dualität zweier Netzwerke erkennt und Lösung eines Netzwerkes kennt, so kann man also auch die Lösung des dualen Netzwerkes.

Die Dualität lässt sich auch ausnützen, wenn ein Teil eines Netzwerkes dual zu einem bereits analysierten Teilnetzwerk ist. Das ist insbesondere dann der Fall, wenn Serie- oder Parallelschaltungen von gleichartigen Zweipolen vorhanden sind, wie im vorhergehenden Abschnitt.

Die folgende Liste zeigt die wichtigsten zueinander dualen Grössen und Elemente:

Strom	↔	Spannung
Stromquelle	↔	Spannungsquelle
Leitwert	↔	Widerstand
Leerlauf	↔	Kurzschluss
Offener Schalter	↔	geschlossener Schalter
Kapazität	↔	Induktivität
Masche	↔	Knoten
Parallelschaltung	↔	Serieschaltung
Dreiecksschaltung	↔	Sternschaltung

Beispiel



Beide Schaltungen sind dual. Ihre Gleichungen, und deshalb auch ihre Lösungen, haben die gleiche mathematische Form und lassen sich durch Vertauschen der Grössen ineinander überführen.

3.5. Netzwerke

Die Berechnung von Netzwerken bei zeitabhängigen Strömen und Spannungen verläuft ganz ähnlich wie im Gleichstromfall. Zunächst versucht man, Mehrpole durch Ersatzschaltungen mit Zweipolen darzustellen, weil Netzwerke, welche nur Zweipole enthalten, leichter zu analysieren sind. Anschliessend werden die Netzwerke durch Umformungen und Zusammenfassen von Bauelementen so stark wie möglich vereinfacht. Für diese Netzwerke notiert man alle unabhängigen Knoten- und Maschengleichungen, sowie die Zweipolgleichungen. Das erhaltene Gleichungssystem wird schliesslich analytisch oder numerisch gelöst.

Sind mehrere Quellen in einem Netzwerk vorhanden, so wird versucht, die Quellen so umzuformen und zusammenzufassen, dass möglichst wenige, gleichartige Quellen (entweder Strom- oder Spannungsquellen) vorkommen. Gelingt die simultane Behandlung mehrerer Quellen nicht, so wird das Superpositionsprinzip verwendet, d.h. es wird für jede Quelle eine separate Schaltung analysiert, in der alle andern Spannungsquellen durch Kurzschlüsse und alle andern Stromquellen durch Leerläufe ersetzt werden. Anschliessend werden die Lösungen der einzelnen Quellen superponiert. Es reicht deshalb, eine Methode zur Analyse von Netzwerken mit einer einzigen, idealen Quelle zu beherrschen. Die *Quelle* wird übrigens stets zur *Inhomogenität* in den auftretenden Differentialgleichungen. *Passive* Netzwerke, welche keine Quellen enthalten, werden deshalb immer mit *homogenen* Differentialgleichungen beschrieben.

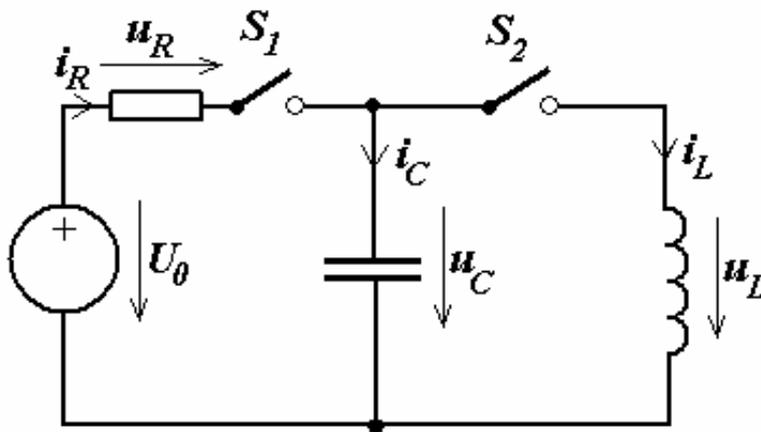
Enthält ein Netzwerk nur lineare Zweipole und ausserdem entweder keine Kapazitäten oder keine Induktivitäten, so erhält man einfache Differentialgleichungen *erster Ordnung*, welche nach der bereits skizzierten Methode gelöst werden. Dabei werden zuerst die Anfangsbedingungen (meist mit Gleichstromanalyse) bestimmt. Dann wird eine spezielle Lösung

der inhomogenen Differentialgleichungen und die allgemeine Lösung der zugehörigen homogenen Differentialgleichungen gesucht, die Lösungen überlagert und die Unbekannten im Lösungsansatz so bestimmt, dass die Anfangsbedingungen erfüllt werden.

Enthält ein Netzwerk sowohl Kapazitäten als auch Induktivitäten, so können sich Differentialgleichungen *zweiter Ordnung* ergeben, welche etwas schwieriger zu lösen sind. Typische Beispiele sind Schwingkreise, welche im nächsten Kapitel ausführlicher untersucht werden.

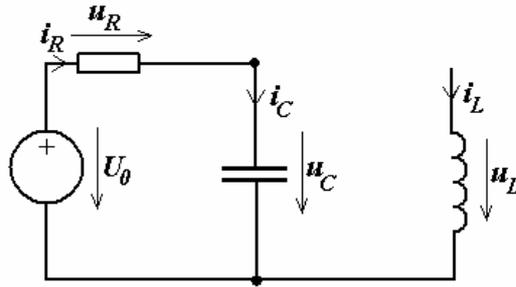
Sind mehrere Schalter vorhanden, welche ihren Zustand zu verschiedenen Zeiten ändern, so muss schrittweise — längs der Zeitachse — vorgegangen werden. Dabei kann sich für jeden Schaltzustand ein unterschiedliches Netzwerk ergeben, welches während des i -ten Zeitintervalls $t_i^a < t < t_i^e$ "in Betrieb" ist. Der Anfangszustand (zum Zeitpunkt t_i^a) ist gleichzeitig der Endzustand (zum Zeitpunkt t_{i-1}^e) des Netzwerks, welches während des vorhergehenden, $i-1$ -ten Zeitintervalls "in Betrieb" war und zuvor analysiert wurde.

Zur Verdeutlichung betrachten wir eine einfache Schaltung gemäss Figur 39 mit einer idealen Spannungsquelle mit konstanter Spannung U_0 , einem Widerstand R , einer Kapazität C , einer Induktivität L und zwei Schaltern S_1 und S_2 , welche zu unterschiedlichen Zeiten betätigt werden und zwar wie folgt: S_1 wird zur Zeit $t = 0$ geschlossen und bei $t = t_1$ wieder geöffnet. Danach wird S_2 zur Zeit $t = t_2$ geschlossen. Ausserdem seien die Kapazität und die Induktivität für $t < 0$ ungeladen, d.h. für $t < 0$ sind alle Ströme und Spannungen Null, ausser der Spannung über der Quelle, welche natürlich U_0 beträgt. Dies sind die Anfangsbedingungen für das erste Zeitintervall $0 \leq t < t_1$ vom Schliessen bis zum Öffnen von S_1 .



Figur 39: Schaltung mit zwei Schaltern S_1 und S_2 .

1) Erstes Zeitintervall $0 \leq t < t_1$: Da die Spule nicht angeschlossen ist, bleiben i_L und u_L Null und es muss lediglich der linke Teil der Schaltung berechnet werden:



Figur 39a: Schaltung mit S_1 geschlossen und S_2 offen.

Dieser Schaltung sind wir in Figur 26 bereits begegnet. Dabei wird die Kapazität aufgeladen und es gilt:

$$u_C(t) = U_0 - U_0 \exp(-t/RC) = U_0[1 - \exp(-t/RC)].$$

Daraus erhalten wir von der Zweitorgleichung für die Kapazität sofort die Ströme

$$i_R(t) = i_C(t) = C du_C/dt = U_0/R \exp(-t/RC)$$

und aus der Zweitorgleichung für den Widerstand schliesslich auch

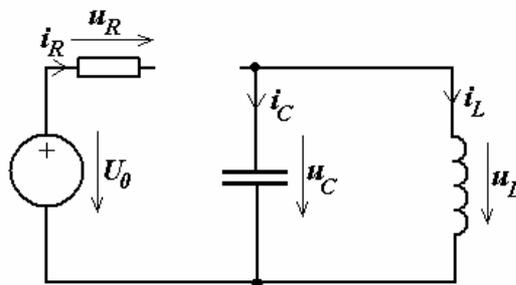
$$u_R(t) = R i_R(t) = U_0 \exp(-t/RC).$$

2) Zweites Zeitintervall $t_1 \leq t < t_2$: Beide Schalter sind wieder offen. Also fließen keine Ströme und die Spannung über der Kapazität bleibt konstant, nimmt also den Endwert des vorhergehenden Zeitintervalls an:

$$u_C(t) = u_C(t_1) = U_0 - U_0 \exp(-t_1/RC) = U_0[1 - \exp(-t_1/RC)].$$

Dieser Wert ist auch der Endwert dieses Zeitintervalls und der Anfangswert des folgenden Zeitintervalls.

3) Drittes Zeitintervall $t_2 < t < t_3$: Nun bleibt S_1 offen, während S_2 geschlossen ist:



Figur 39b: Schaltung mit S_1 offen und S_2 geschlossen.

Die Spannungsquelle und der Widerstand sind offenbar isoliert und es gilt $i_R = u_R = 0$. Wir müssen also noch den rechten Teil der Anordnung berechnen. Dieser besteht aus einer Kapazität und einer Induktivität, welche einen einfachen LC Schwingkreis bilden. Dieser wird durch die Knotengleichung

$$i_L = -i_C = i$$

die Maschengleichung

$$u_L = u_C = u$$

und die Zweitorgleichungen

$$Cdu/dt = -i; \quad Ldi/dt = u$$

beschrieben. In dieser Form haben wir übrigens die Knoten- und die Maschengleichung bereits berücksichtigt. Offenbar haben wir zwei gekoppelte Differentialgleichungen erster Ordnung für die Unbekannten u und i . Um diese zu entkoppeln, leiten wir die eine Gleichung nach der Zeit ab und setzen die andere ein. Dies ist übrigens das Standardvorgehen bei gekoppelten Differentialgleichungen. Damit erhalten wir folgende Differentialgleichung für u :

$$u = Ldi/dt = -LCd^2u/dt^2 \text{ oder: } LCD^2u/dt^2 + u = 0.$$

Dies ist eine einfache, homogene Differentialgleichung zweiter Ordnung (eine homogene Schwingungsdifferentialgleichung) mit der allgemeinen Lösung:

$$u(t) = A\cos(\omega t) + B\sin(\omega t),$$

wobei $\omega = 1/\sqrt{LC}$ die Kreisfrequenz der Schwingung ist. Wir müssen nun noch die unbekannt Amplituden A und B aus der Anfangsbedingungen $u(t_2) = U_0(1 - U_0 \exp(-t_1/RC))$ und $i(t_2) = 0$ berechnen. Dazu ist es sinnvoll, die Substitution $t' = t - t_2$ zu verwenden und die obige allgemeine Lösung durch

$$u(t') = A'\omega\cos(\omega t') - B'\omega\sin(\omega t'),$$

zu ersetzen. Dann gilt für den Strom

$$i(t') = -Cdu/dt' = CA'\omega\sin(\omega t') + CB'\omega\cos(\omega t').$$

Die Anfangsbedingung für den Strom ist also mit $B' = 0$ erfüllt. Aus der Anfangsbedingung für die Spannung erhält man dann sofort $A' = u(t'=0) = u(t_1)$.

Im eben betrachteten LC Schwingkreis verlaufen Strom und Spannung offenbar sinusförmig, solange keiner der beteiligten Schalter seinen Zustand ändert. In der Energie- und in der Nachrichtentechnik werden sinusförmige Größen sehr häufig verwendet. Diese entstehen nicht nur in Schwingkreisen, sondern auch auf natürliche Weise in Wechselstromgeneratoren, bei denen im Wesentlichen eine mechanische Drehbewegung „automatisch“ einen

sinusförmigen Zeitverlauf ergibt. Im folgenden Kapitel werden derartige Zeitverläufe genauer untersucht und Hilfsmittel vorgestellt, welche die Analyse so vereinfachen, dass das explizite Lösen von Differentialgleichungen entfällt.

4. Wechselstrom

Um 1900 gab es zwei wesentliche Anwendungen für die Elektrotechnik: 1) Die Energietechnik, bei der man versuchte, Gasbeleuchtungen durch elektrische Lampen und Dampfmaschinen durch elektrische Motoren abzulösen und 2) die Nachrichtentechnik, bei der man versuchte, elektrische Signale über Drähte (Telegraphen, Telefon) oder mit Antennen über den freien Raum zu übertragen.

Der bedeutendste Pionier der Elektrotechnik war zweifellos Edison, welcher nicht nur die elektrische Glühlampe erfand, sondern auch mit enormem Einsatz nach weiteren, mehr oder weniger nutzbringenden Anwendungen der elektrischen Phänomene suchte und diese in einer stark wachsenden „High Tech“ Firma, der heute noch existierenden „General Electric Company“, ausbeutete bzw. vermarktete. Diese Anwendungen erhöhten natürlich den Bedarf an elektrischer Energie. Die einzige Möglichkeit, effizient elektrische Energie zu erzeugen bestand damals in der Umformung mechanischer Energie in elektrische Energie mittels Generatoren. Diese bestehen im Wesentlichen aus Spulen und Permanentmagneten und nutzen das Induktionsgesetz. Die mechanische Drehbewegung führt dazu, dass in den Spulen eine Spannung mit sinusförmigem Zeitverlauf $u(t) = \dot{U} \sin \omega t$ induziert wird. Um Gleichstrom zu erhalten, wurden auf den Achsen der Generatoren Umschalter (sogenannte Kommutatoren) aufgebracht, welche dafür sorgten, dass die Ausgangsspannung gleichgerichtet wurde, d.h. $u(t) = \dot{U} |\sin \omega t|$. Tesla erkannte, dass dadurch verschiedene Nachteile entstanden und schlug vor, für die Energieversorgung Wechselstromgeneratoren zu verwenden. Dies führte zu einem langen, erbitterten Streit zwischen Edison und Tesla. Schliesslich setzte sich die Energieversorgung mit Wechselstrom fast überall durch.

Der sinusförmige Zeitverlauf hat nicht nur technische, sondern auch mathematische Vorteile. Erinnern wir uns daran, dass wir im vorangehenden Kapitel Differentialgleichungen mit ersten und zweiten Ableitungen nach der Zeit begegnet sind und dass die Behandlung der resultierenden Gleichungssysteme im Vergleich zu den Gleichungssystemen bei Gleichstrom sehr mühsam ist. Dies mag Edisons Vorliebe für Gleichstrom erklären. Sind alle zeitveränderlichen Grössen $f(t)$ sinusförmig, so kann man diese gemäss

$$f(t) = A \sin(\omega t + \varphi)$$

mit drei skalaren, reellen Grössen, der Amplitude A , der Kreisfrequenz ω und der Phase φ beschreiben. Die Beschreibung von sinusförmigen Grössen ist deshalb zwar etwas schwieriger als die Beschreibung von konstanten Grössen im Gleichstromfall, aber doch erheblich einfacher als die Beschreibung beliebiger zeitveränderlicher Grössen. Ausserdem erweist sich die zeitliche Ableitung

$$df(t)/dt = A \omega \cos(\omega t + \varphi) = A \omega \sin(\omega t + \varphi + \pi/2)$$

als sehr einfach.

In diesem Kapitel werden wir nun Hilfsmittel zur bequemen und effizienten Behandlung von Wechselstromgrößen vorstellen, das Wechselstromverhalten der wichtigsten Zweipole betrachten und einfache Netzwerke von Zweipolen analysieren.

4.1. Fourieranalyse

4.1.1. Fourierintegrale

Im Prinzip lässt sich jede zeitveränderliche Funktion $f(t)$, welche für $-\infty < t < +\infty$ definiert ist und über t integrierbar ist, durch eine Überlagerung sinusförmiger Anteile mit unterschiedlichen Frequenzen beliebig genau approximieren, bzw. mittels Fourierintegralen darstellen. Es existieren verschiedene Formen von Fourierintegralen. Ein Beispiel ist:

$$f(t) = \int_0^{\infty} A(\omega) \cos(\omega t + \varphi(\omega)) d\omega.$$

Dabei sind die Amplitude und die Phase Funktionen der Frequenz. Die Bestimmung des Phasengangs, d.h., die Berechnung von $\varphi(\omega)$ ist etwas mühsam. Man verwendet deshalb das aus der Goniometrie bekannte Additionstheorem $\cos(\alpha - \beta) = \cos(\alpha)\cos(\beta) + \sin(\alpha)\sin(\beta)$ (mit $\alpha = \omega t$ und $\beta = \varphi(\omega)$) zur Umformung des obigen Integrals zu

$$f(t) = \int_0^{\infty} a(\omega) \cos(\omega t) - b(\omega) \sin(\omega t) d\omega.$$

Dabei ist offenbar $a(\omega) = A(\omega)\cos(\varphi(\omega))$, $b(\omega) = A(\omega)\sin(\varphi(\omega))$. Sind die Funktionen $a(\omega)$ und $b(\omega)$ bekannt, so erhält man daraus den Amplitudengang

$$A(\omega) = \sqrt{a^2(\omega) + b^2(\omega)}$$

und den Phasengang

$$\varphi(\omega) = \arctg(b(\omega)/a(\omega)).$$

Zur Berechnung der Funktionen $a(\omega)$ und $b(\omega)$ ist es bequem, die aus der komplexen Analysis bekannte, komplexe Erweiterung auszunützen. Fassen wir $a(\omega)$ und $b(\omega)$ als Real- bzw. Imaginärteil einer komplexen Funktion $F(\omega) = a(\omega) + jb(\omega)$ auf, so gilt

$$a(\omega) = \operatorname{Re}(F(\omega)); \quad b(\omega) = \operatorname{Im}(F(\omega)).$$

Ausserdem gilt bekanntlich $\exp(\pm jx) = \cos(x) \pm j \sin(x)$, wobei $j = \sqrt{-1}$. Wir können also den obigen Integranden wie folgt umschreiben:

$$a(\omega) \cos(\omega t) - b(\omega) \sin(\omega t) = \operatorname{Re}(F(\omega) \exp(j\omega t)).$$

Daraus erhalten wir die komplexe Form des Fourierintegrals:

$$f(t) = \operatorname{Re} \int_0^{\infty} F(\omega) \exp(j\omega t) d\omega = \frac{1}{2} \int_{-\infty}^{\infty} F(\omega) \exp(j\omega t) d\omega .$$

Dabei ist $F(\omega)$ der komplexe Frequenzgang, für welchen

$$F(\omega) = \frac{1}{\pi} \int_{-\infty}^{\infty} f(t) \exp(-j\omega t) dt$$

gilt. Daraus erhalten wir schliesslich den Amplitudengang und den Frequenzgang der gegebenen Funktion $f(t)$.

N.B. bei der obigen Herleitung wurde die Symmetrie $a(\omega) = a(-\omega)$ bzw. die Antisymmetrie $b(\omega) = -b(-\omega)$ von Real- bzw. Imaginärteil von $F(\omega)$ ausgenützt.

Die Theorie der Fourierintegrale ist nicht sehr einfach. Ausserdem sind neben den hier verwendeten Formen der Integrale, weitere Formen mit zum Teil anderen Normierungsfaktoren in der Literatur anzutreffen, was verwirrend ist. Wichtig ist, dass die *komplexe Erweiterung* eine formale Vereinfachung ergibt, welche das Auffinden der Lösung erleichtert. Von der komplexen Erweiterung werden wir auch in der Schaltungsanalyse profitieren.

4.1.2. Fourier-Reihen

Ist die gegebene Zeitfunktion $f(t)$ periodisch, d.h. gilt $f(t+T) = f(t)$, wobei T die Periodendauer ist, so vereinfacht sich die Fourier-Analyse erheblich. Zunächst muss nur noch über eine einzige Periode statt vom $-\infty$ bis $+\infty$ integriert werden. Aus den Orthogonalitätsrelationen für \sin und \cos ergibt sich ausserdem, dass das Fourier-Spektrum $F(\omega)$ nicht mehr *kontinuierlich*, sondern *diskret* wird. D.h. $F(\omega)$ wird Null, ausser für bestimmte Frequenzen, welche ein ganzzahliges Vielfaches der *Grundfrequenz*

$$\omega_0 = 2\pi/T$$

sind. Statt der Fourierintegrale erhält man Fourier Reihen, für welche ebenfalls verschiedene Formen verwendet werden. Z.B. gilt für eine T periodische Funktion:

$$f(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos(k\omega_0 t) - b_k \sin(k\omega_0 t)).$$

Die entsprechenden Fourier-Koeffizienten erhält man z.B. aus den Integralen

$$a_k = \frac{2}{T} \int_0^T f(t) \cos(k\omega_0 t) dt ; \quad b_k = -\frac{2}{T} \int_0^T f(t) \sin(k\omega_0 t) dt .$$

Die Fourier-Koeffizienten für viele periodische Funktionen sind übrigens tabelliert.

Sowohl Fourierintegrale als auch Fourier Reihen ermöglichen es uns, die explizite Lösung der Differentialgleichungssystem im *Zeitbereich*, welche das Zeitverhalten der Netzwerke beschreiben, zu umgehen. Enthält ein Netzwerk eine (oder mehrere) Quelle(n) mit gegebenem Zeitverlauf $f(t)$, so werden zunächst die Fourier-Koeffizienten von $f(t)$ bestimmt. Anschliessend werden die Lösungen für jede auftretende Frequenz ω im *Frequenzbereich* gesucht. Dabei profitiert man von den Beziehungen:

$$\frac{d}{dt} \sin(\omega t) = \omega \cos(\omega t) ; \frac{d}{dt} \cos(\omega t) = -\omega \sin(\omega t) ; \frac{d}{dt} \exp(\omega t) = \omega \exp(\omega t)$$

welche es im Wesentlichen ermöglichen, die zeitlichen Ableitungen durch einfache Multiplikationen mit ω zu ersetzen.

Im Folgenden setzen wir voraus, dass alle Quellen eines Netzwerkes ein- und denselben, harmonischen Zeitverlauf haben, also durch drei reelle Grössen, die Amplitude, die Phase und die Kreisfrequenz vollständig beschrieben werden können. Ist dies nicht der Fall, so können im periodischen Falle die Fourier Reihen und im allgemeinen Falle Fourierintegrale verwendet werden, um den gegebenen Zeitverlauf in seine harmonischen Komponenten zu zerlegen.

Zu beachten ist, dass sämtliche Spannungen und Ströme in einem linearen Netzwerk sinusförmig werden und dieselbe Frequenz haben, wenn alle Quellen einen sinusförmigen Verlauf mit einer bestimmten Frequenz aufweisen. Damit werden alle Ströme und Spannungen bei bestimmter Frequenz durch zwei reelle Grössen, die Amplitude und die Phase beschrieben. Dies wird im Folgenden vorausgesetzt.

4.2. Komplexe Darstellung und Zeiger

Wir haben bereits im vorhergehenden Abschnitt die komplexe Erweiterung verwendet um Fourierintegrale formal zu vereinfachen. Das wollen wir hier etwas ausführlicher betrachten. Sei

$$f(t) = \hat{A} \cos(\omega t + \varphi)$$

eine harmonische Funktion mit bestimmter Kreisfrequenz ω , Amplitude \hat{A} und Phase φ . Da die Amplitude gleichzeitig der Maximalwert oder Spitzenwert der Funktion f ist, kennzeichnen wir diesen Wert mit einem Dach. In einem Netzwerk kann f die Spannung $u(t)$ oder auch der Strom $i(t)$ sein. Mit dem Additionstheorem

$$\cos(\alpha + \beta) = \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta) \text{ (mit } \alpha = \omega t \text{ und } \beta = \varphi(\omega))$$

erhalten wir die phasenfreie Form

$$f(t) = a \cos(\omega t) - b \sin(\omega t)$$

mit den zwei Amplituden

$$a = \hat{A} \cos \varphi, \quad b = \hat{A} \sin \varphi.$$

Sind a und b bekannt, so erhält man daraus

$$\hat{A} = \sqrt{a^2 + b^2}, \quad \varphi = \arctg(b/a).$$

Nun können wir a in der *2-dimensionalen xy Ebene* in horizontaler (x) Richtung und b in vertikaler (y) Richtung auftragen, d.h. das reelle Wertepaar (a,b) als Komponenten eines Vektors in der 2-dimensionalen Ebene auffassen. Die Länge dieses Vektors entspricht offenbar der Amplitude \hat{A} , der Winkel welchen dieser Vektor mit der x Achse bildet, hingegen der Phase φ . \hat{A} und φ sind also die Polarkoordinaten des Punktes (a,b) in der xy Ebene.

Die 2-dimensionale xy Ebene können wir ebenso gut als *komplexe Ebene* mit $z = x+jy$ auffassen. In dieser Ebene beschreibt die komplexe Zahl $\underline{C} = a+jb$ die Amplituden a und b unserer harmonischen Funktion und es gilt

$$\hat{A} = \sqrt{\operatorname{Re}(\underline{C})^2 + \operatorname{Im}(\underline{C})^2} = |\underline{C}|, \quad \varphi = \arctg(\operatorname{Im}(\underline{C})/\operatorname{Re}(\underline{C})).$$

Zur Verdeutlichung haben wir übrigens die *komplexe Zahl \underline{C} unterstrichen* und ausserdem einen *Grossbuchstaben* verwendet, da diese Zahl nicht zeitvariabel sondern *konstant* ist.

Wegen

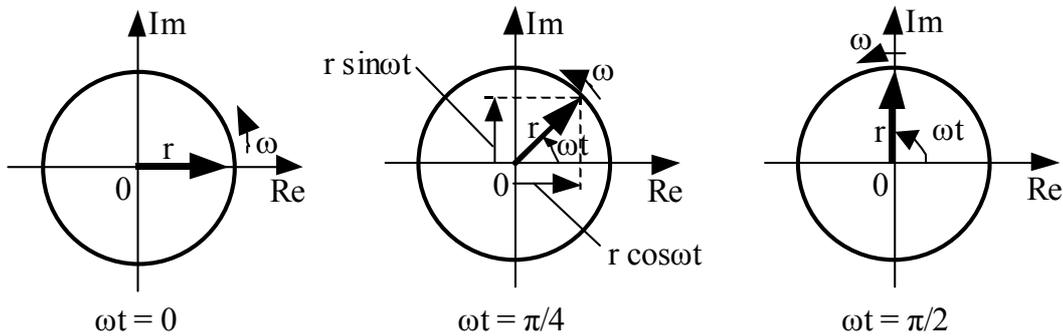
$$\exp(j\omega t) = \cos(\omega t) + j\sin(\omega t)$$

macht es Sinn, die gegebene Zeitfunktion folgendermassen zu notieren:

$$f(t) = a\cos(\omega t) - b\sin(\omega t) = \operatorname{Re}((a+jb)\exp(j\omega t)) = \operatorname{Re}(\underline{C}\exp(j\omega t)).$$

D.h. $f(t)$ ist der Realteil eines Produkts einer komplexen Zahl \underline{C} mit der einfachen, komplexen Zeitfunktion $\exp(j\omega t)$. Stellt man diese Funktion in der komplexen Ebene dar, so ergibt sich ein Einheitskreis um den Ursprung der komplexen Ebene. Zu einer bestimmten Zeit t ist $\exp(j\omega t)$ eine komplexe Zahl mit Betrag 1 und kann somit durch einen Einheitsvektor dargestellt werden. Dieser Vektor rotiert mit der Zeit t gleichmässig um den Ursprung der komplexen Ebene. Seine Drehgeschwindigkeit ist proportional zur Kreisfrequenz ω .

Stellen wir eine beliebige, harmonische Funktion $f(t)$ in der komplexen Ebene dar, so erhalten wir einen Kreis mit Radius \hat{A} . Der Momentanwert dieser Funktion wird durch einen gleichmässig rotierenden Vektor (Drehzeiger) der Länge $|\hat{A}|$ dargestellt. Zur Zeit $t = 0$ ist $\exp(j\omega t) = 1$ und somit der Drehzeiger gleich dem Zeiger, welcher \underline{C} darstellt. Figur 40 veranschaulicht den Spezialfall $\underline{C} = r$ d.h. $\varphi = 0$.



Figur 40: Komplexe Zeiger für die harmonische Zeitfunktion $f(t) = r \exp(j \omega t)$.

Der Vorteil der komplexen Darstellung liegt darin, dass man die Zeitabhängigkeit $\exp(j \omega t)$ abspalten kann (sofern alle Quellen in einem Netzwerk dieselbe Kreisfrequenz ω haben). Ist dies nicht der Fall, wird die Methode der Quellenüberlagerung verwendet!) und dann nur noch mit den komplexen Amplituden rechnen muss. Für diese Rechnungen gelten die relativ einfachen Regeln der komplexen Analysis.

4.3. Blind-, Wirk-, und Scheinleistung

Betrachten wir einen linearen Zweipol, über dem die Spannung

$$u(t) = \text{Re}(\underline{U} \exp(j \omega t)) = \hat{U} \cos(\omega t + \varphi_U)$$

liegt und durch den der Strom

$$i(t) = \text{Re}(\underline{I} \exp(j \omega t)) = \hat{I} \cos(\omega t + \varphi_I)$$

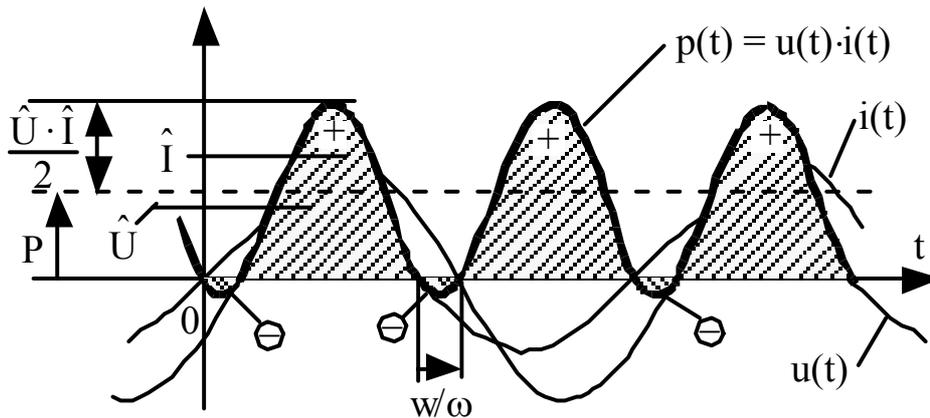
fließt. Dabei gilt gemäss vorhergehendem Abschnitt

$$\hat{U} = \sqrt{\text{Re}(\underline{U})^2 + \text{Im}(\underline{U})^2} = |\underline{U}|, \quad \varphi_U = \text{arctg}(\text{Im}(\underline{U}) / \text{Re}(\underline{U})),$$

$$\hat{I} = \sqrt{\text{Re}(\underline{I})^2 + \text{Im}(\underline{I})^2} = |\underline{I}|, \quad \varphi_I = \text{arctg}(\text{Im}(\underline{I}) / \text{Re}(\underline{I})).$$

Für den Zeitwert der Leistung des Zweipols gilt $p = ui$, also

$$p(t) = \hat{U} \hat{I} \cos(\omega t + \varphi_U) \cos(\omega t + \varphi_I).$$



Figur 41: Verlauf von $u(t)$, $i(t)$, $p(t)$ für $\varphi > 0$ (allg. $\varphi = \varphi_u - \varphi_i$)

Die goniometrische Umformung dieser Formel mit dem Additionstheorem ist recht mühsam. Es ergibt sich schliesslich

$$p(t) = \hat{U}\hat{I} \{ \cos(\varphi_u - \varphi_i)[1 + \cos 2(\omega t + \varphi_u)] + \sin(\varphi_u - \varphi_i)[\sin 2(\omega t + \varphi_u)] \} / 2.$$

Mit der Definition der *Scheinleistung*

$$S = \hat{U}\hat{I}/2$$

erhalten wir

$$p(t) = S \{ \cos(\varphi_u - \varphi_i)[1 + \cos 2(\omega t + \varphi_u)] + \sin(\varphi_u - \varphi_i)[\sin 2(\omega t + \varphi_u)] \}$$

und erkennen, dass $p(t)$ zwei Summanden enthält, welche im Wesentlichen proportional zum Kosinus bzw. Sinus der Phasendifferenz

$$\varphi = \varphi_u - \varphi_i$$

von Spannung und Strom sind. Der erste Term wird *Zeitwert der Wirkleistung*, der zweite Term *Zeitwert der Blindleistung* genannt¹. Man schreibt:

$$p_W(t) = S \cos(\varphi) [1 + \cos 2(\omega t + \varphi_u)]; \quad p_B(t) = S \sin(\varphi) [\sin 2(\omega t + \varphi_u)].$$

Vor allem in der Energietechnik und in der Messtechnik sind die Mittelwerte von grosser Bedeutung. Wegen der Periodizität der harmonischen Schwingungen, muss nur über eine Periode (z.B. von $t = 0$ bis $t = T = 2\pi/\omega$) gemittelt, d.h. integriert werden:

¹ Man erkennt auch, dass $p(t)$ aus einem zeitlich konstanten und einem zeitharmonischen Teil (mit doppelter Frequenz) zusammengesetzt ist. Dass die Wirkleistung nicht nur den konstanten Term, sondern auch einen zeitharmonischen Anteil beinhaltet, ist etwas eigenartig und wohl nur dadurch motiviert, dass so die Formel für die Blindleistung einfacher wird, bzw. dass damit die Blindleistung zu $\sin\varphi$ proportional wird.

$$P = \frac{1}{T} \int_0^T p(t) dt.$$

Man sieht sofort, dass der *Mittelwert der Blindleistung* $P_B = 0$ wird und dass für den *Mittelwert der Wirkleistung*

$$P_W = S \cos \varphi = P$$

gilt. Man nennt daher P kurzerhand *Wirkleistung*. Damit ist $p_W(t) = S \cos(\varphi) [\dots] = P [\dots]$. Nun führen wir eine ähnliche Abkürzung in der obigen Formel für $p_B(t)$ ein:

$$Q = S \sin \varphi.$$

Q ist offenbar die Amplitude der zeitabhängigen Blindleistung und wird oft kurzerhand Blindleistung genannt. Dabei ist zu beachten, dass Q nicht dem Mittelwert der Blindleistung entspricht.

Vergleicht man die Definition der Scheinleistung mit der Definition der Leistung bei Gleichstrom, so „stört“ der Faktor $1/2$. In der Elektrotechnik werden daher oft statt der Spitzenwerte \hat{U} und \hat{I} von Strom und Spannung die *Effektivwerte* angegeben:

$$U_{eff} = \hat{U} / \sqrt{2}; \quad I_{eff} = \hat{I} / \sqrt{2}$$

angegeben, wobei der Index *eff* meist weggelassen wird. Damit gilt für Schein-, Wirk- und Blindleistung:

$$S = UI; \quad P = UI \cos \varphi; \quad Q = UI \sin \varphi.$$

Übungshalber verwenden wir nun noch die *komplexe Notation* für $p(t)$:

$$p(t) = \operatorname{Re}(\underline{U} \exp(j\omega t)) \operatorname{Re}(\underline{I} \exp(j\omega t)).$$

Aus der komplexen Analysis weiss man, dass für zwei komplexe Zahlen z_1 und z_2

$$\operatorname{Re}(z_1) \operatorname{Re}(z_2) = \operatorname{Re}(z_1 z_2 + z_1 z_2^*) / 2$$

gilt, wobei z_2^* die zu z_2 konjugiert komplexe Zahl ist. Also ist

$$p(t) = \operatorname{Re}(\underline{U} \underline{I} \exp(j2\omega t) + \underline{U} \underline{I}^* \exp(0)) / 2 = \operatorname{Re}(\underline{U} \underline{I} \exp(j2\omega t)) / 2 + \operatorname{Re}(\underline{U} \underline{I}^*) / 2.$$

Offenbar kann $p(t)$ in einen Teil, welcher mit der doppelten Kreisfrequenz 2ω schwingt und einen konstanten Teil zerlegt werden. Für den zweiten Teil erhalten wir sofort:

$$P = \operatorname{Re}(\underline{U} \underline{I}^*) / 2 = \operatorname{Re}(\hat{U} \hat{I} \exp(j\varphi)) / 2 = \hat{U} \hat{I} \cos(\varphi) / 2.$$

Wird das Verbraucherpfilsystem verwendet und ist die Wirkleistung P eines Zweipols positiv, so nimmt der Zweipol im Zeitmittel Energie auf. Diese Energie wird beispielsweise von Kraftwerken den Kunden verrechnet. Die Blindleistung ist ein Mass für die Energie-

menge, welche ein Zweipol während jeder Schwingungsperiode aufnimmt und wieder abgibt². Wir werden gleich sehen, dass ideale Kapazitäten und Induktivitäten Bauelemente sind, deren Wirkleistung Null ist, deren Blindleistung hingegen positiv ist.

4.4. Bauelemente

Wir betrachten dieselben Bauelemente wie im vorhergehenden Kapitel, nehmen nun aber an, der Zeitverlauf sei harmonisch, werde also durch eine bestimmte Kreisfrequenz ω beschrieben.

Da ideale Schalter entweder durch einen Kurzschluss oder einen Leerlauf ersetzt werden können, verzichten wir auf eine explizite Beschreibung von Schaltern.

Nichtlinearitäten von Bauelementen führen dazu, dass die Ströme und Spannungen „verzerrt“ werden, d.h. die Ströme und Spannungen in Netzwerken, welche nichtlineare Bauelemente enthalten, sind i.a. nicht harmonisch. Dadurch wird eine einfache Rechnung mit komplexen Grössen verunmöglicht. Aus diesem Grunde konzentrieren wir uns in diesem Kapitel auf *lineare* Bauelemente.

Da es für die Bauelemente keine Rolle spielt, ob die Spannungen und Ströme harmonisch sind oder nicht, werden dieselben Schaltsymbole verwendet wie im vorhergehenden Kapitel.

4.4.1. Quellen

Offenbar lässt sich eine ideale harmonische *Spannungsquelle* durch die Kreisfrequenz ω und einen komplexen Spannungswert \underline{U} beschreiben. Für den Zeitverlauf gilt dann in Anlehnung an die Darstellung von $f(t)$ aus dem vorhergehenden Abschnitt:

$$u(t) = \operatorname{Re}(\underline{U} \exp(j\omega t)).$$

\underline{U} beinhaltet die Amplitude \hat{U} und die Phasenlage zur Zeit $t = 0$:

$$\hat{U} = \sqrt{\operatorname{Re}(\underline{U})^2 + \operatorname{Im}(\underline{U})^2} = |\underline{U}|, \quad \varphi_u = \operatorname{arctg}(\operatorname{Im}(\underline{U}) / \operatorname{Re}(\underline{U})).$$

Enthält das angeschlossene Netzwerk nur lineare Bauteile und Quellen mit derselben Kreisfrequenz ω , so ist auch der Strom der Spannungsquelle harmonisch und es gilt:

$$i(t) = \operatorname{Re}(\underline{I} \exp(j\omega t)).$$

² Diese Aussage ist mit Vorsicht zu geniessen, da ja auch die Wirkleistung einen zeitharmonischen Teil enthält, was bereits in der vorhergehenden Fussnote erwähnt wurde.

$$\hat{I} = \sqrt{\operatorname{Re}(\underline{I})^2 + \operatorname{Im}(\underline{I})^2} = |\underline{I}|, \quad \varphi_i = \arctg(\operatorname{Im}(\underline{I}) / \operatorname{Re}(\underline{I})).$$

Je nach Belastung der Spannungsquelle durch das Netzwerk ergeben sich unterschiedliche Ströme. Damit kann auch die von der Quelle abgegebene Wirkleistung unterschiedlich gross sein. Wird die Quelle nur durch ein Netzwerk mit Ohm'schen Widerständen belastet, so wird die abgegebene Wirkleistung positiv und die Blindleistung $Q = 0$.

Ähnliches lässt sich über ideale harmonische Stromquellen sagen, nur ist bei diesen \underline{I} gegeben, während sich \underline{U} aus der Belastung durch das angeschlossene Netzwerk ergibt.

4.4.2. Widerstand

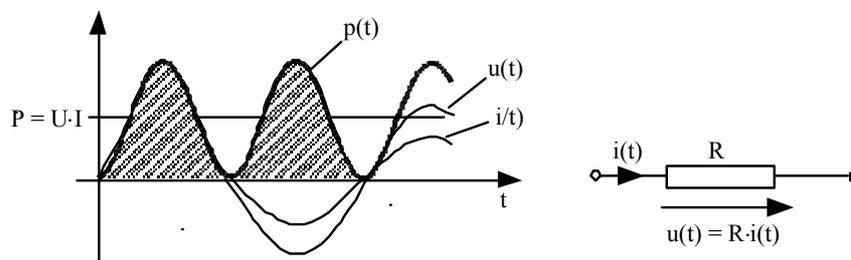
Für den Ohm'schen Widerstand gilt $u(t) = Ri(t)$ und somit

$$U = RI; \quad \hat{U} = R\hat{I}; \quad \varphi_U = \varphi_i; \quad \underline{U} = R\underline{I}.$$

Offensichtlich verbraucht der Widerstand die Leistung

$$P = \hat{U}\hat{I}/2 = UI = RI^2 = U^2/R.$$

Die Blindleistung des Widerstandes verschwindet weil die Phasendifferenz von Strom und Spannung verschwindet. Der Ohm'sche Widerstand ist also auch bei Wechselstrom der ideale Verbraucher.



Figur 42: Spannung, Strom und Leistung eines Widerstandes

4.4.3. Kapazität

Für die Kapazität gilt $i(t) = Cdu(t)/dt$. Werden sinusförmige Verläufe von Strom und Spannung mit der Kreisfrequenz ω vorausgesetzt, so ergibt sich sofort

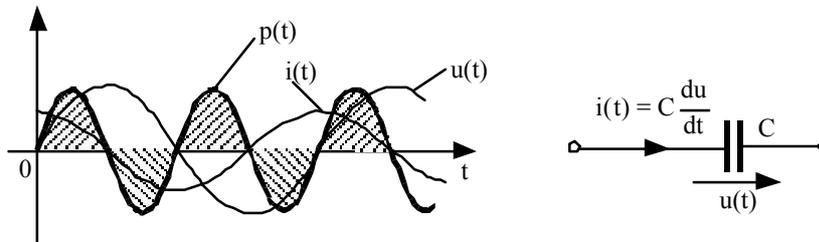
$$\underline{I} = j\omega C \underline{U}; \quad I = \omega CU; \quad \hat{I} = \omega C\hat{U}; \quad \varphi = -\pi/2.$$

Die letzte Gleichung besagt, dass der Strom der Spannung um 90° nachläuft.

Offensichtlich verbraucht die Kapazität keine Wirkleistung:

$$P = 0.$$

Der Betrag der Blindleistung der Kapazität ist übrigens maximal, da die Phasendifferenz von Strom und Spannung 90° beträgt. Wird eine Kapazität an eine ideale Quelle angeschlossen, so pendelt die gesamte Energie zwischen Quelle und Kapazität hin und her.



Figur 43: Spannung, Strom und Leistung einer Kapazität

4.4.4. Induktivität

Für die Induktivität gilt $u(t) = L di(t)/dt$. Werden sinusförmige Verläufe von Strom und Spannung mit der Kreisfrequenz ω vorausgesetzt, so ergibt sich sofort

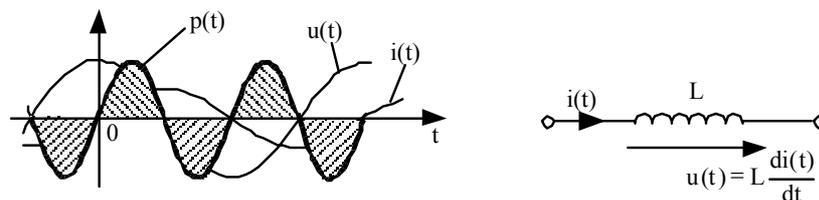
$$\underline{U} = j\omega L \underline{I}; \quad U = \omega LI; \quad \hat{U} = \omega L \hat{I}; \quad \varphi = +\pi/2.$$

Die letzte Gleichung besagt, dass der Strom der Spannung um 90° vorläuft.

Offensichtlich verbraucht die Induktivität keine Wirkleistung:

$$P = 0.$$

Der Betrag der Blindleistung der Induktivität ist maximal, da die Phasendifferenz von Spannung und Strom 90° beträgt. Wird eine Induktivität an eine ideale Quelle angeschlossen, so pendelt Energie zwischen Quelle und Induktivität hin und her.



Figur 44: Spannung, Strom und Leistung einer Induktivität

Induktivität und Kapazität sind übrigens für die Elektrizitätswerke unangenehm, da die Ströme in den Zuleitungen zu Verlusten führen, welche das Elektrizitätswerk dem Konsumenten nicht verrechnen kann.

4.5. Impedanz und Admittanz

Im vorhergehenden Abschnitt haben wir für Widerstand und Induktivität die folgenden komplexen Gleichungen angegeben:

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

Bei einer Serieschaltung von Widerstand und Induktivität summieren sich die Spannungen und wir erhalten

$$\underline{U} = (R + j\omega L) \underline{I} = \underline{Z} \underline{I}.$$

Dabei ist $\underline{Z} = R + j\omega L$ die (komplexe) *Impedanz*. Man schreibt auch $\underline{Z} = R + jX$. Dabei bezeichnet X die *Reaktanz*.

Sind mehrere Impedanzen in Serie geschaltet, so summieren sich die Impedanzen in gleicher Weise wie in Serie geschaltete Widerstände bei Gleichstrom.

Einen Widerstand R können wir alternativ durch den Leitwert $G = 1/R$ beschreiben. Diese Beschreibung ist bei Parallelschaltungen vorteilhaft, da sich dann die Leitwerte addieren. Wir definieren deshalb die *Admittanz* als den Kehrwert der Impedanz:

$$\underline{Y} = 1/\underline{Z}.$$

Bei Parallelschaltungen summieren sich die Admittanzen, da

$$\underline{I} = \underline{Y} \underline{U}$$

gilt und sich die Ströme bei Parallelschaltungen summieren.

Die zur Serieschaltung von Widerstand und Induktivität duale Schaltung ist eine Parallelschaltung von Leitwert und Kapazität. Für diese gilt:

$$\underline{I} = (G + j\omega C) \underline{U} = \underline{Y} \underline{U}.$$

Da wir die Admittanz als Kehrwert einer Impedanz und diese wiederum als Serieschaltung eines Widerstandes mit einer Induktivität auffassen können, erhalten wir mit einer komplexen Erweiterung:

$$1/(G + j\omega C) = (G - j\omega C)/(G^2 + \omega^2 C^2) = (R + j\omega L).$$

Der Vergleich von Real- und Imaginärteil zeigt, dass eine Serieschaltung eines Widerstandes R mit einer Induktivität L von aussen gesehen exakt dieselben elektrischen Eigenschaften hat wie die Parallelschaltung eines Leitwertes G mit einer *negativen* Kapazität C .

4.6. Schaltungsanalyse

Die Analyse von Netzwerken bestehend aus linearen Zweipolen (Quellen, Widerständen, Kapazitäten und Induktivitäten) bei Wechselstrom, bzw. sinusförmigem Zeitverlauf ist — bis auf die komplexe Erweiterung der Ströme und Spannungen sowie der Einführung der Impedanzen und Admittanzen anstelle der Widerstände und Leitwerte — dieselbe wie die Analyse von Netzwerken bestehend aus Quellen und Widerständen bei Gleichstrom. Dabei können Impedanzen bzw. Admittanzen auch Induktivitäten und Kapazitäten beschreiben. Nach der komplexen Erweiterung ergeben sich formal dieselben Gleichungssysteme und es können deshalb auch dieselben Methoden angewendet werden.

Soll ein Netzwerk von Hand analysiert werden, so wird dieses zunächst durch Quellenumformungen, Quellenüberlagerung, Berücksichtigung von Serie- und Parallelschaltungen, Stern-Dreiecksumwandlungen etc. so weit wie möglich vereinfacht. Anschliessend werden die Zweipolgleichungen und die linear unabhängigen Knoten- und Maschengleichungen notiert und die so erhaltenen, linearen Gleichungssysteme aufgelöst. Für die Knoten- und Maschengleichungen gelten die bereits bekannten Kirchhoff'schen Gesetze in komplexer Form:

$$\sum_{k=1}^K I_k = 0; \quad \sum_{k=1}^K U_k = 0$$

für einen Knoten mit K Strömen bzw. eine Masche mit K Spannungen.

Selbstverständlich lassen sich auch die systematischen Methoden der Knoten- bzw. Maschenanalyse vom Gleichstrom- auf den Wechselstromfall übertragen. Beispiele zur Schaltungsanalyse bei Wechselstrom sind auf S. A33 ff. zu finden.

4.7. Frequenzgang und Ortskurven

Die Energietechnik ist wohl der einzige grössere Anwendungsbereich der Elektrotechnik, in dem sinusförmige Spannungs- und Stromverläufe mit einer bestimmten Frequenz vorausgesetzt werden können. In der Nachrichtentechnik werden oft die informationstragenden Signale einer „Trägerfrequenz“ „aufmoduliert“. Dabei ist die Trägerfrequenz beispielsweise die Frequenz eines Nachrichtensenders. Für ein „*amplitudenmoduliertes*“ Signal schreibt man dann beispielsweise:

$$f(t) = f_{mod}(t) \sin(\omega t + \varphi),$$

wobei ω und φ die Kreisfrequenz und die Phase des „Trägers“ bezeichnen. Ähnlich gilt für *frequenzmodulierte* Signale:

$$f(t) = A \sin(\omega(t) t + \varphi)$$

und für *phasenmodulierte* Signale

$$f(t) = A \sin(\omega t + \varphi(t))$$

Ist f_{mod} bei der Amplitudenmodulation zeitlich konstant, so lassen sich die bisherigen Methoden zur Netzwerkanalyse bei Wechselstrom mit einer einzigen Frequenz anwenden. Dies bedeutet gleichzeitig, dass keinerlei Information übertragen wird. Bei allen interessanten nachrichtentechnische Anwendungen ist das Signal deshalb nicht völlig harmonisch.

Die oben skizzierten, „modulierten“ Funktionen lassen sich mit Hilfe der Fourieranalyse untersuchen. Dabei ergibt sich ein Fourierspektrum, welches üblicherweise nur innerhalb eines *Frequenzbandes* $\omega_{min} < \omega < \omega_{max}$ nennenswerte Anteile besitzt. Die verwendeten elektronischen Schaltungen können deshalb innerhalb eines bestimmten Frequenzbandes analysiert werden.

Die Abhängigkeit der Ströme und Spannungen von der Frequenz wird *Frequenzgang* genannt. Da Ströme und Spannungen durch komplexe Größen beschrieben werden, ist es sinnvoll, den Frequenzgang in der komplexen Ebene graphisch aufzuzeichnen. Betrachtet man einen beliebigen Zweipol, so sind auch dessen Impedanz und Admittanz i.a. frequenzabhängige komplexe Größen, welche in der komplexen Ebene graphisch aufgezeichnet werden können. Die entsprechenden Kurven werden Ortskurven genannt. Ortskurven können übrigens auch für Mehrpole angegeben werden. Z.B. kann mittels Ortskurven das Verhalten eines Verstärkers angegeben werden. Der Einfachheit halber beschränken wir uns hier auf Zweipole.

Prinzipiell kann ein Zweipol aus einem bekannten Netzwerk bestehen. Er kann aber auch eine „black box“ mit unbekanntem Innenleben sein. Im ersten Falle können Frequenzverhalten bzw. Ortskurven aus der Netzwerkanalyse bestimmt werden. Im zweiten Falle kann man versuchen, ein äquivalentes Netzwerk mit demselben Frequenzverhalten innerhalb des interessierenden Frequenzbandes zu finden. Dadurch wird es möglich, die „black box“ in einem Netzwerk mit gängigen Analyseverfahren zu untersuchen. Das Auffinden eines äquivalenten Netzwerkes ist ein Syntheseproblem, welches ungleich schwieriger zu lösen ist als die reinen Analyseprobleme. Wir beschränken uns hier deshalb auf letztere. Ausserdem beschränken wir uns auf Zweipole, welche lediglich aus einem *RLC Netzwerk* bestehen, welche also nur mit Widerständen, Kapazitäten und Induktivitäten aufgebaut werden.

Zunächst betrachten wir die einfachsten *RLC Netzwerke*, welche nur ein Element enthalten. Offenbar gibt es folgende 3 Fälle:

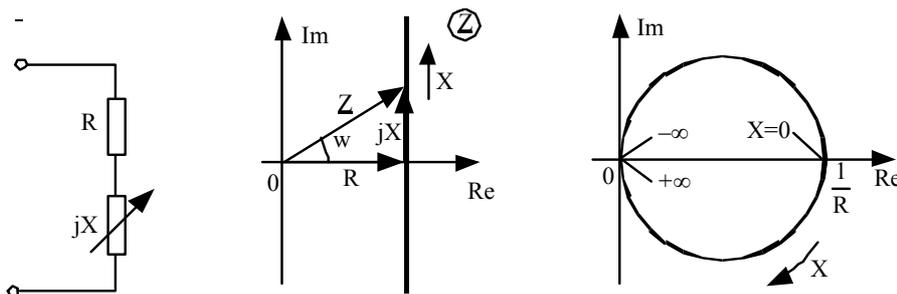
- 1) Für einen *Widerstand* gilt $\underline{Z} = R$. Folglich ist die Ortskurve des Widerstandes in der komplexen Impedanzebene ein einfacher Punkt auf der reellen Achse: Die Impedanz eines Widerstandes ist frequenzunabhängig. Dasselbe gilt natürlich für einen Leitwert und auch für die Darstellungen von Widerstand und Leitwert in der komplexen Admittanzebene.
- 2) Für eine Induktivität gilt $\underline{Z} = j\omega L$. Dadurch wird die positive imaginäre Achse der Impedanzebene beschrieben. Als Spezialfälle erhält man für $\omega = 0$ den Ursprung der Ebene $\underline{Z} = 0$ und für $\omega = \infty$ den Punkt auf der imaginären Achse im Unendlichen. Auch in der komplexen Admittanzebene ergibt sich aus $\underline{Y} = 1/j\omega L = -j/\omega L$ eine gerade

Ortskurve, diesmal jedoch die negative imaginäre Achse mit den speziellen Punkten $-j\infty$ für $\omega = 0$ und $\underline{Y} = 0$ für $\omega = \infty$.

- 3) Es erstaunt sicher nicht, dass auch die Ortskurven der Kapazität in der Impedanz- bzw. Admittanzebene mit der negativen bzw. positiven imaginären Achse zusammenfallen, gilt doch $\underline{Z} = 1/j\omega C = -j/\omega C$ bzw. $\underline{Y} = j\omega C$.

Etwas komplizierter wird die Situation, wenn ein Zweipol aus Serie- oder Parallelschaltungen von Widerständen mit Kapazitäten oder Induktivitäten besteht. Die entsprechenden Formeln für die Impedanz und Admittanz des Zweipols lassen sich oft ohne Probleme finden. Die anschließende graphische Darstellung braucht etwas Geschick, ist aber oft recht einfach. Zur Veranschaulichung betrachten wir eine Serieschaltung eines Widerstandes und einer Induktivität.

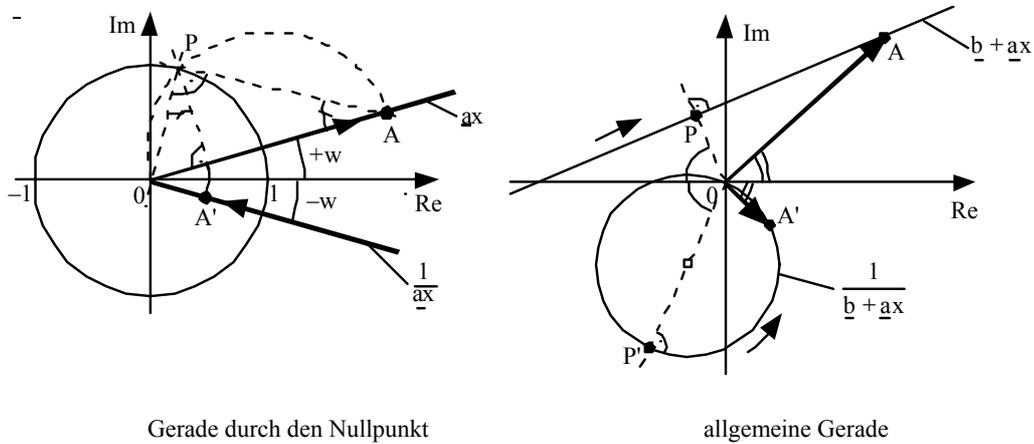
Figur 45 zeigt die Ortskurve für die Impedanz einer Serieschaltung eines Widerstandes und einer Induktivität ($\underline{Z} = R + j\omega L$) und die zugehörige Ortskurve für die Admittanz ($\underline{Y} = 1/(R + j\omega L) = (R - j\omega L)/(R^2 + \omega^2 L^2)$). Dabei kann man sich zu Nutze machen, dass die Reaktanz $X = \omega L$ proportional zur Frequenz ist. Man kann deshalb X statt ω als Kurvenparameter verwenden.



Figur 45: Ortskurven für die Impedanz $\underline{Z} = R + jX$ und die Admittanz $\underline{Y} = 1/\underline{Z}$ für $-\infty < X < \infty$

Dass sich in der Impedanzebene ein um R verschobener Strahl parallel zur y Achse ergibt, ist leicht nachvollziehbar. Dass die Ortskurve in der Admittanzebene hingegen einen Kreis (genauer einen Halbkreis) beschreibt, ist weniger offensichtlich. Wie bereits erwähnt, spielt die dabei vorzunehmende Umformung $\underline{Y} = 1/\underline{Z}$ auch eine wichtige Rolle bei der Vereinfachung von Parallel- bzw. Serieschaltungen. Es ist daher sinnvoll, die Inversion etwas genauer anzusehen. Man kann $\underline{Y} = 1/\underline{Z}$ als komplexe Abbildung verstehen und die zugehörigen Sätze aus der komplexen Analysis beherrsigen. Wichtig ist, dass es sich um eine *konforme* Abbildung handelt, bei welcher Kreise stets auf Kreise abgebildet werden. Dabei sind Geraden als Spezialfälle von Kreisen anzusehen. Konforme Abbildungen sind winkeltreu, d.h. die Winkel bleiben gleich. Um einen Kreis eindeutig festzulegen, reicht es, drei Punkte auf dem Kreis oder zwei Punkte und ein Winkel bei einer der Punkte zu bestimmen. Sinnvollerweise wählen wir die Endpunkte $X = 0$, $X = \infty$, sowie einen beliebigen Punkt dazwischen, z.B. $X = 1$.

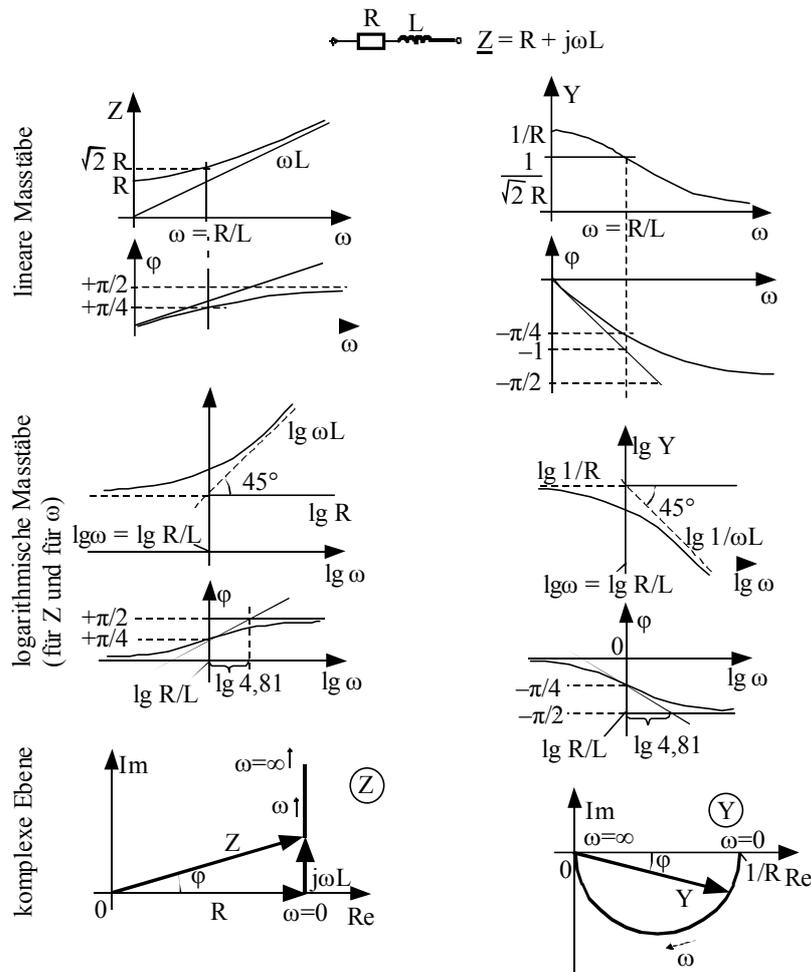
Die Inversion $\underline{Y} = 1/\underline{Z}$ kann in der komplexen Ebene als Inversion am Einheitskreis verstanden werden. Diese lässt sich mit geometrischen Konstruktionen (mit Zirkel und Lineal) bewerkstelligen. Figur 46 veranschaulicht diesen Sachverhalt.



Figur 46: Einfache Inversionen in der komplexen Ebene

Eine häufig verwendete Alternative zu den Ortskurven in der komplexen Ebene ist die Darstellung von *Amplitudengang* und *Frequenzgang*, d.h., das Aufzeichnen der Kurven $A(\omega)$ und $\varphi(\omega)$. Wenn ein breiter Frequenzbereich dargestellt werden muss oder wenn die Amplituden in einem breiten Bereich variieren, kann es sinnvoll sein, logarithmische oder doppelt logarithmische Darstellungen zu verwenden. Figur 47 veranschaulicht diese Darstellungen für eine Serieschaltung von einem Widerstand und einer Induktivität. Dabei bezeichnet \lg den Zehnerlogarithmus, für den oft auch ausführlicher \log_{10} geschrieben wird. Der Zehnerlogarithmus wird übrigens dem natürlichen Logarithmus in der Elektrotechnik meist vorgezogen, da wir üblicherweise das Dezimalsystem für Zahlenangaben verwenden.

Solange ein Zweipol aus Widerständen und entweder Kapazitäten oder Induktivitäten besteht, unterscheiden sich die entsprechenden Ortskurven nicht wesentlich von den bisher betrachteten Kurven. Eine wesentliche Änderung ergibt sich wenn Induktivitäten und Kapazitäten miteinander kombiniert werden. Derartige Schaltungen haben spezielle Resonanzen, welche insbesondere zur Erzeugung von harmonischen Schwingungen von Interesse sind. Im folgenden Abschnitt werden wir einige einfache Fälle skizzieren.



Figur 47: Verschiedene Darstellungsarten der Frequenzgänge und Ortskurven von $\underline{Z} = R + j\omega L$ und $\underline{Y} = 1/\underline{Z}$ für $0 \leq \omega < \infty$.

4.8. Spezielle Schaltungen

Wir haben gesehen, dass durch Kombination der Zweipole Widerstand (R), Kapazität (C) und Induktivität (L) neue Zweipole entstehen, deren Impedanz bzw. Admittanz recht unterschiedliche Frequenzgänge aufweisen. In der Praxis gibt es verschiedene Aufgabenstellungen, bei denen ganz bestimmte Frequenzgänge wünschenswert sind. Beispielsweise soll in einem Nachrichtensender eine harmonische Schwingung mit einer präzisen Sendefrequenz erzeugt werden. Dazu werden Schwingkreise aufgebaut, welche im folgenden Unterabschnitt skizziert werden. Ein Radio ist ein Nachrichtenempfänger, welcher üblicherweise Nachrichten von unterschiedlichen Sendern mit unterschiedlichen Sendefrequenzen empfängt. Um einen bestimmten Sender auswählen zu können, werden Filter verwendet, welche aus dem gesamten empfangenen Frequenzspektrum einen bestimmten Teil „herausfiltern“. Filter werden im zweiten Unterabschnitt skizziert.

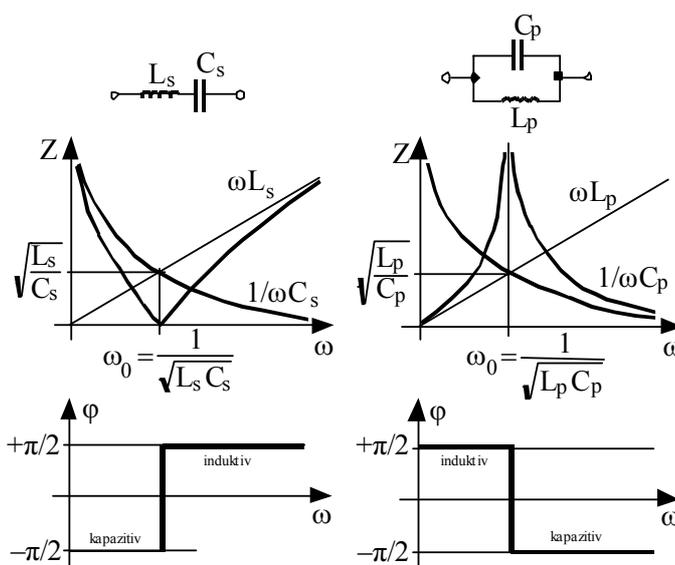
Die Synthese von Schwingkreisen und Filtern mit ganz bestimmten, gewünschten Charakteristiken ist eine recht anspruchsvolle „inverse“ Aufgabe, auf die hier nicht näher eingegangen werden kann. Wir beschränken uns darauf, die einfachsten Schaltungen, welche in der Nachrichtentechnik oft gebraucht werden, kurz vorzustellen.

4.8.1. Schwingkreise

Wir haben bereits gesehen, dass die Impedanz einer Induktivität L und einer Kapazität C rein imaginär sind und zwar erstere positiv und letztere negativ. Da sich Impedanzen von Serieschaltungen addieren, ist es klar, dass eine LC Serieschaltung die Impedanz Null haben kann. Da die Impedanzen von L bzw. C — je nach Frequenz — die gesamte positive bzw. negative imaginäre Achse überstreichen, verschwindet für jede LC Serieschaltung die Impedanz bei einer ganz bestimmten Frequenz ω_0 . Es gilt dann

$$j\omega_0 L - j\frac{1}{\omega_0 C} = 0 \Rightarrow \omega_0 = 1/\sqrt{LC}.$$

Wird die Impedanz Null, so wird natürlich die Admittanz unendlich. Natürlich lässt sich durch eine LC Parallelschaltung erreichen, dass bei der Frequenz ω_0 die Admittanz verschwindet und die Impedanz unendlich wird. Wie man leicht nachprüft, ergibt sich dieselbe Formel für ω_0 . Am Ende des vorhergehenden Kapitels haben wir bereits die Differentialgleichung einer LC Parallelschaltung analysiert und gesehen, dass die Lösung im Wesentlichen eine harmonische Schwingung mit der Kreisfrequenz ω_0 ist. Die LC Parallelschaltung wird deshalb *Parallelschwingkreis* oder *Parallelresonanzkreis* und die dazu duale LC Serieschaltung *Seriewschwingkreis* oder *Serieresonanzkreis* genannt. Für beide ist die *Resonanzfrequenz* ω_0 charakteristisch. Figur 48 zeigt die entsprechenden Ortskurven.



Figur 48: Ortskurve des Betrages Z der Impedanz \underline{Z} und des Phasenwinkels φ beim verlustlosen Serie- bzw. Parallelresonanzkreis als Funktion von ω für $0 \leq \omega < \infty$.

In der Praxis lassen sich ideale Induktivitäten bzw. Kapazitäten nicht realisieren. Reale Spulen und Kondensatoren weisen Ohm'sche Verluste auf, werden also durch Serie- bzw. Parallelschaltungen von Induktivitäten bzw. Kapazitäten mit Widerständen approximiert. Es ist klar, dass diese Widerstände gerade in der Nähe der Resonanzfrequenz dominant werden.

Um die Umgebung der Resonanzfrequenz zu analysieren, wird zunächst die *Verstimmung*

$$\eta = \frac{\omega}{\omega_0} - \frac{\omega_0}{\omega} = \frac{\omega^2 - \omega_0^2}{\omega \omega_0}$$

eingeführt. Damit lässt sich die Formel für die Impedanz des Serieschwingkreises umformen:

$$\underline{Z}_s = j(\omega L_s - 1/\omega C_s) = j\eta\sqrt{L_s/C_s}$$

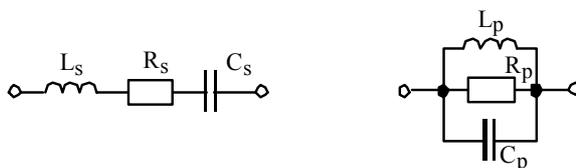
Analog ergibt sich für die Admittanz des Parallelschwingkreises

$$\underline{Y}_p = j\eta\sqrt{C_p/L_p}.$$

Für eine kleine Abweichung $\Delta\omega$ von der Resonanzfrequenz ω_0 ergibt sich $\eta \approx 2\cdot\Delta\omega/\omega_0$ und damit

$$\underline{Z}_s \approx j 2\Delta\omega/\omega_0 \sqrt{L_s/C_s} = j \Delta\omega 2L_s; \quad \underline{Y}_p \approx j 2\Delta\omega/\omega_0 \sqrt{C_p/L_p} = j \Delta\omega 2C_p.$$

Nun führen wir einen Widerstand in Serie- und Parallelschwingkreis ein um den Ohm'schen Verlusten von Induktivität und Kapazität Rechnung zu tragen. Selbstverständlich wäre es einleuchtend, jeweils zwei Widerstände — einen in Serie zur Induktivität und einen parallel zur Kapazität — einzuführen. Die resultierenden Netzwerke sind aber etwas komplizierter zu analysieren. Wir konzentrieren uns daher zunächst auf Schwingkreise mit nur einem Widerstand gemäss folgender Figur.



Figur 49: Ersatzschaltbild für verlustbehafteten Serie- bzw. Parallelresonanzkreis (es wird zwischen *s* und *p* unterschieden, weil $R_s \ll R_p$ ist).

Zunächst definieren wir die Resonanzfrequenz genau so wie im verlustfreien Fall (ohne Widerstand):

$$\omega_0 = \frac{1}{\sqrt{L_s C_s}} = \frac{1}{\sqrt{L_p C_p}}$$

Für die Impedanz bzw. Admittanz finden wir sofort

$$\underline{Z}_s = R_s + j\eta\sqrt{L_s/C_s}; \quad \underline{Y}_p = 1/R_p + j\eta\sqrt{L_p/C_p}.$$

Bei der Resonanzfrequenz, d.h. Verstimmung $\eta = 0$ ist also die Impedanz des Serieschwingkreises gleich dem Verlustwiderstand R_s und die Admittanz des Parallelschwingkreises gleich dem Verlustleitwert $1/R_p$. Offenbar ist die Qualität der Schwingkreise von R_s bzw. $1/R_p$ abhängig. Um die Qualität zu messen, wird die *Güte* Q wie folgt definiert:

$$Q = 2\pi (\text{max. Energie in } L \text{ oder } C) / (\text{pro Periode in } R \text{ in Wärme umgesetzte Energie}).$$

Dabei ist zu beachten, dass die maximalen Energien, welche L bzw. C speichern, bei Resonanzfrequenz einander gleich sind. Beispielsweise gilt für den Serieschwingkreis

$$w_{L_s}(t) = (L_s I^2 / 2) (1 - \cos 2\omega t), \text{ mit } W_{L_s \text{ max}} = L_s I^2$$

$$w_{C_s}(t) = I^2 / (2\omega^2 C_s) (1 - \cos 2\omega t), \text{ mit } W_{C_s \text{ max}} = I^2 / \omega^2 C_s.$$

Setzt man $\omega = \omega_0 = 1/\sqrt{L_s C_s}$ ein, so sieht man, dass $W_{L_s \text{ max}} = W_{C_s \text{ max}}$ gilt. Es spielt also keine Rolle, welcher der Energiewerte in der Formel für Q eingesetzt wird. Da nur der Widerstand Energie in Wärme umsetzt, folgt

$$Q_s = 2\pi (L_s I^2) / (R_s I^2 T) = \omega_0 L_s / R_s = \sqrt{L_s / C_s} / R_s.$$

Dabei wurde berücksichtigt, dass für die Periodendauer $T = 2\pi/\omega_0$ gilt. Der etwas seltsame Faktor 2π in der Definition von Q wird damit klar. In ähnlicher Manier findet man für die Güte eines Parallelschwingkreises

$$Q_p = 2\pi (C_p U^2) / (TU^2/R_p) = \omega_0 C_p R_p = R_p \sqrt{C_p / L_p}.$$

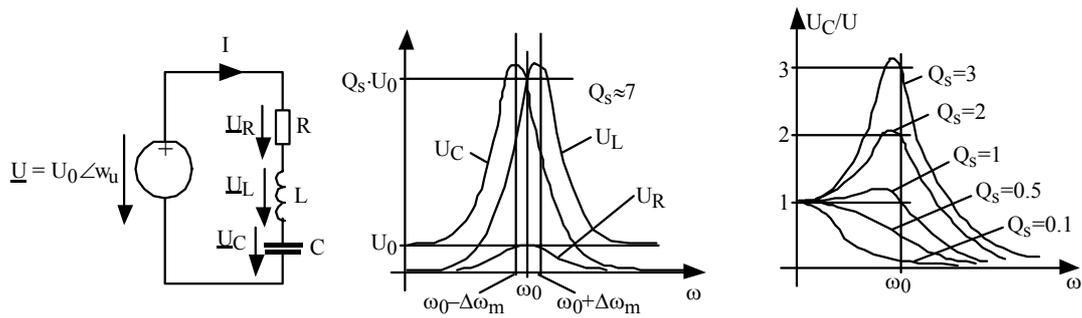
Mit der Güte lassen sich nun die Formeln für die Impedanz bzw. Admittanz wie folgt vereinfachen:

$$\underline{Z}_s = R_s(1 + j\eta Q_s); \quad \underline{Y}_p = 1/R_p (1 + j\eta Q_p).$$

Ausserdem gilt:

$$U_{C_s} = U_{L_s} = Q_s U_{R_s} = Q_s U_0; \quad I_{C_p} = I_{L_p} = Q_p I_{R_p} = Q_p I.$$

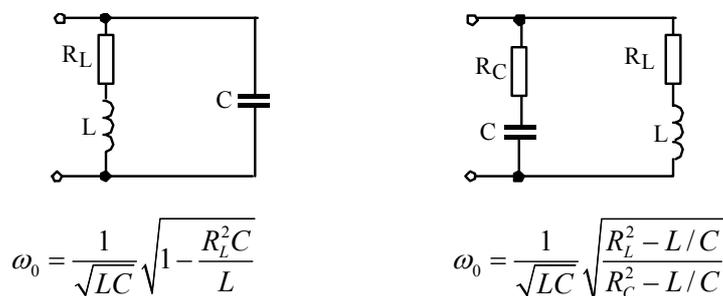
Die Gütezahl gibt also auch an, wieviel mal grösser die Spannung (bzw. der Strom im Parallelresonanzkreis) an der Induktivität und an der Kapazität bei Resonanz ist als die am Resonanzkreis angelegte Spannung (bzw. der Strom im Parallelresonanzkreis).



Figur 50: Verlauf der Spannungen U_R , U_L und U_C eines Serieresonanzkreises für $Q_s \approx 7$ und zum Vergleich von U_C für verschiedene Werte von Q_s (U_R , U_L , U_C sind Effektivwerte).

Wie bereits erwähnt, führt die Berücksichtigung der Verluste in der Induktivität bzw. in der Kapazität in der Praxis zu komplizierteren Ersatzschaltungen für Resonanzkreise. Figur 51 zeigt zwei typische Beispiele mit den entsprechenden Resonanzfrequenzen ω_0 . Ähnlich wie bei den einfacheren Resonanzkreisen von vorhin gilt für die Berechnung der Resonanzfrequenz folgende Regel:

Bei der Resonanzfrequenz ist der Imaginärteil der Impedanz bzw. der Admittanz gleich Null.



Figur 51: Andere Ersatzschaltungen für Parallelresonanzkreise

Selbstverständlich können nicht nur RLC Schwingkreise mit mehr als einem Widerstand, sondern auch solche mit mehr als einer Kapazität oder Induktivität hergestellt und analysiert werden. Dadurch können kompliziertere Frequenzgänge synthetisiert werden. Eine weitergehende Analyse würde aber den Rahmen dieser Vorlesung sprengen.

4.8.2. Zweitore — Transformatoren

Bisher haben wir uns auf die Betrachtung der Frequenzabhängigkeit von Zweipolen konzentriert, wobei wir durch Kombination der Elemente Widerstand, Kapazität und Induktivität Zweipole mit unterschiedlichen Frequenzgängen konstruieren konnten. Kapazität und Induktivität können zwar elektrische bzw. magnetische Energie speichern, sind aber —

ebenso wie der Widerstand — passive Elemente, die keine Energie erzeugen bzw. in ein Netzwerk einspeisen. Die aus passiven Elementen zusammengesetzten Zweipole sind natürlich ebenso passiv. In einem völlig passiven Netzwerk sind die Ströme und Spannungen üblicherweise überall Null, weil nirgendwo Energie ins Netzwerk eingespeisen wird. Dieser Fall ist natürlich ohne praktisches Interesse. Um zu interessanten Netzwerken zu gelangen, brauchen wir mindestens eine Quelle und mindestens einen Verbraucher, d.h. Nutzer der von der Quelle gelieferten Energie.

Wir betrachten nun den einfachsten Fall mit einer einzigen Quelle und einem einzigen Verbraucher und einem beliebigen Netzwerk dazwischen. Dieses Netzwerk hat offenbar vier Anschlüsse, ist also ein *Vierpol*. Zwei Pole verbinden das Netzwerk mit der Quelle und zwei Pole verbinden es mit dem Verbraucher. Erstere werden als *Eingangstor* (kurz Eingang) und letztere als *Ausgangstor* (kurz Ausgang) des Netzwerks bezeichnet. Man nennt ein derartiges Netzwerk auch *Zweitoren*. Viele elektronische Schaltungen bestehen aus hintereinander geschalteten Zweitoren.

Die mathematische Beschreibung von Zweitoren ist natürlich etwas aufwendiger als die Beschreibung von Zweipolen aber doch noch relativ einfach. Im Prinzip wird dabei ein Zusammenhang zwischen Eingangsgrösse und Ausgangsgrösse hergestellt und dieser Zusammenhang kann als mathematische *Transformation* beschrieben werden. Ein typisches Beispiel eines Zweitores ist der *Transformator*, welcher aus zwei gekoppelten Spulen besteht (siehe Figur 31). Dabei wird die erste, primäre Spule mit N_p Windungen mit dem Eingangstor und die zweite, sekundäre Spule mit N_s Windungen mit dem Ausgangstor verbunden. Im Unterabschnitt 3.1.4. haben wir bereits die Ersatzschaltung eines Transformators mit zwei Induktivitäten und einer Gegeninduktivität und den zugehörigen Differentialgleichungen gesehen. Da im Wechselstromfall (nach der komplexen Erweiterung) die zeitlichen Ableitungen durch ein Produkt mit $j\omega$ ersetzt werden können, vereinfachen sich die Gleichungen, welche den Transformator beschreiben, erheblich:

$$\underline{U}_a = j\omega (L_a \underline{I}_a + M \underline{I}_b)$$

$$\underline{U}_b = j\omega (L_b \underline{I}_b + M \underline{I}_a).$$

Statt diese Gleichungen explizite zu lösen, machen wir einige einfache Überlegungen, welche für praktische Transformatoren zutreffen und die Lösung vereinfachen.

Wird die Primärseite mit einer Wechselspannungsquelle \underline{U}_p verbunden und die Sekundärseite mit einem Ohm'schen Widerstand belastet, so fliesst primärseitig ein Strom \underline{I}_p , dessen Amplitude zur Amplitude von \underline{U}_p proportional ist. Dieser Strom erzeugt in der primären Spule einen magnetischen Fluss $\underline{\Phi}_p$, welcher proportional zur Windungszahl N_p und zum Strom \underline{I}_p ist. Damit ist die Amplitude von $\underline{\Phi}_p$ auch proportional zur Amplitude von \underline{U}_p . Transformatoren werden normalerweise so gebaut, dass praktisch der gesamte magnetische Fluss durch die sekundäre Spule fliesst, d.h. dass $\underline{\Phi}_s = \underline{\Phi}_p$ gilt. $\underline{\Phi}_s$ induziert in der sekundären Spule eine Spannung \underline{U}_s induziert, welche proportional zur Windungszahl N_s ist. Ausserdem sind die Amplituden von \underline{U}_s und $\underline{\Phi}_s$ zueinander proportional, so dass schliesslich

folgt, dass die Amplituden von Eingangs- und Ausgangsspannung zueinander proportional sind und der Quotient der Windungszahlen der Proportionalitätsfaktor ist:

$$\hat{U}_s = N_s/N_p \hat{U}_p.$$

Für die Ströme findet man übrigens die umgekehrte Proportionalität:

$$\hat{I}_s = N_p/N_s \hat{I}_p.$$

Ist eine Wechselspannungsquelle mit einer bestimmten Spannungsamplitude \hat{U}_p gegeben, so kann man diese also auf eine beliebige Spannung transformieren. Auf der Sekundärseite „sieht“ der Verbraucher also eine Spannungsquelle mit der Amplitude \hat{U}_s . Ist der Verbraucher ein Ohm'scher Widerstand R_s , so gilt natürlich $R_s = \hat{U}_s/\hat{I}_s$. Daraus folgt, dass die primärseitige Quelle am Eingang des Transformators einen Verbraucher mit dem Widerstand

$$R_p = \hat{U}_p/\hat{I}_p = (N_p^2/N_s^2) \hat{U}_s/\hat{I}_s = (N_p/N_s)^2 R_s.$$

„sieht“.

Realistische Transformatoren haben diverse Mängel, welche eine genaue Beschreibung erschweren. So werden beispielsweise Eisenkerne verwendet, welche den Fluss der primären Spule durch die sekundäre Spule führen. Eisen ist aber elektrisch leitend und der magnetische Fluss induziert deshalb im Eisenkern selbst Spannungen, welche zu sogenannten Wirbelströmen führen. Diese erhitzen den Kern und entziehen deshalb dem elektrischen Netz Energie. Darüber hinaus hat Eisen Eigenschaften, welche mathematisch sehr schwer zu beschreiben sind (Nichtlinearität, Sättigungseffekte, Hysterese). Diese Eigenschaften sind nicht nur störend beim Bau von Transformatoren mit Eisenkernen, sie können auch genutzt werden. Wichtigste Nutzung ist wohl der Bau von Speicherelementen. In der Frühzeit des Computerbaus wurden übrigens Ringkernspeicher verwendet, welche im Wesentlichen aus kleinen Transformatoren bestanden, deren Kerne zwei verschiedene Magnetisierungszustände annehmen konnten.

In der Energietechnik werden die Eisenkerne von Transformatoren lamelliert um so die Wirbelströme klein zu halten. Bei höheren Frequenzen reicht die Lamellierung nicht aus und es werden teurere Materialien (z.B. Ferrit) verwendet. Bei noch höheren Frequenzen verzichtet man schliesslich ganz auf Materialien, welche den magnetischen Fluss kanalisieren und nimmt Streuflüsse in Kauf, welche nicht sowohl die primäre als auch die sekundäre Spule durchfliessen.

Da Transformatoren in der modernen Elektronik eher selten verwendet werden, verzichten wir auf eine ausführlichere Beschreibung.

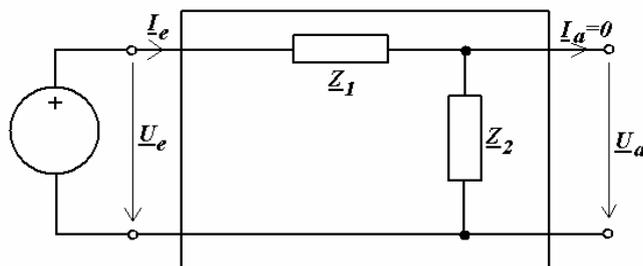
4.8.2. Zweitore — Passive Filter

Nachdem wir mit dem Transformator ein Zweitor, welches im Wesentlichen auf der Gegeninduktivität aufbaut, betrachtet haben, wenden wir uns Zweitoren zu, welche aus *RLC* Netzwerken bestehen. Um diese Zweitore zu charakterisieren, betrachten wir die Aus-

gangsgrößen als Funktionen der Eingangsgrößen und ihre Abhängigkeit von der Frequenz.

Zur Messung der Frequenzabhängigkeit wird normalerweise eine Quelle mit einstellbarer Frequenz mit den Eingangsklemmen verbunden und der Ausgang mit einem Ohm'schen Widerstand belastet. In der Hochfrequenztechnik ist dabei die Impedanzanpassung üblich, d.h. es wird eine Quelle mit vorgeschriebenem Innenwiderstand (z.B. 50 Ohm) verwendet und der Lastwiderstand ist gleich diesem Innenwiderstand. Die Impedanzanpassung wird bei der Hochfrequenztechnik nötig weil auch die erforderlichen Zuleitungen zu den Ein- und Ausgangsklemmen des Zweitors eine charakteristische Impedanz haben, wie wir später noch sehen werden.

Bei nicht sehr hohen Frequenzen arbeitet man der Einfachheit halber ohne Impedanzanpassung und setzt ideale Quellen an den Eingangsklemmen voraus. Wir beschränken und hier auf diesen Fall und nehmen ausserdem an, es handle sich um eine Spannungsquelle. Damit ist die Eingangsspannung eine bekannte, konstante Grösse. Die Ausgangsspannung hängt offenbar nicht nur von der Eingangsspannung und dem Zweipol selber, sondern auch von der angeschlossenen Last ab. Wir beschränken uns im Folgenden auf einen unbelasteten Ausgang (*Leerlauf* am Ausgang), d.h. wir setzen $I_a = 0$. Figur 52 zeigt die entsprechende Schaltung.



Figur 52: Passives Zweitor bestehend aus zwei Impedanzen mit idealer Spannungsquelle am Eingang und Leerlauf am Ausgang.

Da alle beteiligten Komponenten lineare Zweipole sind, muss Ausgangsspannung proportional zur Eingangsspannung sein:

$$\underline{U}_a = \underline{v} \underline{U}_e.$$

Dabei ist die komplexe, frequenzabhängige Grösse \underline{v} die *Verstärkung* oder *Übertragungsfunktion* des Zweitors. Deren Betrag ergibt den *Amplitudengang* und deren Phase den *Phasengang* des Zweitors.

Wir betrachten nun den besonders einfachen Fall eines Zweitors bestehend aus zwei Impedanzen \underline{Z}_1 und \underline{Z}_2 gemäss Figur 52. Dabei handelt es sich um einen einfachen *Spannungsteiler*, für den

$$v = \frac{U_a}{U_e} = \frac{Z_2}{Z_1 + Z_2}$$

gilt. Sind die Impedanzen Z_1 und Z_2 Ohm'sche Widerstände R_1 bzw. R_2 , so erhalten wir offensichtlich eine frequenzunabhängige Verstärkung mit Betrag zwischen 0 und 1 und Phase 0. Ein ähnlich „langweiliges“ Verhalten ergibt sich, wenn beide Impedanzen Kapazitäten oder Induktivitäten sind. Interessant sind hingegen die Kombinationen von einem Widerstand mit einer Kapazität oder einer Induktivität und die Kombinationen einer Kapazität mit einer Induktivität.

Ist Z_1 eine Kapazität C und Z_2 ein Widerstand R , so erhalten wir

$$y = U_a/U_e = R/[1/(j\omega C) + R]$$

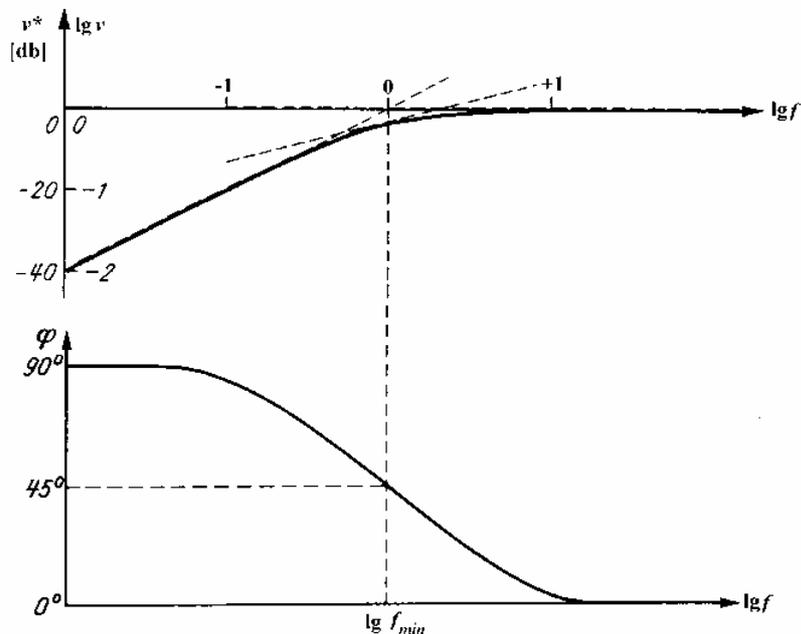
und somit für den Betrag der Verstärkung

$$v = |y| = R / \sqrt{R^2 + 1/(\omega C)^2}$$

und für die Phase

$$\varphi = \arctan(1/\omega RC).$$

Für tiefe Frequenzen ($\omega \rightarrow 0$) ist die Verstärkung nahezu 0, für hohe Frequenzen ($\omega \rightarrow \infty$) hingegen nahezu 1. Aus diesem Grund wird der Zweitor *Hochpass* genannt.



Figur 53: Bode-Diagramm, Amplitudengang (oben) und Phasengang (unten) eines RC Hochpasses mit Grenzfrequenz $f_{min} = 1\text{Hz}$.

In der logarithmischen Darstellung des Amplitudengangs erkennt man zwei nahezu lineare Abschnitte, deren Tangenten sich bei v_{min} kreuzen. Die zugehörige Frequenz $f_{min} = \omega_{min}/2\pi$ ist die untere *Grenzfrequenz*. Es gilt:

$$v_{min} = \frac{1}{\sqrt{2}} = \frac{R}{\sqrt{R^2 + \frac{1}{\omega_{min} C}}}$$

Daraus folgt nach kurzer Umrechnung

$$\omega_{min} = 2\pi f_{min} = \frac{1}{RC}.$$

Oberhalb der Grenzfrequenz ist der Betrag der Verstärkung des RC-Hochpasses näherungsweise 1, die Ausgangsspannung also nahezu gleich der Eingangsspannung. Unterhalb der Grenzfrequenz nimmt der Betrag der Verstärkung in der doppelt logarithmischen Darstellung nahezu linear ab. Charakteristisch für diesen Bereich ist die entsprechende Steigung. Für tiefe Frequenzen, d.h. für $\lg(\omega) \rightarrow -\infty$ erhalten wir

$$\lg v = \lg \left(R / \sqrt{R^2 + 1/(\omega C)^2} \right) = \lg R - \frac{1}{2} \lg \left(R^2 + 1/(\omega C)^2 \right) \approx \lg R - \frac{1}{2} \lg \left(1/(\omega C)^2 \right) = \lg(\omega RC).$$

Mit $\lg v = \lg(f/f_{min})$ findet man, dass die asymptotische Steigung $m = 1$ ist. Die Asymptote geht ausserdem durch den Punkt $(\lg(f_{min}), 0)$ in der doppelt logarithmischen Darstellung.

In der Elektrotechnik ist es üblich, anstelle von $\lg v$ die dazu proportionale Grösse

$$v^* = 20 \cdot \lg v \text{ [db]}$$

zu verwenden³. Die zugehörige Einheit db (*Dezibel*) ist dimensionslos. Dadurch ergibt sich eine Skalierung der Werte in vertikaler Richtung in unserer doppelt logarithmischen Darstellung, welche sich in der Praxis bewährt hat, da Angaben in Dezibel (ohne zusätzliche Kommastellen) für die meisten praktischen Anwendungen ausreichend genau sind.

Als Einheit für die horizontale Richtung in unserer doppelt logarithmischen Darstellung verwendet man oft die *Dekade*. Das ist nichts anderes als der Faktor 10, der sich ergibt, wenn $\lg f$ um 1 zunimmt. Damit ergibt sich für die Steigung $m = 1$ auch der Wert $m^* = 20$ db/Dekade⁴.

³ Dieselbe Definition wird auch verwendet, wenn v ein Quotient von Stromamplituden ist. Ist v hingegen ein Quotient von Leistungen, so schreibt man $v^* = 10 \cdot \lg v$ weil Leistungen Produkte von Strömen und Spannungen sind.

⁴ In der Musik bevorzugt man bekanntlich die Oktave, welche dem Faktor 2 entspricht. Die Steigung 20 db/Dekade ist nahezu 6 db/Oktave. D.h. bei einer Frequenzverdoppelung steigt die Verstärkung um 6db, was wiederum nahezu dem Faktor 2 entspricht.

Der Amplitudengang in der doppelt logarithmischen Darstellung (*Bode-Diagramm* genannt) lässt sich durch die beiden Asymptoten für tiefe und hohe Frequenzen und die Grenzfrequenz f_{min} bzw. dem Schnittpunkt der beiden Asymptoten schon recht gut approximieren. Die Steigungen dieser Asymptoten sind 20 db/Dekade bzw. 0 db/Dekade. Diese Approximation stimmt offenbar in der Nähe der Grenzfrequenz, d.h. beim Schnittpunkt der beiden Asymptoten relativ schlecht. Man kann deshalb den Wert der Verstärkung $v_{min} = 1/\sqrt{2}$ bei f_{min} hinzuziehen und die Tangente im entsprechenden Punkt bestimmen. Man findet $v^*_{min} \approx -3\text{db}$ und die Tangentensteigung $m^* = 10 \text{ db/Dekade}$. Dies ist der Mittelwert der beiden asymptotischen Steigungen.

Setzen wir für das Zweitor bestehend aus zwei Impedanzen $Z_1 = R$ und $Z_2 = L$, so erhalten wir einen *RL Hochpass*, dessen Frequenzgang demjenigen des RC Hochpasses sehr ähnlich ist. Weil die Impedanz einer Induktivität etwas einfacher als diejenige einer Kapazität ist, werden die entsprechenden Formeln sogar etwas einfacher als beim RC Hochpass. In der Praxis ist die Herstellung von ausreichend grossen Induktivitäten bei tiefen Frequenzen sehr kostspielig. Daher werden RL Hochpässe selten verwendet. Wir verzichten deshalb auf eine explizite Diskussion.

Vertauschen wir Widerstand und Kapazität beim RC Hochpass, d.h. setzen wir $Z_1 = R$ und $Z_2 = C$, so erhalten wir einen *RC Tiefpass*. Bei der Analyse des RC Tiefpasses geht man formal genau gleich vor wie beim RC Hochpass. Zunächst wird die komplexe Verstärkung

$$v = \frac{1}{1 + j\omega RC}$$

berechnet und daraus Betrag

$$v = \frac{1}{\sqrt{1 + (\omega RC)^2}}$$

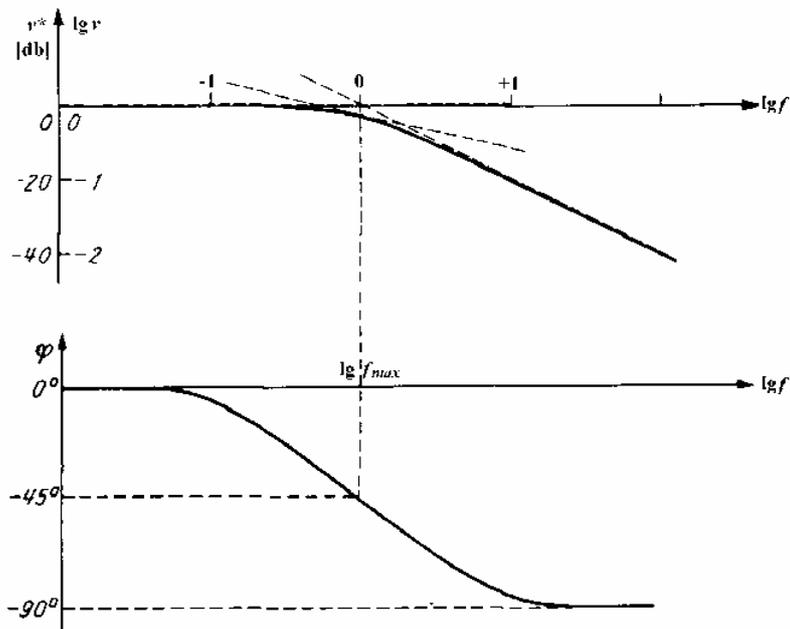
und Phase

$$\varphi = -\arctan(\omega RC)$$

sowie die obere Grenzfrequenz

$$\omega_{max} = 2\pi f_{max} = \frac{1}{RC}$$

bestimmt. Wie man auch in Figur 54 sieht, weisen die Frequenzgänge von RC Hoch- und Tiefpass starke Ähnlichkeiten auf. Die Formeln für die Grenzfrequenzen sind gar identisch. Auch der RC Tiefpass lässt sich durch zwei Asymptoten (mit den Steigungen 0 db/Dekade und -20 db/Dekade), welche sich bei der Grenzfrequenz schneiden, gut approximieren. Bei der Grenzfrequenz hat man wieder eine Verstärkung $v^*_{max} \approx -3\text{db}$ und eine Tangente mit der Steigung $m^* = -10 \text{ db/Dekade}$.



Figur 54: Bode-Diagramm, Amplitudengang (oben) und Phasengang (unten) eines RC Tiefpasses mit Grenzfrequenz $f_{\min} = 1\text{Hz}$.

Das Vorgehen bei der Analyse eines *RL Tiefpasses* mit $\underline{Z}_1 = L$ und $\underline{Z}_2 = R$, ist dasselbe wie in den betrachteten Fällen. Wir verzichten deshalb auf eine explizite Darstellung.

Setzen wir $\underline{Z}_1 = L$ und $\underline{Z}_2 = C$, so erhalten wir einen *LC Tiefpass*. Setzen wir hingegen $\underline{Z}_1 = C$ und $\underline{Z}_2 = L$, so erhalten wir einen *LC Hochpass*. Da das Vorgehen wieder dasselbe ist, wie in den vorhergehenden Beispielen, verzichten wir auf eine weitere Ausführung, weisen aber darauf hin dass *LC Hoch- und Tiefpass* in den Sperrbereichen eine verdoppelte Steigung der Asymptote ($\pm 40\text{ dB/Dekade}$) aufweisen. Das heisst dass ein LC Hochpass Signale mit tiefen Frequenzen stärker unterdrückt als ein RC Hochpass.

Weil die bisher betrachteten Hoch- und Tiefpässe einen Teil der Frequenzen mehr oder weniger stark dämpfen, bzw. herausfiltern, nennt man derartige *Zweipole Filter*. Grob gesprochen besteht jeder Hoch- und Tiefpass aus einem Durchlass- und einem Sperrbereich, wobei die beiden Bereiche durch die Grenzfrequenz getrennt werden. In der Umgebung der Grenzfrequenz von Durchlass- und einem Sperrbereich erscheint zumindest im einfachen RC Fall allerdings etwas kühn. Für viele Anwendungen ist die Separation von Sperr- und Durchlassbereich einfacher Hoch- und Tiefpässe bestehend aus zwei Zweipolen nicht ausreichend. Darüber hinaus braucht man für viele Anwendungen Filter mit komplizierteren Charakteristiken. Die einfachsten davon sind *Bandpass* und *Bandsperr*. Der Bandpass weist einen endlichen Durchgangsbereich — eingerahmt von Sperrbereichen — und die Bandsperr einen endlichen Sperrbereich — eingerahmt von Durchgangsbereichen — auf. Einfache Bandpässe und Bandsperr lassen sich durch Hintereinanderschalten von je einem Hoch- und einem Tiefpass konstruieren. Aufgabe des Filterdesigners ist es, durch geeignete RLC Netzwerke Zweitore zu entwerfen, welche vorgegebene Spezifikationen er-

füllen. Diese Aufgabe ist allerdings recht anspruchsvoll und kann im Rahmen dieser Vorlesung nicht behandelt werden.

5. Halbleiterschaltungen

Halbleiter sind spezielle, kristalline Werkstoffe, deren elektrische Eigenschaften durch gezielte Verunreinigung (Dotierung) stark verändert werden können. Insbesondere an Kontaktflächen unterschiedlich dotierter Halbleiter können interessante Quanteneffekte auftreten, welche die Herstellung von Halbleiterbauelementen genutzt werden. In dieser Vorlesung verzichten wir der Einfachheit halber auf eine Beschreibung dieser mikroskopischen Effekte. Wir beschränken uns also auf die schaltungstechnischen Aspekte und beschreiben die Bauelemente durch ihre Kennlinien oder vereinfachte Ersatzschaltungen.

1948 wurde der erste Transistor als eine Erweiterung der bereits bekannten Halbleiterdiode untersucht. Dabei wurde die Halbleiterdiode, welche aus einer sogenannten n- und einer p-dotierten Schicht aus Germanium bestand, durch Addition einer weiteren Schicht erweitert und es entstand eine pnp Struktur mit interessanten Eigenschaften. Danach wurden laufend technologische Verbesserungen der Materialien sowie der Halbleiterbauelemente gemacht, welche deren Eigenschaften (Abmessungen, Verstärkung, Verlustleistung, Schaltverhalten, Frequenzgang, Temperaturempfindlichkeit, usw.) enorm verbesserten. Selbstverständlich wurden auch weitere Schichten zu pnp Struktur hinzugefügt. So entstand der Thyristor mit einer pnpn Struktur, welcher in der Leistungselektronik Verwendung fand. Für die elektronische Schaltungstechnik blieben aber Dioden und Transistoren zentral, so dass wir uns hier auf diese beiden Klassen von Halbleiterbauelementen beschränken werden.

Obwohl die technologische Verbesserung der Eigenschaften der Bauelemente sehr wichtig ist, bietet doch die Schaltungstechnik oft einfachere Ansätze um störende Effekte zu vermindern. So ist beispielsweise die Nichtlinearität der Kennlinien von Transistoren beim Bau von Verstärkern sehr störend. Durch geeignete Zusammenschaltung von mehreren Transistoren und Widerständen erhält man jedoch ausgezeichnete Verstärker mit hoher Verstärkung und geringer Nichtlinearität. Angesichts der enormen Vielfalt von Halbleiterschaltungen müssen wir uns im folgenden auf die wichtigsten Aspekte und auf die einfachsten Grundsaltungen beschränken.

Der Wunsch nach Miniaturisierung und die Möglichkeit, Mängel von Bauelementen schaltungstechnisch auszumerzen, führte auf die Herstellung integrierter Schaltungen (integrated circuits — IC), welche heute in enorm vielen Formen erhältlich sind. Die meisten elektronischen Schaltungen bestehen deshalb aus mehreren ICs, während nicht integrierte Transistoren immer seltener eingesetzt werden. Ein typisches Beispiel aus der Analogtechnik sind sogenannte Operationsverstärker, welche für viele Anwendungen — insbesondere für Analogrechner — sehr nützlich sind. Noch viel weiter verbreitet sind integrierte Schaltungen in der Digitaltechnik. Hier ist die Vielfalt so gross, dass wir uns auf die einfachsten, logischen Schaltungen beschränken müssen.

5.1. Dioden

Halbleiterdioden bestehen aus einer einfachen pn Struktur, d.h. aus einer p (positiv) und n (negativ) dotierten Halbleiterschicht. Die ersten Halbleiterdioden bestanden aus Selen und Germanium. Heute sind Silizium und Galliumarsenid viel weiter verbreitet. Wir haben bereits im Abschnitt 2.1.10 unterschiedlichste Varianten von Dioden skizziert. Diese ermöglichen den vielfältigen Einsatz von Dioden für unterschiedlichste Anwendungen. Wir werden uns im folgenden auf die wichtigsten Anwendungen beschränken.

Zunächst betrachten wir die typische Kennlinie einer Halbleiterdiode etwas genauer. Figur 55 zeigt die prinzipielle Kennlinie und die wichtigsten charakteristischen Grössen.



Figur 55: Kennlinie einer Halbleiterdiode mit den charakteristischen Grössen: Maximale Sperrspannung $U_{\text{sperr max}}$, maximaler Durchlassstrom I_{max} und Durchlassspannung U_D bei $I_D = I_{\text{max}}/10$.

Für die Diodenkennlinie bei nicht zu grossen Strömen kann man in guter Näherung¹

$$I = I_S(T) (\exp(U/U_T) - 1)$$

schreiben, wobei, T die Temperatur, I_S der Sperrstrom im Wendepunkt der Sperrkennlinie und

$$U_T = kT/e_0$$

ist. Dabei ist k die Boltzmann-Konstante und e_0 die Elementarladung. Bei Zimmertemperatur ist $U_T \approx 25$ mV, I_S für Germaniumdioden etwa $10\mu\text{A}$ und für Siliziumdioden etwa 10nA . Für $U_T = 25\text{mV}$ ergibt sich damit bei einem typischen Durchlassstrom von 10mA eine Durchlassspannung U_D von etwa 0.27V für Germaniumdioden und etwa 0.56V für Siliziumdioden.

¹ Unter Vernachlässigung von Lichteinstrahlung, dynamischen Effekten, etc.

Die maximale Sperrspannung beträgt bei Zenerdioden 1...200V. Für Gleichrichterioden kann die maximale Sperrspannung noch höhere Werte aufweisen. Der maximale Durchlassstrom hängt stark von der Bauform² der Diode ab und kann einige Ampère betragen.

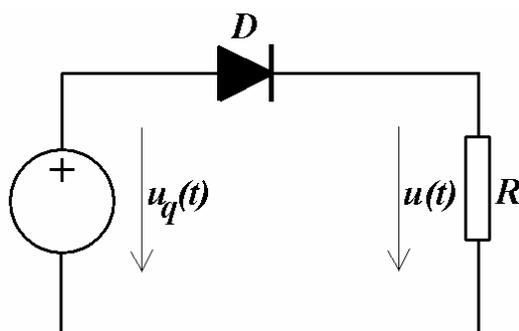
Je nach Anwendung und gewünschter Genauigkeit kann die oben angegebene Diodenkennlinie vereinfacht werden. Da die Nichtlinearität der Diodenkennlinie die Rechnung erheblich erschwert, linearisiert man die Kennlinie gerne in mindestens zwei Bereichen. Stark vereinfacht kann man im Sperrbereich ($-U_{sperr\ max} < U < U_D$) $I = 0$ und im Durchlassbereich ($U > U_D$) $U = U_D$ schreiben.

5.1.1. Gleichrichterschaltungen

Die wichtigste Eigenschaft der Diode ist ihre gleichrichtende Wirkung. Die ideale Diode ist ein idealer *Gleichrichter*, d.h. sie lässt Strom nur in einer Richtung durch. Für Gleichrichter gibt es verschiedene Anwendungen.

Die wichtigste Anwendung ist die Umwandlung von Wechsel- in Gleichstrom. Dabei ist zu beachten, dass unser Stromnetz mit Wechselstrom arbeitet, weil bei Wechselstrom die Spannung mittels Transformatoren herauf- und heruntergesetzt werden kann, während elektronische Schaltungen mit Gleichstrom versorgt werden müssen. Nehmen wir an, es stehe eine (nahezu ideale) Wechselspannungsquelle $u_q(t) = \hat{U}_q \sin(\omega t)$ — z.B. die Steckdose oder der Ausgang eines Transformators, welcher primärseitig an die Steckdose angeschlossen wurde — zur Verfügung. Verbinden wir einen Anschluss der Wechselspannungsquelle mit einer idealen Diode gemäss Figur 56, so erhalten wir am Ausgang der Diode eine zeitlich veränderliche Spannung $u(t)$, deren Vorzeichen stets positiv ist. Es gilt:

$$u(t) = u_q(t), \text{ falls } u_q(t) > 0 \text{ und andernfalls } u(t) = 0.$$



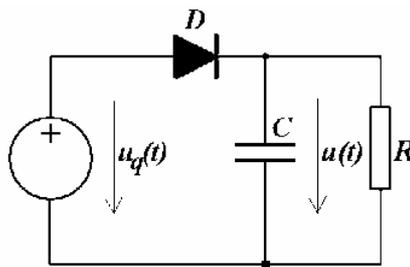
Figur 56: Gleichrichtung mit einer Diode D und einem Lastwiderstand R .

² Bei hohen Durchlassströmen ergeben sich hohe Verlustleistungen, welche die Diode erwärmen und schliesslich zerstören. Bei Dioden mit hohen Durchlassströmen muss man deshalb in erster Linie dafür sorgen, dass die Wärmeenergie gut abgeführt werden kann.

Die Fourieranalyse dieser Zeitfunktion zeigt, dass $u(t)$ tatsächlich einen zeitlich konstanten Teil aufweist, daneben aber auch Fourierkomponenten mit der Kreisfrequenz ω und ganzzahligen Vielfachen davon. Es gilt:

$$u(t) = \hat{U}_q \left[\frac{1}{\pi} + \frac{1}{2} \sin(\omega t) - \frac{2}{\pi} \left(\frac{\cos(2\omega t)}{1 \cdot 3} + \frac{\cos(4\omega t)}{3 \cdot 5} + \frac{\cos(6\omega t)}{5 \cdot 7} + \dots \right) \right].$$

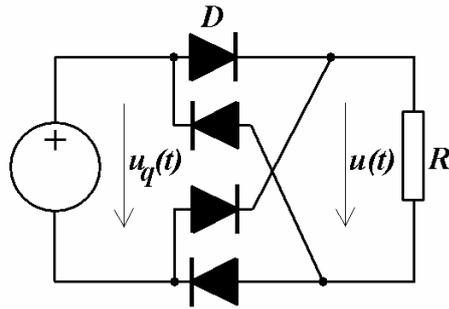
Will man die zeitabhängigen Anteile unterdrücken, so kann man einen Tiefpass nachschalten. Dessen Grenzfrequenz sollte möglichst weit unterhalb der Grundfrequenz ω liegen, damit die erste Fourierkomponente, welche doch eine recht grosse Amplitude $\hat{U}_q/2$ aufweist, genügend gut unterdrückt wird. Wie wir bereits gesehen haben, besteht ein einfacher Tiefpass aus einer Kapazität C und einem Widerstand R . Um eine möglichst tiefe Grenzfrequenz zu erhalten, setzen wir $C = \infty$ (siehe Figur 57)³. Damit erhalten wir die Grenzfrequenz 0, d.h. eine optimale Filterwirkung für jede beliebige Kapazität und die konstante Ausgangsspannung $u(t) = \hat{U}_q$. Wird unsere Gleichrichterschaltung mit einem Ohm'schen Widerstand R_L belastet, so entlädt sich die Kapazität immer dann, wenn die Eingangsspannung kleiner als die Spannung über der Kapazität ist.



Figur 57: Gleichrichtung mit einer Diode D und nachfolgender Glättung mit einer Kapazität C , Lastwiderstand R .

Ein offensichtlicher Mangel der sogenannten *Einweggleichrichtung* mit einer einzigen Diode ergibt sich aus der Tatsache, dass die Wechselspannungsquelle nur während der positiven Halbperiode belastet wird. Abhilfe schafft der Brückengleichrichter, welcher aus 4 Dioden besteht (siehe Figur 58).

³ Praktisch ergeben natürlich die nötigen Drähte einen Widerstand >0 .



Figur 58: Gleichrichtung mit einem Brückengleichrichter bestehend aus 4 Dioden D mit Lastwiderstand R .

Für den idealen Brückengleichrichter gilt einfach:

$$u(t) = |u_q(t)|.$$

Die entsprechende Fourier-Zerlegung ergibt:

$$u(t) = \hat{U}_q \left[\frac{2}{\pi} - \frac{4}{\pi} \left(\frac{\cos(2\omega t)}{1 \cdot 3} + \frac{\cos(4\omega t)}{3 \cdot 5} + \frac{\cos(6\omega t)}{5 \cdot 7} + \dots \right) \right].$$

Dabei ist beachtenswert, dass die Amplitude des Gleichspannungsterms — gegenüber der Einweggleichrichtung — verdoppelt wurde und dass die erste harmonische Komponente fehlt, was die Filterung der zeitveränderlichen Komponenten natürlich erleichtert. Brückengleichrichter sind vorteilhaft, wenn der Lastwiderstand nicht sehr klein ist. Aus diesem Grunde werden in Netzteilen zur Energieversorgung von elektronischen Geräten fast ausschliesslich Brückengleichrichter eingesetzt. Bei hoher Belastung und bei hohen Anforderungen an die Qualität der gelieferten Gleichspannung sind trotzdem sehr grosse Kapazitäten erforderlich und diese sind sowohl teuer als auch voluminös und schwer. Bei modernen Netzgeräten verzichtet man daher auf eine saubere Glättung mit einer grossen Kapazität und verwendet stattdessen kompliziertere elektronische Schaltungen mit kleineren Kapazitäten in Kombination mit Widerständen, Transistoren und Zenerdioden.

In der Nachrichtentechnik wurde die gleichrichtende Wirkung von Dioden schon zu Beginn des 20. Jahrhunderts zur Detektion von amplitudenmodulierten Signalen in Radioempfängern eingesetzt. Damals kamen übrigens Kristalldetektoren und etwas später Röhrendioden zum Einsatz. Das Prinzip dabei ist dasselbe wie bei der Einweggleichrichtung. Dabei wird zunächst das Antennensignal zu einem Schwingkreis geführt, welcher den gewünschten Sender „herausfiltert“. Am Ausgang des Schwingkreises erhält man dann ein amplitudenmoduliertes Signal der Form

$$u(t) = a(t)\sin(\omega t).$$

Mit der Einweggleichrichtung erhält man näherungsweise

$$u(t) = a(t) \left[\frac{1}{\pi} + \frac{1}{2} \sin(\omega t) - \frac{2}{\pi} \left(\frac{\cos(2\omega t)}{1 \cdot 3} + \frac{\cos(4\omega t)}{3 \cdot 5} + \frac{\cos(6\omega t)}{5 \cdot 7} + \dots \right) \right],$$

wenn sich $a(t)$ zeitlich wesentlich langsamer ändert als $\sin(\omega t)$. Filtert man die höheren Frequenzen mit einem Tiefpass heraus, so bleibt das eigentliche Nachrichtensignal

$$u(t) = a(t)/\pi.$$

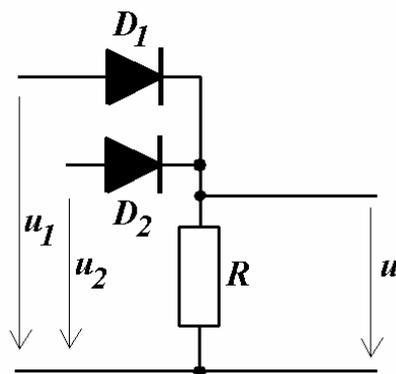
Dieses Verfahren wird im Prinzip auch heute noch zur Detektion, d.h. Demodulation amplitudenmodulierter Signale verwendet.

Neben den hier skizzierten Einweg- und Brückengleichrichtern wurden verschiedene weitere Gleichrichterschaltungen entwickelt, welche aus mindestens 2 Dioden bestehen, hier aber nicht explizit vorgestellt werden. Dazu gehören insbesondere Spannungsvervielfacher, welche in der Hochspannungstechnik zum Erzeugen hoher Gleichspannungen (bei geringem Strombedarf) verwendet werden.

5.1.2. Logische Schaltungen

Bei der elektronischen Realisierung binärer logischer Schaltungen wird meist der logische Wert 1 durch eine Spannung oberhalb einer bestimmten Schwelle U_1 und der logische Wert 0 durch eine Spannung unterhalb einer bestimmten Schwelle U_0 (mit $U_0 < U_1$) beschrieben. Die Wahl dieser Schwellen hängt von der verwendeten Technologie ab.

Logische Schaltungen lassen sich mit Netzwerken von Schaltern realisieren. In besonders einfachen Fällen reicht der Einsatz von Dioden als Schalter. Figur 59 zeigt eine ODER-Schaltung mit zwei Dioden und einem Widerstand.



Figur 59: ODER-Schaltung mit zwei Dioden und einem Widerstand.

Für ideale Dioden ergibt sich offenbar

$$u = \max(u_1, u_2),$$

wenn an den beiden Eingängen ideale Spannungsquellen vorausgesetzt werden. Dies gilt auch in guter Näherung wenn der Innenwiderstand der Quellen am Eingang viel kleiner als der Widerstand R in dieser Schaltung ist. Stellt u_1 oder u_2 eine logische 1 dar, d.h. gilt

$$(u_1 > U_1) \text{ ODER } (u_2 > U_1),$$

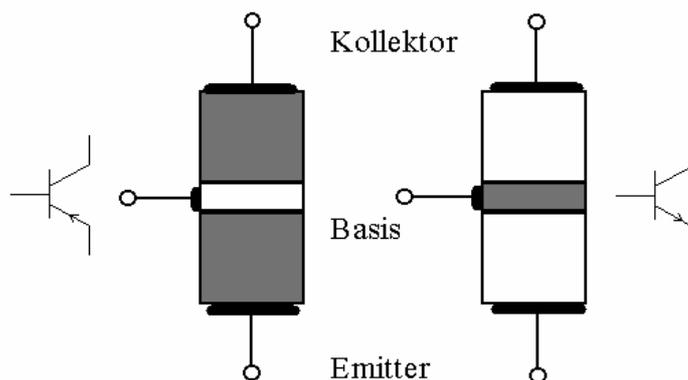
so wird auch $u > U_1$. Also ist dann auch am Ausgang eine logische 1. Umgekehrt verifiziert man leicht, dass am Ausgang eine logische Null erhalten wird, wenn an beiden Eingängen logische Nullen gesetzt sind.

Reale Dioden „schalten“ nicht bei der Spannung $0V$, sondern erst bei einer Schwellenspannung, welche von den verwendeten Materialien abhängt und bei Silizium ca. $0.6V$ beträgt. Sind die Spannungen an den Eingängen der ODER-Schaltung nur knapp oberhalb der Schwelle U_1 , so liegt die Ausgangsspannung etwas unterhalb U_1 , was nicht akzeptabel ist. Um diesen Fehler zu beheben, werden Transistoren eingesetzt. Die daraus entstehenden logischen Schaltungen nennt man DTL (Dioden-Transistor-Logik).

Da ohnehin Transistoren für logische Schaltungen erforderlich sind und die Herstellung integrierter Schaltungen um so einfacher wird je geringer die Anzahl unterschiedlicher Bauelemente ist, wurde die DTL Technik durch Techniken, welche ohne Dioden auskommen, abgelöst. Auf diese wird später eingegangen.

5.2. Bipolare Transistoren

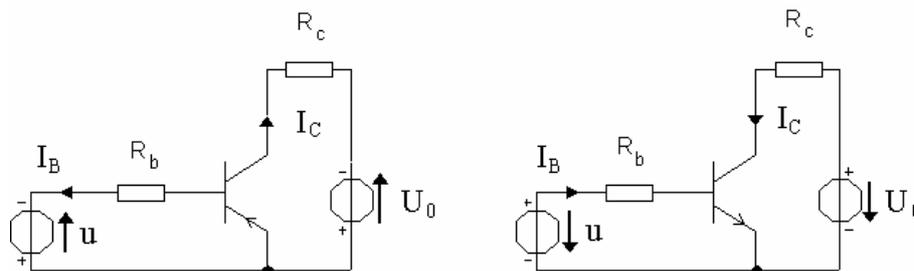
Bipolar-Transistoren bestehen aus drei Schichten, entweder in der Reihenfolge pnp oder npn wie in Figur 60 schematisch dargestellt. Die beiden äusseren Schichten werden *Kollektor* bzw. *Emitter* genannt, die mittlere Schicht ist die *Basis*.



Figur 60: pnp- (links) und npn- (rechts) Transistor mit zugehörigem Symbol

Auf den ersten Blick wirkt ein Transistor wie eine Serieschaltung von entgegengesetzt orientierten Dioden, welche an der Basis miteinander verbunden sind. Man könnte daher vermuten, dass kein Strom von Kollektor zu Emitter fließen kann. Diese Vermutung ist aber falsch, weil die mittlere Zone sehr dünn und nur schwach dotiert ist, so dass — je nach Basisstrom — auch ein Strom von Kollektor zu Emitter fließen kann, welcher sogar viel grösser als der Basisstrom ist.

Üblicherweise ist der obere pn-Übergang in Sperrichtung gepolt, der untere in Durchlassrichtung. Dieses prinzipielle Anschlussschema ist in Figur 61 wiedergegeben.



Figur 61: Anschlussschema und Polaritäten für den pnp- (links) und den npn- (rechts) Transistor. U_0 ist die Speisespannung für den Ausgangskreis, u symbolisiert die Eingangsspannung.

Wir betrachten nun den npn Transistor etwas ausführlicher. Im Hauptkreis (Kollektor-Emitter) liegt am Kollektor der positive Pol der Versorgungsspannung (Spannungsquelle rechts). Damit ist der obere pn-Übergang gesperrt. Im Steuerkreis (Basis-Emitter) liegt an der Basis (p-Zone) eine positive Spannung, d.h. der pn-Übergang Basis-Emitter ist in Durchlassrichtung gepolt und es fließt ein Strom I_B der durch die Steuerspannung u , sowie u_{BE} und R_b bestimmt wird.

In der Übergangszone zwischen Emitter und Basis wird der Strom primär von Elektronen getragen, da die Emitter-n-Schicht vergleichsweise hoch dotiert ist. Sobald die aus dem Emitter injizierten Elektronen in die Basis gelangen, geraten sie dort zusätzlich in den Bereich des elektrischen Feldes, welches vom gesperrten Kollektor herrührt; die Elektronen 'sehen' dabei den am Kollektor anliegenden Pluspol der Speisung und werden dorthin gezogen.

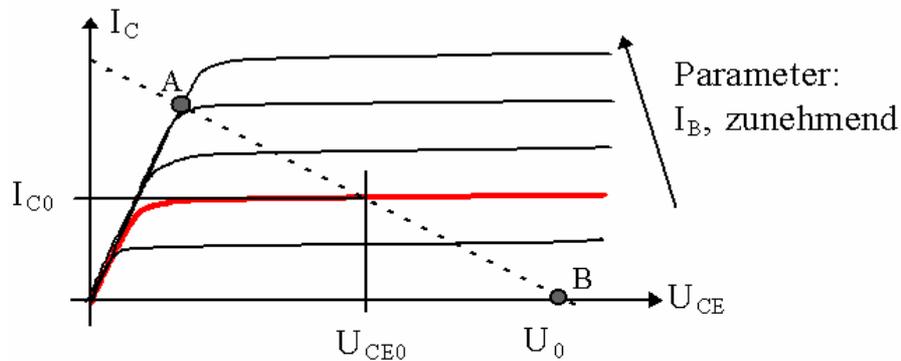
Als Ergebnis fließt der grösste Teil des Stromes I_E , welcher aus dem Emitter austritt, über den Kollektor als Kollektorstrom I_C zurück; nur ein kleiner Teil fließt als Basisstrom I_B .

Zwischen I_B und I_C herrscht in guter Näherung Proportionalität:

$$I_C = h_{fE} I_B.$$

h_{fE} ist als *statische Stromverstärkung* des Transistors bekannt. Sie beträgt zwischen 10 und einigen 100 bei den üblichen Transistoren. Diese Gleichung gilt natürlich nur dann, wenn tatsächlich die aus dem Emitter austretenden Elektronen in ein starkes Feld vom Kollektor

gelangen, also bei einer nicht zu niedrigen Kollektor-Emitter-Spannung U_{CE} . Die Ausgangskennlinie $I_C = f(U_{CE})$ ist in Figur 62 dargestellt. Man sieht, dass ab einer minimalen Spannung U_{CE} der Kollektorstrom I_C praktisch nicht mehr von U_{CE} abhängt, sondern nur noch vom Basisstrom I_B .

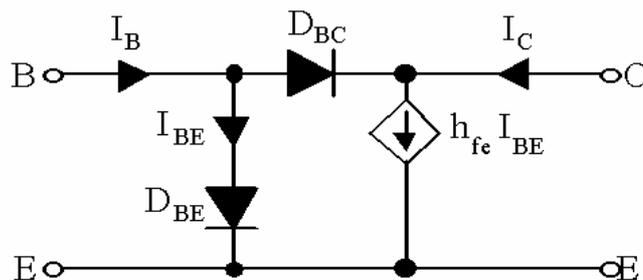


Figur 62: Kennlinienfeld eines Transistors mit eingezeichneter Arbeitsgeraden und zwei verschiedenen Arbeitspunkten A und B

Mit diesem Kennlinienfeld lässt sich nun die Ausgangsspannung U_{CE} berechnen. Es gilt für eine Schaltung nach Figur 61

$$U_{CE} = U_0 - R_C I_C (= U_0 - R_C h_{FE} I_B)$$

Ein einfaches Modell, welches das Verhalten eines bipolaren npn-Transistors bei Gleichstrom (und bei langsamen Vorgängen) in erster Näherung beschreibt, ist das folgende



Figur 63: Einfaches Grosssignal-Ersatzschema des npn bipolaren Transistors. Beim pnp-Transistor sind die Stromrichtungen und die Polaritäten der Dioden umgekehrt.

Wie man sieht, kommt zu den beiden entgegengesetzt gerichteten Dioden D_{BE} (Basis-Emitter Diode) und D_{BC} (Basis-Kollektor Diode) eine gesteuerte Stromquelle hinzu, welche dem vom Emitter in den Kollektor injizierten Kollektorstrom Rechnung trägt. Beim pnp-Transistor sind die Richtungen der Dioden, der Ströme und der Spannungen umgekehrt.

Es können folgende Betriebsarten unterschieden werden:

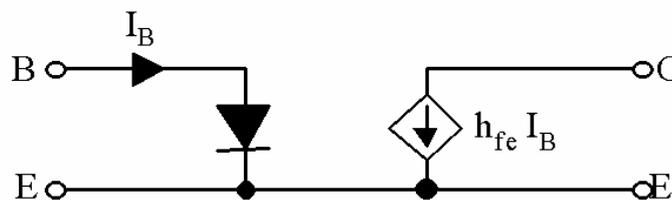
Cutoff-Bereich: Beide Dioden sind gesperrt, insbesondere die Basis-Emitter Diode: U_{BE} ist positiv bei pnp-, negativ bei npn-Transistoren. Es fließt kein Strom vom Emitter in die Basis und daher werden keine Träger von der Basis in den Kollektor injiziert. Deshalb fließt auch kein Kollektorstrom; man befindet sich im Punkt B auf der Kennlinie (Figur 62).

Sättigungs-Bereich: Beide Dioden sind leitend. Die Kollektor-Emitterspannung kann zwar nicht negativ werden, sie ist aber sehr klein ($\approx 0.3V$) und man befindet sich im Punkt A auf der Kennlinie. Es gilt

$$I_C < h_{FE} I_B$$

Ein Teil des Kollektorstromes fließt durch die Stromquelle, der andere durch die Diode zur Basis (der Anteil, welcher zum Emitter injiziert wird, ist bei modernen Transistoren sehr klein).

Aktiver Bereich: Die Basis-Emitter Diode ist leitend, während die Kollektor-Basis Diode gesperrt ist. In diesem Arbeitsbereich funktioniert der Transistor als Verstärker. Die CB-Diode kann weggelassen werden und die Ersatzschaltung vereinfacht sich zu derjenigen von Figur 64.



Figur 64: Ersatzschema des npn bipolaren Transistors im aktiven Bereich

Transistoren werden meist entweder als Schalter oder als Verstärker eingesetzt.

5.2.1. Der Transistor als Schalter

In dieser Funktionsart wird der Transistor entweder im Cutoff (Transistor ausgeschaltet, Punkt B) oder in der Sättigung (Transistor eingeschaltet, Punkt A) betrieben, siehe Figur 62. Somit ist die Verlustleistung im Transistor sehr klein, da im Cutoff der Strom sehr klein und in der Sättigung die Spannung sehr klein ist. Der aktive Bereich wird beim Umschalten sehr schnell überquert, u.a. um die Verlustleistung klein zu halten. Die beiden wesentlichen Anwendungen in dieser Funktionsart sind:

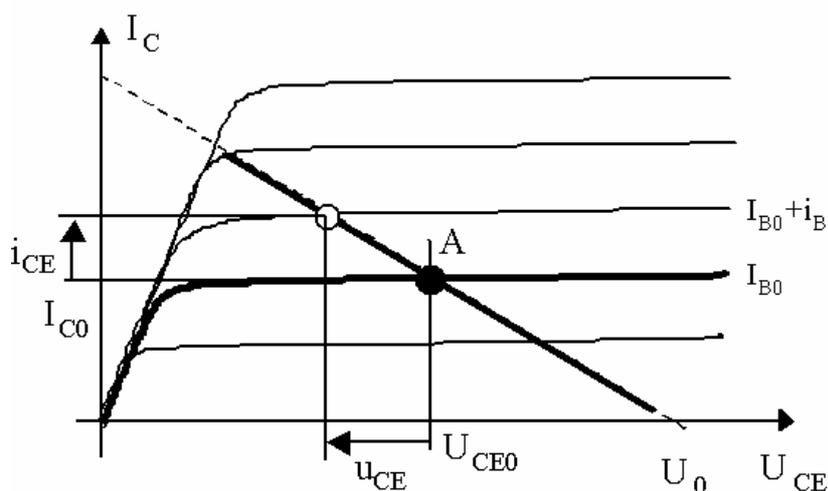
- Analogschalter, um analoge Spannungen zu schalten, z.B. in Choppern, Abtastglieder, usw.

- Digitaltechnik: Hier nimmt die Ausgangsspannung prinzipiell zwei Zustände an: $\approx 0V$ (Zustand "0") und $\approx U_0$ (Zustand "1"). Diese Funktionsart wird in einem späteren Abschnitt ausführlicher behandelt.

5.2.2. Der Transistor als Verstärker

In der Verstärkertechnik geht es darum, kleine Spannungsänderungen, bzw. kleine Wechselspannungen zu vergrössern, d.h. zu verstärken.

Bei jedem Wert des Eingangsstromes I_B muss der Transistor dabei im aktiven Bereich bleiben. Gerät der Transistor in den Cutoff- oder in den Sättigungsbereich, so wird die Ausgangsspannung (in der Amplitude) abgeschnitten und damit das Ausgangssignal verzerrt. Um dies zu verhindern, muss man einen *Arbeitspunkt*, bzw. einen Gleichstromwert definieren, um den herum die Strom- bzw. Spannungsänderungen erfolgen. Dieser Arbeitspunkt sollte möglichst in der Mitte des aktiven Bereichs liegen, wie in Figur 65 gezeigt.



Figur 65: Arbeitskennlinie eines npn Transistors im aktiven Bereich

Der Transistor ist normalerweise ein nichtlineares Element, vor allem wegen der BE-Diode. Das macht die Berechnung der Beschaltung eines Transistors schwierig. Betrachtet man das dynamische Verhalten, so wird die Lage sogar noch komplizierter, weil dann den Ladungen im Transistor Rechnung getragen werden muss und die Ersatzschaltung des Transistors durch Kapazitäten ergänzt werden muss. Ebenso erschweren thermische Effekte die Analyse. Schliesslich sind Transistoren ebenso wie die Dioden lichtempfindlich. Üblicherweise wird die Lichtempfindlichkeit durch lichtundurchlässige Gehäuse unterdrückt. Bei Phototransistoren macht man sich die Lichtempfindlichkeit jedoch zu Nutze. Wie dem auch sei, eine ausführliche Beschreibung und Analyse von Transistoren würde den Rahmen dieser Vorlesung sprengen. Wir beschränken uns deshalb im folgenden auf die wichtigsten und einfachsten Kenngrössen.

Solange die Eingangsspannungsänderung an der BE-Diode klein ist (für die meisten Anwendungsfälle trifft dies zu — sogenanntes *Kleinsignalverhalten*), kann die Kennlinie um den Arbeitspunkt (I_{C0} , U_{CE0}) herum als annähernd linear betrachtet werden und durch ihre Tangente an diesem Punkt beschrieben werden. Diese Linearisierung bringt eine drastische Vereinfachung für die Analyse, gibt dem Transistor aber ein zu perfektes Erscheinungsbild. In der Praxis muss man deshalb den Abweichungen von der Linearität Rechnung tragen.

Die Steigung der Tangenten im Arbeitspunkt bezeichnet man als *differentielle Kenngrößen* oder *Kleinsignalparameter*. Wir skizzieren kurz einige dieser Größen.

Zunächst definieren wir die differentiellen Widerstände über zwei der drei Klemmen eines Transistors als Ableitungen der entsprechenden Spannungen nach den entsprechenden Strömen. So gilt für den Steuerungskreis Basis-Emitter⁴:

$$r_{BE} = \left. \frac{\partial U_{BE}}{\partial I_B} \right|_{U_{CE}=\text{const}}$$

Um den differentiellen Widerstand r_{BE} zu berechnen, setzen wir die Diodenkennlinie der Basis-Emitter Diode ein:

$$I_B = I_S(T) (\exp(U_{BE}/U_T) - 1)$$

und finden für nicht zu kleine Spannungen U_{BE} :

$$1/r_{BE} = \frac{\partial I_S(T) (\exp(U_{BE}/U_T) - 1)}{\partial U_{BE}} = \frac{I_S(T)}{U_T} (\exp(U_{BE}/U_T) - 1) \approx \frac{I_B}{U_T}$$

Also gilt:

$$r_{BE} \approx U_T / I_B$$

Der differentielle Widerstand r_{BE} ist also näherungsweise umgekehrt proportional zum Basisstrom, wobei die Proportionalitätskonstante von der Bauart des Transistors unabhängig ist und bei Zimmertemperatur etwa 25 mV beträgt.

In ähnlicher Weise definiert man den differentiellen Widerstand r_{CE} des gesteuerten Kreises⁵:

⁴ Da es sich um den differentiellen Widerstand des Steuerkreises handelt, spricht man oft auch vom Eingangswiderstand. Für einfache Transistorschaltungen kann dieser Widerstand tatsächlich gleich dem Eingangswiderstand der Schaltung sein, für viele Transistorschaltungen ist dies aber keineswegs der Fall.

⁵ Diesen Widerstand nennt man oft auch Ausgangswiderstand, obwohl der Ausgangswiderstand vieler Transistorschaltungen von r_{CE} verschieden ist.

$$r_{CE} = \left. \frac{\partial U_{CE}}{\partial I_C} \right|_{U_{BE} = \text{const}}$$

Dieser Wert bezeichnet die Steigung der Ausgangskennlinie im Arbeitspunkt. Oberhalb der Sättigungsspannung wird r_{CE} recht hoch (typischerweise einige $k\Omega$ bis $M\Omega$).

Bei der Einführung der Transistoren, waren Röhren die wichtigsten Elemente zum Bau von Verstärkern. Während Transistoren im wesentlichen stromgesteuerte Stromquellen sind, welche durch die Stromverstärkung h_{fE} charakterisiert werden, sind Röhren spannungsgesteuerte Stromquellen, welche durch die *Steilheit* S charakterisiert werden. Natürlich ändert sich der Kollektorstrom indirekt auch wenn die Basis-Emitter Spannung geändert wird. Damit lässt sich die Steilheit auch für Transistoren angeben:

$$S = \left. \frac{\partial I_C}{\partial U_{BE}} \right|_{U_{CE} = \text{const}}$$

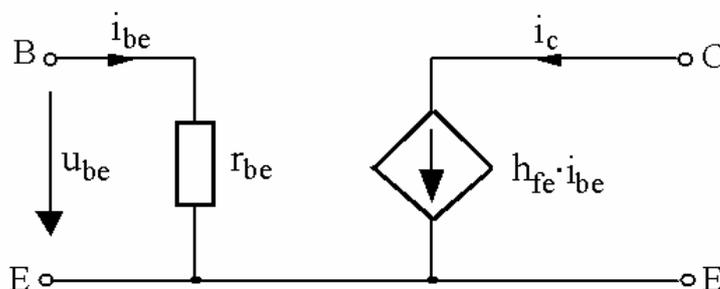
Setzen wir $I_C = h_{fE} I_B$ ein, so wird klar, dass die Steilheit zur Stromverstärkung proportional ist. Mit der Definition des differentiellen Widerstandes r_{BE} finden wir schliesslich den einfachen Zusammenhang

$$S = h_{fE} / r_{BE}.$$

Nun liegt es natürlich nahe, auch noch die *Spannungsverstärkung* v des Transistors in analoger Weise zu definieren. Für die verschiedenen Beschaltungen der Transistoren kann die Spannungsverstärkung ebenso wie die Ein- und Ausgangswiderstände sehr stark variieren. Obwohl oft eine Spannungsverstärkung erwünscht ist, gibt es auch Anwendungen mit $v \approx 1$, wie wir noch sehen werden.

Neben den erwähnten charakteristischen Grössen sind für die Verstärkertechnik verschiedene weitere Grössen von Bedeutung, auf welche wir hier nicht eingehen können. Diese kennzeichnen insbesondere Verzerrungen, Rausch- und Temperaturverhalten.

Da für das *Kleinsignalverhalten* des Transistors annähernd lineare Verhältnisse vorausgesetzt werden können, wird oft das folgende Ersatzschaltbild verwendet.



Figur 66: Kleinsignalersatzschaltbild des npn-Transistors

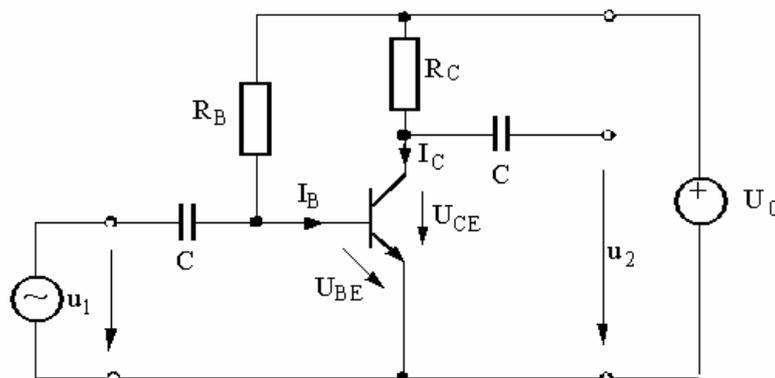
Die Berechnung von Transistorschaltungen vereinfacht sich dadurch wesentlich, weil dank der Linearität der Ersatzschaltung das Superpositionsprinzip angewendet werden kann: Quellen können separat betrachtet werden. Nimmt man eine Grundbeschaltung wie in Figur 67 an, so kann die Berechnung der externen Bauelemente nach folgendem Schema erfolgen:

1. Man setzt die Signalquelle $u_1 = 0$, berechnet den Arbeitspunkt mit dem Grosssignal-Ersatzschema und bestimmt die Kleinsignalparameter für den Arbeitspunkt.
2. Man setzt die Speisespannung $U_0 = 0$ und berechnet die Wechselstromgrössen mit dem Kleinsignal-Ersatzschema.

Dieses Vorgehen wird im Folgenden anhand einer einfachen Verstärkerschaltung erklärt.

5.2.3. Emitterschaltung

Prinzipiell gibt es drei verschiedene Typen von Transistorschaltungen, je nachdem, welcher der drei Anschlüsse des Transistors gleichstrommässig auf konstantes Potential gelegt wird. Die am meisten verwendete Schaltung ist die Emitterschaltung mit dem Emitter auf Masse, wie in Figur 67. Zur Festlegung des Arbeitspunktes und zur Einstellung der Verstärkung gibt es übrigens verschiedene Varianten der Emitterschaltung. Dabei kann entweder der Arbeitspunkt über den Basisstrom I_B oder über die Basis-Emitter Spannung U_{BE} festgelegt werden. Wir betrachten hier einen einfachen Fall mit Festlegung von I_B .⁶



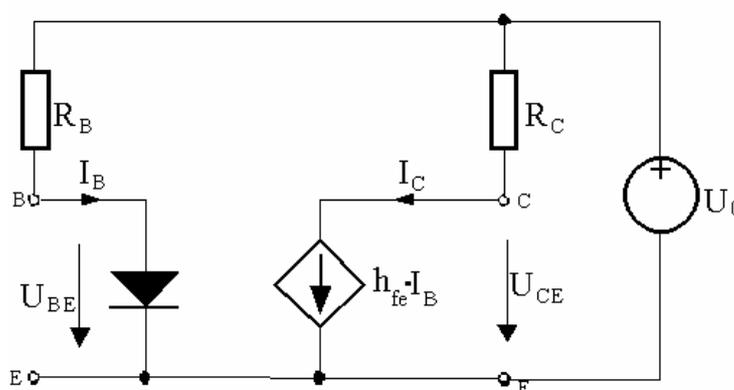
Figur 67: Emitterschaltung als Grundbeschaltung eines bipolaren Transistors. Die beiden Kapazitäten dienen der Ein- bzw. Auskopplung des Wechselspannungssignals. Sie werden üblicherweise so gross dimensioniert, dass sie bei der Berechnung des Kleinsignalverhaltens weggelassen werden können.

⁶ Zur Festlegung des Arbeitspunktes über die Basis-Emitter Spannung verwendet man einen Spannungsteiler bestehend aus zwei Widerständen über der Versorgungsspannungsquelle U_0 , also einen zusätzlichen Widerstand von der Basis zur Masse, welcher in Figur 67 fehlt.

Die Arbeitspunktgrößen werden im folgenden mit Grossbuchstaben und Index 0 (z.B. I_{B0}) gekennzeichnet. Wechselstromgrößen (Kleinsignalgrößen) werden mit Kleinbuchstaben gekennzeichnet. Wir schreiben also für die Gesamtströme bzw. -spannungen⁷:

$$I_B = I_{B0} + i_B, \quad U_{CE} = U_{CE0} + u_{CE}, \quad \text{usw.}$$

a) Berechnung des Arbeitspunktes: Die Wechselspannungsquelle u_1 wird Null gesetzt. Somit sind nur die Gleichströme zu berücksichtigen. Die Kapazitäten werden als Unterbruch betrachtet. Unter Berücksichtigung des Grosssignal-Ersatzschemas erhält man folgende Schaltung, wenn man beachtet, dass die Basis-Kollektor Diode sperrt und deshalb in guter Näherung weggelassen werden kann:



Figur 68: Ersatzschaltung für Schaltung nach Abb. 10.20 für den Gleichstromfall.

Weil die Basis-Emitter Diode im Durchlassbereich arbeitet, können wir in guter Näherung $U_{BE} = 0.7V$ schreiben und erhalten:

$$I_{B0} = (U_0 - 0.7V)/R_B.$$

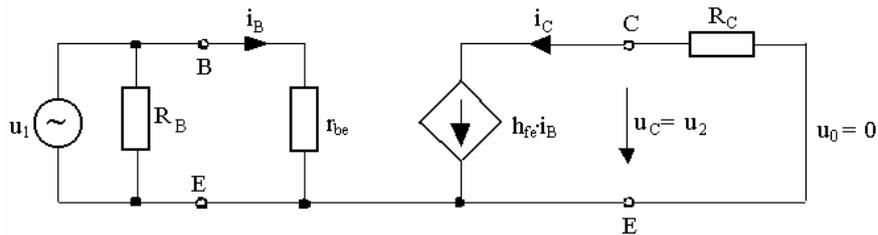
Bei vorgegebenen Arbeitspunkt, z.B. $I_{B0} = 60 \mu A$ kann diese Gleichung auch nach R_B aufgelöst werden. Auf diese Weise wird der Widerstand R_B , welcher den Basisstrom und damit den Arbeitspunkt festlegt, dimensioniert.

Die Kollektor-Emitter Spannung ergibt sich aus folgender Spannungsteilerrechnung:

$$U_{CE0} = U_0 - R_C I_{C0} = U_0 - R_C h_{FE} I_{B0}.$$

b) Berechnung der Wechselspannungsgrößen: Hier wird die Speisespannungsquelle U_0 Null (= Kurzschluss) gesetzt. Setzen wir für den Transistor die Kleinsignal-Ersatzschaltung ein, so erhalten wir die Ersatzschaltung gemäss Figur 69.

⁷ Dabei werden inkonsequenterweise Grossbuchstaben für die zeitabhängigen Gesamtströme bzw. -spannungen verwendet!



Figur 69: Ersatzschaltung für den Wechselspannungsfall (Kleinsignalverstärkung)

Für den Eingangskreis sowie die gesteuerte Stromquelle im Ausgangskreis gilt:

$$i_B = u_1 / r_{BE}, \quad i_C = h_{fE} i_B, \quad u_{CE} = -R_C i_C.$$

Durch Einsetzen folgt die Ausgangsspannung

$$u_2 = u_{CE} = -u_1 h_{fE} R_C / r_{BE}.$$

Daraus ergibt sich schliesslich die Spannungsverstärkung

$$v = u_2 / u_1 = -h_{fE} R_C / r_{BE}.$$

Es sieht so aus, als ob die Spannungsverstärkung der Emitterschaltung durch Einsatz eines grossen Widerstandes R_C fast beliebig gross gemacht werden könnte. Dabei ist zu beachten, dass wir in der obigen Berechnung den differentiellen Widerstand r_{CE} der Einfachheit halber vernachlässigt haben, weil dieser normalerweise sehr gross (im Vergleich zu R_C) ist. Ist dies nicht der Fall, so muss R_C durch die Parallelschaltung von R_C mit r_{CE} ersetzt werden. Damit wird v durch die Transistordaten begrenzt.

In der Praxis führt der Versuch, die Spannungsverstärkung eines einzelnen Transistors so gross wie möglich zu machen, zu unerwünschten Effekten, wie Signalverzerrungen.

5.2.4. Gegenkopplung

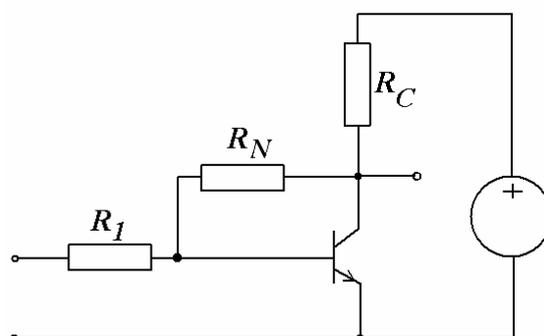
Signalverzerrungen, welche durch Nichtlinearitäten verursacht werden, sind für viele Anwendungen sehr störend. Deshalb arbeitet man mit moderaten Spannungsverstärkungen einzelner Transistoren und kaskadiert mehrere Transistoren, um grössere Verstärkungen zu erhalten.

Um die Linearität einer einzelnen Verstärkerstufe und die Stabilität des Arbeitspunktes zu verbessern, wird gerne ein Teil des Ausgangssignals an den Eingang zurückgeführt und zwar so, dass die Phase des zurückgeführten Ausgangssignals gegenüber der Phase des Eingangssignals um 180 Grad verschoben ist. Damit wirkt das rückgeführte Ausgangssignal dem Eingangssignal entgegen und vermindert die Verstärkung. Dies wird Gegenkopplung genannt. Ohne die erwähnte Phasenverschiebung ergibt sich eine Rückkopplung, wel-

che das Eingangssignal solange verstärkt, bis die Verstärkerstufe übersteuert wird, was zu enormen Verzerrungen und zu meist unerwünschten Effekten (Rückkopplungspfeifen) führt.

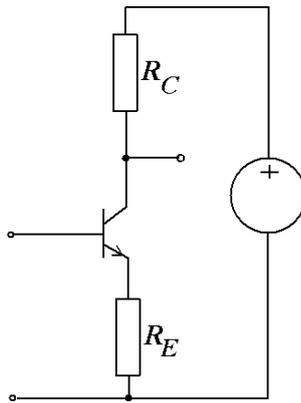
Da das Signal aus Strom und Spannung besteht, kann man sowohl den Strom als auch die Spannung gegenkoppeln. Im Prinzip wird natürlich immer Strom und Spannung gegengekoppelt, man unterscheidet jedoch zwei verschiedene Grundschaltungen, welche *Spannungsgegenkopplung* und *Stromgegenkopplung* genannt werden.

In der Formel für die Spannungsverstärkung der Emitterschaltung weist das negative Vorzeichen darauf hin, dass die Ausgangsspannung der Emitterschaltung gegenüber der Eingangsspannung um 180 Grad phasenverschoben ist. Um eine Gegenkopplung zu erhalten, ist also keine weitere Phasendrehung nötig. Man kann also lediglich einen Teil der Ausgangsspannung mit einem Spannungsteiler vom Kollektor abnehmen und an die Basis zurückführen. Im wesentlichen wird diese sogenannte Spannungsgegenkopplung durch einen Widerstand vom Kollektor zum Emitter bewerkstelligt (siehe Figur 70). Der zum Spannungsteiler gehörige Widerstand von der Basis auf Masse kann weggelassen werden, da zwischen Basis und Masse bereits der Widerstand r_{BE} liegt. Damit gleicht die Spannungsgegenkopplung der Arbeitspunkteinstellung mit einem Widerstand von der Gleichspannungsquelle zur Basis, welcher den Basisstrom festlegt.



Figur 70: Prinzip der Spannungsgegenkopplung mit dem Widerstand R_N bei der Emitterschaltung.

Bei der Stromgegenkopplung würde man wohl einen Stromteiler am Kollektor erwarten, welcher einen Teil des Ausgangsstromes an die Basis zurückführt. Ein solcher Stromteiler würde aus dem Lastwiderstand am Ausgang einerseits und einem Widerstand vom Kollektor zur Basis andererseits bestehen. Damit hätte die Stromgegenkopplung dieselbe Form wie die Spannungsgegenkopplung gemäss Figur 70. Tatsächlich aber besteht die Stromgegenkopplung bei der Emitterschaltung aus einem Widerstand R_E zwischen Emitter und Masse, wie in Figur 71 skizziert.



Figur 71: Prinzip der Stromgegenkopplung mit dem Widerstand R_E bei der Emitterschaltung.

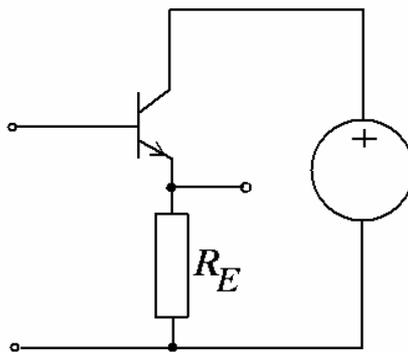
Die Wirkungsweise der Stromgegenkopplung ist die folgende: Nimmt die Eingangsspannung zu, so nimmt der Kollektorstrom und damit der Emittterstrom zu, welcher nahezu gleich dem Kollektorstrom ist. Damit steigt die Spannung über R_E und die Basis-Emitter Spannung wird reduziert weil

$$U_{Ein} = U_{BE} + U_{RE}$$

gilt.

5.2.5. Kollektorschaltung, Emitterfolger

Lässt man bei der stromgegekoppelten Emitterschaltung den Kollektorwiderstand R_C weg, so erhält man die Kollektorschaltung gemäss Figur 72. Diese ist also gewissermassen vollkommen gegengekoppelt und hat eine Spannungsverstärkung v , welche nahezu 1 ist. Dadurch ergeben sich sehr kleine Verzerrungen.



Figur 72: Prinzip der Kollektorschaltung.

Zunächst scheint eine Spannungsverstärkung $v = 1$ wenig attraktiv. Interessant ist diese Schaltung wegen ihres hohen Eingangswiderstandes

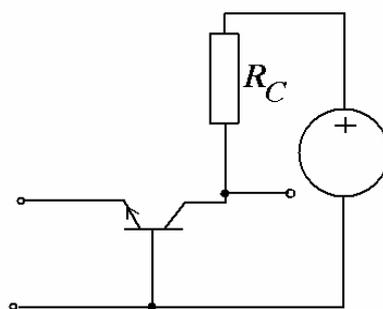
$$r_e = r_{BE} + h_{fE} R_E$$

bei relativ niedrigem Ausgangswiderstand in der Größenordnung von $R_E || (R_Q + r_{BE}) / h_{fE}$. Man kann die Kollektorschaltung deshalb als Impedanzwandler verwenden. Um den Wert dieser Schaltung schätzen zu lernen, erinnere man sich an die RC Filter, bei denen ein niedriger Eingangswiderstand und ein hoher Ausgangswiderstand vorausgesetzt wurden. Um solche Filter problemlos hintereinander zu schalten, sind Impedanzwandler nötig.

Weil bei der Kollektorschaltung das Emitterpotential dem Basispotential nachfolgt, wird diese Schaltung meist *Emitterfolger* genannt.

5.2.6. Basisschaltung

Bei der Basisschaltung (Figur 73) liegt die Basis auf Masse. Wie bei der Emitterschaltung wird der Kollektor über einen Widerstand mit der Gleichspannungsquelle verbunden. Ausserdem ist u_{BE} — ebenfalls wie bei der Emitterschaltung — die Eingangsspannung. Die Basisschaltung hat auch dieselbe Spannungsverstärkung wie die Emitterschaltung, die Ausgangsspannung ist jedoch in Phase mit der Eingangsspannung. Der wichtigste Unterschied zur Emitterschaltung ist der geringe Eingangswiderstand, da die Signalquelle mit dem Emitterstrom, d.h. der Summe von Basis- und Kollektorstrom belastet wird. In der Niederfrequenztechnik werden meist hohe Eingangswiderstände gewünscht und deshalb die Emitterschaltung der Basisschaltung vorgezogen. Wie wir noch sehen werden, arbeitet man in der Hochfrequenztechnik meist mit Impedanzanpassung und entwirft Verstärker so, dass Ein- und Ausgangswiderstand der Leitungsimpedanz (meist 50 oder 60 Ω) der verwendeten Anschlusskabel entsprechen. Die Basisschaltung ist deshalb vor allem in der Hochfrequenztechnik anzutreffen.



Figur 73: Prinzip der Basisschaltung.

5.2.7. Praktische Hinweise

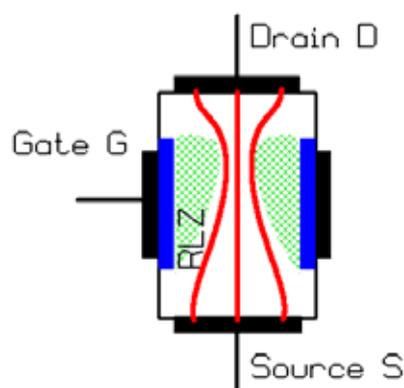
Bipolartransistoren werden als Einzelelemente angeboten. Sie unterscheiden sich durch:

- zulässige maximale Spannung U_{CEmax} (einige 10V bis über 1000V)
- zulässigen maximalen Strom (einige mA bis mehrere 100A)
- zulässiger Verlustleistung (mW bis einige 100W)
- Verstärkung h_{fE} (zwischen etwa 10 und knapp 1000)
- Frequenzbereich in welchem sie arbeiten können (MHz bis GHz)

Je nach Leistung sind verschiedene Gehäuseformen üblich, entweder Kunststoff oder Metall. Daneben findet man Transistoren mit sehr kleinen Nennströmen natürlich in jeder integrierten Schaltung.

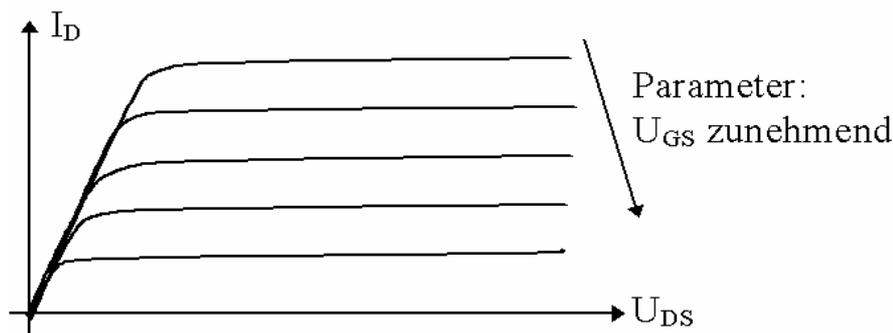
5.3. Feldeffekttransistoren

Ein radikal anderer Weg als beim Bipolartransistor wird beim Feldeffekttransistor beschritten. Dieser beruht darauf, dass beim Anlegen einer Spannung in Sperrichtung an einem pn-Übergang die beweglichen Ladungsträger abgesaugt werden und eine praktisch nicht leitende Zone übrig bleibt. Die Eindringtiefe dieser nichtleitenden Zone wird durch die angelegte Spannung gegeben und ist damit variabel. Diesen Effekt kann man benutzen, um (in einer ersten Annäherung) einen variablen Widerstand zu erzeugen, wie in Figur 74 angedeutet. Durch die Raumladungszone RLZ des ringförmig angebrachten pn-Überganges wird, bei angelegter Spannung zwischen den beiden Hauptanschlüssen D (Drain) und S (Source), eine Verengung des leitenden Pfades verursacht und damit eine Modulation des Widerstandes.



Figur 74: Widerstandsmodulation über Vorspannung eines pn-Überganges, Strompfade sind rot eingezeichnet

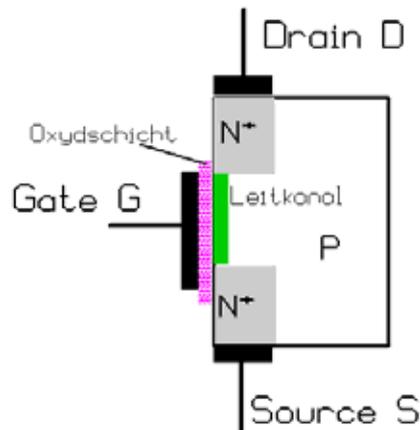
Diese Modellvorstellung ist allerdings nur bei kleinen Spannungen U_{DS} zwischen Drain und Source korrekt; bei höheren Spannungen kommt es zu einer Interaktion zwischen dem Drain-Source-Feld und demjenigen, welches durch die Gate-Source-Spannung verursacht wird. Bei einer konstanten Gate-Source-Spannung wird mit zunehmender Drain-Source-Spannung der Leitkanal verkleinert; der Widerstand nimmt also zu. Dies hat zur Folge, dass der zwischen diesen Anschlüssen fliessende Strom weitgehend konstant bleibt und es ergibt sich eine Ausgangskennlinie gemäss Figur 75.



Figur 75: Kennlinienfeld eines Feldeffekttransistors

Der eben beschriebene Feldeffekttransistor ist der Typ mit Sperrschichtsteuerung, auf Englisch *Junction Field Effect Transistor*, J-FET. Er kann sowohl mit n-leitendem als auch mit p-leitendem Ausgangsmaterial hergestellt werden. Zur Ansteuerung müssen lediglich die entsprechenden Polaritäten von U_{GS} beachtet werden. Zu beachten ist, dass J-FET's von Haus aus leitend sind. Erst durch Anlegen der GS-Spannung wird der DS Strom reduziert!

Eine Erweiterung des FETs erhält man, wenn der pn-Übergang durch eine Isolierschicht ersetzt wird, auf welche ein metallischer Kontakt aufgebracht wird. Die Struktur sieht dann prinzipiell wie in Figur 76 aus. Je nach dem Aufbringen und den Eigenschaften der Oxydschicht sind die resultierenden FET entweder von Haus aus sperrend (durch Anlegen von U_{DS} wird der Strom vergrössert, "Anreicherungstyp", engl.: enhancement) oder von Haus aus leitend (durch Anlegen von U_{DS} wird der Strom verkleinert, "Verarmungstyp", engl.: depletion). Selbstverständlich existieren auch wieder n-Kanal und p-Kanal Typen. Seinen Namen MOSFET hat diese Familie durch die Art der Erzeugung des Gates erhalten: **Metal-Oxide-Semiconductor**.



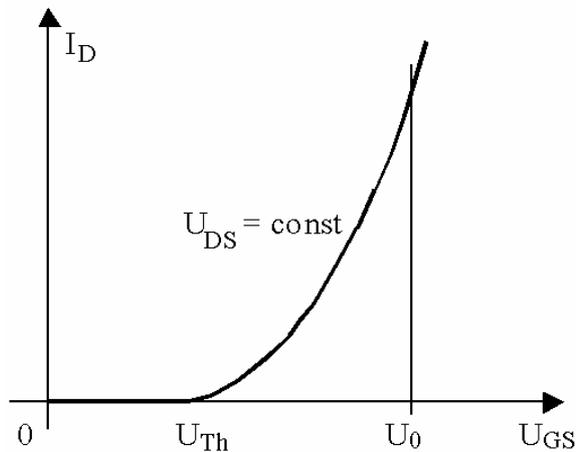
Figur 76: MOSFET (eine mögliche Variante)

Der MOSFET vom Anreicherungstyp ist heute das Grundelement der meisten digitalen integrierten Schaltungen. Früher wurden Schaltungen fabriziert, welche nur n-Kanal Transistoren (NMOS) besaßen. Heute werden fast ausschliesslich Schaltungen verwendet, welche beide Typen aufweisen (CMOS, Complementary MOS). Die Schwellenspannung (threshold voltage), d.h. die Gate-Spannung bei welcher ein Strom zwischen Source und Drain zu fließen beginnt, wird zwischen der Speisespannung und Null ausgelegt. Somit sperrt der MOSFET wenn die Gate-Spannung Null ist, während er voll leitet wenn diese bei der Speisespannung liegt. Für Einzelheiten sei auf die Vorlesung "Digitaltechnik" verwiesen.

Für die Berechnung der Schaltungen mit FET muss nun eine Formulierung gefunden werden, welche den Zusammenhang zwischen U_{GS} und dem Strom I_D angibt (siehe Figur 77). Sobald die Spannung U_{DS} einen Wert erreicht hat, bei welchem der Drainstrom in etwa konstant ist, kann er in guter Näherung durch eine spannungsgesteuerte Stromquelle ersetzt werden, d.h.

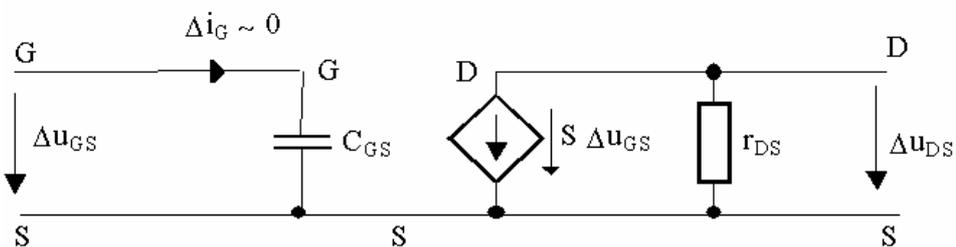
$$I_D = S U_{GS}.$$

Wir begegnen hier auf ganz natürliche Weise der Steilheit S , welche bereits für Bipolartransistoren definiert wurde. S heisst Englisch "transconductance" und hat die Dimension Strom/Spannung (meist von der Grössenordnung mA/V).



Figur 77: Kennlinie eines Enhancement MOS-Transistors

Für Feldeffekttransistoren lässt sich ein Ersatzschema angeben, welches demjenigen der Bipolartransistoren ähnelt, anstelle des Eingangswiderstandes jedoch eine Eingangskapazität C_{GS} aufweist. Für Übersichtsrechnungen kann C_{GS} übrigens weggelassen werden (Verbindung G-S offen):



Figur 78: Ersatzschema für einen FET

Damit können Verstärkerschaltungen mit FETs im wesentlichen mit der gleichen Methode wie die Bipolartransistoren berechnet werden.

Ähnlich wie bei den Bipolartransistoren wird auch bei den FET's eine ganze Palette von Typen angeboten. Zusätzliches Unterscheidungsmerkmal ist der minimale Durchlasswiderstand R_{DS} .

5.4. Operationsverstärker

Die Verstärkerschaltungen mit einem einzigen Transistor, welche wir im vorherigen Abschnitt skizziert haben, weisen verschiedene Unzulänglichkeiten auf. Es wurde bereits darauf hingewiesen, dass Verbesserungen durch Gegenkopplung, Kaskadierung mehrerer Ver-

stärkerstufen und andere Massnahmen erzielt werden können. Die Integrierte Schaltungstechnik ermöglicht es, Schaltungen mit sehr vielen Transistoren auf kleinstem Raum unterzubringen und gewissermassen durch verschwenderischen Umgang mit Transistoren Verstärkerschaltungen zu entwickeln, welche dem Idealbild eines Verstärkers sehr nahe kommen.

Ein wichtiges Problem bei den integrierten Schaltungen sind die Kapazitäten. Einerseits entstehen kleine Kapazitäten meist unerwünschterweise „von selbst“, andererseits können grössere Kapazitäten nicht integriert werden. Dasselbe gilt auch für Induktivitäten. Bei den skizzierten Transistorverstärkern wurden Koppelkondensatoren verwendet um das Wechselnungssignal ein- bzw. auszukoppeln. Diese Kondensatoren ergeben zusammen mit den Ein- und Ausgangswiderständen Hochpassfilter. Damit der entsprechende Verstärker auch bei tiefen Frequenzen arbeitet, müssen die Koppelkondensatoren relativ gross sein, können also nicht integriert werden. Aus diesem Grunde ist es vorteilhaft, bei der Entwicklung integrierter Verstärker auf Koppelkondensatoren zu verzichten und die integrierten Verstärkerstufen direkt zu verbinden. Das allfällige Hinzufügen von Hochpassfiltern wird dann dem Anwender überlassen.

Der Verzicht auf Koppelkondensatoren bringt zwar einige Schwierigkeiten beim Entwurf der Verstärkerschaltungen aber auch den Vorteil, dass die entsprechenden Schaltungen bei beliebig tiefen Frequenzen — bis hinunter zu Gleichstrom — betrieben werden können.

Die Kapazitäten, welche in allen elektronischen Schaltungen „von selbst“ entstehen ergeben normalerweise eine Tiefpasswirkung, verhindern also, dass die Schaltung bei hohen Frequenzen richtig arbeitet. Durch Miniaturisierung werden diese unerwünschten Kapazitäten meist automatisch reduziert. Integrierte Schaltungen arbeiten deshalb oft bei recht hohen Frequenzen. Trotzdem ist der Frequenzbereich immer nach oben begrenzt, wobei die obere Frequenzgrenze von der verwendeten Technologie abhängt.

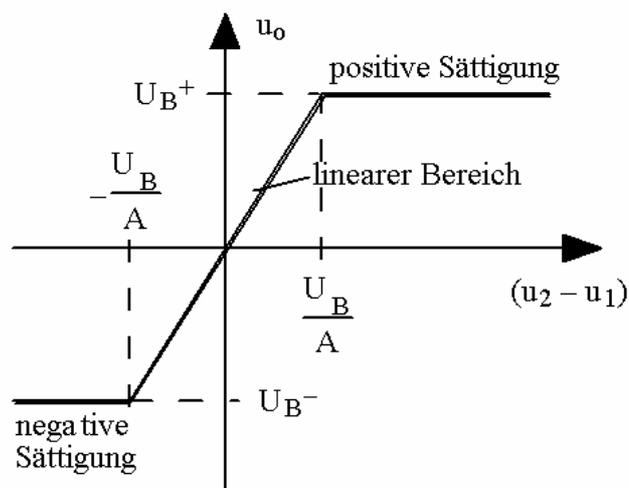
Wir überlegen uns nun, welche Eigenschaften ein „idealer“ Verstärker haben sollte. Zunächst bevorzugt man bei nicht sehr hohen Frequenzen mit Spannungen zu arbeiten, d.h. die Signale durch entsprechende Spannungen zu charakterisieren. Man ist deshalb an einer möglichst grossen Spannungsverstärkung interessiert. Selbstverständlich soll das Ausgangssignal möglichst genau proportional zum Eingangssignal sein, d.h. der Verstärker soll möglichst linear sein. Prinzipiell liegt am Eingang immer eine Signalquelle mit Innenwiderstand. Damit diese möglichst wenig belastet wird, soll der Eingangswiderstand möglichst hoch sein. Dies gilt vor allem bei Messverstärkern. Andererseits wird der Ausgang durch eine Last belastet. Damit die Ausgangsspannung möglichst unabhängig von dieser Last ist, ist ein niedriger Ausgangswiderstand wünschenswert.

Die eben skizzierten Wünsche an „ideale“ Verstärker sind recht vage und können sich je nach Anwendung stark unterscheiden. So ist beispielsweise ein Eingangswiderstand von $1\text{M}\Omega$ für viele Anwendungen gross, für manche Anwendungen hingegen klein. Aus diesem Grund gibt es nicht einfach eine integrierte Schaltung eines idealen Verstärkers sondern viele verschiedene integrierte Verstärkerschaltungen für unterschiedliche Anwendungen. Das gilt auch für den Spezialfall des Operationsverstärkers, den wir im folgenden genauer

betrachten werden, da er sehr vielseitig in der Anwendung und gleichzeitig relativ einfach in der Schaltungsanalyse ist. In der Analogtechnik sind deshalb Operationsverstärker sehr weit verbreitet, z.B. in Messgeräten als Messverstärker, in aktiven Filterschaltungen, in Analog/Digital- und Digital/Analog-Wandlern usw.

Der Name Operationsverstärker (Englisch Operational Amplifier, kurz OpAmp) kommt daher, dass diese Verstärker für Analogcomputer entwickelt wurden und die Realisierung mathematischer Operationen (Addieren, Subtrahieren, Invertieren, Skalieren aber auch Differenzieren und Integrieren) ermöglichen. Wir werden die entsprechenden Schaltungen im folgenden skizzieren.

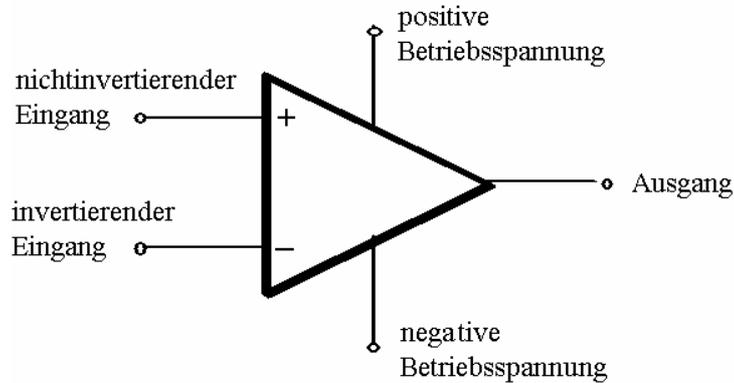
Die Verstärkerschaltungen im letzten Abschnitt wurden mit einer Gleichspannungsquelle gespeist. Es ist klar, dass die Ausgangsspannung dieser Schaltungen zwischen 0V und der Speisespannung liegen muss. Bei Operationsverstärkern will man normalerweise sowohl positive als auch negative Signalspannungen verarbeiten können. Aus diesem Grunde werden Operationsverstärker normalerweise mit einer positiven und einer negativen Speisespannung versorgt. Diese Spannungen sind entgegengesetzt gleich und limitieren den Arbeitsbereich, wie Figur 79 zeigt.



Figur 79: Idealisierte DC-Übertragungscharakteristik eines Operationsverstärkers

Wir haben gesehen, dass bei einer Transistorstufe in Emitterschaltung die Spannungsverstärkung v einen negativen Wert aufweist. Derartige Verstärker nennt man invertierende Verstärker. Ein invertierender Verstärker mit $v = -1$ ist ein analoger Inverter, welcher das Vorzeichen des zu verarbeitenden Signals ändert. Weil die Inversion eine wichtige Operation ist, verfügt der Operationsverstärker über einen invertierenden Eingang. Zusätzlich verfügt der Operationsverstärker über einen nicht invertierenden Eingang. Die Ausgangs-

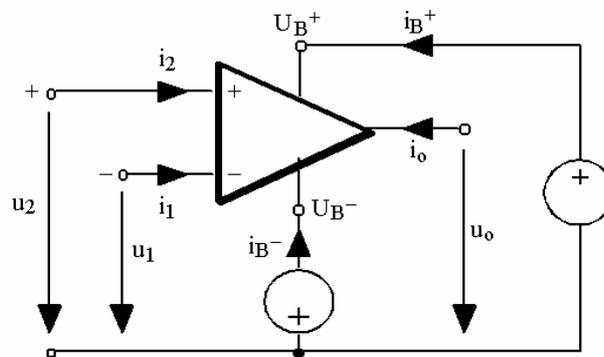
spannung ist proportional zur Differenz der beiden Eingangsspannungen, d.h. der Operationsverstärker ist ein *Differenzverstärker*. Figur 80 zeigt das Blockschaltbild.



Figur 80: Schaltbild eines Operationsverstärkers. Bei Schaltungen mit Operationsverstärkern werden die Betriebsspannungen meist nicht extra eingezeichnet.

5.4.1. Grundlegende Überlegungen

Bei der Bezeichnung der Anschlüsse schreibt man in der Schaltungstechnik nicht deren Bedeutung an (siehe Figur 80), sondern versteht sie mit Symbolen, wie in Figur 81.



Figur 81: Klemmenspannungen und -ströme eines Operationsverstärkers.

Für die folgenden Rechnungen gehen wir immer davon aus, dass die Ströme — wie in Figur 81 eingezeichnet — in den Operationsverstärker hinein fließen. Damit wir das Verhalten eines Operationsverstärkers mit externer Beschaltung berechnen können, müssen wir die Bedingungen an den Klemmen im unbeschalteten Zustand kennen.

Zunächst muss natürlich die Ausgangsspannung im Bereich $U_B^- \dots U_B^+$, d.h. zwischen den beiden Versorgungsspannungen liegen. Damit der Operationsverstärker im linearen Bereich und nicht in der Sättigung arbeitet, muss die Differenz der Eingangsspannungen genügend klein sein, weil für die Ausgangsspannung

$$u_o = A (u_2 - u_1)$$

gilt, wobei die Spannungsverstärkung ohne externe Beschaltung (Englisch open-loop gain) A sehr gross ist. Normalerweise ist A sogar so gross (meist 10'000 und mehr), dass man für die Berechnung in guter Näherung

$$u_2 = u_1$$

setzen kann.

Weil die Eingangswiderstände der beiden Eingänge sehr hoch (meist deutlich mehr als $1\text{M}\Omega$) sind, kann man ausserdem die Eingangsströme meist vernachlässigen, d.h.

$$i_2 = i_1 = 0$$

setzen.

5.4.2. Beschalteter Operationsverstärker

Je nach externer Beschaltung mit Widerständen, Kapazitäten, Induktivitäten, oder andern Bauelementen, kann der Operationsverstärker eine recht unterschiedliche Rolle spielen. Wichtig ist dabei, dass meist eine *Gegenkopplung* wie bei den einfachen Transistorverstärkern eingeführt wird, d.h., das Ausgangssignal wird über einen Zweipol oder über ein Netzwerk an den invertierenden Eingang zurück geführt. Geschieht die Gegenkopplung über einen einfachen Widerstand, so wird mit diesem im wesentlichen die Verstärkung eingestellt.

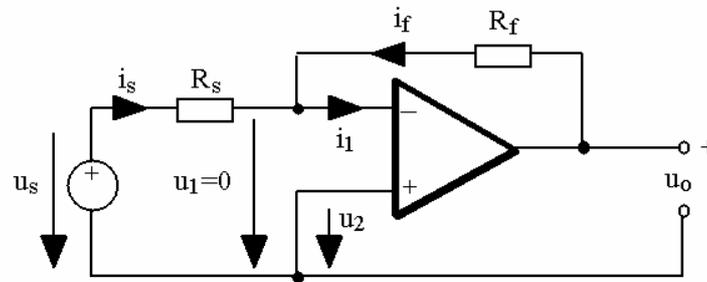
Bei vielen Anwendungen wird der nicht invertierende Eingang auf Masse gelegt und das Eingangssignal über einen Zweipol oder ein Netzwerk dem invertierenden Eingang zugeführt. Dadurch wird die Schaltungsanalyse besonders einfach, weil man für den invertierenden Eingang $i_1 = 0$ und $u_1 = 0$ setzen kann und den nicht invertierenden Eingang für die weitere Rechnung ignorieren kann.

In den folgenden Unterabschnitten skizzieren wir die einfachsten Schaltungen mit einem Operationsverstärker.

5.4.3. Invertierender Verstärker

Setzen wir den nicht invertierenden Eingang auf Masse und führen wir das Eingangssignal dem invertierenden Eingang über einen Ohm'schen Widerstand R_s zu, so erhalten wir einen

invertierenden Verstärker mit praktisch frequenzunabhängiger Verstärkung (innerhalb des Frequenzbereiches des betreffenden Operationsverstärkers). Mit Hilfe eines Gegenkopplungswiderstandes R_f können wir die Spannungsverstärkung auf einen nahezu beliebigen Wert einstellen.



Figur 82: Invertierender Verstärker.

Wegen $i_1 = 0$ gilt offenbar $i_s = -i_f$. Setzen wir das Ohm'sche Gesetz für die beiden Widerstände ein, so folgt wegen $u_1 = 0$ sofort

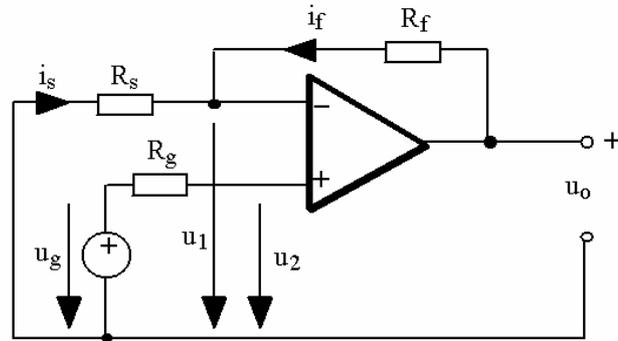
$$v = u_o / u_s = -R_f / R_s.$$

Selbstverständlich gilt diese vereinfachte Rechnung nur dann, wenn R_s deutlich kleiner als der Eingangswiderstand des Operationsverstärkers ist und wenn R_f den Ausgang des Operationsverstärkers nicht zu stark belastet, d.h. deutlich grösser als der Ausgangswiderstand ist. Ausserdem muss natürlich die Signalspannung u_s so klein sein, dass der lineare Bereich des Operationsverstärkers nicht verlassen wird.

Für eine analoge Inversion, d.h. Vorzeichenumkehr des Eingangssignals setzt man $R_f = R_s$ und erhält mit $v = -1$ sofort $u_o = -u_s$.

5.4.4. Nicht invertierender Verstärker

Es liegt auf der Hand, dass für einen nicht invertierenden Verstärker die Signalspannungsquelle u_g mit dem Innenwiderstand R_g mit dem nicht invertierenden Eingang verbunden werden muss, dieser Eingang also nicht mehr auf Masse gelegt werden kann. Wie beim invertierenden Verstärker muss aber die Spannungsverstärkung v mit einem Gegenkopplungswiderstand R_f vom Ausgang zum invertierenden Eingang eingestellt werden. Würden wir den invertierenden Eingang direkt auf Masse legen, so ergäbe sich mit $u_2 = u_1 = 0$ auch $u_o = 0$, also $v = 0$. Um dies zu vermeiden, legen wir den invertierenden Eingang über den Widerstand R_s auf Masse, wie in Figur 83 angegeben.



Figur 83: Nicht invertierender Verstärker.

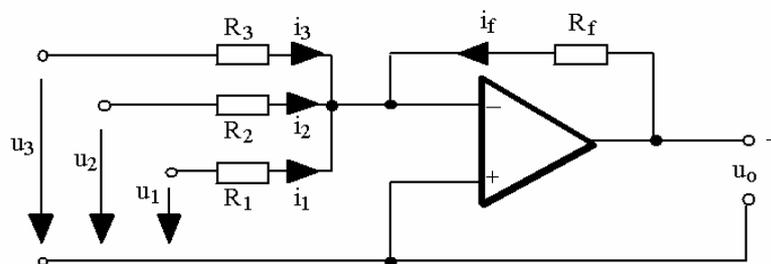
Wegen $i_2 = 0$ folgt sofort, dass keine Spannung über R_g liegt und somit $u_2 = u_g$. Nun ist natürlich auch $u_2 = u_1$ und somit $u_1 = u_g$. Wie beim invertierenden Verstärker gilt ausserdem $i_s = -i_f$. Setzen wir wieder das Ohm'sche Gesetz für die beiden Widerstände R_s und R_f ein, so ergibt sich die Spannungsverstärkung

$$v = \frac{u_o}{u_g} = \frac{u_o}{u_1} = \frac{i_f R_f - i_s R_s}{-i_s R_s} = \frac{R_f + R_s}{R_s} = 1 + \frac{R_f}{R_s}$$

5.4.5. Addierer

Verbindet man mehrere Eingangsspannungsquellen über Widerstände mit dem invertierenden Eingang und verwendet man ausserdem dasselbe Schema, wie beim invertierenden Verstärker (siehe Figur 84), so ergibt sich dieselbe Rechnung wie beim invertierenden Verstärker, wenn man i_s durch die Summe der Eingangsströme ersetzt. Z.B. gilt für drei Eingänge:

$$i_s = i_1 + i_2 + i_3 .$$



Figur 84: Addierer mit invertiertem Vorzeichen.

Daraus ergibt sich

$$u_o = -R_f(u_1/R_1 + u_2/R_2 + u_3/R_3).$$

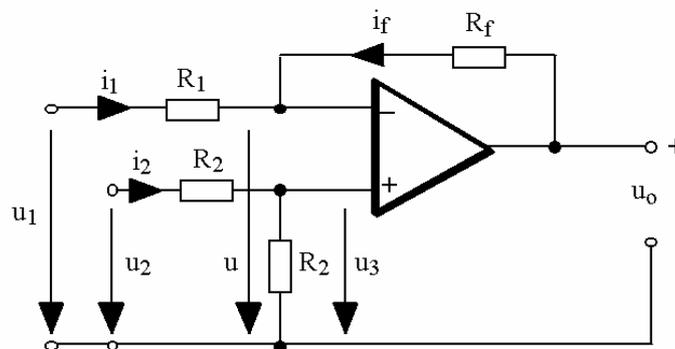
Setzt man

$$R_1 = R_2 = R_3 = R_f$$

So wird offenbar die Ausgangsspannung bis auf das Vorzeichen gleich der Summe der Eingangsspannungen. Um das Vorzeichen zu korrigieren, ist also nur noch ein invertierender Verstärker mit Spannungsverstärkung $v = 1$ erforderlich. Zur analogen Addition von beliebig vielen Grössen reichen also zwei Operationsverstärker.

5.4.6. Subtrahierer

Die Subtraktion ($a-b$) kann als Addition und Inversion ($a+(-b)$) aufgefasst werden. Zur analogen Subtraktion könnte man deshalb auf die Inversion und die Addition zurückgreifen. Für eine Subtraktion wären damit drei Operationsverstärker erforderlich. Weil ohnehin die Differenz der Eingangsspannungen verstärkt wird, reicht allerdings ein einziger Operationsverstärker mit Schaltung gemäss Figur 85.



Figur 85: Subtrahierer.

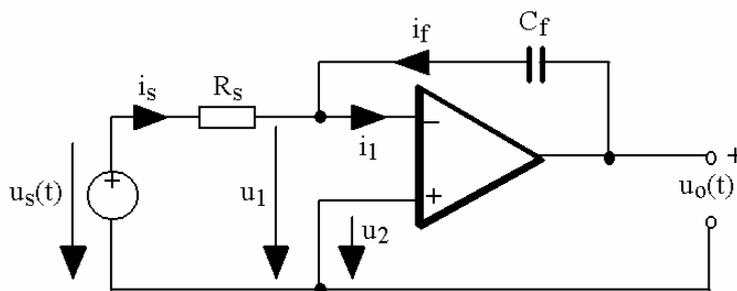
Für $R_2 = R_1 = R_3 = R_f$ finden wir mit einer ähnlichen Rechnung wie in den vorhergehenden Unterabschnitten

$$u_o = u_2 - u_1.$$

5.4.7. Analoge Integration

Wie wir wissen, gilt für eine Kapazität $i = C \partial u / \partial t$. Die Spannung u über der Kapazität ist somit das Integral über den Strom i . Da der Strom wiederum mit einem Ohm'schen Wider-

stand in eine Spannung umgewandelt werden kann, reicht an sich ein einfaches RC Glied, um eine analoge Spannungsintegration durchzuführen, wie wir bereits im Kapitel 3 gesehen haben. Ein solcher RC Integrator ist allerdings nicht „belastbar“, d.h. es darf kein Ausgangsstrom fließen, weil dieser die Kapazität entladen und damit das Resultat verfälschen würde. Wir „bereichern“ deshalb das RC Glied mit einem Operationsverstärker, wie in Figur 86.



Figur 86: Integration.

Offensichtlich ist dies dieselbe Schaltung, wie der invertierende Verstärker mit einer Kapazität statt eines Widerstandes zur Gegenkopplung. Wir brauchen also lediglich die Gleichung $u_o = R_f i_f$ durch

$$i_f = C \frac{\partial u_o}{\partial t}$$

zu ersetzen und erhalten nach Integration über die Zeit

$$u_o(t) = u_o(t_0) - \frac{1}{R_s C_f} \int_{t_0}^t u_s(t') dt'$$

Wenn der Kondensator C_f zur Zeit t_0 ungeladen ist, ergibt sich

$$u_o(t) = -\frac{1}{R_s C_f} \int_{t_0}^t u_s(t') dt'$$

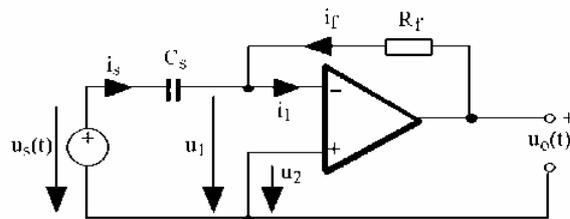
D.h. die Ausgangsspannung ist das mit dem Faktor $f = -1/R_s C_f$ skalierte Integral der Eingangsspannung. Natürlich kann man die Kapazität und den Widerstand so dimensionieren, dass $f = -1$ wird. Um das negative Vorzeichen loszuwerden kann man — wie schon beim Addierer — einen Inverter nachschalten.

Eine Alternative zur analogen Integration ist im Prinzip die RL Kombination einer Induktivität (anstelle von R_s) und eines Widerstandes (anstelle von C_f) mit einem Operationsver-

stärker. In der Schaltungstechnik werden Kapazitäten aus praktischen Gründen den Induktivitäten vorgezogen. Wir verzichten hier deshalb auf die Darstellung der RL Variante.

5.4.8. Analoge Differentiation

Auch für die analoge Differentiation kann man entweder eine Kapazität oder eine Induktivität in Kombination mit einem Widerstand verwenden. Figur 87 zeigt die RC Variante. Offenbar haben dabei gegenüber der RC Integration R und C ihre Plätze getauscht.



Figur 87: Differentiation.

Zur Berechnung geht man in analoger Weise wie bei dem Integrator und wie bei dem invertierenden Verstärker vor und findet

$$u_o(t) = -C_s R_f \frac{\partial u_s}{\partial t}.$$

Wie schon beim Integrator hat man also einen negativen Skalierungsfaktor, welchen man durch geeignete Wahl von Kapazität und Widerstand und einen nachfolgenden Inverter loswerden kann.

5.4.9. Weitere Schaltungen für Analogcomputer

Neben den bisher vorgestellten Schaltungen mit Operationsverstärker für Analogcomputer existieren verschiedene weitere Schaltungen, welche die Kennlinien nichtlinearer Bauelemente ausnützen. Beispielsweise kann man sich zu Nutze machen, dass die Kennlinien von Dioden und Transistoren einen recht grossen Bereich mit nahezu exponentiellem Verhalten aufweisen. Ersetzt man R_s beim invertierenden Verstärker durch ein solches Bauelement, so erhält man einen einfachen Exponentialfunktionsgenerator, bzw. einen Potenzierer. Ersetzt man stattdessen R_f durch eine Diode oder einen Transistor, so erhält man einen Logarithmierer. Diese Bausteine kann man wiederum mit einem Addierer kombinieren um eine Multiplikation auszuführen. Durch kompliziertere Netzwerke, welche oft aus Dioden und Widerständen bestehen kann man schliesslich verschiedenartige Funktionen (Parabeln, Sinus etc.) recht gut approximieren.

5.4.10. Filter

Ersetzen wir die Signalspannungsquelle $u_s(t)$ beim invertierenden Verstärker durch eine sinusförmige Wechselspannungsquelle

$$u_s(t) = \hat{U}_s \cos(\omega t)$$

so können wir vorteilhafterweise mit komplexen Grössen rechnen. Im wesentlichen bleibt die Rechnung aber unverändert. Verallgemeinern wir die Schaltung geringfügig, indem wir die beiden Widerstände durch komplexe Impedanzen ersetzen, so ändert sich nichts am Rechengang, wir erhalten aber die komplexe, frequenzabhängige Verstärkung:

$$\underline{v} = \underline{U}_o / \underline{U}_s = -\underline{Z}_f / \underline{Z}_s.$$

Der Vergleich mit den passiven RLC Filtern zeigt, dass wir auch mit dieser *aktiven* Schaltung⁸ Hochpass- und Tiefpass-Filter erhalten können, wenn beispielsweise eine der beiden Impedanzen ein Widerstand und die andere eine Kombination eines Widerstandes mit einer Kapazität ist. Ersetzt man die Impedanzen durch kompliziertere RC Netzwerke, so erhält man aktive Filter höherer Ordnung, welche sehr unterschiedliche Frequenzgänge haben können.

Die wichtigsten Nachteile der passiven Filter (Betrag der Verstärkung < 1 , geringe Belastbarkeit des Ausgangs bzw. Lastabhängigkeit des Frequenzganges, geringer Innenwiderstand des Innenwiderstandes der Spannungsquelle am Eingang) können offensichtlich mit Hilfe des Operationsverstärkers vermieden oder wenigstens stark vermindert werden.

5.4.11. Operationsverstärker in Digitalschaltungen

Obwohl Operationsverstärker für Analogcomputer entwickelt wurden und in der Analogtechnik weit verbreitet sind, gibt es auch einige interessante Anwendungen für die Digitaltechnik, bei denen die Operationsverstärker so betrieben werden, dass die Ausgangsspannung im wesentlichen nur zwei verschiedene Werte annimmt. Beispiele sind 1) Spannungskomparatoren, bei denen die Ausgangsspannung anzeigt, ob die Eingangsspannung grösser oder kleiner als eine Referenzspannung ist, 2) Schmitt-Trigger, bei denen die Ausgangsspannung eine logische „1“ anzeigt, sobald die Eingangsspannung eine obere Schwellenspannung übersteigt und eine logische „0“, sobald die Eingangsspannung eine untere Schwellenspannung unterschreitet, 3) Multivibratoren, welche regelmässig von „0“ auf „1“ und zurück auf „0“ umschalten und als Taktgeneratoren verwendet werden können, 4) Univibratoren, welche einen einzelnen „1“ Puls mit vorgegebener Zeitdauer ausgeben, wenn am Eingang ein kurzer Puls mit nicht sauber definierter Form eintrifft.

⁸ Der Operationsverstärker wird als aktives Element betrachtet, da er den Speisespannungsquellen Energie entnimmt und diese in die elektronische Schaltung einspeist. Die gesamte elektronische Schaltung, welche die Speisespannungsquellen umfasst, ist natürlich trotzdem passiv.

Da übliche Operationsverstärker im Vergleich zu modernen Digitalschaltungen langsam schalten, sind die erwähnten Schaltungen kaum in Computern⁹ anzutreffen. Schnelle Komparatoren, Schmitt-Trigger, Multivibratoren und Univibratoren lassen sich übrigens recht einfach mit jeweils zwei Transistoren aufbauen. Die Verwendung eines Operationsverstärkers anstelle von zwei Transistoren hat für diese Schaltungen vor allem den Vorteil der einfacheren Analyse.

5.5. Digitale Schaltungen

Da eine spezielle Vorlesung über Digitaltechnik angeboten wird, streifen wir digitale Halbleiterschaltungen nur ganz kurz. Im Gegensatz zu den analogen Schaltungen werden dabei die Halbleiter nicht im linearen (aktiven) Bereich betrieben, sondern entweder im Sättigungsbereich oder im Cutoff-Bereich. Dadurch wird eine genaue Analyse erschwert. Um die Funktionsweise der entsprechenden Schaltungen zu verstehen genügen aber oft relativ einfache, logische Überlegungen. Dies wollen wir im folgenden mit den drei einfachsten logischen Schaltungen demonstrieren. Dabei verwenden wir die sogenannte Widerstands-Transistor-Logik (RTL), welche auf Kombinationen von bipolaren Transistoren mit Widerständen beruht und einfach zu verstehen ist. Daneben existieren verschiedene andere Techniken, auf die wir hier nicht eingehen:

Dioden-Transistor-Logik (DTL), Transistor-Transistor-Logik (TTL), Emitter-Coupled-Logic (ECL), MOS-Logik, komplementäre MOS-Logik (C-MOS), usw.

Logische Schaltungen werden gerne durch Tabellen beschrieben, welche die Ausgangszustände für alle möglichen Eingangszustände angeben. Es ist klar, dass diese Tabellen um so umfangreicher werden, je mehr Ein- und Ausgänge vorhanden sind. Wir beschränken uns hier auf die einfachsten Fälle mit einem Ausgang und ein bis zwei Eingängen. Ausserdem beschränken wir uns auf die gängige Binärlogik, welche nur zwei Zustände unterscheidet und setzen voraus, dass die logische „0“ durch eine geringe Signalspannung $<U_{low}$, die logische „1“ durch eine hohe Signalspannung $>U_{high}$ dargestellt werde.

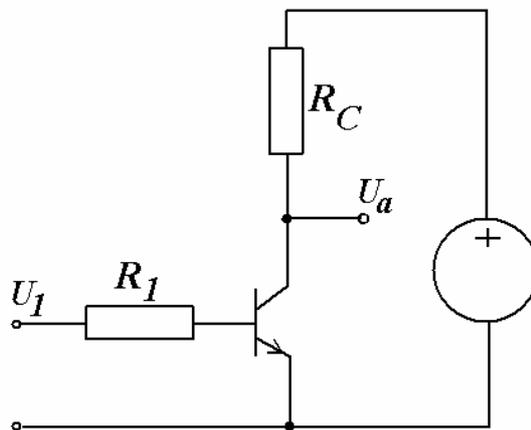
Theoretisch existieren vier binär-logische Schaltungen mit einem Ein- und einem Ausgang. Zwei davon sind ohne grosses Interesse, weil sie den Ausgangszustand „0“ bzw. „1“ unabhängig vom Eingangszustand annehmen. Von den übrigen zwei Schaltungen ist die sogenannte *Identität* ebenfalls nicht sehr spannend, weil ihr Ausgangszustand gleich dem Eingangszustand ist. Diese drei Schaltungen lassen sich leicht durch einfache Drähte realisieren.

⁹ Digitaltechnik beschränkt sich nicht auf den Bau von Computern. Sie ist auch für Datenübertragung und die digitale Mess- und Regeltechnik wichtig.

5.5.1. Inverter

Theoretisch existieren vier binär-logische Schaltungen mit einem Ein- und einem Ausgang. Zwei davon sind ohne grosses Interesse, weil sie den Ausgangszustand „0“ bzw. „1“ unabhängig vom Eingangszustand annehmen. Von den übrigen zwei Schaltungen ist die sogenannte *Identität* ebenfalls nicht sehr spannend, weil ihr Ausgangszustand gleich dem Eingangszustand ist. Diese drei Schaltungen lassen sich leicht durch einfache Drähte realisieren.

Aus schaltungstechnischer Sicht existiert nur eine interessante binär-logische Schaltung mit einem Ein- und einem Ausgang: der Inverter. Dieser negiert den Eingangszustand und kann als übersteuerter analoger Inverter konzipiert werden. Dazu eignet sich offenbar die Emitter-schaltung, welche wir für den linearen Betrieb ausführlich beschrieben haben. Um diese Schaltung möglichst stark zu übersteuern, verzichten wir auf jegliche Gegenkopplung und lassen ausserdem Widerstände zur Arbeitspunkteinstellung weg. Damit erhalten wir eine sehr einfache Schaltung gemäss Figur 88.



Figur 88: Einfacher Inverter in RTL Ausführung.

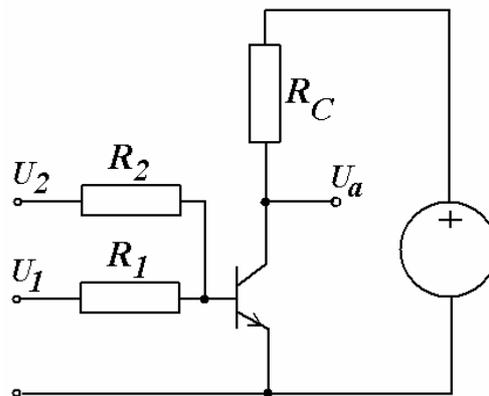
Ist die Eingangsspannung U_1 kleiner als etwa 0.6V bei einem Siliziumtransistor, so sperrt dieser, d.h. der Kollektorstrom wird klein und U_a wird nahezu gleich der Versorgungsspannung (sofern der Ausgang nicht zu stark belastet wird), d.h. gross. Ist umgekehrt U_1 gross (ähnlich gross wie die Versorgungsspannung), so leitet der Transistor und U_a wird klein.

Wegen der Basis-Emitter-Spannung von ca. 0.6V bei Siliziumtransistoren können derartige Schaltungen natürlich nur bei Versorgungsspannungen betrieben werden, welche einiges grösser als 0.6V sind.

5.5.2. NAND und NOR

Theoretisch existieren diverse binär-logische Schaltungen mit zwei Eingängen und einem Ausgang. Dazu gehören solche, deren Ausgangszustand von keinem oder nur von einem

Eingang abhängig ist. Interessant sind natürlich nur jene Schaltungen, welche auf beide Eingänge „reagieren“. Dazu gehören die bekannten logischen UND und ODER Verknüpfungen, sowie die Entweder-Oder Verknüpfung. Schaltungstechnisch einfacher zu realisieren sind die negierten UND und ODER Verknüpfungen, NAND und NOR. Prinzipiell kann man alle logischen Verknüpfungen durch passende Kombinationen von NAND oder NOR erzeugen. Die NAND und NOR Gatter sind deshalb die wichtigsten Grundbausteine. In der RTL Technik erhält man diese leicht durch einfache Erweiterungen des Inverters. Figur 89 zeigt eine einfache Version eines NOR Gatters¹⁰.



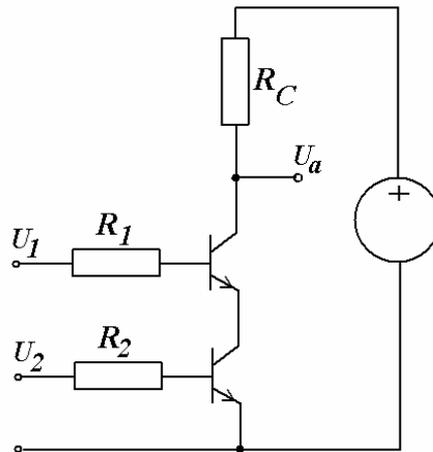
Figur 89: Einfaches NOR Gatter in RTL Ausführung.

Offenbar leitet der Transistor wenn wenigstens eine der beiden Eingangsspannungen hoch ist, so dass die Ausgangsspannung dann tief wird. Die Ausgangsspannung ist nur dann hoch, wenn beide Eingangsspannungen tief sind.

Offenbar erhält man übrigens aus dem NOR Gatter sofort einen Inverter, wenn man beide Eingänge miteinander elektrisch verbindet oder wenn man einen Eingang mit Masse (Spannung 0V) verbindet.

Ein NAND Gatter entsteht im Prinzip durch Serieschaltung zweier Transistoren, wie in Figur 90.

¹⁰ In integrierten Schaltungen werden kompliziertere Schaltungen mit mehr als einem Transistor vorgezogen.



Figur 90: NAND Gatter in RTL Ausführung.

Ist beispielsweise U_2 hoch, so leitet der über R_2 angesteuerte Transistor. Damit ein Kollektorstrom fließen kann, muss aber auch der andere Transistor leiten. U_a wird also nur dann tief, wenn beide Eingangsspannungen hoch sind.

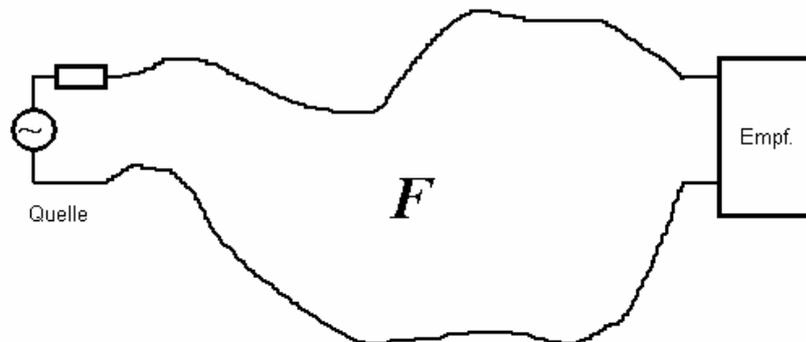
Aus dem NOR Gatter erhält man einen Inverter, wenn man beide Eingänge miteinander elektrisch verbindet oder wenn man einen Eingang mit der Speisespannungsquelle verbindet (d.h. auf „1“ setzt).

Die skizzierten RTL Schaltungen „funktionieren“ anscheinend für recht willkürlich gewählte Widerstandswerte. Tatsächlich ist eine geschickte Wahl der Widerstände wichtig, damit die Schaltungen sowohl zuverlässig arbeiten als auch nicht zu viel Energie verbrauchen bzw. in den Widerständen in Wärme umsetzen. Wegen den unumgänglichen und relativ grossen Wärmeverlusten wird die RTL Logik heute nicht mehr verwendet. Moderne logische Schaltungen werden in der Vorlesung über Digitaltechnik ausführlich behandelt.

6. Leitungen

Bisher sind wir davon ausgegangen, dass die einzelnen Bauteile eines Netzwerkes über Drähte so miteinander verbunden werden können, dass 1) der Strom auf der ganzen Länge jedes Drahtes konstant ist, 2) die Drähte keine Ladungen speichern, 3) die Spannung bezüglich Masse auf jedem Draht konstant ist, 4) Signale in nebeneinander liegenden Drähten einander nicht beeinflussen, 5) keine Spannungen in die Drähte induziert werden, 6) die geometrische Form der Drähte keine Rolle spielt usw. Wir haben also die Drähte so vollkommen idealisiert, dass wir sie in unseren Rechnungen gar nicht berücksichtigen mussten und dies, obwohl wir wissen, dass die gemachten Annahmen nicht zutreffen. Dieses Vorgehen ist einerseits durch die starke Vereinfachung der Rechnungen und andererseits dadurch begründet, dass die nicht berücksichtigten Effekte klein sind, wenn die Abmessungen der Schaltungen klein im Vergleich zu den auftretenden Wellenlängen sind. Einerseits führt die Miniaturisierung elektronischer Schaltungen zu einer Reduktion der Abmessungen, andererseits führt der Bau immer schnellerer Schaltungen zu einer Reduktion der auftretenden Wellenlängen. Vor allem bei komplexen und damit relativ grossen schnellen Schaltungen führen die erwähnten Effekte zu Fehlfunktionen, die nicht tolerierbar sind. In diesem Kapitel werden wir versuchen, die wichtigsten Effekte mit einem vereinfachten Formalismus in den Griff zu bekommen.

Zunächst wollen wir abschätzen, wann das Arbeiten mit den idealisierten Drähten fragwürdig wird. Wir betrachten dazu den einfachst möglichen Fall einer Verbindung einer Signalquelle mit einem Empfänger gemäss Figur 91.



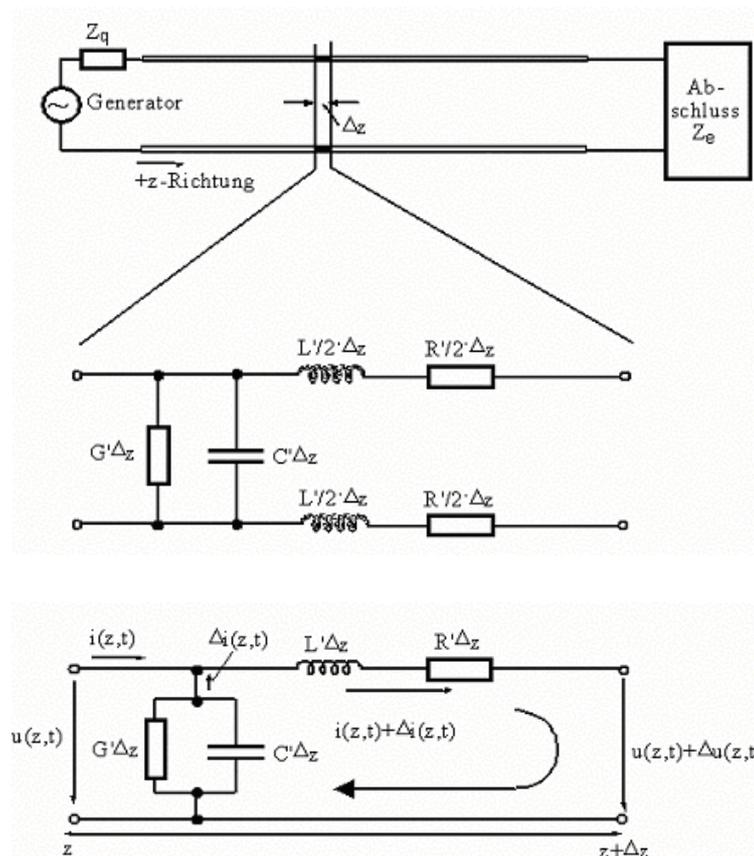
Figur 91: Signalquelle, welche über Drähte mit einem Empfänger verbunden ist.

Zunächst ist klar, dass wegen des Induktionsgesetzes — das wir schon bei der Behandlung von Induktivität und Gegeninduktivität verwendet haben — zeitliche Änderungen des magnetischen Flusses durch die Fläche F , welche durch die Drähte berandet wird, Spannungen in die Drähte induziert, welche sich der eigentlichen Signalspannung überlagern. Dadurch wird das Signal gestört und der Empfänger erhält nicht die gewünschte Information. Da die induzierte Spannung zur Fläche F proportional ist, kann man diese Störung

dadurch vermeiden, dass man die Drähte möglichst kurz macht und sie ausserdem möglichst parallel mit kleinem Drahtabstand führt. Dadurch entsteht eine Zweidrahtleitung, deren Eigenschaften wir im folgenden genauer betrachten wollen.

6.1. Zweidrahtleitung

Figur 92 zeigt eine Zweidrahtleitung, welche eine Signalquelle mit einem Empfänger verbindet. Dabei wird die Quelle durch eine Spannungsquelle mit Impedanz Z_q und der Empfänger durch eine Impedanz Z_e approximiert. Diese Approximation ist in der Hochfrequenztechnik für die meisten Anwendungen gut brauchbar. Dabei ist zu beachten, dass die Impedanzen frequenzabhängige Grössen sein können. Ist das Signal nicht sinusförmig, so macht es deshalb meist Sinn, das Signal in seine Fourierkomponenten zu zerlegen und die Leitung im Frequenzbereich zu analysieren.



Figur 92: Signalquelle, welche über eine symmetrische Zweidrahtleitung von konstantem Querschnitt mit einer Abschlussimpedanz verbunden ist.

6.1.1. Leitungsbeläge

Für eine genaue Berechnung der Ausbreitung elektromagnetischer Wellen auf einer Zweidrahtleitung müssten wir von den Maxwell Gleichungen ausgehen, welche das Fundament der elektromagnetischen Phänomene bilden, welche hier von Bedeutung sind. Das heisst, wir müssten die elektromagnetischen Felder berechnen, welche den gesamten Raum ausfüllen. Weil diese Berechnung recht anspruchsvoll ist, beschränken wir uns auf eine Näherungsrechnung, indem wir lediglich die Haupteffekte berücksichtigen, welche sich durch die uns bekannten Ströme und Spannungen, sowie durch die Zweipole Widerstand, Induktivität und Kapazität beschreiben lassen.

Betrachten wir einen kleinen Ausschnitt der Leitung mit der Länge Δz , so können wir die Ladungen, welche sich allenfalls auf den Drähten aufhalten berücksichtigen, indem wir annehmen, diese gehören zu einer Kapazität

$$\Delta C = C' \Delta z.$$

Wenn sich längs der Leitung die Querschnittsform nicht ändert, so ist der sogenannte *Kapazitätsbelag* C' (Kapazität pro Längeneinheit) für die ganze Leitung konstant und lediglich abhängig von der Querschnittsgeometrie und vom Material (Isolator) zwischen den Drähten.

Obwohl man normalerweise gute Isolatoren zwischen den Drähten verwendet, können kleine Ströme von einem Draht zum andern fließen. Diese Ströme sind natürlich proportional zur Spannung zwischen den beiden Drähten und der zugehörige Effekt lässt sich gut durch einen Leitwert

$$\Delta G = G' \Delta z$$

beschreiben, welcher parallel zur Kapazität ΔC geschaltet ist. Wie schon die Kapazität ist auch der Leitwert proportional zur Länge Δz und lediglich von Querschnittsgeometrie und Material (Isolator) zwischen den Drähten abhängig. Auch der *Leitwertsbelag* G' ist also eine charakteristische Grösse für die Zweidrahtleitung.

Um der endlichen Leitfähigkeit der Drähte Rechnung zu tragen, kann man einen Widerstandsbelag R' mit

$$\Delta R = R' \Delta z$$

einführen. Dabei ist zu beachten, dass beiden Drähten je ein Widerstand ΔR_1 bzw. ΔR_2 zugeordnet werden kann, welche sich jedoch schaltungstechnisch zu einem einzigen Widerstand ΔR zusammenfassen lassen. Dadurch wird der Rechengang vereinfacht. Die beiden Widerstände hängen vom betreffenden Drahtquerschnitt und vom Drahtmaterial ab. Zudem führen elektrodynamische Effekte dazu, dass die Ströme mit zunehmender Frequenz auf die Oberfläche gedrängt werden (Skinneffekt) und nicht gleichmässig über den Drahtquerschnitt verteilt sind. Ausserdem beeinflusst der Nachbardraht die Stromverteilung im betrachteten Draht um so stärker, je kleiner der Drahtabstand ist (Proximityeffekt). Diese beiden Effekte

erhöhen den Widerstandsbelag. Der Widerstandsbelag ist deshalb von der gesamten Querschnittsgeometrie, den Materialeigenschaften und der Frequenz abhängig.

Schliesslich berücksichtigen wir das Induktionsgesetz, indem wir einen Induktivitätsbelag L' in ähnlicher Weise wie den Widerstandsbelag R' einführen und

$$\Delta L = L' \Delta z$$

schreiben. Auch hier fassen wir die Induktivitäten der beiden Drähte zu einer einzigen zusammen. Übrigens führen Skin- und Proximityeffekt auch zu einer Frequenzabhängigkeit von L' .

Die feldtheoretische Berechnung der vier *Leitungsbeläge* C' , G' , R' , L' ist recht schwierig. Eine exakte analytische Lösung existiert selbst für den einfachen Fall einer symmetrischen Doppeldrahtleitung mit kreiszylindrischen Drähten nicht. Es existieren aber verschiedenste numerische Methoden und messtechnische Verfahren, um die Leitungsbeläge mit ausreichender Genauigkeit zu bestimmen. Wir gehen deshalb im folgenden davon aus, dass C' , G' , R' , L' für die betrachtete Leitung bekannte Grössen sind.

6.1.2. Telegraphengleichungen

Da wir ein Ersatzschaltbild für ein Stück der Zweidrahtleitung haben, können wir dieses Stück als Zweitor bestehend aus einem kleinen Netzwerk auffassen und mit unseren gewohnten Methoden analysieren. Wir notieren also die Kirchhoff'schen Gesetze.

Für die Masche, welche aus dem Leitungsstück gebildet wird gilt:

$$u(z, t) - u(z + \Delta z, t) = R' \Delta z i(z, t) + L' \Delta z \frac{\partial}{\partial t} i(z, t).$$

Dabei haben wir angenommen, dass Δz so klein sei, dass die Ströme über das Drahtstück nahezu konstant sind. Dividieren wir diese Gleichung durch Δz , so erhalten wir im Limes $\Delta z \rightarrow 0$ die Differentialgleichung

$$-\frac{\partial u}{\partial z} = R' i + L' \frac{\partial i}{\partial t}.$$

Notieren wir die Knotengleichung für einen der beiden Drähte, so folgt mit demselben Vorgehen

$$-\frac{\partial i}{\partial z} = G' u + C' \frac{\partial u}{\partial t}.$$

Diese beiden Differentialgleichungen sind gekoppelte Differentialgleichungen erster Ordnung und werden *Telegraphengleichungen* genannt, weil im 19. Jahrhundert die Telegraphenverbindungen mit Zweidrahtleitungen realisiert wurden.

Prinzipiell löst man gekoppelte Differentialgleichungssysteme erster Ordnung in zwei Schritten. Im ersten Schritt werden die Gleichungen entkoppelt, in dem man eine Gleichung ableitet und die andere einsetzt. Die entkoppelten Gleichungen sind deshalb Differentialgleichungen zweiter Ordnung. Sie werden im zweiten Schritt gelöst.

Leiten wir die erste Telegraphengleichung nach z ab, so erhalten wir nach Einsetzen der zweiten Telegraphengleichung und einer kleinen Umformung

$$\frac{\partial^2 u}{\partial z^2} = R'G'u + (R'C' + L'G') \frac{\partial u}{\partial t} + L'C' \frac{\partial^2 u}{\partial t^2}.$$

Diese Differentialgleichung ist entkoppelt, weil sie i nicht mehr enthält. Analog enthalten wir die entkoppelte Differentialgleichung für i – welche u nicht mehr enthält:

$$\frac{\partial^2 i}{\partial z^2} = R'G'i + (R'C' + L'G') \frac{\partial i}{\partial t} + L'C' \frac{\partial^2 i}{\partial t^2}.$$

6.1.3. Leitungswellen

Da die Differentialgleichungen für u und i formal identisch sind, ergeben sich für die Ströme und Spannungen auf der Zweidrahtleitung auch formal dieselben Lösungen. Offenbar sind diese Lösungen Funktionen des Ortes z und der Zeit t . Um diese zu finden, versucht man normalerweise die beiden Variablen mit einem *Produktansatz* der Form

$$F(z,t) = Z(z) \cdot T(t)$$

zu *separieren*. Dieser Ansatz führt dann zu je einer Differentialgleichung für die Ortsabhängigkeit Z und für die Zeitabhängigkeit T . Für letztere sind tatsächlich die aus der Fourieranalyse bekannten harmonischen Funktionen Lösungen. Da wir bereits wissen, dass wenigstens zwei der vier Leitungsbeläge frequenzabhängig sind, ist es ohnehin sinnvoll, im Frequenzbereich zu arbeiten und sinusförmige Zeitverläufe vorauszusetzen. Wir arbeiten bequemerweise mit komplexen Grössen und schreiben

$$i(z,t) = \text{Re}(\underline{I}(z) \cdot \exp(j\omega t)); \quad u(z,t) = \text{Re}(\underline{U}(z) \cdot \exp(j\omega t)).$$

Wir können damit jede zeitliche Ableitung durch eine einfache Multiplikation ersetzen und erhalten für die Telegraphengleichungen im Frequenzbereich

$$-\frac{\partial \underline{U}}{\partial z} = (R' + j\omega L') \underline{I},$$

$$-\frac{\partial \underline{I}}{\partial z} = (G' + j\omega C') \underline{U}$$

sowie die entkoppelten Gleichungen

$$\frac{\partial^2 \underline{U}}{\partial z^2} = (R'G' + j\omega(R'C' + L'G') - \omega^2 L'C')\underline{U} = (R' + j\omega L')(G' + j\omega C')\underline{U},$$

$$\frac{\partial^2 \underline{I}}{\partial z^2} = (R'G' + j\omega(R'C' + L'G') - \omega^2 L'C')\underline{I} = (R' + j\omega L')(G' + j\omega C')\underline{I}.$$

Diese einfachen Differentialgleichungen zweiter Ordnung haben Lösungen der Form

$$\underline{U}(z) = \underline{U}_0^+ e^{-\gamma z} + \underline{U}_0^- e^{+\gamma z}; \quad \underline{I}(z) = \underline{I}_0^+ e^{-\gamma z} + \underline{I}_0^- e^{+\gamma z}.$$

Dabei ist die komplexe Grösse $\gamma = \alpha + j\beta$ die *Fortpflanzungskonstante*. Setzen wir obige Lösungen für Spannung und Strom in den entkoppelten Differentialgleichungen ein, so finden wir

$$\gamma = \alpha + j\beta = \sqrt{(R' + j\omega L')(G' + j\omega C')} \quad .$$

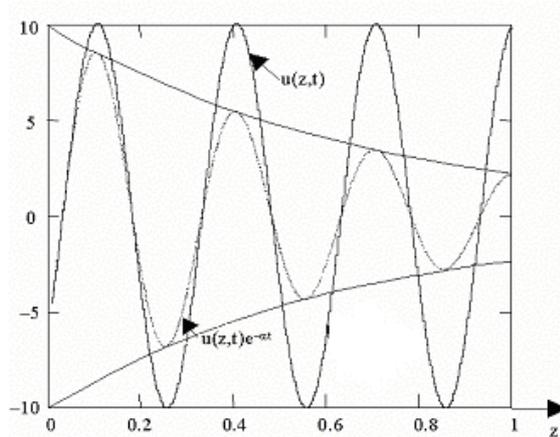
Der Realteil α der Fortpflanzungskonstanten wird *Dämpfungskonstante*, deren Imaginärteil β *Phasenkonstante* genannt. Diese Bezeichnungen erklären sich aus der Analyse der Lösung, welche gleich folgt. Etwas seltsam mag erscheinen, dass der obere Index bei den Amplituden nicht mit den Vorzeichen im Exponenten überein stimmt. Um dies zu verstehen, setzen wir die Zeitabhängigkeit ein und erhalten

$$u(z, t) = \text{Re}(\underline{U}_0^+ e^{j\omega t - \gamma z} + \underline{U}_0^- e^{j\omega t + \gamma z}); \quad i(z, t) = \text{Re}(\underline{I}_0^+ e^{j\omega t - \gamma z} + \underline{I}_0^- e^{j\omega t + \gamma z}).$$

Wie man leicht nachprüft, beschreiben die Terme mit positivem oberem Index eine Welle, welche sich mit zunehmender Zeit t in die *positive* z Richtung ausbreitet sofern die Phasenkonstante, d.h. der Imaginärteil von γ positiv ist. Wählt man die z Richtung so, dass die z Achse von der Signalquelle zum Verbraucher zeigt, so kennzeichnen die positiven oberen Indices die *hinlaufenden Wellen*, die negativen Indices hingegen die *rückläufigen Wellen*. Ist die Dämpfungskonstante positiv, so nehmen die hinlaufenden Wellen exponentiell mit zunehmendem z ab. Dies sieht man, wenn man $\gamma = \alpha + j\beta$ einsetzt und obige Gleichungen leicht umschreibt. Für die Spannung der hinlaufenden Welle findet man beispielsweise

$$u(z, t) = e^{-\alpha z} \text{Re}(\underline{U}_0^+ e^{j(\omega t - \beta z)}).$$

Der Faktor $e^{-\alpha z}$ beschreibt dabei die exponentielle Dämpfung. Figur 93 illustriert diesen Verlauf für einen bestimmten Zeitpunkt.



Figur 93: Spannungsverlauf einer hinlaufenden Leitungswelle zu einer bestimmten Zeit.

Die Analyse der rückläufigen Wellen verläuft ganz analog. Dabei sieht man, dass diese sich in $-z$ Richtung ausbreiten und zwar mit derselben Phasenkonstanten und demselben Dämpfungsfaktor wie die hinlaufenden Wellen.

6.1.4. Phasen- und Gruppengeschwindigkeit

Die Phase der Welle wird in der obigen Formel durch die zweite Exponentialfunktion mit rein imaginärem Argument beschrieben. Setzt man dieses Argument Null, so findet man die Bedingung für Orte gleicher Phase:

$$\beta z = \omega t \text{ oder } z = (\omega / \beta) t.$$

Diese Orte bewegen sich mit der *Phasengeschwindigkeit*

$$v_p = \omega / \beta$$

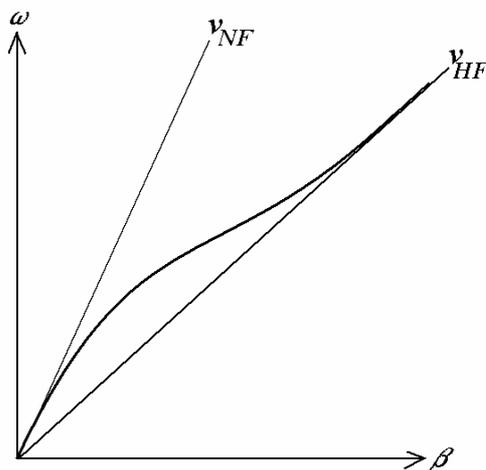
in z Richtung. Dasselbe gilt natürlich auch für die Ströme, weil für Ströme und Spannungen formal dieselben Gleichungen gelten.

Um Information auf einer Leitung zu übertragen, muss man Wellen mit unterschiedlichen Frequenzen überlagern. So kann man beispielsweise Wellen mit pulsartigem Zeitverlauf erzeugen. Ist das zugehörige Frequenzspektrum (Fourierspektrum) ausgedehnt, so kann die Analyse recht aufwendig werden. Ist das Spektrum jedoch schmal genug, so kann man den Frequenzgang linearisieren und annehmen, die Phasenkonstante β sei eine lineare Funktion von ω . Mit einem derartigen schmalbandigen Spektrum kann man eine Wellengruppe erzeugen, welche eine pulsartige Envelope hat. Diese bewegt sich mit der *Gruppengeschwindigkeit*

$$v_g = d\omega / d\beta.$$

Auf dispersiven Leitungen sind Phasen- und Gruppengeschwindigkeiten unterschiedlich. Als Folge davon ergeben sich Verformungen des Signalverlaufs längs der Leitungen. Dies ist insbesondere auch für die Pulsübertragung in der Digitaltechnik wichtig.

Eine rasche Übersicht vermittelt das Phasendiagramm gemäss Figur 94. Dabei charakterisiert die Tangentensteigung in jedem Punkt der Kurve die Gruppengeschwindigkeit. Die Phasengeschwindigkeit entspricht hingegen der Steigung der Verbindungsgeraden des Punktes auf der Kurve mit dem Nullpunkt. Bei guten Leitungen ist die Differenz der beiden Asymptoten für niedere und hohe Frequenzen relativ gering und ausserdem gehen beide Asymptoten durch den Nullpunkt. Das heisst, dass bei niederen und hohen Frequenzen Phasen- und Gruppengeschwindigkeit nahezu gleich sind und damit die Dispersion gering ist. Der Einfachheit halber werden wir die Zweidrahtleitung im folgenden meist als dispersionsfrei voraussetzen, d.h. annehmen, das Phasendiagramm bestehe aus einer einzigen Geraden durch den Nullpunkt. Dadurch erübrigt sich eine Fourier-Zerlegung der Signale auf der Leitung. Für grobe Übersichtsrechnungen ist diese Näherung akzeptabel.



Figur 94: Phasendiagramm für eine Zweidrahtleitung mit den beiden Asymptoten für niedrige Frequenzen (NF) und hohe Frequenzen (HF).

6.1.5. Charakteristische Impedanz Z_0

Da die z -Abhängigkeit der hinlaufenden Welle durch $e^{-\gamma z}$ beschrieben wird, ergibt sich bei der Ableitung nach z in den Telegraphengleichungen lediglich eine Multiplikation mit $-z$ und wir erhalten beispielsweise mit der zweiten Telegraphengleichung¹¹

$$\gamma \underline{I} = (G' + j\omega C') \underline{U}.$$

¹¹ Man kann für die folgende Herleitung auch die erste Telegraphengleichung heranziehen.

Dividieren wir beidseitig durch I und setzen wir die Formel für γ ein, so finden wir

$$\underline{Z}_0 = \frac{U}{I} = \sqrt{\frac{R' + j\omega L'}{G' + j\omega C'}}.$$

Dieser Wert hat die Dimension einer Impedanz und ist charakteristisch für jede Zweidrahtleitung. Man nennt \underline{Z}_0 deshalb *charakteristische Impedanz* oder *Wellenimpedanz* obwohl sich dieser Wert nicht mit einem Impedanzmessgerät messen lässt.

Für die rücklaufende Welle ergibt sich dieselbe charakteristische Impedanz wie für die hinlaufende Welle, wie man leicht nachprüft.

Von grosser Bedeutung ist die charakteristische Impedanz an den Leitungsenden und an Übergängen von Leitungen mit unterschiedlichen Impedanzen, wie wir später sehen werden.

6.1.6. Verlustfreie Leitungen

Verluste in den Drähten drücken sich im Widerstandsbelag R' , Verluste in den Isolatoren hingegen im Leitwertsbelag G' nieder. Beide Verlustarten entnehmen dem System Energie und dämpfen die Signale. Ersteres ist vor allem bei Energieübertragungsleitungen unerwünscht und wird in der Energietechnik dadurch vermieden, dass Freileitungen (die Luft ist ein qualitativ recht guter Isolator mit geringen Verlusten) mit grossen Drahtabständen, hohen Spannungen und relativ geringen Strömen bei tiefen Frequenzen betrieben werden. In der Informationstechnologie ist man meist gezwungen, bei hohen Frequenzen und niedrigen Spannungen zu arbeiten. Signalverstärker und Digitaltechnik ermöglichen es, gewisse Verluste hinzunehmen. Dabei ist allerdings zu beachten, dass bei schnellen Signalen die Leitungen oft viel grösser als die Wellenlängen sind, was bei Energieversorgungsleitungen kaum der Fall ist. Zudem ergeben sich andere, weit unangenehmere Effekte aus der Frequenzabhängigkeit der Leitungsbeläge. So folgt natürlich eine Frequenzabhängigkeit der Fortpflanzungskonstanten, welche zu Dispersion und damit zu Signalverzerrungen führt. Aus diesem Grund bemüht man sich, möglichst verlustfreie Leitungen herzustellen und zu verwenden. Praktisch kann man Leitungen mit sehr guten Isolatoren und damit vernachlässigbarem Leitwertsbelag G' fabrizieren, der Widerstandsbelag R' lässt sich hingegen nicht beliebig reduzieren, weil die Leitfähigkeit der besten Leiter (Silber, Kupfer) nicht verbessert werden kann und weil die praktisch ideal leitenden Supraleiter aus finanziellen und technischen Gründen (erforderliche Kühlung) nicht in Frage kommen. Der Widerstandsbelag kann deshalb vor allem durch Vergrösserung des Drahtquerschnitts reduziert werden, was dem Wunsch nach Miniaturisierung und geringen Kosten entgegensteht. Ausserdem ist zu beachten, dass die Dämpfung auch bei kleinen Dämpfungskonstanten exponentiell mit der Leitungslänge zunimmt. Damit kann die Dämpfung nicht generell für eine bestimmte Leitung, sondern nur für genügend kurze Leitungen vernachlässigt werden.

Sind die Leitungsverluste vernachlässigbar, so kann man $R' = 0$ und $G' = 0$ einsetzen und erhält wesentlich vereinfachte Gleichungen. So gilt für die Fortpflanzungskonstante

$$\gamma = j\beta = j\omega\sqrt{L'C'} .$$

Die Dämpfungskonstante α verschwindet also (wie hoffentlich erwartet). Die entsprechenden Wellen breiten sich also ungedämpft aus und die Amplituden am Leitungsende sind ebenso gross wie am Leitungsanfang. Für die Phasengeschwindigkeit erhält man sofort

$$v_p = \omega/\beta = 1/\sqrt{L'C'} .$$

Auf den ersten Blick scheint die Phasengeschwindigkeit konstant und gleich der Gruppengeschwindigkeit zu sein, was heissen würde, dass die verlustlose Zweidrahtleitung auch dispersionsfrei ist. Zu beachten ist allerdings, dass die Phasen- und Gruppengeschwindigkeit frequenzabhängig ist, weil L' i.a. frequenzabhängig ist. Bei niedrigen und ebenso bei hohen Frequenzen ist L' immerhin in guter Näherung konstant und die Leitung somit wenigstens in gewissen Frequenzbereichen nahezu dispersionsfrei (siehe Figur 94).

Des weiteren wird auch die charakteristische Impedanz vereinfacht:

$$Z_0 = \frac{U}{I} = \sqrt{\frac{L'}{C'}} .$$

Offensichtlich wird diese Grösse reell — wie ein Ohm'scher Widerstand (obwohl keine Verluste auftreten).

6.1.7. Phasengeschwindigkeit und Wellenlänge

Wir haben bereits gesehen, dass sich die Leitungswellen mit der Phasengeschwindigkeit

$$v_p = \omega/\beta$$

längs der Leitung ausbreiten. Für die Phasengeschwindigkeit gilt ausserdem

$$v_p = \lambda/T,$$

wobei λ die Wellenlänge und T die Periodendauer der Schwingung ist. Für letztere gilt

$$T = 1/f = 2\pi/\omega ,$$

wobei f die Frequenz bezeichnet. Damit ergibt sich der einfache Zusammenhang zwischen Phasenkonstante und Wellenlänge:

$$\lambda = 2\pi/\beta.$$

Dieser ist insofern wichtig, als wir früher gesagt haben, unsere Idealisierungen der Drähte einer Schaltung seien zulässig, wenn die Drahtlängen klein im Vergleich zur Wellenlänge seien. Hier haben wir gewissermassen eine Formel zur Berechnung der Wellenlänge nachgeliefert. Weil sich die involvierte Phasenkonstante als Imaginärteil der Fortpflanzungs-

konstante (ausser bei verlustfreien Leitungen) nicht sehr einfach berechnen lässt und die Kenntnis der Leitungsbeläge voraussetzt, ist die exakte Berechnung der Wellenlänge gar nicht einfach. Für grobe Abschätzungen verwendet man deshalb die einfache Freiraumwellenlänge λ_0 , welche sich aus der Lichtgeschwindigkeit c (ca. $3 \cdot 10^8$ m/s) gemäss

$$\lambda_0 = c/f = 2\pi c/\omega$$

berechnen lässt.

Bei praktischen Leitungen ist die Wellenlänge insbesondere dann kleiner, wenn ein Isolator die beiden Drähte trennt. Wird der Isolator durch die relative Dielektrizitätskonstante ϵ_r und die relative Permeabilitätskonstante μ_r beschrieben, so gilt meist in guter Näherung

$$\lambda = \frac{\lambda_0}{\sqrt{\epsilon_r \mu_r}}.$$

Für die meisten Isolatoren ist übrigens μ_r nahezu 1, so dass vor allem ϵ_r für eine Verkürzung der Wellenlänge sorgt. Für die meisten Isolatoren ist ϵ_r in der Grössenordnung von 2 und nur selten grösser als 12.

6.2. Mehrdrahtleitungen

Will man die Datenraten steigern, so kann man zu Leitungsformen übergehen, welche bei höheren Frequenzen¹² arbeiten oder *Mehrdrahtleitungen* verwenden. Um N Signale auf einer Mehrdrahtleitung parallel zu übertragen braucht es übrigens nicht $2N$ Drähte (d.h. N Drahtpaare) sondern lediglich $N+1$ Drähte, weil auf $N+1$ Drähten N verschiedene Leitungswellen ausbreitungsfähig sind, welche sich durch unterschiedliche Stromverteilungen und unterschiedliche Fortpflanzungskonstanten auszeichnen. Man kann also auf einem N -adrigen Flachbandkabel $N-1$ Signale übertragen¹³. Da die zugehörige Theorie recht auf-

¹² Mit zunehmenden Frequenzen treten verschiedene Probleme auf, die spezielle Massnahmen und eine sehr präzise Fertigung von Datenleitungen für schnelle Pulse erforderlich machen. Beispielsweise werden die Drähte verdreht (Englisch: twisted pair). Obwohl Zweidrahtleitungen geradezu antik anmuten, sind moderne Zweidrahtleitungen High-Tech Produkte, welche nur von wenigen spezialisierten Firmen fabriziert werden können.

¹³ Zu den $N-1$ Leitungswellen existiert übrigens auch bei tiefen Frequenzen ein weiterer Wellentyp, der durch die Leitungstheorie nicht erfasst wird. Bei diesem „Gleichtakt“-Wellentyp fliessen die Ströme in allen Drähten in dieselbe Richtung. Im Gegensatz zu den Leitungswellen ist also die Summe der Leitungsströme durch einen Leitungsquerschnitt nicht Null. Die „Rückleitung“ erfolgt bei diesem speziellen Wellentyp über die Verschiebungsströme, welche in der Leitungstheorie vernachlässigt werden. Verschiebungsströme sind proportional zur zeitlichen Ableitung des elektrischen Feldes und erzeugen — wie die Leitungsströme — ein verwirbeltes Magnetfeld. Verschiebungsströme treten auch in Isolatoren auf und sind zentral für Wellen auf optischen Fasern ebenso wie für die von Antennen abgestrahlten freien Wellen. Die Gleichtaktwelle existiert übrigens auch auf einem einzelnen Draht (Sommerfeld-Draht) und wird in dieser Form manchmal als Speisung von Antennen verwendet. Wird der Einzeldraht mit einem Isolator ummantelt, nennt man die

wendig ist, verzichten wir auf eine ausführliche Darstellung und betrachten lediglich ein besonders einfaches, symmetrisches Flachbandkabel mit 3 identischen Drähten bei einer bestimmten Frequenz. Selbstverständlich verwenden wir die komplexe Darstellung der Ströme und Spannungen.

Aus Symmetriegründen existieren auf diesem Kabel ein symmetrischer Wellentyp mit der Stromverteilung $-\underline{I}_s, 2\underline{I}_s, -\underline{I}_s$ und ein antisymmetrischer Wellentyp mit der Stromverteilung $-\underline{I}_a, 0, +\underline{I}_a$. Die zugehörigen (komplexen) Fortpflanzungskonstanten γ_s und γ_a sind voneinander verschieden und können näherungsweise aus der Ersatzschaltung der Dreidrahtleitung berechnet werden. Der Einfachheit halber nehmen wir ausserdem an, die Fortpflanzungskonstanten seien nicht frequenzabhängig.

In der Digitaltechnik werden Flachbandkabel meist so verwendet, dass der erste Draht auf Masse liegt und die übrigen Drähte je ein Signal übertragen sollen. In unserm einfachen Fall haben wir also am Anfang der Leitung die Stromverteilung $-\underline{I}_{10}-\underline{I}_{20}, \underline{I}_{10}, \underline{I}_{20}$, wenn die beiden Signale (in den Drähten 2 und 3 — Die Indices bezeichnen die Signalnummern und nicht die Drahtnummern!) durch \underline{I}_{10} und \underline{I}_{20} dargestellt werden. Damit werden der symmetrische und der antisymmetrische Wellentyp folgendermassen angeregt:

$$\underline{I}_{s0} = \underline{I}_{10} / 2, \quad \underline{I}_{a0} = \underline{I}_{20} + \underline{I}_{10} / 2.$$

Wie man sieht, werden auch dann beide Wellentypen angeregt, wenn $\underline{I}_{20} = 0$ ist. Ist hingegen $\underline{I}_{10} = 0$, so wird nur die antisymmetrische Welle angeregt.

Am Ende der Leitung erhalten wir

$$\underline{I}_{se} = \underline{I}_{s0} e^{-\gamma_s L}, \quad \underline{I}_{ae} = \underline{I}_{a0} e^{-\gamma_a L},$$

wobei L die Länge der Leitung ist. Daraus finden wir die Ströme in den beiden Signalleitungen, d.h. in den Drähten 2 und 3. Auf der ganzen Leitung und damit auch am Leitungsende gilt:

$$\underline{I}_1 = 2 \underline{I}_s, \quad \underline{I}_2 = \underline{I}_a - \underline{I}_s.$$

Setzen wir die Werte am Leitungsende ein, so finden wir:

$$\underline{I}_{1e} = \underline{I}_{10} e^{-\gamma_s L}, \quad \underline{I}_{2e} = (\underline{I}_{20} + \underline{I}_{10} / 2) e^{-\gamma_a L} - (\underline{I}_{10} / 2) e^{-\gamma_s L}.$$

Offenbar wird bei fehlendem Eingangssignal 2 ($\underline{I}_{2e} = 0$) am Leitungsende $\underline{I}_{2e} \neq 0$ — ausser wenn die beiden Fortpflanzungskonstanten gleich gross sind, was aber normalerweise nicht

entsprechende Welle auch Goubau Welle. Gleichtakt-Wellen werden relativ leicht durch freie Wellen, welche auf die Leitung auftreffen angeregt. Dadurch wird die Übertragung natürlich gestört. Wegen der grossen Störfähigkeit werden Gleichtaktwellen praktisch nicht ausgenützt. Zur Unterdrückung der störenden Gleichtaktwellen werden an den Enden von Zwei- und Mehrdrahtkabeln oft Ferritkerne angebracht. Bei sehr hohen Frequenzen werden weitere Wellen ausbreitungsfähig, welche ebenfalls nicht durch die Leitungstheorie beschrieben werden. Bei so hohen Frequenzen werden Mehrdrahtleitungen allerdings nicht betrieben, so dass sich eine Beschreibung der höheren Wellentypen erübrigt.

der Fall ist. Auf Draht 3 bzw. Signalleitung 2 empfängt man also nicht nur das Signal 2, sondern auch einen Teil des Signals 1. Dies nennt man *Übersprechen*. Dieser Effekt ist bei kurzen Leitungen vernachlässigbar. Will man längere Mehrdrahtleitungen betreiben, so muss man dafür sorgen, dass jedes Signal nur einen bestimmten Wellentyp anregt.

6.3. Andere Leitungsformen

Offene Mehrdrahtleitungen sind nicht nur anfällig auf Übersprechen, sondern auch auf externe Störungen. Um die Störanfälligkeit zu reduzieren wird oft die Massenleitung so ausgelegt, dass sie die Datenleitungen umgibt und gegen äussere Einflüsse abschirmt. Der einfachste Fall mit einer einzigen Datenleitung ist das Koaxialkabel mit „koaxialen“, kreisförmigen Innen- und Aussenleitern. Koaxialkabel werden in der Messtechnik gerne verwendet und ebenso in der analogen Hochfrequenztechnik (z.B. Antennenkabel). Auch für Computernetzwerke wurden Koaxialkabel eine Weile lang bevorzugt, inzwischen aber meist durch Twisted-Pair und optische Fasern abgelöst.

Koaxialkabel lassen sich leitungstheoretisch gleich wie Zweidrahtleitungen behandeln, unterscheiden sich von diesen jedoch durch die Leitungsbeläge und die anderen charakteristischen Grössen. Eine ausführliche Darstellung erübrigt sich deshalb.

Lässt man bei einem Koaxialkabel den Innenleiter weg, so entsteht ein Hohlleiter. Es scheint so, dass damit keine Signale übertragen werden können, da keine Signalleitung vorhanden ist. Tatsächlich werden auf Hohlleitern erst bei hohen Frequenzen verschiedene Wellentypen ausbreitungsfähig, mit denen Information übertragen werden. Die Berechnung dieser Wellentypen basiert jedoch nicht auf der Leitungstheorie, sondern auf den Maxwell Gleichungen. Hohlleitern werden bei sehr hohen Frequenzen zur Übertragung von Signalen mit grossen Leistungen verwendet (Sendeanlagen). Wir verzichten hier auf eine Darstellung weil Informatiker kaum mit Hohlleitern zu tun haben.

Lässt man auch den äusseren Leiter weg, so erhält man dielektrische Wellenleiter, welche wie die Hohlleiter bei sehr hohen Frequenzen verwendet werden. Zu den dielektrischen Wellenleitern gehören auch die optischen Fasern, denen wir uns später zuwenden werden.

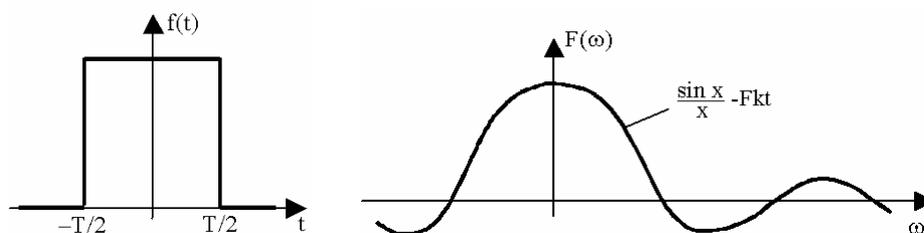
6.4. Impulsausbreitung

Hauptverwendungszweck von Zwei- und Mehrdrahtleitungen in der Digitaltechnik ist die Übertragung von Impulsen. Der Einfachheit halber beschränken wir uns im folgenden auf Zweidrahtleitungen. Bei Mehrdrahtleitungen kommen weitere Effekte¹⁴ hinzu, deren Behandlung aber mathematisch recht hohe Anforderungen stellt. Ausserdem ist die präzise

¹⁴ Wellentypenzerlegung, Übersprechen der verschiedenen Signale etc.

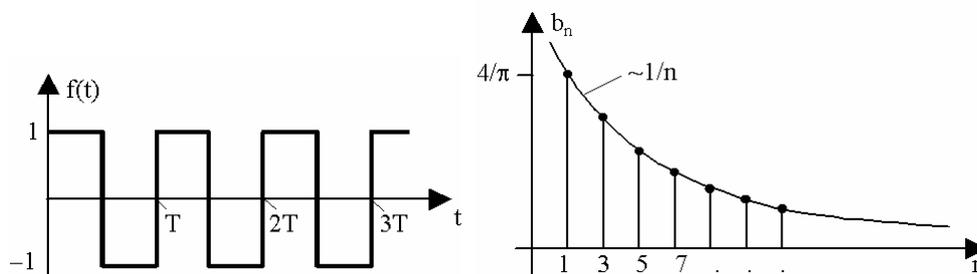
Berechnung der Impulsausbreitung auf Zweidrahtleitungen schon recht schwierig, weil die Frequenzabhängigkeit der Leitungsbeläge zu Dispersion führt, welche sich in Verformungen der Pulsform längs der Leitung niederschlägt. Da diese Berechnung eine gewisse Gewandtheit im Umgang mit Fourier Transformationen voraussetzt, beschränken wir uns auf eine grobe, qualitative Darstellung im folgenden Unterabschnitt. Danach vernachlässigen wir die Dispersion und betrachten das Impulsverhalten an den Leitungsenden etwas ausführlicher.

Zur Berechnung eines einzelnen Pulses muss — wegen der Frequenzabhängigkeit der Leitungsparameter — die Pulsfunktion mit der Fouriertransformation vom Zeitbereich in den Frequenzbereich transformiert werden. Für einfache Zeitfunktionen, wie die Rechteckfunktion gemäss Figur 95 können die entsprechenden Fourierintegrale analytisch gelöst werden und man erhält das entsprechende, kontinuierliche Fourierspektrum. Anschliessend muss die Wellenausbreitung längs der Leitung für jede Frequenz des Spektrums berechnet werden. Schliesslich ergibt sich die gesuchte Lösung aus der Überlagerung aller Fourierkomponenten. Analytisch ist dies nur in einfachen Fällen mit zusätzlichen, vereinfachenden Annahmen möglich. Es ist daher normalerweise sinnvoll, numerische Methoden einzusetzen.



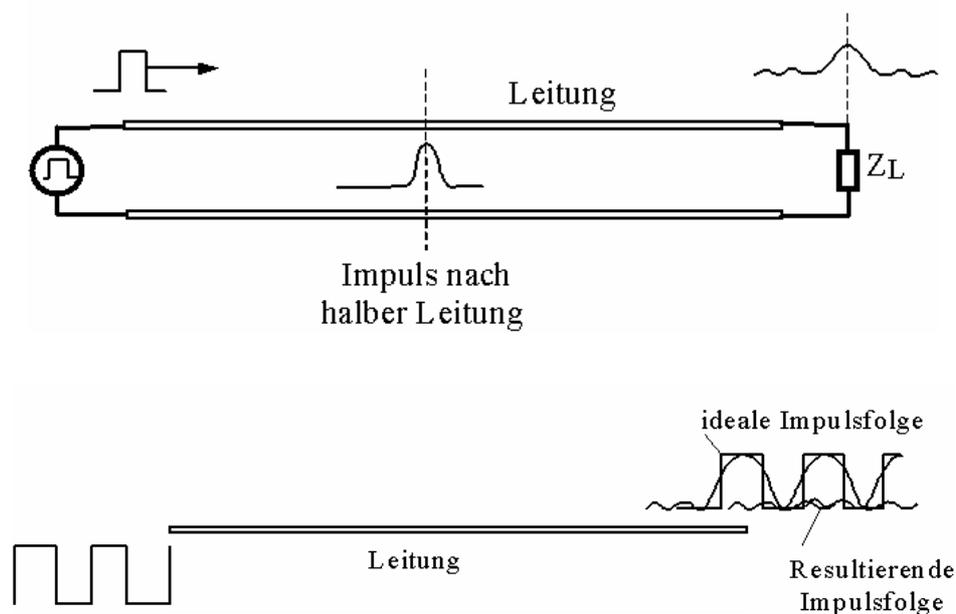
Figur 95: Rechteckimpuls im Zeitbereich und zugehöriges Fourier Spektrum

Etwas einfacher wird die Analyse bei regelmässigen Pulsfolgen (streng periodischen Signalen), da dann das Fourierspektrum diskret wird und „nur“ abzählbar unendlich viele Fourier Komponenten berechnet werden müssen. Figur 96 zeigt beispielsweise eine Rechteckschwingung mit dem zugehörigen Spektrum.



Figur 96: Rechteckschwingung im Zeitbereich und zugehöriges Fourier Spektrum

Zu beachten ist, dass die Fourierspektren unendlich ausgedehnt sind, d.h. unendlich hohe Frequenzen beteiligt sind. Wie wir wissen, nehmen die technischen Probleme ebenso wie die analytischen Probleme mit zunehmender Frequenz zu. So treten bei hohen Frequenzen Effekte auf, die mit der Leitungstheorie nicht erfasst werden. Insbesondere können Wellentypen ausbreitungsfähig werden, die durch das zugehörige elektromagnetische Feld und nicht einfach durch Leitungsströme und -spannungen beschrieben werden können. Zudem nimmt (glücklicherweise?) die Dämpfung mit zunehmender Frequenz zu. Dies hat zur Folge, dass die Leitungen wie Tiefpässe wirken und die hochfrequenten Anteile des Fourierspektrums unterdrückt werden. Für die Übertragungstechnik ergeben sich daraus unerwünschte Impulsverzerrungen. Für die numerische Berechnung reduziert sich allerdings der Aufwand, da hohe Frequenzen mit gutem Gewissen ausser Acht gelassen werden können. Figur 97 illustriert die Pulsformen am Ein- und Ausgang einer Zweidrahtleitung. Dabei ist übrigens angenommen, dass an den Leitungsenden keine Reflexionen auftreten. Wie Reflexionen vermieden werden können, werden wir im folgenden sehen.



Figur 97: Dispersion von Einzelimpuls und Impulsfolge auf einer reflexionsfreien Zweidrahtleitung

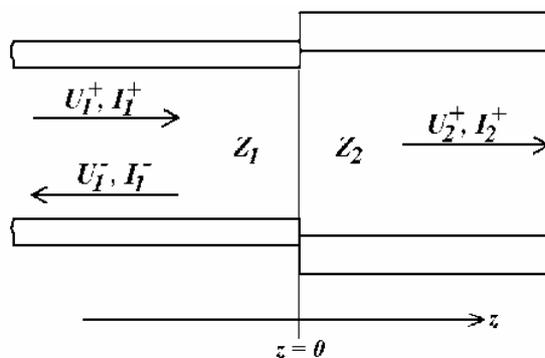
6.5. Stossstellen und Abschlüsse

In der Praxis verwendet man oft nicht nur eine einzelne Leitung zwischen Sender und Empfänger. So hat man beispielsweise an einer Antenne einen symmetrischen Anschluss mit einer einfachen Doppeldrahtleitung, will aber ein Koaxialkabel anschliessen, um Störeinflüsse zu vermeiden. Verbindet man unterschiedliche Leitungen miteinander, so treten vor allem am Übergang (Stossstelle) Phänomene auf, die wir nun betrachten wollen. Diese Phänomene sind im Prinzip für alle Leitungen dieselben. Deshalb beschränken wir uns der

Einfachheit halber zunächst auf die Zweidrahtleitungen, welche wir bisher am ausführlichsten beschrieben haben. Dabei vernachlässigen wir die Dispersion, nehmen also an, die Fortpflanzungskonstanten der beteiligten Leitungen seien nicht frequenzabhängig. Dies ist eine grobe Vereinfachung, welche praktisch selten gut erfüllt ist, es uns aber ermöglicht, den Haupteffekt mit geringem Aufwand zu verstehen. Was wir im folgenden ausführen gilt also streng genommen für sinusförmige Signale mit einer bestimmten Frequenz, in guter Näherung für Pulse mit begrenztem Spektrum in einem Frequenzbereich, über den sich die Fortpflanzungskonstanten nur wenig ändern und nur grob für Pulse mit ausgedehntem Spektrum.

6.5.1. Stossstellen von Zweidrahtleitungen

Eine Verbindung zweier Zweidrahtleitungen mit unterschiedlicher Querschnittsgeometrie braucht natürlich etwas Raum. Wir gehen aber davon aus, dass die Abmessungen des Überganges viel kleiner als die Leitungslängen und auch viel kleiner als die kürzesten Wellenlängen sind, welche für die Datenübertragung eine Rolle spielen. Wir idealisieren nun einen solchen Übergang, indem wir seine Abmessungen vernachlässigen und annehmen, die Querschnittsgeometrie ändere sich bei $z = 0$ abrupt (siehe Figur 98), aber so, dass die Ströme und Spannungen von der ersten Leitung ($z < 0$) auf die zweite Leitung ($z > 0$) kontinuierlich übergehen und dass die Verbindungen der Leitungsdrähte ideale Knoten seien, welche keine Ladung speichern. Wir behandeln die Stossstelle also so, wie eine normale Verbindung in einem Netzwerk. Ausserdem beginnen wir mit einer Wechselstrombetrachtung und beschreiben die Ströme und Spannungen mit komplexen Grössen.



Figur 98: Stossstelle zweier Zweidrahtleitungen. Die Pfeile bezeichnen die Laufrichtungen der beteiligten Wellen. Konvention für Strom- und Spannungspfeile (nicht eingezeichnet): Alle Spannungspfeile zeigen von der oberen zur unteren Leitung. Die Strompfeile der oberen Drähte zeigen in die Laufrichtungen, in den unteren Drähten entgegengesetzt dazu.

Zunächst sei eine in $+z$ Richtung laufende Welle auf der ersten Leitung ($z < 0$) vorhanden. Wir nennen diese *einfallende Welle* und bezeichnen die zugehörige Spannung und den zugehörigen Strom mit $\underline{U}_1^+(z)$, $\underline{I}_1^+(z)$. Diese Welle regt an der Stossstelle ($z = 0$) eine *transmittierte Welle* mit $\underline{U}_2^+(z)$, $\underline{I}_2^+(z)$ an, welche auf der zweiten Leitung in $+z$ Richtung weiter läuft. Sind die Wellenimpedanzen \underline{Z}_1 und \underline{Z}_2 der beiden Leitungen unterschiedlich, so kön-

nen die Kontinuitätsgleichungen an der Stossstelle $\underline{U}_1^+(0) = \underline{U}_2^+(0)$, $\underline{I}_1^+(0) = \underline{I}_2^+(0)$ nicht gleichzeitig mit den Impedanzgleichungen $\underline{Z}_1 = \underline{U}_1^+(0) / \underline{I}_1^+(0)$, $\underline{Z}_2 = \underline{U}_2^+(0) / \underline{I}_2^+(0)$ erfüllt werden. Um diese Gleichungen erfüllen zu können, müssen wir zusätzlich auf der Leitung 1 eine *reflektierte Welle* mit $\underline{U}_1^-(z)$, $\underline{I}_1^-(z)$ ansetzen. Diese Welle soll in $-z$ Richtung laufen und wir setzen deshalb fest, dass die positive Zählrichtung für den Strom dieser Welle umgekehrt sei als für die einfallende Welle. Es gelten deshalb an der Stossstelle die Kontinuitätsgleichungen

$$\underline{U}_1^+(0) + \underline{U}_1^-(0) = \underline{U}_2^+(0),$$

$$\underline{I}_1^+(0) - \underline{I}_1^-(0) = \underline{I}_2^+(0).$$

Mit Hilfe der Wellenimpedanzen können wir in der ersten Gleichung die Spannungen durch Ströme ausdrücken und erhalten

$$\underline{Z}_1 (\underline{I}_1^+(0) + \underline{I}_1^-(0)) = \underline{Z}_2 \underline{I}_2^+(0).$$

Nun können wir entweder \underline{I}_1^- oder \underline{I}_2^+ eliminieren und erhalten den Strom der transmittierten oder der reflektierten Welle in Abhängigkeit vom Strom der einfallenden Welle:

$$\underline{I}_2^+(0) = \underline{T} \underline{I}_1^+(0), \text{ wobei } \underline{T} = 2\underline{Z}_1 / (\underline{Z}_1 + \underline{Z}_2);$$

$$\underline{I}_1^-(0) = \underline{\Gamma} \underline{I}_1^+(0), \text{ wobei } \underline{\Gamma} = (\underline{Z}_2 - \underline{Z}_1) / (\underline{Z}_1 + \underline{Z}_2).$$

Dabei ist \underline{T} der Transmissionskoeffizient und $\underline{\Gamma}$ der Reflexionskoeffizient. Man prüft sofort nach, dass $\underline{T} = 1$ und $\underline{\Gamma} = 0$ genau dann gilt, wenn die beiden Wellenimpedanzen \underline{Z}_1 und \underline{Z}_2 gleich sind. Ist dies der Fall, so sind die Leitungen *angepasst* und es entsteht keine Reflexion an der Stossstelle. Dabei ist zu beachten, dass Leitungen mit unterschiedlicher Querschnittsgeometrie durchaus dieselbe Wellenimpedanz haben können und dass andererseits Leitungen mit identischer Querschnittsgeometrie aber unterschiedlichen Materialeigenschaften auch unterschiedliche Wellenimpedanzen haben.

Da sich die transmittierten Wellen mit der Fortpflanzungskonstanten γ_2 in $+z$ Richtung ausbreiten, gilt für diese

$$\underline{I}_2^+(z) = \underline{T} \underline{I}_1^+(0) \exp(-\gamma_2 z).$$

Die reflektierten Wellen hingegen breiten sich in $-z$ Richtung aus, so dass

$$\underline{I}_1^-(z) = \underline{\Gamma} \underline{I}_1^+(0) \exp(+\gamma_1 z)$$

gilt.

In der Praxis werden Leitungen so gebaut, dass die Verluste möglichst klein sind, so dass die Widerstands- und Leitwertbeläge vernachlässigt werden können und die Dämpfungskonstanten α verschwinden. Wie wir bereits gesehen haben, werden die Wellenimpedanzen auf derartigen, *verlustfreien* Leitungen reell. Daraus ergibt sich sofort, dass auch die Reflexions- und Transmissionskoeffizienten von Stossstellen verlustfreier Leitungen reell wer-

den. Dadurch wird insbesondere die Berechnung der Impulsausbreitung vereinfacht, wenn wir annehmen, die Leitungen seien *dispersionsfrei*, d.h. die Wellenausbreitung auf den aneinander stossenden Leitungen werde durch frequenzunabhängige Fortpflanzungskonstanten beschrieben. In diesem Fall kann die Fourieranalyse der Pulse umgangen werden und anstelle der obigen komplexen Gleichungen ergeben sich die einfachen zeitabhängigen Gleichungen

$$i_2^+(0,t) = T i_1^+(0,t), \text{ wobei } T = 2Z_1 / (Z_1 + Z_2);$$

$$i_1^-(0,t) = \Gamma i_1^+(0,t), \text{ wobei } \Gamma = (Z_2 - Z_1) / (Z_1 + Z_2).$$

Die an der Stossstelle (bei $z = 0$) erzeugten, transmittierten Pulse wandern dann mit der Geschwindigkeit v_{p2} in $+z$ Richtung, die reflektierten Pulse hingegen mit der Geschwindigkeit v_{p1} in $-z$ Richtung.

6.5.2. Leitungsabschluss

Prinzipiell endet jede Datenleitung beim Sender und beim Empfänger. Üblicherweise findet man dort einen Stecker, welcher die Leitung mit einem Netzwerk innerhalb des Senders bzw. Empfängers verbindet. Von einer Zweidrahtleitung aus „sieht“ man im Prinzip einen Zweipol, welcher recht komplizierte Eigenschaften haben kann. Wir betrachten im folgenden nur den einfachen Fall eines linearen Zweipols. Im Wechselstromfall nehmen wir also an, der Sender könne durch eine Spannungsquelle \underline{U}_q mit Impedanz \underline{Z}_q und der Empfänger durch eine einfache Lastimpedanz \underline{Z}_L dargestellt werden. Das Vorgehen ist nun identisch mit dem Vorgehen bei der Stossstelle zweier Leitungen, d.h. wir idealisieren die Drahtverbindungen von Leitung und Empfänger bzw. Sender durch Knoten und rechnen mit den üblichen Netzwerkgleichungen. Dadurch ergeben sich im wesentlichen dieselben Gleichungen, wie bei der Stossstelle, wenn wir die Impedanz der zweiten Leitung durch die Lastimpedanz bzw. die Quellenimpedanz¹⁵ ersetzen. Es gilt also für die Reflexionskoeffizienten bei der Last bzw. Quelle

$$\underline{\Gamma}_L = (\underline{Z}_L - \underline{Z}) / (\underline{Z}_L + \underline{Z}), \quad \underline{\Gamma}_q = (\underline{Z}_q - \underline{Z}) / (\underline{Z}_q + \underline{Z}),$$

wobei \underline{Z} die Wellenimpedanz der Leitung ist.

Damit ergibt sich offenbar an den Leitungsenden keine Reflexion, wenn die Impedanzen *angepasst* sind, d.h. für $\underline{Z}_L = \underline{Z}$ bzw. $\underline{Z}_q = \underline{Z}$.

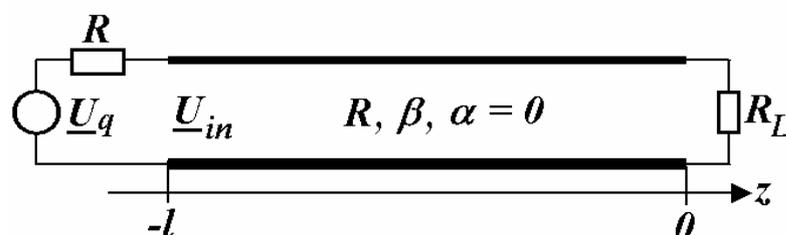
Neben dem Idealfall der Anpassung ohne jede Reflexion, sind die Extremfälle mit maximalen Reflexionskonstanten (*Totalreflexion*) beachtenswert. So ergibt sich offensichtlich

¹⁵ Dass die Spannungsquelle des Senders für ein Signal, welches von der Leitung auf den Sender einfällt keinen Einfluss hat, ergibt sich aus der Quellenüberlagerung. Die Quelle des einfallenden Signals wird am andern Leitungsende (beim Verbraucher) gedacht. Tatsächlich entsteht dieses Signal üblicherweise durch Reflexion am Leitungsende (beim Verbraucher).

$\underline{\Gamma} = -1$ für $\underline{Z}_L = 0$ bzw. $\underline{Z}_q = 0$ (Kurzschluss) und $\underline{\Gamma} = +1$ für $\underline{Z}_L = \infty$ bzw. $\underline{Z}_q = \infty$ (Leerlauf). Dabei ist wichtig, dass eine ideale Spannungsquelle (Kurzschluss) ebenso wie eine ideale Stromquelle (Leerlauf) Totalreflexionen ergeben, also in dieser Hinsicht keineswegs ideal sind. Der Innenwiderstand von Hochfrequenzquellen ist deshalb normalerweise ein festgelegter Wert, welcher der Impedanz handelsüblicher Kabel entspricht (häufig 50Ω).

6.5.3. Eingangsimpedanz

Wir betrachten nun eine verlustfreie Leitung ($\alpha = 0$, $\gamma = j\beta$) mit Wellenimpedanz $\underline{Z} = R$, welche auf der einen Seite durch eine sinusförmige Spannungsquelle angeregt wird und auf der andern Seite mit einer Impedanz $\underline{Z}_L = R_L$ belastet wird. Ausserdem sei die Leitung eingangsseitig angepasst, d.h. der Innenwiderstand der Quelle sei ebenfalls R . Figur 99 illustriert dies.



Figur 99: Verlustlose Leitung angeregt durch eine angepasste Wechselspannungsquelle, belastet mit einem beliebigen Widerstand.

Damit wird der Reflexionsfaktor am Leitungsende reell. Da es üblich ist, das Leitungsende als Referenzebene mit Längskoordinate $z = 0$ zu bezeichnen, schreiben wir:

$$\underline{\Gamma}_L = \Gamma(0) = \frac{R_L - R}{R_L + R} = |\underline{\Gamma}_L| e^{j\Phi},$$

wobei die Phase Φ des Reflexionsfaktors Null ist. Die Reflexion am Lastwiderstand erzeugt eine rückläufige Welle mit \underline{U}^- , \underline{I}^- , welche sich der hinlaufenden Welle mit \underline{U}^+ , \underline{I}^+ überlagert. Damit ergibt sich für die Spannung und den Strom längs der Leitung

$$\underline{U}(z) = \underline{U}^+ e^{-j\beta z} + \underline{U}^- e^{+j\beta z} = \underline{U}^+ (e^{-j\beta z} + \underline{\Gamma}_L e^{+j\beta z}) = \underline{U}^+ e^{-j\beta z} (1 + \underline{\Gamma}_L e^{2j\beta z}),$$

$$\underline{I}(z) = \frac{\underline{U}^+ e^{-j\beta z} - \underline{U}^- e^{+j\beta z}}{\underline{Z}} = \underline{U}^+ \frac{e^{-j\beta z} - \underline{\Gamma}_L e^{+j\beta z}}{\underline{Z}} = \frac{\underline{U}^+}{\underline{Z}} e^{-j\beta z} (1 - \underline{\Gamma}_L e^{2j\beta z}).$$

Diese Gleichungen gelten natürlich auch am Eingang der Leitung, bei $z = -l$, wo man diese Größen gut messen kann. Dort gilt:

$$\underline{U}(-l) = \underline{U}^+ (e^{+j\beta l} + \underline{\Gamma}_L e^{-j\beta l}),$$

$$\underline{I}(-l) = \frac{\underline{U}^+}{\underline{Z}} (e^{+j\beta l} - \underline{\Gamma}_L e^{-j\beta l}).$$

Damit messen wir am Eingang die *Eingangsimpedanz*

$$\underline{Z}_{in} = \frac{\underline{U}(-l)}{\underline{I}(-l)} = \frac{\underline{U}^+ (e^{+j\beta l} + \underline{\Gamma}_L e^{-j\beta l})}{\underline{I}^+ (e^{+j\beta l} - \underline{\Gamma}_L e^{-j\beta l})} = \underline{Z} \frac{e^{+j\beta l} + \underline{\Gamma}_L e^{-j\beta l}}{e^{+j\beta l} - \underline{\Gamma}_L e^{-j\beta l}}.$$

Die Eingangsimpedanz ist offenbar nicht nur vom Reflexionsfaktor, sondern auch von den Leitungseigenschaften und insbesondere von der Leitungslänge abhängig. Mit Zweidrahtleitungen können deshalb Impedanzen „erzeugt“ werden, welche sich schaltungstechnisch verwenden lassen. Ersetzen wir noch den Reflexionsfaktor durch die Leitungsimpedanz \underline{Z} und die Abschlussimpedanz \underline{Z}_L ($\underline{\Gamma}_L = (\underline{Z}_L - \underline{Z})/(\underline{Z}_L + \underline{Z})$), so finden wir nach kurzer Umrechnung:

$$\underline{Z}_{in} = \underline{Z} \cdot \frac{\underline{Z}_L + j\underline{Z} \tan(\beta l)}{\underline{Z} + j\underline{Z}_L \tan(\beta l)}.$$

Meist wird nun noch die Eingangsimpedanz auf die Leitungsimpedanz bezogen und die *normalisierte Reaktanz*

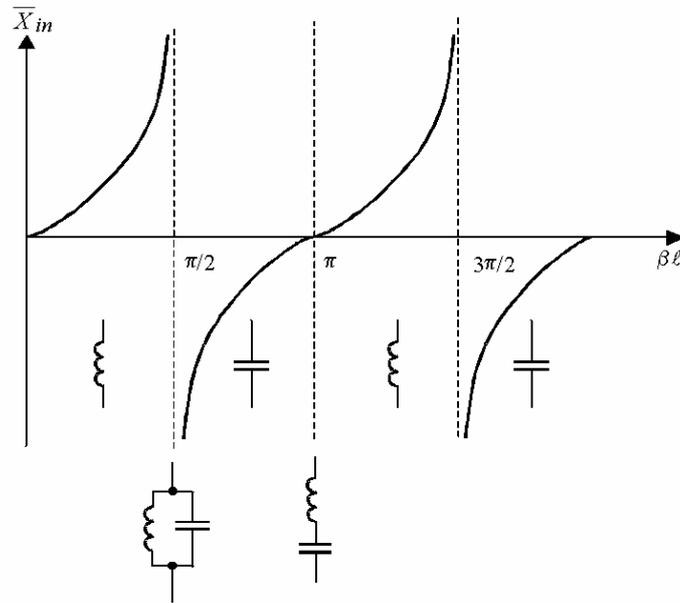
$$\overline{X}_{in} = -j \frac{\underline{Z}_{in}}{\underline{Z}} = -j \frac{\underline{Z}_L + j\underline{Z} \tan(\beta l)}{\underline{Z} + j\underline{Z}_L \tan(\beta l)}$$

eingeführt.

Besonders interessant sind nun die Spezialfälle der kurzgeschlossenen Leitung ($\underline{Z}_L = 0$) und der offenen Leitung ($\underline{Z}_L = \infty$) mit Totalreflexion am Leitungsende. Mit $\underline{Z}_L = 0$ finden wir

$$\overline{X}_{in} = -j \cdot \frac{j\underline{Z} \tan(\beta l)}{\underline{Z}} = \tan(\beta l).$$

Figur 100 zeigt den Verlauf der normalisierten Reaktanz.



Figur 100: Normalisierte Reaktanz einer kurzgeschlossenen Leitung in Funktion der Leitungslänge (bei fester Frequenz — βl ist proportional zur Frequenz), bzw. der Frequenz (bei fester Länge).

Die genauere Analyse zeigt, dass sich die kurzgeschlossene Leitung — vom Eingang her gesehen — für gewisse Werte von βl wie eine Induktivität und für andere Werte wie eine Kapazität benimmt. Ist βl ein ganzzahliges Vielfaches von $\pi/2$, so benimmt sich die Leitung wie ein Schwingkreis. Dieser Sachverhalt ist in Figur 100 angedeutet. Innerhalb eines Frequenzbereichs kann die verlustfreie Leitung (bei fester Länge, Wellenwiderstand R) also beispielsweise als frequenzabhängige Impedanz mit $L = R \tan(\beta l)/\omega$ verwendet werden.

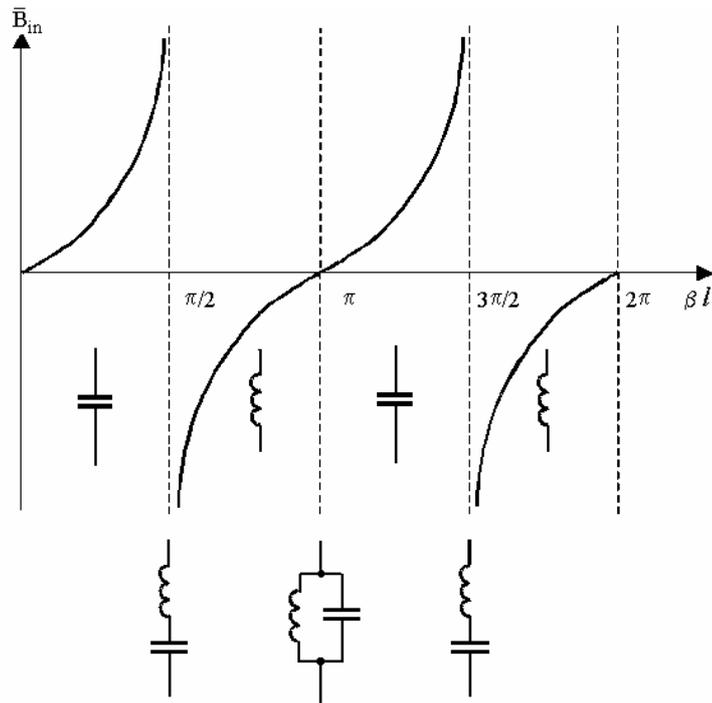
Das Vorgehen bei der offenen Leitung ist ganz analog, man erhält zunächst

$$\underline{Z}_{in} = -j\underline{Z} \cot(\beta l) = -j\underline{Z} / \tan(\beta l).$$

Nun führt man aber die *normalisierte Suszeptanz*

$$\overline{B}_{in} = -j\underline{Z} / \underline{Z}_{in} = \tan(\beta l)$$

ein. Figur 101 zeigt den Verlauf der normalisierten Reaktanz.



Figur 101: Normalisierte Suszeptanz einer offenen Leitung in Funktion der Leitungslänge (bei fester Frequenz — β ist proportional zur Frequenz), bzw. der Frequenz (bei fester Länge).

6.5.4. Stehende Wellen

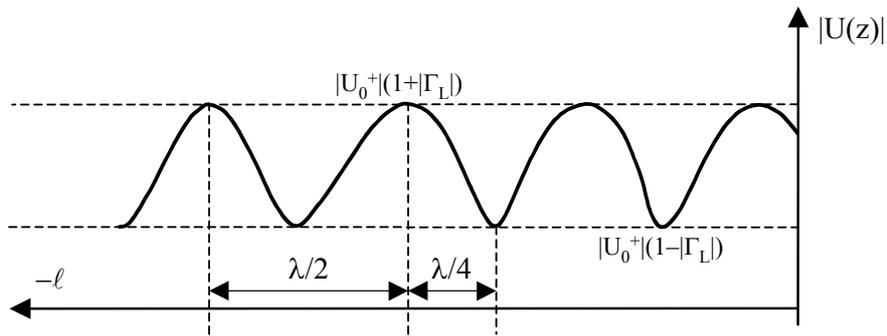
Wir betrachten nun noch den Verlauf von Strom und Spannung längs der Zweidrahtleitung (wo sich Messungen etwas schwerer durchführen lassen als am Eingang). Die Formeln für die komplexen Strom- und Spannungswerte haben wir bereits angegeben. Da sich die Beträge viel leichter messen lassen als die komplexen Werte, berechnen wir nun die Beträge. Für die Spannung finden wir:

$$|\underline{U}(z)| = \left| \underline{U}^+ e^{-j\beta z} (1 + \underline{\Gamma}_L e^{2j\beta z}) \right| = \left| \underline{U}^+ \right| \left| 1 + \underline{\Gamma}_L e^{2j\beta z} \right| = \left| \underline{U}^+ \right| \left| 1 + |\underline{\Gamma}_L| e^{j(\Phi - 2\beta z)} \right|.$$

Auf analoge Weise finden wir für den Strom:

$$|\underline{I}(z)| = \left| \underline{U}^+ / \underline{Z} \right| \left| 1 - |\underline{\Gamma}_L| e^{j(\Phi - 2\beta z)} \right|.$$

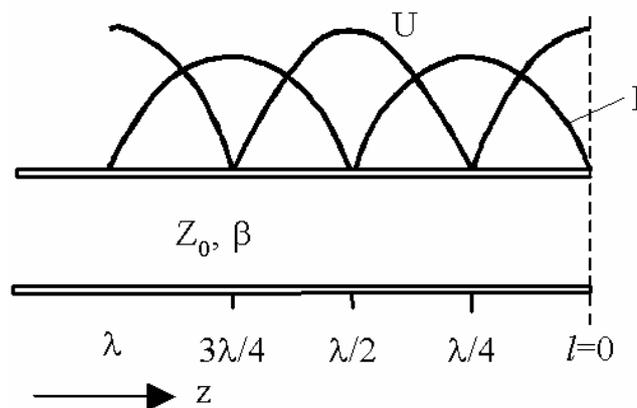
Spannungsmaxima längs der Leitung treten offenbar dann auf, wenn $\exp(j(\Phi - 2\beta z)) = 1$ wird und Spannungsminima wenn $\exp(j(\Phi - 2\beta z)) = -1$. Figur 102 illustriert den Verlauf des Spannungsbetrags bei einer verlustfreien Leitung mit komplexer Abschlussimpedanz.



Figur 102: Verlauf des Spannungs Betrags auf einer verlustfreien Leitung mit komplexer Abschlussimpedanz.

Aus der Distanz des ersten Spannungsmaximums vom Leitungsende lässt sich offenbar die Phase Φ des Reflexionsfaktors ablesen. Ist die Abschlussimpedanz — ebenso wie die Impedanz der Leitung — reell, so wird der Reflexionsfaktor reell und somit $\Phi = 0$ wenn der Reflexionsfaktor positiv ist und $\Phi = \pi$ wenn der Reflexionsfaktor negativ ist. Damit findet man also am Leitungsende entweder ein Spannungsmaximum oder ein Spannungsminimum. Aus dem Ort der Spannungsmaxima lässt sich aber auch die Wellenlänge auf der Leitung ablesen. Für die Distanz d benachbarter Maxima (oder Minima) gilt nämlich $2\beta d = 2\pi$ und damit $\lambda = 2d$.

Interessant sind nun wieder die Spezialfälle. Zunächst ergeben sich bei Totalreflexion, d.h. Kurzschluss und Leerlauf mit Reflexionsfaktor -1 bzw. $+1$ Vereinfachungen. Man sieht, dass die Spannungsminima $= 0$ werden (und ebenso die Stromminima). Die z Abhängigkeiten der Strom- und Spannungs beträge werden im wesentlichen durch die Funktionen $|\sin|$ und $|\cos|$ beschrieben. Figur 103 zeigt den Fall einer offenen Leitung. Da die rückläufige Welle ebensoviel Energie transportiert wie die hinlaufende Welle, wird längs der Leitung überhaupt keine Energie transportiert. Die Spannungs- und Stromwerte an jedem Ort z sind zwar zeitlich nicht konstant, es scheint aber so, dass sich die Welle nicht bewegt, d.h. am Ort stehen bleibt und man spricht daher von einer *stehenden Welle*.



Figur 103: Verlauf des Spannungs- und Strom Betrags auf einer offenen, verlustfreien Leitung.

Ist die Leitung ideal abgeschlossen, so existiert keine reflektierte Welle. Die hinlaufende Welle transportiert dann die gesamte Energie vom Sender zum Empfänger. Zudem werden die Spannungs- und Strombeträge längs der gesamten Leitung konstant.

Für einen nicht idealen Abschluss ergibt sich im Verbraucher eine Überlagerung der absorbierten Welle und einer stehenden Welle. Letztere äussert sich in einem bestimmten Verhältnis von Spannungsmaximum zu Spannungsminimum (Stehwellenverhältnis, Englisch: Standing Wave Ratio = SWR). Man schreibt:

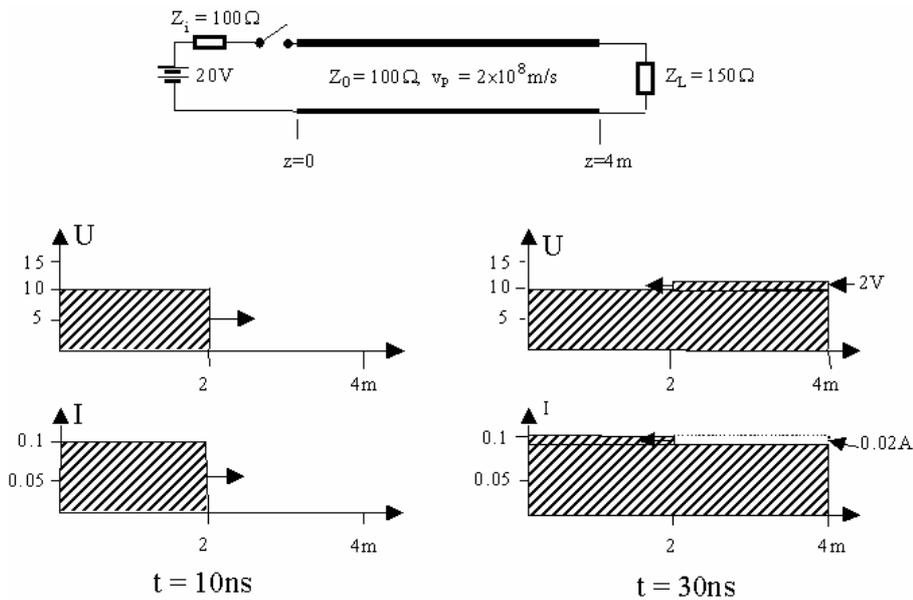
$$SWR = \frac{U_{\max}}{U_{\min}} = \frac{1 + |\Gamma_L|}{1 - |\Gamma_L|}.$$

Das Stehwellenverhältnis ist positiv reell und liegt zwischen 1 (idealer Abschluss) und unendlich (Totalreflexion am Leitungsende). Aus dem Stehwellenverhältnis findet man den Betrag des Reflexionsfaktors:

$$|\Gamma_L| = \frac{SWR - 1}{SWR + 1}.$$

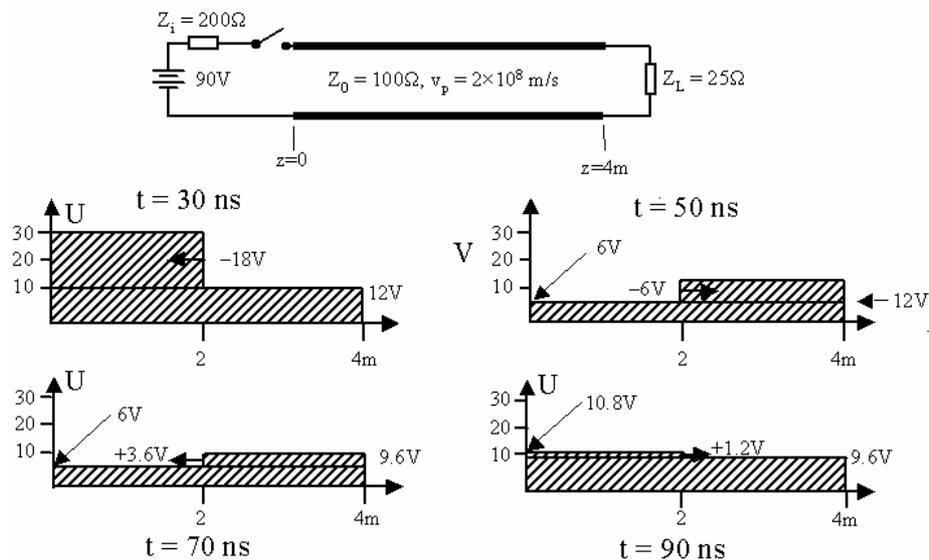
6.5.5. Mehrfachreflexionen

Reflexionen an Stossstellen von Leitungen und an Leitungsenden sind in der Analogtechnik ebenso wie in der Digitaltechnik unerwünscht. Ist lediglich das Leitungsende beim Empfänger nicht angepasst, so wird dadurch das Empfangssignal abgeschwächt, was für den Empfänger keine Probleme ergibt, wenn der Transmissionskoeffizient nicht sehr klein ist. Figur 104 illustriert dies für den Einschaltvorgang bei einer verlustfreien, dispersionslosen Leitung, wobei die Dispersion vernachlässigt wurde.



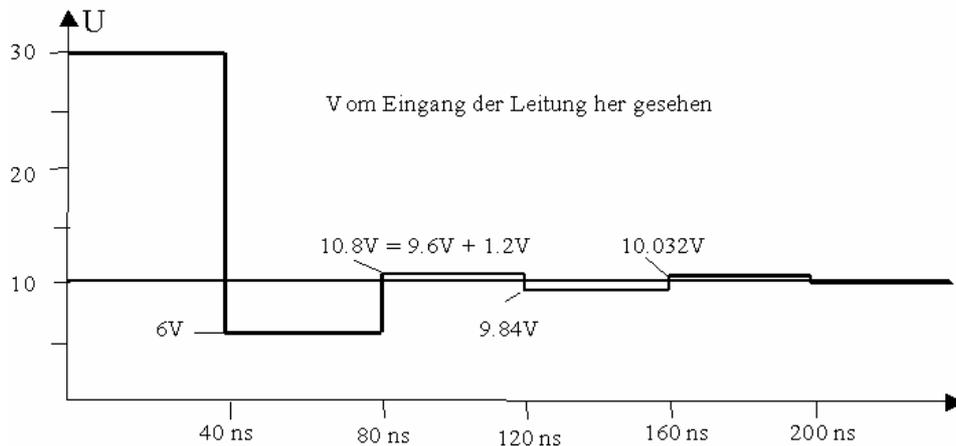
Figur 104: Einschaltvorgang einer einseitig angepassten Leitung, welche am Ende nicht angepasst ist.

Sind hingegen an beiden Leitungsenden Fehlanpassungen vorhanden, so ergeben sich Mehrfachreflexionen, d.h. wiederholt hin- und herlaufende Wellen (siehe Figur 105). Dasselbe gilt natürlich auch, wenn Leitungen mit verschiedenen Impedanzen verwendet werden oder wenn die unpräzise Fertigung eines Kabels dazu führt, dass die Impedanz des Kabels nicht über seine gesamte Länge konstant ist.



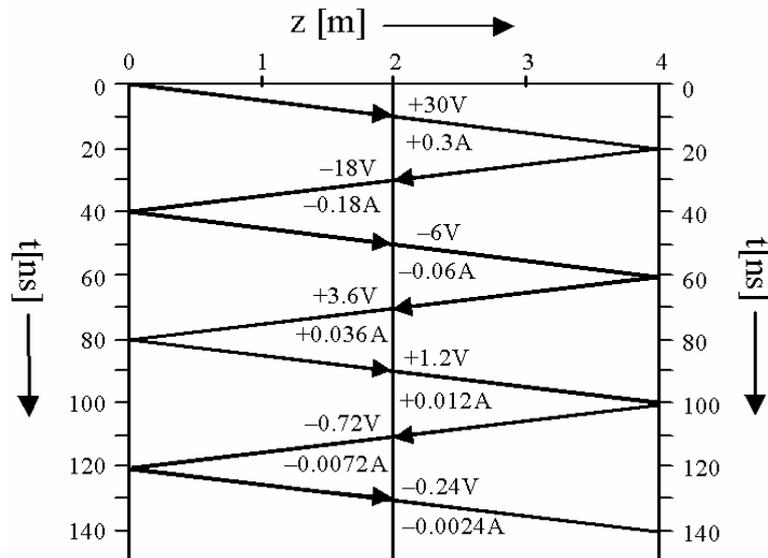
Figur 105: Einschaltvorgang einer beidseitig fehlangepassten Leitung.

Die Signale auf Leitungen mit Mehrfachreflexionen können mit unendlichen Reihen dargestellt werden. Da die Beträge der Reflexionskoeffizienten normalerweise kleiner als 1 sind, nehmen die Werte mit jeder zusätzlichen Reflexion ab. Oft genügt es, nur wenige Terme zu berücksichtigen, da die Reihen rasch konvergieren, wenn die Leitungen einigermassen gut angepasst sind. Figur 106 illustriert dies für die Eingangsspannung derselben Anordnung wie in Figur 105.



Figur 106: Eingangsspannung beim Einschaltvorgang gemäss Figur 105.

Um eine bessere Übersicht zu erhalten, ist ein *Raum-Zeit-Diagramm*, wie in Figur 107, vorteilhaft.



Figur 107: Raum-Zeit-Diagramm für den Einschaltvorgang gemäss Figur 105.

Die Behandlung von Mehrfachreflexionen bei dispersiven Leitungen ist viel schwieriger als im dispersionsfreien Fall. Man kommt dann um eine Fourieranalyse nicht mehr herum und hat ausserdem frequenzabhängige Reflexionskoeffizienten.

6.6. Frequenzmultiplex

Wie erwähnt, sind bei Hohlleitern die ausbreitungsfähigen Wellentypen frequenzmässig nach unten begrenzt. Obwohl theoretisch auf optischen Fasern bei beliebig tiefen Frequenzen ausbreitungsfähige Wellentypen existieren, können auch diese Wellenleiter in der Praxis nur innerhalb eines begrenzten Frequenzbereichs $f_{min} \dots f_{max}$ betrieben werden. Das Frequenzspektrum des Rechteckimpulses gemäss Figur 94 lässt vielleicht vermuten, dass zur Übertragung von Pulsen der niederfrequente Spektralanteil wichtig sei. Tatsächlich haben streng zeitbegrenzte Signale¹⁶ immer ein unendlich ausgedehntes Frequenzspektrum. Man kann in der Digitaltechnik aber auch mit Pulsen arbeiten, welche ein endliches Frequenzspektrum haben und dafür nicht zeitbegrenzt sind. Aus der Ähnlichkeit der Fouriertransformation und der zugehörigen Rücktransformation kann man folgern, dass beispielsweise ein Signal mit rechteckförmigem Frequenzgang im Zeitbereich im wesentlichen die Form $Asin(kt)/(kt)$ aufweist. Man kann also in der Figur 94 einfach die beiden Figuren miteinander vertauschen.

Die Funktion $Asin(kt)/(kt)$ hat zwar unendlich viele Nebenmaxima aber nur ein einziges globales Maximum. Ist die Amplitude A bekannt, so kann man elektronisch mit einfachen Mitteln einen Rechteckpuls zuordnen. Dazu setzt man beispielsweise eine Schwelle $A/2$ und stellt mit einem Komparator fest, ob das Signal grösser als der Schwellenwert ist. Trifft dies zu, so ordnet man eine logische "1" zu, andernfalls eine "0". Neben $Asin(kt)/(kt)$ existieren unendlich viele Zeitfunktionen mit einem begrenzten Frequenzspektrum. In der Praxis müssen diese Funktionen elektronisch erzeugt werden. So kann man beispielsweise mit Schaltern zunächst zeitbegrenzte Signale erzeugen und diese anschliessend durch ein Bandpassfilter laufen lassen. Da keine idealen Bandpassfilter existieren, ergeben sich natürlich Funktionen deren Frequenzspektrum nur annähernd begrenzt ist. Obwohl diese mathematisch viel schwerer zu beschreiben sind als $Asin(kt)/(kt)$, können sie doch ein recht ähnliches Aussehen im Zeitbereich haben.

Verwendet man Folgen von frequenzbegrenzten Pulsen mit Frequenzspektrum $f_{min} \dots f_{max}$, so muss gewährleistet sein, dass die Abstände zwischen den aufeinanderfolgenden Pulsen gross genug sind. Andernfalls kann es sein, dass die Nebenmaxima mehrerer logischer "1" Werte sich an einem Ort wo eigentlich eine logische "0" anliegt zu einem Wert oberhalb der Schwelle aufsummieren, so dass fälschlicherweise eine "1" detektiert wird. Es braucht deshalb einen gewissen zeitlichen Sicherheitsabstand. Dieser wächst übrigens mit der Signalbandbreite $f_{max} - f_{min}$. Wichtig ist aber zunächst, dass man nahezu frequenzbegrenzte Pulse

¹⁶ Signale die nur innerhalb eines endlichen Zeitintervalls $t_1 \dots t_2$ von Null verschieden sind.

erzeugen kann, welche nur während einer kurzen Zeitspanne eine bestimmte Schwelle überschreiten.

Arbeitet man mit Pulsen deren Frequenzspektrum begrenzt ist, so kann man offenbar auch auf derselben Leitung gleichzeitig mehrere Signale mit unterschiedlichen Frequenzspektren übertragen. Dies nennt man Frequenzmultiplexing. Dieses Verfahren ist übrigens schon sehr alt und wurde seit Beginn der Rundfunktechnik verwendet um mit demselben Empfänger mehrere Sender zu empfangen. Die Frequenzzuteilung (genauer Zuteilung von Frequenzbereichen) ist für den heutigen Mobilfunk ein zentrales Problem, weil die Anwender nach immer höheren Datenraten schreien und höhere Datenraten breitere Frequenzbänder (Bandbreite) erforderlich machen.

Das Vorgehen beim Frequenzmultiplexing besteht im wesentlichen aus den folgenden Schritten: 1) Erzeugung bandbegrenzter Signale mit Bandpassfiltern, 2) Addition der Signale, 3) Übertragung, 4) Demultiplexing, d.h. Zerlegung des Summensignals in die einzelnen Signale mit Bandpassfiltern, 5) Detektion der zugehörigen Information (im digitalen Fall z.B. Detektion der logischen Zustände wie oben angedeutet).

Frequenzmultiplexing ermöglicht es, mehrere Datenkanäle mit demselben Wellentyp zu übertragen. Das führt allerdings nicht dazu, dass insgesamt mehr Daten übertragen werden können, macht also vor allem dann Sinn, wenn die Übertragungskapazität eines Wellentyps viel höher als die gewünschte Datenrate des Datenkanals ist.

7. Antennen

Wie wir im vorhergehenden Kapitel gesehen haben, können Zwei- oder Mehrdrahtleitungen dazu verwendet werden, um elektrische Energie und Information von einem Sender zu einem Verbraucher zu übertragen. Das Medium dieser Übertragung sind elektromagnetische Wellen, welche sich längs der Leitungen ausbreiten. Wir haben ebenfalls gesehen, dass die Wellen an Stossstellen und Leitungsenden — immer dann wenn sich die Wellenimpedanz ändert — zum Teil reflektiert werden, wobei wir davon ausgegangen sind, dass die Energie der einfallenden Wellen sich auf die reflektierten und transmittierten Wellen verteilt. In der Praxis zeigt es sich, dass die Energiebilanz — je nach geometrischer Ausformung des Leitungsüberganges oder -abschlusses ein Defizit aufweist, weil ein Teil der Energie abgestrahlt wird. Bei der Energieübertragung bemüht man sich, die Übergänge so auszulegen, dass nahezu keine Abstrahlung auftritt, weil Energie ein kostbares Gut ist und weil ausserdem die abgestrahlten Wellen elektronische Geräte stören könnten. Das gilt auch für die Nachrichtenübertragung. Hier führen Abstrahlungen längs der Nachrichtenleitung zusätzlich zum unerwünschten Energieverlust dazu, dass die Übertragung leicht abgehört werden kann und dass andere Übertragungsleitungen gestört werden. Aus diesen Gründen werden Leitungsübergänge so ausgelegt, dass die Strahlungsverluste möglichst gering sind.

Die Tatsache, dass Diskontinuitäten von Leitungen zur Abstrahlung elektromagnetischer Wellen führen, kann auch nutzbringend ausgenützt werden. Z.B. wenn ein Nachrichtensender seine Nachrichten¹ jedem Empfänger zukommen lassen möchte, welcher sich für die entsprechenden Nachrichten interessiert und nicht über Kabel erreichbar ist. Die Möglichkeit, Nachrichten elektronisch zu verschlüsseln, macht die freie (nicht durch Leitungen geführte) Wellenausbreitung aber auch für die Übertragung von einem Sender zu einem bestimmten Empfänger interessant und wird im Mobilfunk ausgenützt. Will man nicht geführte Wellen aussenden, so muss man offenbar Diskontinuitäten von Leitungen so auslegen, dass ein möglichst grosser Teil der einfallenden Leitungswelle abgestrahlt wird. Derartige Diskontinuitäten nennen wir Antennen.

Die saubere Beschreibung der freien Wellenausbreitung und der Antennen basiert auf den Maxwell Gleichungen und erfordert ein gutes Verständnis der elektromagnetischen Felder, welches wir hier nicht voraussetzen möchten. Wir beschränken uns deshalb im Folgenden auf eine vorwiegend qualitative Beschreibung und vereinfachte Modelle.

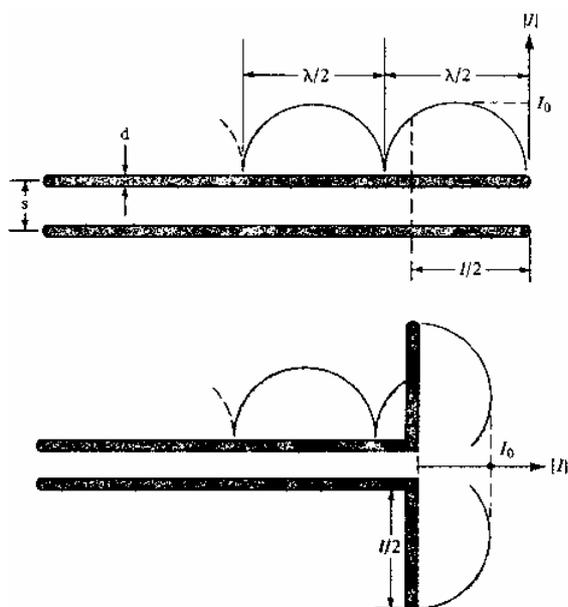
¹ Unter Nachricht verstehen wir hier nicht nur mündliche Nachrichten, sondern auch Musik, Bilder, etc. — jede Art von Information.

7.1. Dipolantennen

Als erstes betrachten wir eine offene (Leerlauf) eine symmetrische Zweidrahtleitung. An der Diskontinuität am Leitungsende wird die einfallende Welle gemäss Leitungstheorie total reflektiert und zwar so, dass der Strom am Leitungsende verschwindet. Betrachten wir der Einfachheit halber den harmonischen Fall, so ergibt sich eine stehende Welle längs der Leitung. Ist s ein Parameter, welcher den Abstand vom Drahtende angibt, so gilt also für die Amplitude des Stromes längs eines Drahtes im Wesentlichen

$$I(s) = A \sin(ks).$$

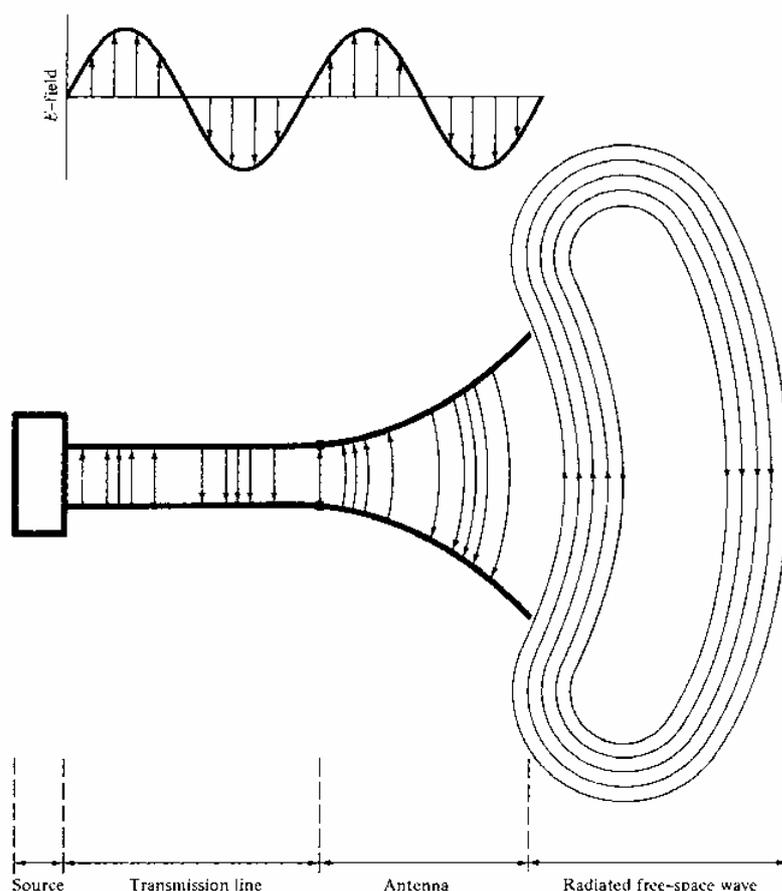
Nun verändern wir die Geometrie der Drahtenden mit der Absicht, möglichst grosse Abstrahlung zu erhalten. Dabei nehmen wir an, dass sich die obige Stromverteilung längs des Drahtes durch die Verformung des Drahtes in erster Näherung nicht ändert, dass wir also die Stromverteilung wenigstens näherungsweise kennen. Figur 108 illustriert dies für rechtwinklig abgeknickten Drahtenden.



Figur 108: Stromverteilung einer leerlaufenden Zweidrahtleitung (oben) und einer Zweidrahtleitung, welche eine symmetrische Dipolantenne speist (unten).

Betrachtet man die Leitung mit den abgeknickten Drahtenden aus einiger Distanz, so „sieht man einen Stab (die Drahtenden) und ein Kabel, welches zur Mitte des Stabes führt. Das elektromagnetische Feld längs des Kabels ist in der Umgebung der Drähte (grösstenteils sogar zwischen den Drähten) konzentriert. Dieses Feld breitet sich praktisch nur längs des Kabels aus. Anders ist die Situation beim Stab (der Dipolantenne). Um zu verstehen, was passiert, muss man das elektromagnetische Feld verstehen. Das folgende ist ein qualitativer Beschreibungsversuch.

Der zeitveränderliche Längsstrom in jedem Draht erzeugt ein magnetisches Feld, welches den Draht umkreist und ebenfalls zeitveränderlich ist. Wie wir wissen, würde ein solches Magnetfeld in einer Drahtschleife einen Strom induzieren. Tatsächlich erzeugt dieses Feld ein verwirbeltes elektrisches Feld, auch wenn keine Drahtschleife vorhanden ist. Das elektrische Feld ist seinerseits wieder zeitveränderlich und erzeugt ein verwirbeltes, zeitveränderliches Magnetfeld usw. Dieser Mechanismus führt zu einer wellenförmigen Ausbreitung des elektromagnetischen Feldes. Im Bereich der Zweidrahtleitung ist die elektromagnetische Welle geführt, d.h. sie bewegt sich längs der Leitung. Beim Antennenstab kommt es hingegen zur „Ablösung“ eines Teils des elektromagnetischen Feldes vom Stab, d.h. ein Teil der elektromagnetischen Welle wird abgestrahlt. Figur 109 illustriert dies auf stark vereinfachte Weise. In Wirklichkeit ist das elektromagnetische Feld in der Nähe der Antenne recht kompliziert.



Figur 109: Vereinfachte Darstellung der Funktionsweise einer Antenne. Die Pfeile stellen das elektrische Feld dar.

Im *Fernfeld*, d.h. in einer Distanz von der Antenne, welche viel grösser als die Wellenlänge und auch viel grösser als die Länge des Antennenstabes ist, erscheint die Antenne wie ein Punkt. Hier bewegt sich die elektromagnetische Welle radial von der Antenne weg. Ausserdem sind das elektrische und das magnetische Feld hier zueinander und zur Ausbrei-

tungsrichtung orthogonal. Da zudem das elektrische Feld zum magnetischen Feld proportional ist, wird diese freie Welle durch folgende Grössen beschrieben: 1) *Ausbreitungsrichtung* (radial), 2) *Polarisation* (Richtung des elektrischen Feldes), 3) *Intensität* (Leistungsdichte).

Wie bereits gesagt, steht das elektrische Feld senkrecht zur (radialen) Ausbreitungsrichtung. Ausserdem liegt der elektrische Feldvektor in der Ebene, welche durch die Richtung des Antennenstabes und des radialen Vektors gebildet wird. Das Magnetfeld hingegen steht senkrecht zu dieser Ebene.

Die Intensität S gibt die Leistung pro Flächeneinheit an. S ist vom Ort abhängig. Zur genaueren Beschreibung ist es natürlich, neben den kartesischen Koordinaten (x,y,z) Kugelkoordinaten (ρ, φ, θ) zu verwenden. Der Ursprung der Koordinatensysteme ($x = y = z = \rho = 0$) wird dort gewählt, wo die Antenne gespiesen wird, also im Falle der Dipolantenne in der Stabmitte. Normalerweise zeigt die z Achse in Richtung des Dipolstabes. Da die Dipolantenne bezüglich ihrer Achse rotationssymmetrisch ist, gibt es keine natürliche Wahl der x und y Richtungen. Diese werden also willkürlich festgelegt. Mit der z Achse wird der Winkel θ eindeutig festgelegt, der Winkel φ ergibt sich hingegen aus der willkürlichen Wahl der x Achse. Integriert man über eine grosse Kugel mit Radius r , welche die Antenne umschliesst, so ergibt sich für die gesamte abgestrahlte Leistung

$$P_{tot} = \oint\!\!\!\oint S(r, \varphi, \theta) dF .$$

Da sich für beliebige Kugelradien dieselbe totale Leistung ergeben muss und die Abstrahlung im Fernfeld in radialer Richtung erfolgt, muss die Dichte S der abgestrahlten Leistung mit dem Abstand im Quadrat abnehmen². Um in grossem Abstand von der Antenne ein Signal zu empfangen, braucht man also entweder eine hohe Sendeleistung oder einen empfindlichen Empfänger.

Als Empfänger verwendet man wieder eine Antenne, welche das empfangene Signal über eine Leitung üblicherweise zu einer elektronischen Schaltung überträgt. Dabei ist nicht nur die Leistungsdichte am Ort der *Empfangsantenne*, sondern auch die Orientierung dieser Antenne wichtig. Im Falle einer Dipolantenne als Sender und einer zweiten Dipolantenne als Empfänger, muss die Empfangsantenne möglichst so ausgerichtet werden, dass sie in Richtung des elektrischen Feldvektors der gesendeten Welle zeigt, d.h. die Empfangsantenne muss senkrecht zum (oben erwähnten) radialen Vektor stehen und in derselben Ebene liegen, wie der radiale Vektor und die Sendeantenne. In der Praxis stellt man Sendeantennen so auf, dass das elektrische Feld entweder senkrecht zum Erdboden (vertikal) oder parallel zum Erdboden (horizontal) polarisiert ist. Die Empfangsantennen müssen dann entsprechend aufgestellt werden. Ist die Sendeantenne vertikal, die Empfangsantenne hingegen

² Dies gilt nicht nur für Dipolantennen, sondern für jede beliebige Antennenanordnung.

horizontal ausgerichtet, so ergibt sich theoretisch kein Empfang³. Ist ϕ der Winkel um den die Empfangsantenne gegenüber der Polarisation der zu empfangenden Welle „verdreht“ montiert wird, so wird die Empfangsleistung um $\cos^2 \phi$ reduziert.

7.2. Strahlungsdiagramme, charakteristische Größen

Da die Dipolantenne nicht kugelsymmetrisch ist, ist es auch naheliegend, dass sie nicht in alle Raumrichtungen gleich stark abstrahlt, d.h. $S(\rho, \varphi, \theta)$ ist nicht nur vom Abstand ρ von der Antenne abhängig. Tatsächlich ist es unmöglich, Antennen zu konstruieren, welche in alle Richtungen gleich stark abstrahlen. Trotzdem arbeitet man gerne mit solchen *Kugelstrahlern* um die Strahlung einer realen Antenne zu charakterisieren. Für einen Kugelstrahler gilt

$$S_{Kugel}(\rho, \varphi, \theta) = P_{tot} / (4\pi\rho^2).$$

Da die Abstandsabhängigkeit einfach und für alle Antennen identisch ist, kann man diese von den Winkelabhängigkeiten separieren indem man die Strahlungscharakteristik auf diejenige des Kugelstrahlers bezieht. Das Verhältnis

$$D = S / S_{Kugel}$$

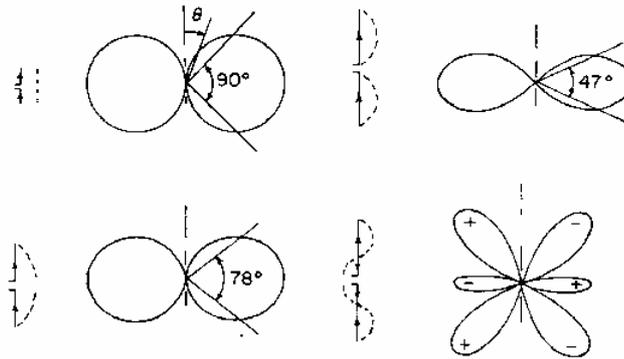
ist für alle Antennen nur winkelabhängig und wird *Richtfaktor* genannt. Meist verwendet man dabei übrigens ein logarithmisches Mass (dB), weil die Werte von D um mehrere Größenordnungen variieren können.

Um sich die Sache nicht zu leicht zu machen, arbeitet man nicht gerne mit der Leistungsdichte, sondern mit dem Betrag der elektrischen Feldstärke. Die Situation ist hier ganz ähnlich wie bei der Schaltungstechnik, wo $P = UI$ gilt. Im Falle eines Ohm'schen Widerstandes R wird dann $P = U^2/R$. Für die abgestrahlte Leistungsdichte im Fernfeld gelten ganz ähnliche Beziehungen. Wie einer Leitung kann man auch dem freien Raum einen Wellenwiderstand⁴ Z_w zuordnen, welcher das Verhältnis von elektrischer Feldstärke E zur magnetischen Induktion H bezeichnet. Es gilt dann analog $S = EH = E^2/Z_w$. Statt S/S_{Kugel} zeichnet man gerne das Strahlungsdiagramm für E/E_{Kugel} .

Da die Dipolantenne rotationssymmetrisch ist, ist auch deren Strahlungsfeld rotationssymmetrisch. Um ihr Strahlungsverhalten zu charakterisieren, reicht deshalb die Angabe der Abhängigkeit vom Winkel θ . Figur 110 zeigt die Strahlungsdiagramme der Dipolantenne für unterschiedliche Stablängen bei einer bestimmten Frequenz.

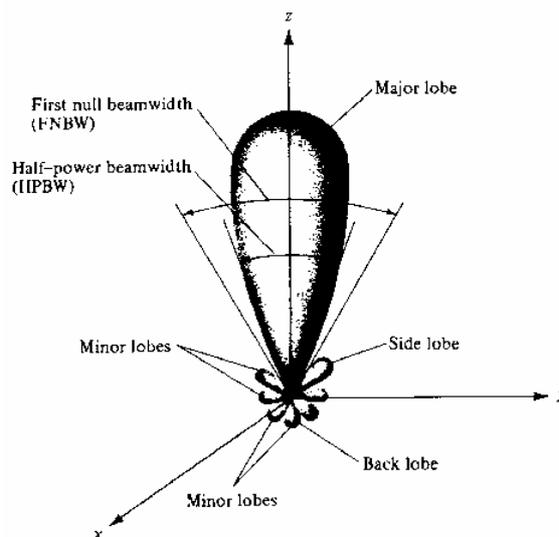
³ In der Praxis verändern bzw. stören Umgebungseinflüsse (Hügel, Gebäude, etc.) das Antennenfeld so, dass die Polarisation nicht streng erhalten bleibt.

⁴ In einem verlustbehafteten Material ergibt sich eine komplexe Wellenimpedanz, im verlustfreien Raum wird die Impedanz — wie bei der verlustfreien Leitung — reell.



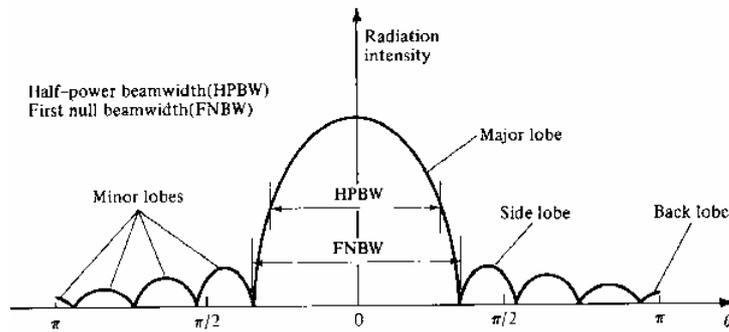
Figur 110: Unterschiedliche Stromverteilungen und zugehörige Strahlungsdiagramme einer symmetrischen Dipolantenne. Die Diagramme in der Ebene senkrecht zur Antenne sind wegen der Rotationssymmetrie der Anordnung einfache Kreise.

Bei komplizierteren Antennen sind die Strahlungsdiagramme nicht rotationssymmetrisch, was deren Darstellung erheblich erschwert. Figur 11 zeigt ein Beispiel.



Figur 111: Räumliche Darstellung der Richtcharakteristik einer Antenne und die wichtigsten charakteristischen Grössen.

Die Schwierigkeit, allgemeine Strahlungsdiagramme zu zeichnen und die dazu erforderliche grosse Datenmenge führt dazu, dass oft nur Strahlungsdiagramme in einer Schnittebene (z.B. xy , xz oder yz Ebene) gezeichnet werden, wie beispielsweise in Figur 112.



Figur 112: Einfache „eindimensionale“ Darstellung der Strahlungsintensität einer Antenne.

Um eine rasche Übersicht über die Abstrahlungseigenschaften einer Antenne zu haben, führt man verschiedene charakteristische Grössen ein, welche in den Figuren 111 und 112 eingetragen sind.

Bisher sind wir davon ausgegangen, dass eine Antenne die ihr zugeführte Leistung in den freien Raum abstrahlt. Tatsächlich ist dies nur zum Teil der Fall. Von der Zuleitung aus gesehen, benimmt sich die Antenne wie eine (frequenzabhängige) Impedanz. Man nennt diese Impedanz auch *Fusspunktimpedanz*. Es ergeben sich damit — je nach Impedanz der Zuleitung — Reflexionen. Ist die Antennenimpedanz bekannt, so berechnet man die Reflexions- und Transmissionsfaktoren genau gleich wie bei Stossstellen von Leitungen oder beim Abschluss einer Leitung.

Der Transmissionsfaktor gibt an, welcher Teil der Leitungswelle tatsächlich in die Antenne hinein läuft. Die entsprechende Leistung ist aber nicht gleich der abgestrahlten Leistung, weil ein Teil dieser Leistung auch in Wärme umgewandelt wird und zwar durch Ohm'sche Verluste in den Antennendrähten oder auch in verlustbehafteten Materialien in der Umgebung der Antenne (z.B. Erdboden). Für die totale, abgestrahlte Leistung gilt also

$$P_{tot} = \eta P_{in},$$

wobei η den *Wirkungsgrad* bezeichnet.

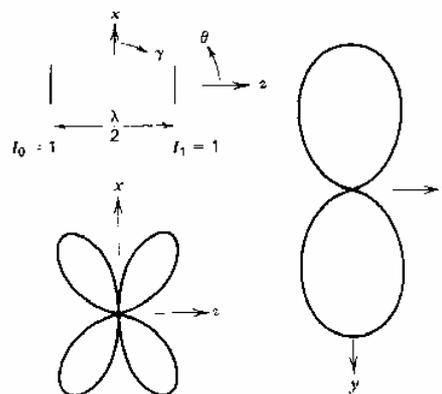
Bei Rundfunkantennen will man meist nach allen Richtungen möglichst gleichmässig abstrahlen. Anders ist das beispielsweise bei Nachrichtenübertragungen mit Satelliten. Hier versucht man, den grössten Teil der abgestrahlten Energie auf den zum Empfang bestimmten Satelliten zu richten. Dies erreicht man mit verschiedenen Methoden, welche im folgenden skizziert werden. Da in diesem Fall die Richtung Sender-Empfänger festgelegt ist, orientiert man die Sendeantenne so, dass der Empfänger die maximale Leistung erhält. Diese ist dann nicht nur proportional zum Wirkungsgrad, sondern auch zum maximalen Wert des Richtfaktors D . Das Produkt ηD_{max} nennt man *Gewinn*.

Bisher haben wir vor allem Sendeantennen betrachtet. Prinzipiell existiert keinerlei Unterschied zwischen Sende- und Empfangsantennen. Empfangsantennen können also gleich wie Sendeantennen charakterisiert werden. Da ein Rundfunkempfänger meist kleine Leistungen

empfängt und die Richtung zum Sender bekannt ist, verwendet man hier meist Antennen mit ausgeprägter Richtcharakteristik. Anders ist das beim Mobilfunk, wo sich der Empfänger nicht immer am gleichen Ort aufhält und nicht jedesmal seine Antenne in Richtung des Senders ausrichten möchte. Je nach Anwendung können die verwendeten Sende- und Empfangsantennen daher sehr unterschiedliche Formen aufweisen. Die Kunst des Antennenentwurfs besteht darin, Antennen zu finden, welche den anwendungsspezifischen Anforderungen genügen. Eine optimale Antenne für alle möglichen Anwendungen gibt es nicht.

7.3. Antennen-Arrays, Smart Antenna

Für viele Anwendungen ist die rotationssymmetrische Strahlungscharakteristik der einfachen Dipolantenne nicht befriedigend und man sucht daher nach komplizierteren Anordnungen mit einer gewissen Richtcharakteristik. Eine relativ einfache Möglichkeit besteht darin, mehrere Antennenstäbe zu verwenden und diese gegenseitig passend anzuordnen. Figur 113 zeigt die Strahlungsdiagramme einer Anordnung von zwei parallelen Dipolantennen im Abstand $\lambda/2$ mit identischer Speisung. Da die Anordnung nicht mehr rotationsymmetrisch in Bezug auf die Antennenachse ist, ergibt sich auch in der Ebene senkrecht zu den Antennenstäben eine Richtcharakteristik. Verwendet man mehrere Antennenstäbe — welche meist regelmässig angeordnet werden — so erhält man ein Antennen-Array. Mit zunehmender Zahl der Stäbe gelingt es dabei, die Richtwirkung nahezu beliebig zu verbessern.

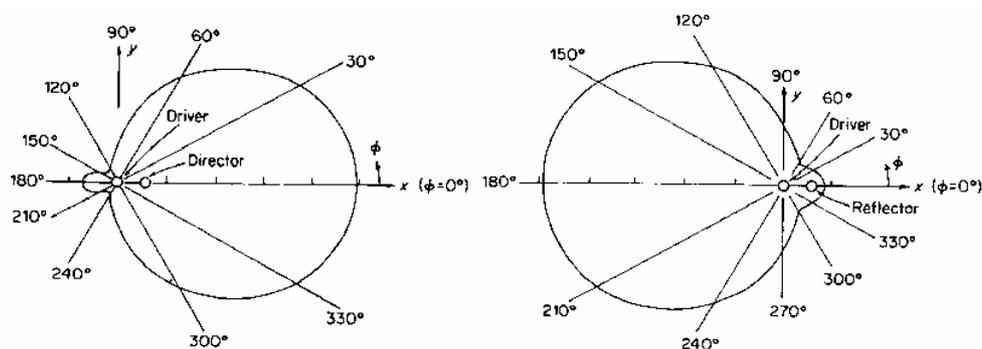


Figur 113: Strahlungsdiagramme in zwei Ebenen für zwei symmetrische Halbwellen-Dipolantennen. Beide Antennen weisen identische Stromverteilungen auf.

Die Richtcharakteristik hängt von verschiedenen geometrischen Faktoren, wie den Positionen, Orientierungen und Längen der Antennenstäbe, aber auch von elektrischen Grössen, wie Amplitude und Phasenlage der Antennenströme, ab. Da sich die Geometrie einer Antenne technisch nur schlecht verändern lässt, ist diese meist bei der Produktion festgelegt. Die elektrischen Grössen lassen sich jedoch mit elektronischen Schaltungen verändern.

Dadurch ergibt sich die Möglichkeit, beispielsweise die Hauptstrahlrichtung einer Antenne während ihres Betriebes elektronisch einzustellen und so entweder den Sender auf einen mobilen Empfänger zu richten oder umgekehrt. Antennen, welche sich selber so einstellen, dass die Übertragung optimiert wird, nennt man Englisch *smart antenna*.

Mehrere Antennenstäbe mit einer passenden Amplitude und Phase des Sendesignals zu versorgen, erfordert einen recht hohen elektronischen Aufwand. In vielen Fällen versucht man deshalb mit einer einzigen Speisung auszukommen. Dabei ist wichtig zu sehen, dass ein (passiver) Antennenstab ohne jede Zuleitung im Feld einer gespiesenen (aktiven) Antenne gewissermassen als Empfangsantenne auftritt und die empfangene Leistung einerseits (zu einem meist kleinen Teil) in Wärme umwandelt und andererseits wieder abstrahlt. Ein Teil der abgestrahlten Leistung wird von der aktiven Antenne empfangen und wieder abgestrahlt usw. Die exakte Berechnung der Wechselwirkung einer aktiven Antenne mit einer oder mehreren passiven Antennen ist recht schwierig und hängt sowohl von der geometrischen Anordnung als auch von der Frequenz bzw. Wellenlänge ab. In den meisten Fällen wird die Länge der aktiven Antenne so gewählt, dass die gesamte Stablänge etwa⁵ die halbe Wellenlänge ist. Wird parallel dazu eine passive Stabantenne aufgestellt, welche etwas (ca. 10%) kürzer ist, so ergibt sich — bei geeignetem Stababstand — ein Strahlungsdiagramm gemäss Figur 114. Wie man sieht, strahlt diese Konfiguration vor allem in Richtung des passiven Stabes, welcher deshalb *Direktor* genannt wird. Ist der passive Stab hingegen etwas länger, so dreht sich die Hauptstrahlrichtung um. Der passive Stab wirkt also ähnlich wie ein Spiegel und wird deshalb *Reflektor* genannt.



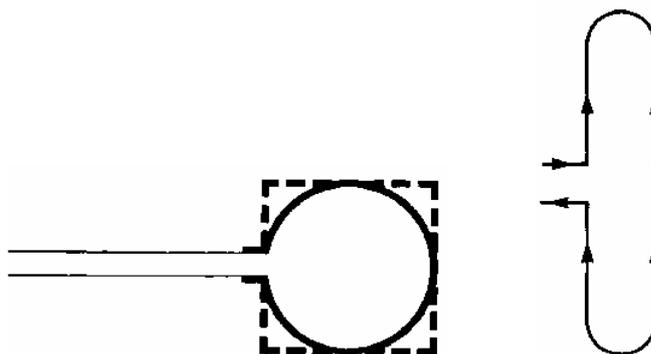
Figur 114: Strahlungsdiagramme einer Dipolantenne mit Direktor (links) und Reflektor (rechts).

Selbstverständlich kann man mehrere Direktoren und Reflektoren verwenden um die Richtcharakteristik weiter zu verbessern.

⁵ Die genaue Länge hängt auch vom Drahtradius und der Geometrie der Drahtenden ab.

7.4. Loop-Antennen

Bisher haben wir Dipolantennen betrachtet, welche im wesentlichen aus einer leerlaufenden Zweidrahtleitung entstanden sind. Eine gebräuchliche Alternative dazu sind Loop-Antennen, welche aus einer kurzgeschlossenen Zweidrahtleitung entstehen. Diese Antennen sehen wie kleine Spulen (meist mit nur einer Windung) aus (siehe Figur 115).

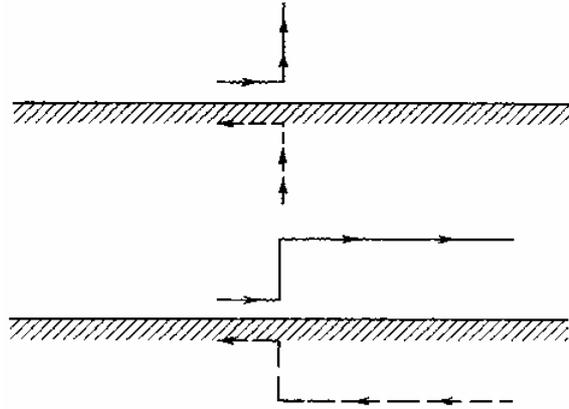


Figur 115: Loop-Antennen (links: kreisförmig oder quadratisch; rechts: gefalteter Dipol), welche als kurzgeschlossene Zweidrahtleitungen aufgefasst werden können.

Oft werden Loop-Antennen mit Dipolantennen kombiniert. Beispielsweise sieht man bei Fernsehantennen oft einen gefalteten Dipol (Figur 115, rechts) als aktive Antenne und mehrerer passive Dipole als Direktoren und Reflektoren.

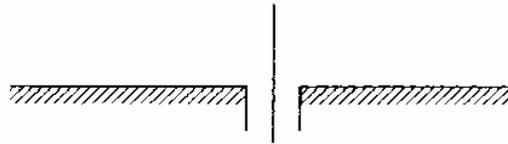
7.5. Spiegelungsprinzip, Grundplatten

Metallische Flächen spiegeln nicht nur Licht, sondern auch niederfrequenterer elektromagnetische Wellen. Wird eine elektrisch gut leitende Ebene in die Nähe einer Antennenanordnung gebracht, so spiegelt sich die Antennenanordnung, d.h. die Stromverteilung (und das elektromagnetische Feld) darin, wie in Figur 116 dargestellt. Dadurch ergibt sich die Möglichkeit, nur die Hälfte einer symmetrischen Antennenanordnung zu realisieren und den symmetrischen Teil durch einen Spiegel zu „simulieren“. Dies ist vor allem auch deshalb interessant, weil in der Schaltungstechnik bei höheren Frequenzen oft mit einer elektrischen Grundplatte (auf Potential 0, d.h. „geerdet“) wird, auf der ein Dielektrikum aufgebracht ist, auf welchem wiederum Leiterbahnen und Bauelemente montiert werden.



Figur 116: Spiegelungsprinzip: Eine leitende Grundplatte „spiegelt“ die gegebene Stromverteilung.

Betrachtet man den oberen Teil von Figur 116, so sieht man, dass die symmetrische Dipolantenne im oberen Halbraum dasselbe Verhalten haben muss, wie eine auf einer Grundplatte vertikal montierte, sogenannte *Monopolantenne*. Für eine solche Antenne ist es naheliegend, den Strom nicht parallel zur Grundplatte, sondern senkrecht dazu — mit einer kreisförmigen Durchführung — einzuspeisen. Dadurch erhält man auf natürliche Weise eine koaxiale Speisung.



Figur 117: Monopolantenne mit koaxialer Zuleitung durch die Grundplatte.

Monopolantennen haben typischerweise eine Länge, welche etwa einem Viertel der Wellenlänge entspricht. Durch die Spiegelung an der Grundplatte wird ihre effektive Länge verdoppelt.

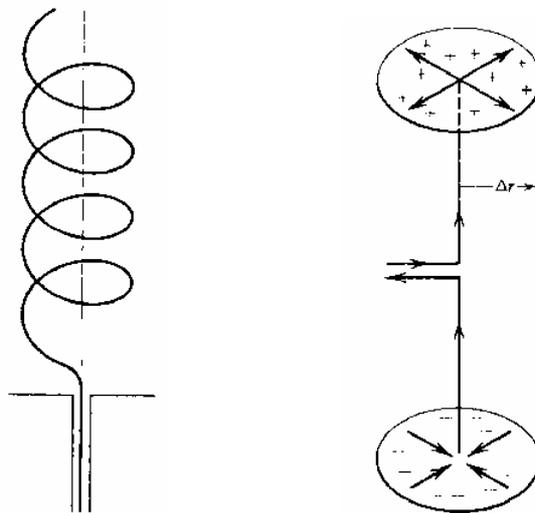
Selbstverständlich kann dasselbe Vorgehen auch auf Loop-Antennen und beliebige, symmetrische Antennenformen angewendet werden.

7.6. Induktivität und Kapazität

Dipolantennen weisen meist eine Länge von etwa einer halben Wellenlänge auf, was bei nicht sehr hohen Frequenzen auf recht unhandliche Antennen führt. Der Grund dafür ist, dass sich bei einer halben Wellenlänge eine Resonanz auf dem Antennenstab ergibt mit einem Strommaximum in der Stabmitte. Dies führt dazu, dass die Antenne relativ gut abstrahlt, d.h. dass der Realteil der Antennenimpedanz dann relativ gross wird. Reduziert man

die Frequenz oder verringert man statt dessen die Stablänge, so nimmt die Abstrahlung stetig ab. Erhöht man umgekehrt die Frequenz (bzw. die Stablänge), so nimmt die Abstrahlung zunächst ebenfalls ab, nimmt dann allerdings wieder zu und erreicht ein weiteres Maximum, wenn die Wellenlänge etwa gleich wie die Stablänge ist, d.h., wenn die zweite Resonanz erreicht wird. Bei höheren Frequenzen (bzw. längeren Stäben) treten weitere Resonanzen auf. Da man die Antennen meist möglichst klein machen möchte, arbeitet man normalerweise bei der ersten Resonanz.

Es stellt sich nun die Frage, ob eine Antenne nicht geometrisch so verändert werden kann, dass die erste Resonanz bei geringeren Abmessungen auftritt. Naheliegender ist es, den Antennendraht einer Dipolantenne spulenförmig aufzuwickeln. Figur 118 zeigt eine solche Helix-Antenne. Derartige Antennen werden gerne in Mobilfunkgeräten verwendet. Ihre Berechnung und Fabrikation ist offensichtlich um einiges aufwendiger als bei Dipolantennen.



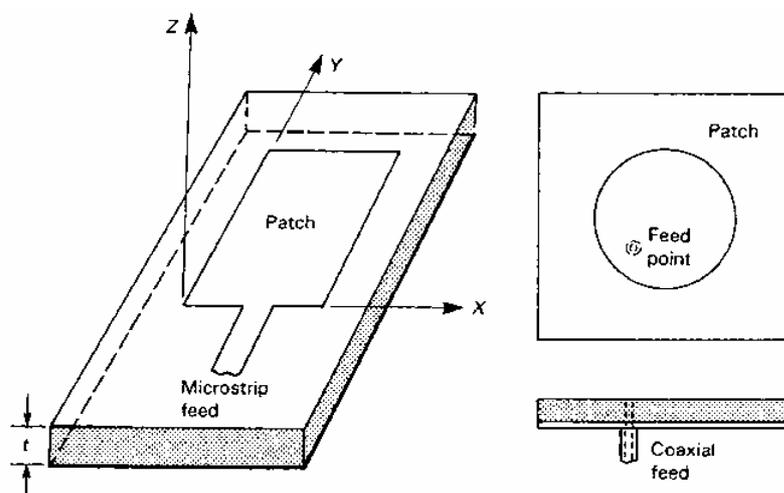
Figur 118: Helix-Antennen mit koaxialer Zuleitung (links) und Dipolantenne mit Dachkapazitäten.

Helix-Antennen erinnern stark an Spulen bzw. Induktivitäten. Ein Alternative zur Verkürzung der Antennenhöhe durch Aufwickeln ergibt sich durch Realisierung einer kleinen Kapazität, wie in Figur 118 angedeutet.

7.7. Patch-Antennen

Die im vorhergehenden Abschnitt skizzierten Antennen mit Dachkapazitäten sind vor allem im Zusammenhang mit gedruckten Schaltungen und Mikrostreifentechnik interessant. Arbeitet man mit einem isolierenden Substrat auf einer Grundplatte, so kann man leicht Leiterflächen (Patch) aufbringen, welche die Rolle der Dachkapazität spielen. Wie Figur 119 zeigt, kann man diese Antennen entweder mit einer Mikrostreifenleitung (im Prinzip

eine Zweidrahtleitung) oder — wie die Monopolantenne — mit einer koaxialen Zuleitung durch die Grundplatte speisen.



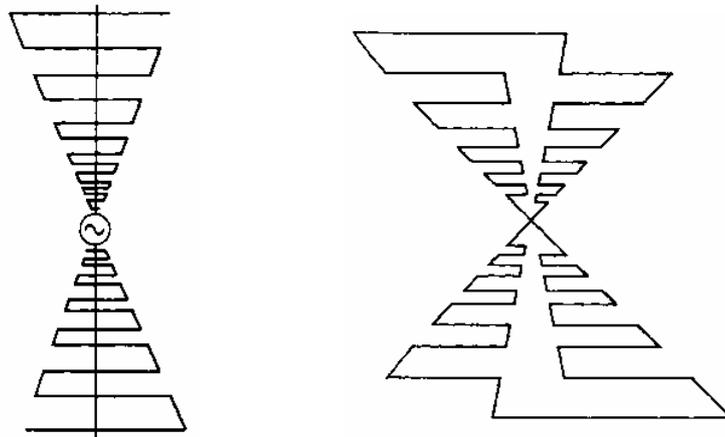
Figur 119: Patch-Antennen mit unterschiedlicher Speisung (rechts: Mikrostreifen, links: Koaxial).

Auch Patch-Antennen werden so betrieben, dass auf dem Patch Resonanzen auftreten. Die technologisch einfache Herstellung verschiedenster Formen bringt einen recht grossen Spielraum und damit verschiedenartigste Formen mit unterschiedlichsten Eigenschaften und Anwendungen.

Die kurz skizzierten Antennen-Arrays lassen sich offensichtlich mit der hier skizzierten Technik gut realisieren. Die Technik der Patch-Antennen lässt sich ausserdem mit verschiedenen Mitteln verfeinern. So kann man mit mehrfach geschichteten Anordnungen arbeiten und dadurch den Spielraum weiter erhöhen.

7.8. Breitbandantennen

Wie erwähnt, nutzen die bisher beschriebenen Antennenformen Resonanzeffekte aus. Diese führen zu relativ hoher Abstrahlung und werden zur Verbesserung der Richtwirkung ausgenutzt. Dadurch entsteht gleichzeitig eine recht starke Frequenzabhängigkeit, welche derjenigen von Bandpassfiltern gleicht. Eine Antenne wird deshalb jeweils nur in einem bestimmten Frequenzband betrieben. Für manche Anwendungen reichen die relativ geringen Bandbreiten einfacher Antennen nicht aus und man sucht deshalb nach komplizierteren Anordnungen mit grösserer Bandbreite. Der Grundgedanke ist dabei die Entwicklung von Formen, welche bei vielen, benachbarten Frequenzen resonante Teile aufweisen. Figur 120 zeigt zwei Beispiele.



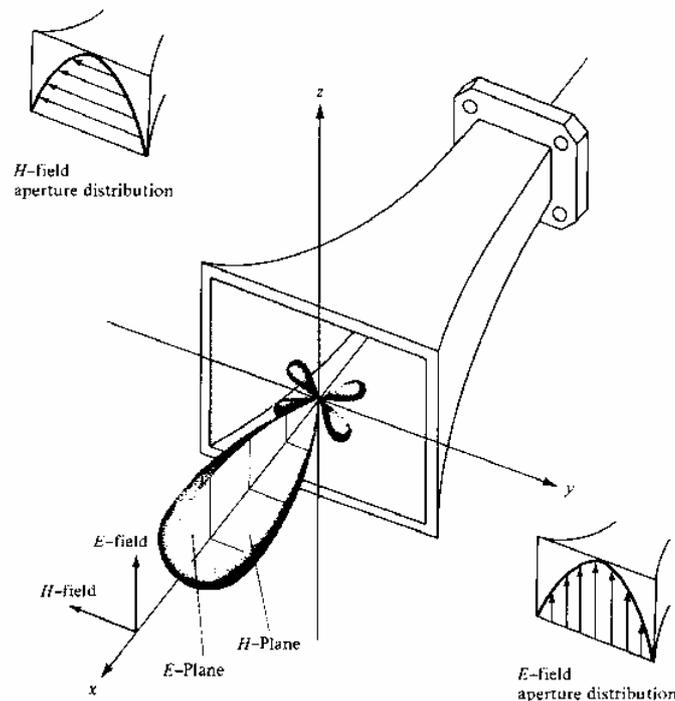
Figur 120: Log-periodische Antennen.

7.9. Andere Antennenformen

Beim Bau von Antennen können verschiedenartigste physikalische Phänomene ausgenutzt werden. Wir haben bereits gesehen, dass eine Grundplatte, aber auch ein einfacher Stab wie Reflektoren für das elektromagnetische Feld wirken. Da es oft wünschenswert ist, dass eine Antenne nur „vorwärts“ strahlt, sind Reflektoren sehr beliebt. Durch passende geometrische Ausformung (parabolisch) können Reflektoren wie Spiegel in der Optik elektromagnetische Wellen auf einen Punkt fokussieren. Da derartige Parabolspiegel gross im Vergleich zur Wellenlänge sein müssen, werden sie vor allem in der Radioastronomie und bei hochfrequenten Signalen verwendet.

Bekanntlich sind Linsen in der Optik eine Alternative zu Parabolspiegeln. Linsen lassen sich ganz ähnlich auch in der Antennentechnik einsetzen, um die Richtcharakteristik einer Antenne zu verändern. Die Möglichkeit, elektromagnetische Wellen an dielektrischen Körpern zu brechen, führt auch zu der Möglichkeit, dielektrische Antennen zu entwickeln, welche durch dielektrische Wellenleiter (in der Optik sind das optische Fibern) gespeist werden. Optische Fibern werden im folgenden Kapitel skizziert, auf eine Beschreibung der in der Mikrowellentechnik verwendeten dielektrischen Wellenleiter verzichten wir hier.

In der Mikrowellentechnik verwendet man bei grossen Leistungen (z.B. in Sendeanlagen) gerne Hohlleiter. Es liegt dann nahe, Antennen zu entwickeln, welche direkt mit einem Hohlleiter verbunden werden können. Durch eine kontinuierliche Aufweitung eines Hohlleiters entstehen auf natürliche Weise Hornantennen, wie in Figur 121 dargestellt. Diese können eine sehr ausgeprägte Richtcharakteristik haben und werden deshalb gerne für Richtstrahlverbindungen und in Messkammern verwendet. Dabei werden die Hornantennen übrigens auch koaxial gespeist.



Figur 121: Horn-Antenne mit Feldverteilung und Strahlungsdiagramm.

7.10. Anwendungen

Obwohl Tesla um 1900 versucht hat, Antennen zur Energieversorgung von Haushalten zu nutzen und obwohl die ersten Radioapparate keine Stromversorgung brauchten und die empfangene Energie ohne jede Verstärkung in akustische Energie umformten, sind Anwendungen, bei denen der Energietransport im Zentrum steht, heute von geringem Interesse. Dazu gehören auch recht fragwürdige Versuche der Krebstherapie durch Bestrahlung mit elektromagnetischen Wellen. Heute wird die Belastung des Organismus durch elektromagnetische Wellen eher gefürchtet. Man versucht daher, die Strahlungsleistungen so weit wie möglich zu reduzieren und lediglich die informationstechnischen Vorteile freier elektromagnetischer Wellen zu nutzen. Die Antennentechnik erlebt daher seit einigen Jahren einen starken Aufschwung.

Zu Beginn des 20. Jahrhunderts wurden Antennen vorwiegend für Rundfunksender und -empfänger verwendet. Dabei ist ein stetiger Trend zu höheren Sendefrequenzen festzustellen. Der Grund dafür liegt in der höheren verfügbaren Bandbreite und dem ständig wachsenden Bedarf an höheren Datenraten, welche nur mit mehr Bandbreite zu realisieren sind. Im Langwellenbereich haben nur wenige Rundfunksender mit schlechter Übertragungsqualität nebeneinander Platz. Man kann deshalb weltweit nur wenige derartige Sender aufstellen. Im Mittelwellenbereich gibt es zwar schon recht viel Platz, weshalb zunächst die meisten „Landessender“ in diesem Bereich arbeiteten. Der schweizerische Landessender

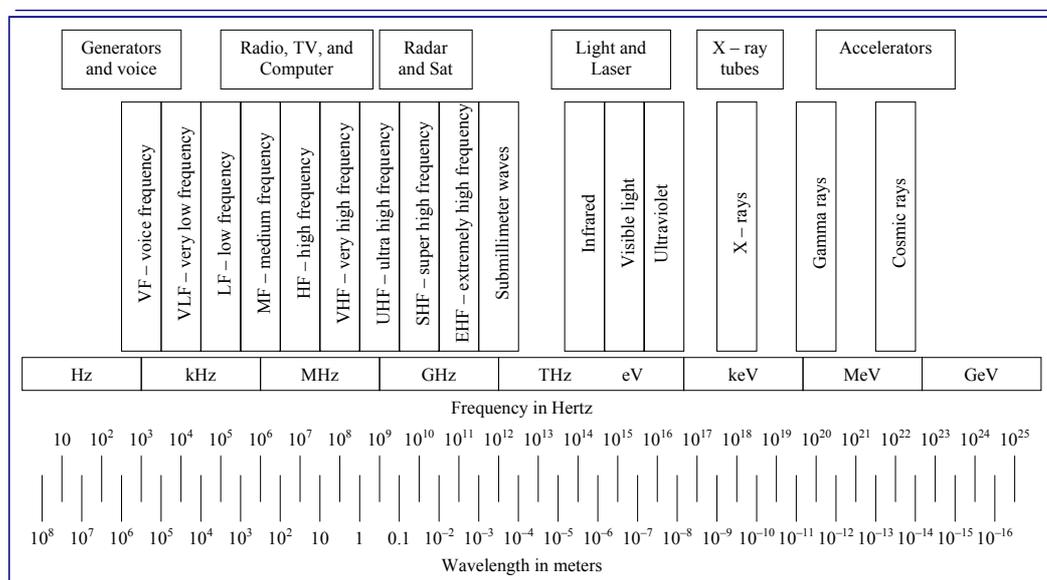
Beromünster wurde aber beispielsweise zunehmend von einem arabischen Sender gestört und deshalb schliesslich eingestellt. Das Kurzwellenband hat zwar schon recht viel Platz und wird deshalb auch von manchen Amateuren genutzt, aber auch einige Eigenheiten (atmosphärische Störungen), welche eine gute Übertragungsqualität erschweren. Die meisten Rundfunksender arbeiten deshalb heute im UKW Bereich. Hier beginnen sich die elektromagnetischen Wellen recht ähnlich wie Licht zu benehmen, d.h. ohne Sichtverbindung wird eine Übertragung mit zunehmender Frequenz schlechter. Damit ist es nicht mehr möglich, mit einem einzelnen Sender mit grosser Reichweite zu arbeiten. Es müssen deshalb über das abzudeckende Empfangsgebiet mehrere Sender aufgestellt werden. Damit sich diese Sender nicht gegenseitig stören, müssen unterschiedliche Frequenzbänder zugeteilt werden. Dasselbe gilt in viel ausgeprägterem Masse beim Mobilfunk, wo mit noch höheren Frequenzen gearbeitet wird.

Im zweiten Weltkrieg wurde mit der Radartechnik eine neue Anwendung für freie elektromagnetische Wellen gefunden: Die Ortung optisch unsichtbarer Objekte. Dabei handelte es sich zunächst um Flugzeuge. Später kamen Anwendungen wie die Suche nach Erdöl, Minen, etc. dazu. Die Möglichkeit, mit Radar nicht nur den Ort, sondern auch die Geschwindigkeit eines Objekts zu messen führte zu Radarfallen, bei denen die Objekte optisch sichtbar sind. Schliesslich führte die Radioastronomie auch zu einer Erweiterung unseres Wissens über das Weltall. Generell ist die Messtechnik neben der Datenübertragungstechnik die wichtigste Anwendung von Antennen.

8. Faseroptische Kommunikation

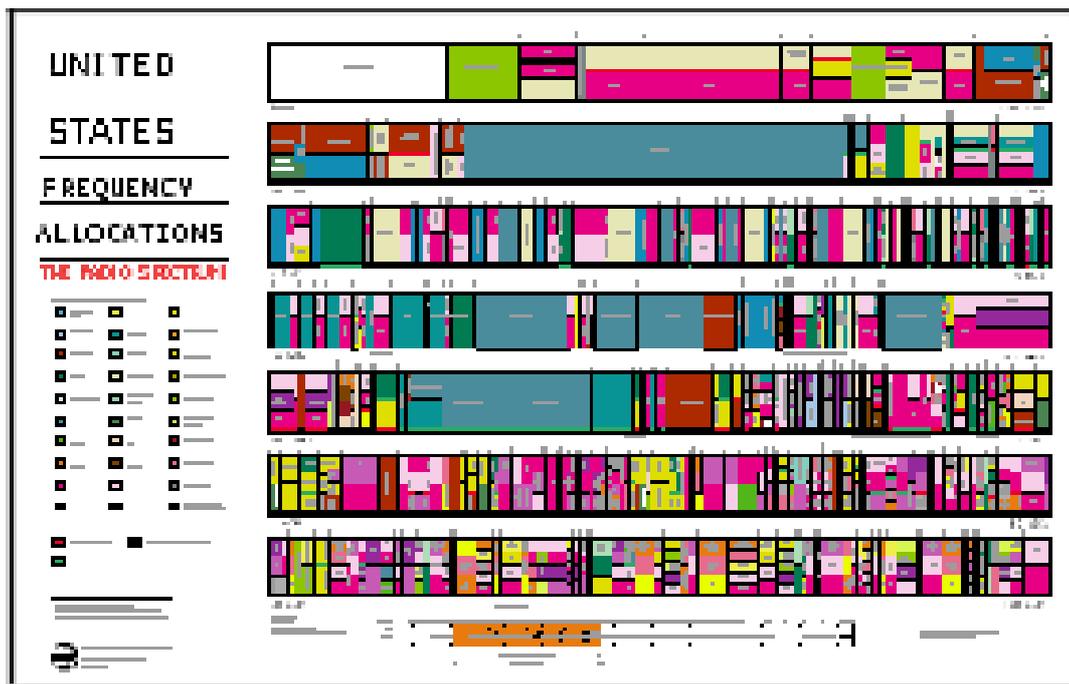
8.1. Das Elektromagnetische Spektrum

Zur Informationsübertragung werden häufig geführte oder sich frei ausbreitende elektromagnetische Wellen verwendet. Allerdings eignet sich dafür praktisch nur ein sehr kleiner Teil des elektromagnetischen Spektrums. Vom gesamten Spektrum (Figur 8.1) werden für die leitungsgebundene Nachrichtenübertragung die Frequenzen bis etwa 1 GHz verwendet, für Funkübertragung Frequenzen von 100 kHz bis 100 GHz, und für die optische Nachrichtenübertragung werden einige wenige Wellenlängen im sichtbaren Bereich des Lichts und im nahen Infrarot benutzt.



Figur 122: Spektrum elektromagnetischer Wellen.

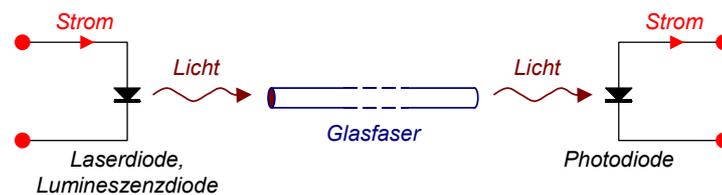
Die Frequenzaufteilung im Radiobereich ist zur Vermeidung von Konflikten streng reglementiert. Die Aufteilung in den USA zeigt Figur 123 (oder übersichtlicher im Grossformat: <http://www.ntia.doc.gov/osmhome/allochrt.pdf>). Die Radiofrequenzen sind heute sehr stark benutzt und es ist kompliziert, Frequenzbänder für neue Dienste zuzuweisen. Die im folgenden Kapitel erläuterte optische Nachrichtenübertragung kann einige Übertragungsbandbreite-Engpässe beheben helfen.



Figur 123 Aufteilung der Radiofrequenzen in den USA.

8.2. Optische Nachrichtenübertragung

Die Figur 1243 zeigt das Prinzip eines optischen Nachrichtensystems. Ein Halbleiterbauelement (Laserdiode oder Lumineszenzdiode) wandelt einen modulierten elektrischen Strom in ein moduliertes Lichtsignal um. Das Licht wird mittels eines dielektrischen Wellenleiters (meistens eine Glasfaser) zum Empfänger übertragen und von diesem wieder in ein moduliertes elektrisches Signal umgewandelt.



Figur 124 Prinzip eines optischen Nachrichtensystems.

Die optische Signalübertragung kann prinzipiell auch durch den freien Raum erfolgen, also ohne die Verwendung einer Glasfaser. Nach diesem Prinzip arbeiten Leuchtturm, Infrarotfernbedienung und bald auch Verbindungen zwischen Satelliten. Im Folgenden wird aber nur auf die optische Übertragung durch Fasern eingegangen.

Der zur Übertragung des Lichtsignals verwendete dielektrische Wellenleiter führt das Licht durch Totalreflexion. Er muss im verwendeten Wellenlängenbereich eine möglichst geringe Dämpfung aufweisen. Während für kurze zu überbrückende Distanzen (etwa einige Millimeter in optischen Computern) eine Vielzahl optisch mehr oder weniger transparenter Materialien zur Verfügung stehen, können für mittlere Distanzen (einige Dutzend Meter) nur Fasern aus Glas oder Kunststoff verwendet werden. Zur Überbrückung grosser Entfernungen (ab etwa 1 km) werden ausschliesslich Fasern aus hochreinem Glas verwendet.

Optische Fasern zeichnen sich im Vergleich mit Kupferleitungen und Radiotransmission aus durch

- grosse Bandbreite
- geringes Gewicht
- Stör- und Abhörsicher, keine Abstrahlung
- Grundmaterialien billiger als Kupfer
- keine Radio-Sendelizenzen vonnöten

Die theoretisch wie praktisch (beinahe) unbegrenzte Bandbreite der Lichtwellenleiter führt dazu, dass, mit Ausnahme der "letzten Meile", künftig fast alle Daten über optische Datenkanäle übertragen werden.

Die zu übertragenden Datenmengen sind durch die Verbreitung des Internet rapide gewachsen. Ein Ende dieser Entwicklung ist nicht absehbar. Ein Beispiel: Der erste komplett digital produzierte Kinofilm, "Star Wars Episode 1", benötigt komprimiert ein Speichervolumen von gut 300 GByte. Heute läuft jeder Fernseher mehrere Stunden täglich, und wer hätte nicht gerne Kinoqualität im Wohnzimmer...

8.2.1. Lichtwellenleiter

8.2.1.1. Prinzip

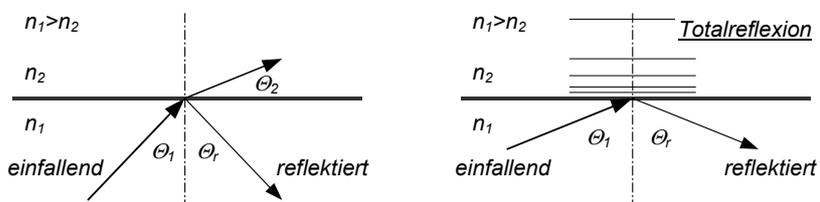
Optische Materialien werden in der Regel durch ihren Brechungsindex n charakterisiert, $n = c/v$, wobei c die Vakuumlichtgeschwindigkeit ist und v bezeichnet die Geschwindigkeit des Lichts im Medium. Beim Übergang eines Lichtstrahles vom Medium 1 (n_1) ins Medium 2 (n_2) kommt es zur Brechung (und zur Reflexion). Es gilt das Brechungsgesetz nach Snellius:

$$n_1 \sin \theta_1 = n_2 \sin \theta_2$$

Wenn ein Lichtstrahl auf die Grenzfläche zu einem optisch dünneren Medium trifft ($n_1 > n_2$), dann kommt es zur Totalreflexion, wenn

$$\theta_1 > \theta_c \quad \text{und} \quad \sin \theta_c = \frac{n_2}{n_1}$$

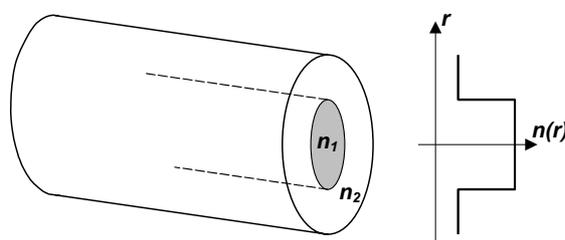
Die Figur 125 veranschaulicht diesen Effekt am Beispiel einer ebenen Grenzfläche. Es ist zu berücksichtigen, dass auch bei Totalreflexion Felder im Medium 2 existieren (diese klingen aber in Richtung "weg von der Grenzfläche" schnell ab). Bei optischen Wellenleitern erfolgt die Führung des Lichtes durch fortgesetzte Totalreflexion.



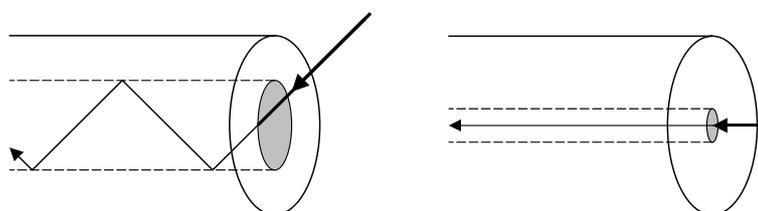
Figur 125: Reflexion am Übergang von Medium 1 (n_1) zum Medium 2 (n_2).

8.2.1.2. Geometrie

Einen typischen Faserquerschnitt zeigt die Figur 126. Der Faserkern (*engl.*: core) hat einen etwas höheren Brechungsindex als der Fasermantel (*engl.*: cladding). Zum Schutz der empfindlichen Faser wird noch eine Kunststoffummantelung hinzugefügt. Man unterteilt Fasern in Monomodefasern und Multimodefasern. Der Durchmesser des Faserkerns der Monomodefaser ist so klein, dass sich gewissermassen nur ein einziger Lichtstrahl in Achsenrichtung ausbreiten kann. Der Kerndurchmesser der Multimodefaser hingegen ist relativ gross, und das Licht kann eine Vielzahl unterschiedlicher "zick-zack-Pfade" zur Ausbreitung nutzen. Die Figur 127 veranschaulicht diesen Sachverhalt.



Figur 126; Faserquerschnitt und Brechzahlprofil einer typischen Stufenprofilfaser.



	Multimodefaser	Monomodefaser
Kerndurchmesser	0.050 ... 0.100 mm	0.008 mm
Aussendurchmesser	0.125 ... 0.200 mm	0.125 mm
Brechzahl Differenz	1% ... 2%	0.1% ... 0.2%

Figur 127: Vergleich von Multimodefaser (links) und Monomodefaser (rechts) sowie typische Parameter beider Fasertypen.

8.2.1.3. Moden

Die Vorstellung, dass sich ein Lichtstrahl in der Faser im Zickzack ausbreitet, stellt eine Vereinfachung dar. Eigentlich muss das Licht, aufgefasst als elektromagnetische Welle, die Maxwell-Gleichungen in der Faser und die entsprechenden Randbedingungen erfüllen. In der optischen Faser gibt es dann entweder eine Feldlösung für diese Gleichungen oder aber (bei grösserem Faserquerschnitt) mehrere Lösungen. Allen diesen Feldlösungen (genannt "Moden") ist gemeinsam, dass sich das Feld nicht konstant, aber stetig über den gesamten Faserquerschnitt verteilt, was klar im Gegensatz zum Konzept des Lichtstrahls steht.

Der Mode und damit die Feldverteilung in einer Monomodefaser lässt sich relativ einfach berechnen. Wenn sich aber in einer Multimodefaser hunderte Moden ausbreiten, dann lassen sich hier die Verhältnisse besser mit der sog. *V-Zahl* beschreiben. Der *V*-Parameter ist definiert durch

$$V = \frac{\pi d}{\lambda} \sqrt{n_1^2 - n_2^2}$$

wobei d der Kerndurchmesser ist und λ die Freiraumwellenlänge bezeichnet. Für $V < 2.405$ kann sich nur ein Mode ausbreiten (eigentlich zwei Moden mit orthogonaler Polarisierung) und es handelt sich um eine Monomodefaser (praktisch muss gelten $2 < V < 2.405$).

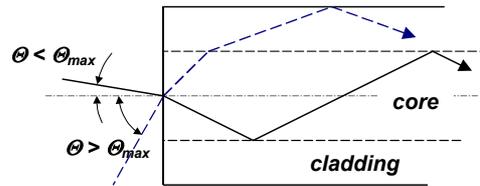
Für eine Multimodefaser gilt $V \gg 2.405$, und die Anzahl der Moden ist etwa $N \approx V^2/2$. Es befindet sich immer auch Feldenergie im Fasermantel. Im Allgemeinen gilt für Multimodefasern, dass Feldenergie im Fasermantel überproportionale Verluste verursacht. Es gilt

$$\frac{P_{cladding}}{P_{core}} \approx \frac{4}{3\sqrt{N}}, \quad (V \text{ gross})$$

8.2.1.4. Apertur

Die Führung des Lichtes erfolgt durch fortgesetzte Totalreflexion. Weil aber der Brechzahlunterschied n_1 zu n_2 oft sehr klein ist, werden nur Strahlen geführt, die ungefähr in Richtung der Faserachse zeigen. Das ist bei der Einkopplung von Licht in die Faser von Bedeutung: Die Lichtquelle muss möglichst paralleles Licht möglichst hoher Intensität liefern (selbst eine Fokussierung mit Linse oder Spiegel verbessert nicht die Parallelität der Lichtstrahlen!). Der Einfallswinkel eines Lichtstrahles sei Θ (Figur 128). Die Totalreflexion findet nur für Winkel $\Theta < \Theta_{max}$ statt, anderenfalls ($\Theta > \Theta_{max}$) wird das Licht in den Fasermantel eingekoppelt und sehr schnell absorbiert. Dieses Verhalten beschreibt der faserspezifische Parameter der *Numerischen Apertur (NA)*

$$NA = \sin \Theta_{max} = \sqrt{n_1^2 - n_2^2}$$



Figur 128: Einkopplung eines Lichtstrahls in eine Stufenindexfaser.

Eine Monomodefaser mit $n_1 = 1.45$, $d = 10 \mu\text{m}$, und einer Brechzahldifferenz von 0.25% hat $n_2 = 1.4464$. Der Grenzwinkel θ_c der Totalreflexion (relativ zur Achse) beträgt 4.05° . Dies und die numerische Apertur widerspiegeln die Tatsache, dass nur beinahe achsenparallele Strahlen in die Faser einkoppeln können, $NA = 0.10$. Sowohl der Grenzwinkel θ_c als auch die numerische Apertur sind unabhängig von der Querschnittsgeometrie der Faser.

Im Vergleich dazu habe eine Multimodefaser mit $n_1 = 1.45$, $d = 62.5 \mu\text{m}$, und einer Brechzahldifferenz von 1.5% hat $n_2 = 1.4282$. Der Grenzwinkel θ_c der Totalreflexion (relativ zur Achse) beträgt 9.94° . Dies und die numerische Apertur widerspiegeln die Tatsache, dass nur mehr oder weniger achsenparallele Strahlen in die Faser einkoppeln können, $NA = 0.25$.

Für die genannte Monomodefaser reicht der zulässige Wellenlängenbereich von $\lambda = 1600 \text{ nm}$ ($V = 2$) bis $\lambda = 1330 \text{ nm}$ ($V = 2.405$). Einer Wellenlänge von $\lambda = 1530 \text{ nm}$ entspricht dann $V = 2.1$.

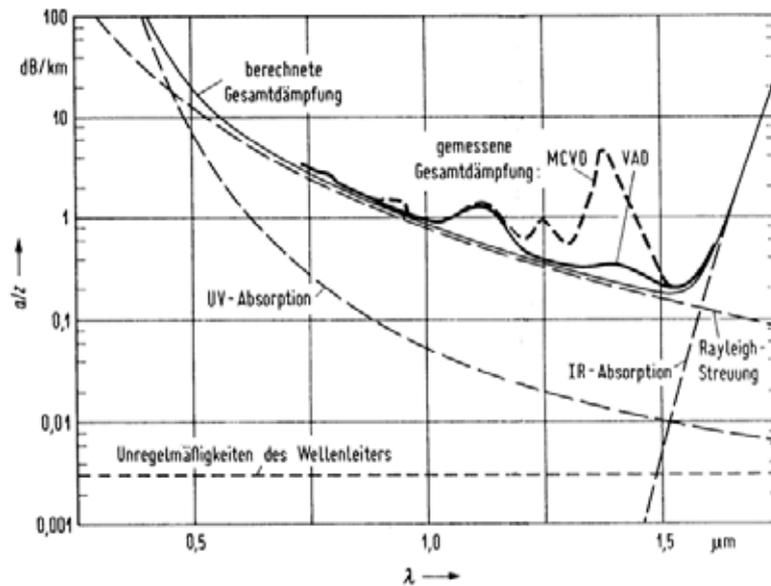
Die erwähnte Multimodefaser führt bei $\lambda = 1300 \text{ nm}$ etwa $N = 2400$ Moden ($V = 69$). Dabei befindet sich etwa 2.7% der Lichtleistung im Fasermantel.

8.2.1.5. Dämpfung

Die Leistung einer Lichtwelle in der Glasfaser nimmt bei der Ausbreitung exponentiell ab. Man definiert einen Dämpfungskoeffizienten durch

$$\alpha = -\frac{10}{\Delta z} \log_{10} \frac{P(z_0 + \Delta z)}{P(z_0)}$$

Verschiedene physikalische Effekte sind für die Verluste verantwortlich. Es gibt Materialabsorption, Absorption durch Verunreinigungen, Streuung aufgrund inhomogener Geometrie, Abstrahlung von engen Kurven. Die Materialabsorption und die Absorption durch einige Verunreinigungen (insbesondere durch OH^-) sind unvermeidlich. Letztere sind besonders stark wellenlängenabhängig. Den spektralen Verlauf der Dämpfung von Quarzglas- (SiO_2) Multimodefasern zeigt Figur 129. Wichtig ist das Dämpfungsminimum bei $\lambda = 1550 \text{ nm}$ und die im Vergleich dazu wesentlich höhere Absorption bei 1300 nm und 850 nm . Die beiden letztgenannten Wellenlängen sind unter anderen Gesichtspunkten technisch vorteilhaft.



Figur 129: Dämpfung einer Monomodefaser (Quarzglas, $(n_1 - n_2)/n_2 = 0.25\%$). VAD und MCVD bezeichnen verschiedene Herstellungsverfahren.

Es ist nicht damit getan, ein Material zu finden, welches eine optische Welle wenig dämpft. Weitere Kriterien sind die preiswerte Herstellung, die Verfügbarkeit von Lichtquellen und Empfängern bei der interessanten Wellenlänge, die mechanische Handhabbarkeit und die Verwendung einigermaßen umweltverträglicher Materialien. Während herkömmliche Glasfasern minimale Verluste von 0.2 dB/km erreichen (siehe Fig. 129), hat man an sog. Halogengläsern (etwa BeF_2) Verluste von 0.001 dB/km bei $\lambda = 3.44 \mu\text{m}$ gemessen. Leider gibt es für diese Wellenlänge keine Halbleiterlaser und keine effizienten Empfänger. Zudem ist Beryllium stark giftig und das Material zieht Wasser stark an (was sehr hohe Verluste zur Folge hätte). Eine andere Möglichkeit wäre die Verwendung von Einkristallen (AgBr , KCl), aber die Herstellung der Kristalle in Faserform erscheint utopisch...

8.2.1.6. Dispersion

Die Handhabung (Verbindungen, Lichtein- und Auskopplung) der Monomodefasern ist ungleich komplizierter als jene der Multimodefasern. Der entscheidende Vorteil der Monomodefasern liegt aber in den wesentlich höheren Datenübertragungsraten für eine gegebene Übertragungsdistanz. Die entsprechende Kennziffer, das Produkt von Datenrate und Distanz, liegt für eine typische Monomodefaser bei $1 \dots 10 \text{ Gbit s}^{-1} \text{ km}$, für eine Multimodefaser um den Faktor 100 tiefer. Der limitierende physikalische Effekt wird Dispersion genannt und bewirkt, dass sich die unterschiedlichen Energieanteile eines "Bits" mit unterschiedlicher Geschwindigkeit ausbreiten. Nach einer gewissen Distanz verwischen aufeinander folgende Bits und die Fehlerrate der Übertragungstrecke (*engl.*: BER, *bit error rate*) wächst stark an.

Die Übertragung in optischen Fasern unterliegt unterschiedlichen Arten von Dispersion. Die *Materialdispersion* entsteht, weil sich die Brechzahl des Glases für verschiedene Wel-

lenlängen des Lichts etwas ändert. Der Einfluss der Materialdispersion ist für Gläser auf Basis SiO_2 gerade null bei Wellenlängen um 1300 nm. In der Monomodefaser kommt hier eine weitere Art der Dispersion zum Tragen, die *Wellenleiterdispersion*. Sie beruht darauf, dass sich mit kleiner werdender Wellenlänge die Lichtenergie zunehmend im Faserkern konzentriert. Weil die Brechzahl des Kerns grösser als jene des Mantels ist, verlangsamt sich die Welle. Aufgrund des geringen Brechzahlunterschiedes ist dieser Effekt klein. Schliesslich existiert in Multimodefasern die sogenannte *Modendispersion*, die viel grösser als die beiden anderen Arten der Dispersion ist. Das Licht wird innerhalb der Multimodefaser mit einer Vielzahl von Moden übertragen, die stark variierende Ausbreitungsgeschwindigkeiten aufweisen. Das folgt anschaulich daraus, dass die Zickzack-Strahlen unterschiedliche Steigungen aufweisen können. Die Modendispersion bewirkt das viel kleinere Datenrate-Distanz-Produkt der Multimodefaser.

Eine typische Multimodefaser hat bei $\lambda = 850$ nm eine Modendispersion von 70 ns km^{-1} , während die Materialdispersion $-86 \text{ ps km}^{-1} \text{ nm}^{-1}$ beträgt. Selbst wenn eine LED (optische Bandbreite 50nm) verwendet wird, so ist die Modendispersion dominant. Eine Datenrate von 3.84 MBit pro Sekunde (Impulslänge 260 ns) sollte dann nicht über mehr als 1...2 km übertragen werden. Die Wellenleiterdispersion beträgt für eine typische Monomodefaser bei 1300 nm beispielsweise $-4 \text{ ps km}^{-1} \text{ nm}^{-1}$.

8.2.1.7. Lichtquellen

Als Lichtquellen kommen praktisch ausschliesslich LED (Leuchtdioden) und LD (Laserdioden) in Betracht. Die Umwandlung elektrischer Energie in Licht beruht bei beiden Bauteilen auf demselben Prinzip: In einem geeigneten Halbleiter werden durch ein elektrisches Feld die gebundenen Elektronen in einen angeregten Zustand gebracht, von dem aus sie unter Emission eines Lichtquants wieder in den Ausgangszustand zurückkehren.

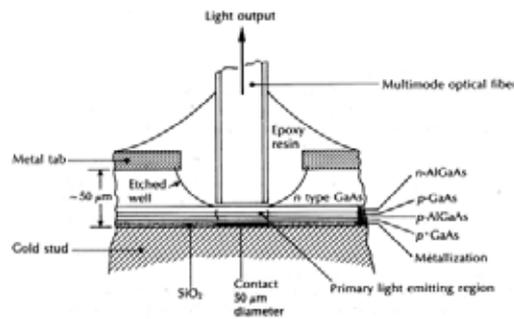
Die Rückkehr in den Ausgangszustand erfolgt spontan nach einer gewissen, materialspezifischen Zeitdauer, sie kann aber auch erzwungen werden, indem man eine Lichtwelle der zu erwartenden Wellenlänge "vorbeischiebt". Diese stimulierte Emission stimmt in Wellenlänge (logischerweise) sowie in Phasenlage und räumlicher Richtung (erstaunlicherweise) mit der anregenden Welle überein, und zwar umso besser, je stärker die anregende Welle ist.

Wenn man erreicht, dass die Mehrheit aller Atome im angeregten Zustand ist (etwa dadurch, dass erstens ein starkes Feld ständig viele Atome anregt und zweites die materialbedingte mittlere Verweildauer im angeregten Zustand ausreichend gross ist), und sich eine starke anregende Welle im Material befindet (Resonatorstruktur nötig), dann hat man einen Laser (*engl.: Light amplification by stimulated emission radiation*). Ein Laser emittiert monochromatisches, kohärentes, nahezu paralleles Licht hoher Intensität. Die Figur 8.9 führt verschiedene Halbleitermaterialien für LED und LD auf.

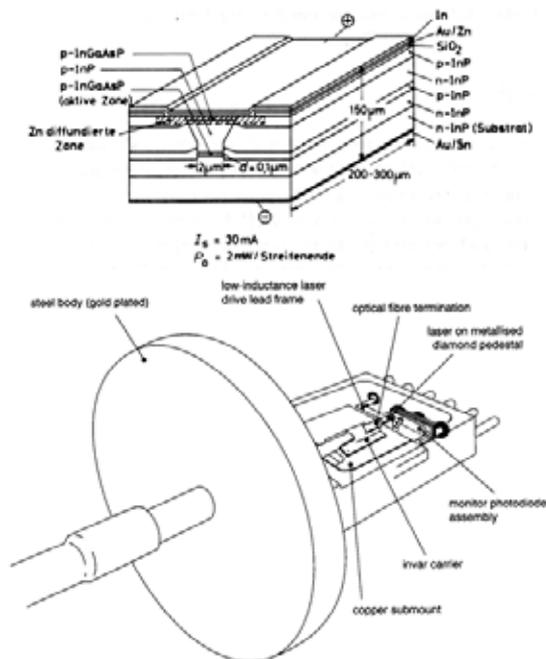
<i>Material</i>	<i>Wellenlänge (nm)</i>	<i>Bereich</i>
GaP	665	Rot
GaAlAs	800 ... 930	Nahes Infrarot
InGaAsP	1000 ... 1700	Infrarot

Figur 130: Spektrale Charakteristik von LED/LD-Halbleitern.

Das lichterzeugende Volumen ist bei der LD extrem klein und der Strom recht gross, entsprechend hoch sind die Anforderungen zur Wärmeableitung (eventuell Temperaturstabilisierung und/oder aktive Kühlung nötig). Weil das Licht der LD ist aber sehr konzentriert und parallel ist, eignet es sich besonders gut zur Einkopplung in Monomodefasern. Die dabei einzuhaltenden mechanischen Toleranzen sind in der Grössenordnung $< 0.1\mu\text{m}$. Die LED strahlt hingegen divergierendes Licht von einer recht grossen Fläche ab. Die LED ist mithin nur zur Speisung von Multimodefasern geeignet. Die Figur 131 zeigt beispielhaft den Aufbau einer LED und die Kopplung LED-Multimodefaser. Die Figur 132 zeigt den Aufbau einer LD und die Kopplung LD-Monomodefaser.



Figur 131: Aufbau einer LED und mechanische Ausführung der Kopplung LED-Multimodefaser.

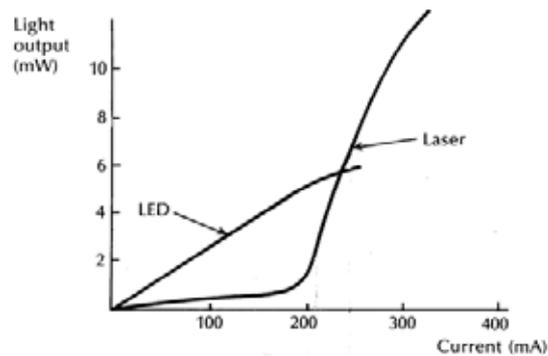


Figur 132: Aufbau einer LD und mechanische Ausführung der Kopplung LD-Monomodefaser.

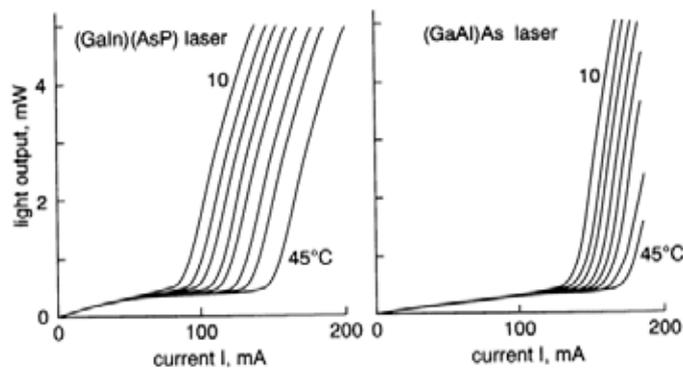
LED und LD sind Sperrschichtbauelemente, sie benötigen daher eine gewisse Spannung, damit ein Strom fließen kann. Diese Spannung ist materialabhängig zwischen 1.3V (GaAs) und 5V (GaN). Oberhalb der jeweiligen Flussspannung steigt der Strom exponentiell an. Die Lichtleistung der LED steigt dann ebenfalls schnell an, während die LD bis zu einem bestimmten Schwellenstrom kaum Licht emittiert. Das ist klar, wenn man berücksichtigt, dass ein gewisser Strom nötig ist, um möglichst viele Atome in den angeregten Zustand zu bringen (Besetzungsinversion). Erst oberhalb des Schwellstromes beginnt die LD zu "lasern". Die Figur 134 zeigt diese Kennlinien.

Weil alle elektronischen Vorgänge in Halbleitern stark temperaturabhängig sind, sind auch die Kennlinien von LED und LD temperaturabhängig (siehe Figur 134).

Aufgrund der hohen Stromdichte in der aktiven Schicht einer LD verändern sich die Eigenschaften der LD mit der Zeit. Es ist daher üblich, in das Gehäuse der LD gleich noch einen Photodetektor einzubauen, der es erlaubt, mittels eines Regelkreises alterungsbedingte Veränderungen zu kompensieren.



Figur 133: Typische Kennlinien von LED und LD.



Figur 134: Typische Temperaturabhängigkeit der Lichtleistung einer LD.

8.2.1.8. Verstärker

Bevor das Lichtsignal in der Glasfaser so schwach wird, dass es im Rauschen verschwindet, muss es verstärkt werden. Dazu verwendet man sog. *Repeater* oder optische Verstärker. Ein *Repeater* besteht aus einem Detektor, der (elektrischen) Signalformung und einem Sendemodul. Der optische Verstärker hingegen vergrößert die Signalamplitude auf optischem Weg.

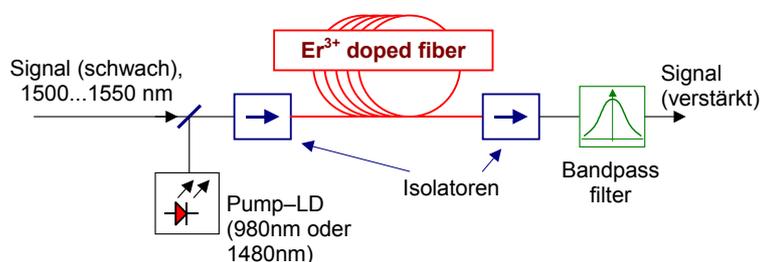
Beim optischen Verstärker unterscheidet man Halbleiterverstärker und Faserverstärker. Beide funktionieren praktisch wie Laser, aber mit dem Unterschied, dass die Lichtwelle im aktiven Medium nicht resonant ist. Wie beim Laser werden durch Zuführung externer Energie die Mehrzahl der Atome in den angeregten Zustand gebracht. Jetzt läuft das schwache, zu verstärkende Lichtsignal durch diese aktive Zone und induziert dabei die Emission von Lichtquanten durch die angeregten Atome. Weil diese stimulierte Emission in Frequenz, Phase und räumlicher Richtung mit der anregenden Welle übereinstimmt, wird das anregende Lichtsignal verstärkt.

Im Halbleiterverstärker wird die Besetzungsinversion (wie in der LD) durch einen hohen elektrischen Strom erreicht. Die aktive Zone ist sehr klein und es entstehen Probleme mit

der verlust- und reflexionsarmen Ankopplung der Glasfasern (wie in der LD). Derzeit haben diese Verstärker keine praktische Bedeutung.

Im Faserverstärker erfolgt die Anregung der potentiell aktiven Atome mit Laserlicht kürzerer Wellenlänge, also höherer Energie, mit dem sog. Pumplaser. Die so angeregten Teilchen geben schnell etwas Energie wieder ab (als Phonon, also Wärme) und bleiben in einem zweiten angeregten Zustand "hängen", wo sie auf das zu verstärkende schwache Lichtsignal warten. Praktisch basiert dieser Verstärker auf einer geeignet dotierten Glasfaser von einigen zehn Metern Länge. Diese aktive Faser kann relativ einfach mit anderen Fasern verbunden werden. Zur Dotierung wird Erbium (für Signalwellenlänge 1500 nm) und Praseodym (für Signalwellenlänge 1300 nm) verwendet. Sogenannte EDFA (*Erbium Doped Fiber Amplifiers*) sind Stand der Technik und werden breit angewendet. Die Figur 135 zeigt das Funktionsschema eines EDFA.

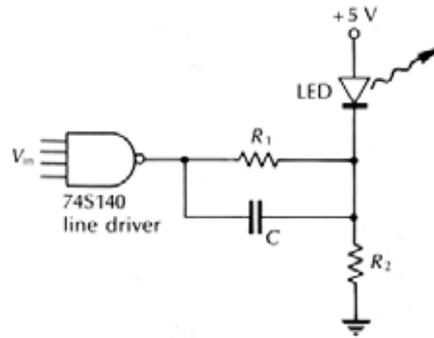
Ein EDFA kompensiert durch Verstärkung auf einigen Metern Länge die Verluste von einigen Kilometern Glasfaser. Solche Verstärker können kaskadiert werden. Weil aber jeder Verstärker dem Signal etwas Rauschen hinzufügt und das bereits vorhandene Rauschen gleichfalls verstärkt wird, ist die maximal überbrückbare Distanz limitiert. Dazu kommt die Dispersion als distanzlimitierende Grösse. Beide Probleme könnten mit einer auf die gesamte Faserlänge gleichmässig verteilten Verstärkung optimal gelöst werden. Diese Verstärkung wird durch nichtlineare Effekte bei hohen Feldstärken erreicht (Raman-Verstärker). Die Dispersion kann in solchen nichtlinearen Fasern praktisch verschwinden. Diese sog. Solitonen-Übertragung ist allerdings noch im Forschungsstadium. Sie verspricht sehr hohe Datenraten (etwa 100 Gb/s) über sehr grosse Distanzen.



Figur 135: Blockschaltbild eines optischen Faserverstärkers (EDFA).

8.2.1.9. Modulatoren

Damit Information übertragen werden kann, muss das Lichtsignal moduliert werden. Am einfachsten geschieht das, indem die Lichtquelle (LED, LD) ein- und ausgeschaltet wird (Figur 136). Dieses Modulationsverfahren nennt man "on-off-keying" ("OOK").



Figur 136: Einfachste Ansteuerung einer LED.

Bei hohen Datenraten ist dieser Vorgang insbesondere bei LD nicht mehr trivial, weil

- starke Ströme schnell geschaltet werden müssen,
- negative Spannungsspitzen die LD zerstören würden,
- thermische Stabilität garantiert werden muss,
- die optische Ausgangsleistung von Temperatur und Alterung abhängig ist

Die LD wird daher vom Hersteller meist mit einer integrierten Photodiode geliefert, damit die optische Ausgangsleistung überwacht werden kann. Die Umwandlung eines Standard-Digitalsignals (z.B. ECL, ggf. parallel) in den LD-Strom übernehmen spezielle Treiberschaltkreise, die gleichzeitig die Kontrollphotodiode und die Temperatursensorik überwachen.

Die OOK-Modulation einer LED erlaubt nur Datenraten bis zu 100 Mb/s, mit der LD erreicht man problemlos 625 Mb/s. Bei höheren Datenraten kommt man aber schnell an die physikalischen Grenzen,

- weil es einige Zeit dauert, bis nach Anlegen der Spannung an die LD die Besetzungsinversion aufgebaut ist und die LD endlich Licht emittiert,
- weil es wieder eine Weile braucht, bis das Licht nach Stromabschaltung wirklich "aus" ist (Abbau der Besetzungsinversion),
- weil am Beginn des Laserprozesses die Wellenlänge noch nicht stabil ist (Dispersion!).

Diesen Problemen kann man ausweichen, wenn die LD nicht ganz ausgeschaltet wird, also ständig mehr oder weniger Licht emittiert.

Eine weitere Methode besteht darin, die LD völlig kontinuierlich strahlen zu lassen und dann diesen Lichtstrahl zu modulieren. Das kann über einen steuerbaren Absorber oder einen Interferenzmodulator geschehen. Der Vorteil der externen Modulation ist, dass die LD stets optimal betrieben werden kann. Prinzipiell ist die externe Modulation sehr schnell (>50 Gb/s). Ein Interferenzmodulator (sog. *MZI*, *Mach-Zehnder-Interferometer*) spaltet das Lichtsignal in zwei Pfade der jeweils halben Leistung auf, verzögert einen Pfad gegenüber dem anderen um eine halbe Wellenlänge (oder nicht), und kombiniert beide Signale wie-

der. Sowohl Absorptionsmodulator als auch *MZI* basieren auf Materialien, deren optische Eigenschaften (Absorption, Brechungsindex) sich durch elektrische Felder beeinflussen lassen. Solche Materialien sind beispielsweise InP (ändert Absorption mit angelegter elektrischer Feldstärke) und LiNbO₃ (ändert Brechungsindex). Weil die Änderung der optischen Parameter aber relativ gering ist, sind derartige Modulatoren recht gross (10...20 mm lang) und damit auch teuer.

Ein Modulator, der eine kontinuierliche Veränderung des Brechungsindex erlaubt, ermöglicht die kontinuierliche Veränderung der Phase des Lichtsignals bei konstanter Amplitude. Damit sind auch bei der optischen Übertragung Phasen- und Frequenzmodulationsverfahren möglich. Diese sog. kohärenten Übertragungsverfahren (man unterscheidet je nach Empfangsmethode noch heterodyne und homodyne Verfahren) zeichnen sich durch verbesserte Wellenlängenselektion und höhere Empfindlichkeit des Empfängers aus.

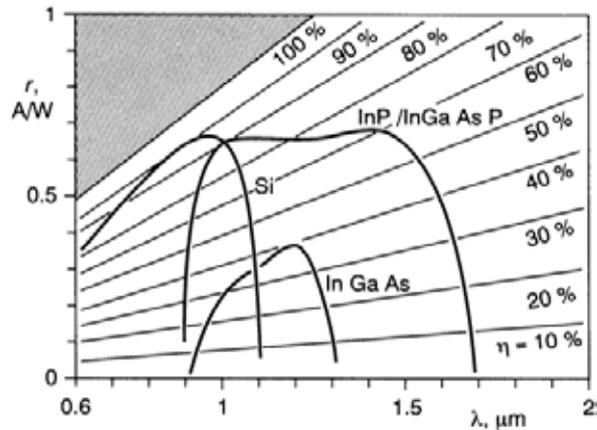
8.2.1.10. Empfänger

Am Ende der optischen Übertragungsstrecke muss das Lichtsignal wieder in ein elektrisches Signal umgewandelt werden. Dafür kommen fast ausschliesslich Photodioden (PD) zum Einsatz, die sich insbesondere durch hohe Geschwindigkeit auszeichnen. Die Photodiode bildet mit der Fasereinkopplung und dem nachgeordneten elektrischen Verstärker eine funktionale und häufig auch bauliche Einheit. Die Optimierung des Empfangsmoduls muss hinsichtlich Wellenlänge, Datenrate, Empfindlichkeit, Rauschen und Kosten erfolgen.

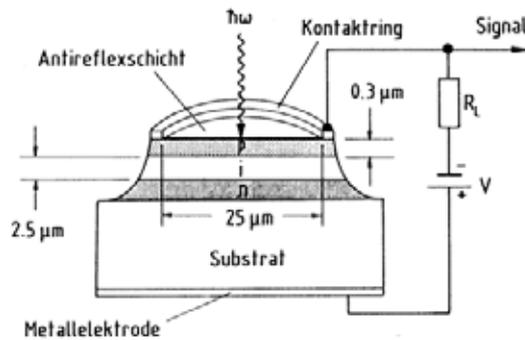
Für verschiedene Materialien zeigt die Figur 137 die spektrale Empfindlichkeit und die Effizienz. Eine Effizienz von 90% etwa bedeutet, dass nur jedes zehnte Photon kein Elektron-Loch-Paar generiert (vorausgesetzt, der absorbierende Bereich ist ausreichend dick).

Die Figur 138 zeigt den Aufbau einer Photodiode. Die lichtabsorbierende Schicht, hier *i*-Schicht genannt, muss möglichst dick sein, um das einfallende Licht möglichst komplett zu absorbieren (hohe Empfindlichkeit). Sie muss andererseits möglichst dünn sein, um optisch generierte Ladungsträger schnell als Strom in der nachfolgenden Elektronik wirksam werden zu lassen (hohe Datenrate).

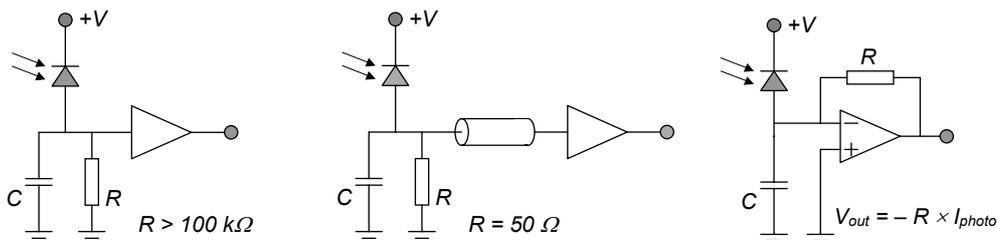
Ein ähnlicher Gegensatz zwischen Empfindlichkeit und Geschwindigkeit gilt für die erste (und entscheidende) Verstärkerstufe. In der Anordnung von Figur 139a erzeugt der Photostrom I eine grosse Spannung $R \cdot I$ am Verstärkereingang. Die Zeitkonstante $R \cdot C$ dieser Schaltung ist jedoch sehr gross und die Datenrate somit gering. Diese Schaltung wird beispielsweise in der Astronomie verwendet, wo es auf hohe Empfindlichkeit ankommt. Die Schaltung nach Figur 139b hingegen ist sehr schnell (RC klein), aber relativ unempfindlich. Einen Kompromiss zwischen Schnelligkeit und Empfindlichkeit stellt der sog. Transimpedanzverstärker nach Figur 139c dar.



Figur 137: Spektrale Empfindlichkeit und Effizienz von Photodioden.



Figur 138: Aufbau einer Photodiode.



Figur 139: Photoempfänger:
 a) empfindlicher Spannungsverstärker;
 b) schneller Spannungsverstärker;
 c) Strom-Spannungs-Wandler (Transimpedanzverstärker).

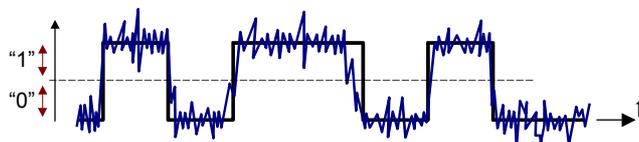
Im Gegensatz zur “normalen” Photodiode (oder “pin-Photodiode”) generiert in der *Avalanche Photodiode* (APD) ein einfallendes Photon *viele* Ladungsträgerpaare. In Wahrheit wird natürlich bestenfalls ein Ladungsträgerpaar erzeugt, dieses aber erfährt durch ein extern angelegtes starkes ($E > 3 \cdot 10^5$ V/cm) elektrisches Feld (*bias*) eine derartige Beschleunigung, dass es zur lawinenartigen Stossionisation kommt. Die Empfindlichkeit der APD ist um

den (Strom-) Multiplikationsfaktor M grösser als jene einer “normalen” PD. Neben der unüblichen hohen Biasspannung von bis zu einigen hundert Volt erzeugt der Empfang mit APD zusätzliches Rauschen.

Für extrem hohe Datenraten werden das absorbierende Volumen herkömmlicher PD extrem klein und die Effizienz schlecht. Dieses Problem lässt sich mit sog. *Traveling wave photodiodes* lösen. Hier ist die optische Absorption über eine grössere Länge (etwa bis zu 1mm) verteilt, und die generierten Ladungsträger bilden auf einer parallellaufenden elektrischen Leitung ein in der Amplitude anwachsendes Signal. Anschaulich gesehen kann *Bit N* gerade den Detektor verlassen, während *Bit N+1* in der Mitte des Detektors ist und *Bit N+2* gerade die Faser verlässt und in den Detektor einkoppelt.

8.2.1.11. Rauschen

In einem binären Übertragungssystem muss in Abständen der Taktzeit (Bitdauer) entschieden werden, ob eine logische Null oder aber eine Eins empfangen wurde. Dazu wird das empfangene und verstärkte Signal mit einer Referenzgrösse verglichen. Weil das Signal statistischen Schwankungen unterliegt, also verrauscht ist, wird mit einer gewissen Wahrscheinlichkeit “falsch entschieden”. Setzt man die Zahl der falsch empfangenen Bits zu den richtig empfangenen Bits ins Verhältnis, dann ergibt sich die Bitfehlerwahrscheinlichkeit (engl. BER, *Bit Error Rate*). Diese physikalisch bedingte BER des *physical layer* kann auf höheren Ebenen durch den Einsatz von Korrekturcodes verringert werden (aber BER = 0 ist unmöglich). Die Figur 140 zeigt ein typisches verrauschtes digitales Signal.



Figur 140: Verrauschtes digitales Signal im Zeitverlauf und Entscheidungsschwelle.

Obwohl im Folgenden (aus Gründen der Anschaulichkeit) auf binäre, intensitätsmodulierte optische Übertragung Bezug genommen wird, unterliegen analoge und kohärente Übertragungsverfahren ganz ähnlichen Rauscheffekten. Während digitale Übertragungssysteme meist mit der BER charakterisiert werden, benutzt man zur Beschreibung analoger Systeme eher das Signal-Rausch-Verhältnis (engl. SNR, *Signal to Noise Ratio*). Beide Begriffe versuchen, den Einfluss des Rauschens auf das Gesamtsystem zu beschreiben; die Ursachen des Rauschens sind jeweils dieselben.

In einem Übertragungssystem trägt prinzipiell jede Komponente zum Rauschen und damit zur Unkenntlichmachung des Signals bei. Dass bei einer Intensitätsmodulation Amplitudenschwankungen der Signalqualität abträglich sind, ist offensichtlich. Aber auch Phasen- oder Frequenzschwankungen von Träger (also dem Licht) und Signal können sich in Intensitätsschwankungen umwandeln, und zwar

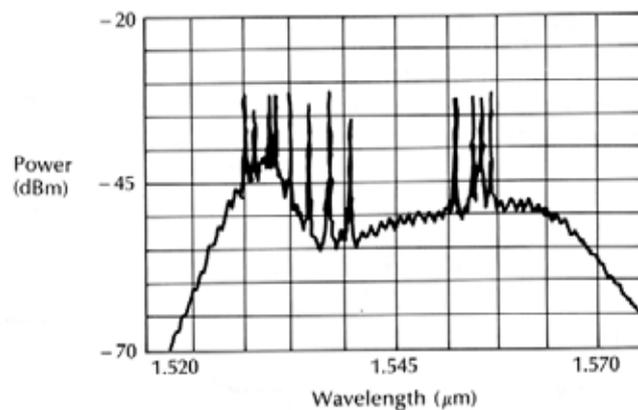
- bei Mehrwegeausbreitung (Interferometern) durch Interferenz bzw. Schwebungen zwischen den einzelnen Signalkomponenten (engl. *beat noise*)
- bei Nichtlinearitäten durch Mischprozesse.

Wenn beispielsweise die LD durch direkte Modulation oder durch Kopplung/Resonanz mit externen Komponenten die Wellenlänge ändert, kann es passieren, dass ein Bit in der einen Hälfte der Bitdauer durch Wellenlänge λ_a übertragen wird, später aber durch Wellenlänge λ_b . Weil aufgrund der Dispersion der Faser verschiedene Wellenlängen unterschiedliche Ausbreitungsgeschwindigkeiten aufweisen, kann dieses Bit am Empfänger durch destruktive Interferenz im schlimmsten Fall “verschwinden”.

Wenn das Lichtsignal optisch verstärkt wird, dann fügt der optische Verstärker dem Signal vielerlei Rauschkomponenten hinzu. Diese werden von gegebenenfalls vorhandenen weiteren optischen Verstärkern mit verstärkt. Ein typisches Spektrum am Ende einer optischen Übertragungsstrecke mit optischen Verstärkern zeigt Figur 141.

Wenn das verrauschte Lichtsignal auf die Photodiode trifft, wird es in ein verrauschtes elektrisches Signal umgewandelt. Die Erzeugung eines Ladungsträgerpaares durch ein Photon ist ein Quantenprozess und unterliegt damit den Gesetzen der Statistik. Selbst ein perfektes optisches Signal führt zu einem statistisch verteilten elektrischen Strom (Poisson-verteilt bei ideal kohärentem Licht). Das so entstandene Rauschen wird Schrotrauschen genannt (engl. *shot noise*). Die sog. Quantenrauschgrenze trifft dann Aussagen wie “19 Photonen pro Bit ergeben ein BER von 10^{-9} ”. In den allermeisten technisch relevanten Systemen wird die Quantenrauschgrenze nicht annähernd erreicht.

Die Photodiode fügt noch selber Rauschanteile hinzu (durch den Dunkelstrom, Leckströme, thermisches Rauschen). Die erste Verstärkerstufe selber rauscht auch und muss daher besonders sorgfältig ausgesucht werden. Wenn möglich werden für dafür FET verwendet, die allerdings den Nachteil haben, langsamer als Bipolartransistoren zu sein. Generell rauschen hochohmige, also langsame Schaltungen weniger als niederohmige, schnelle Schaltungen.



Figur 141: Signale und Rauschspektrum am Ende eine WDM-Übertragungsstrecke mit optischen Verstärkern.

Es lässt sich feststellen, dass die am Empfänger geforderte BER zusammen mit einer Rauschanalyse aller Komponenten und der Dämpfung der Übertragungsstrecke die Leistung des Senders bestimmt. Ein derartiges sog. *power budget* wird im Abschnitt Kommunikationssysteme exemplarisch vorgestellt.

Kohärente Übertragungsverfahren unterliegen allen genannten Rauschquellen ebenso. Allerdings kann durch den Prozess der Frequenzumsetzung (Mischung) das SNR um 3dB...6dB verbessert werden.

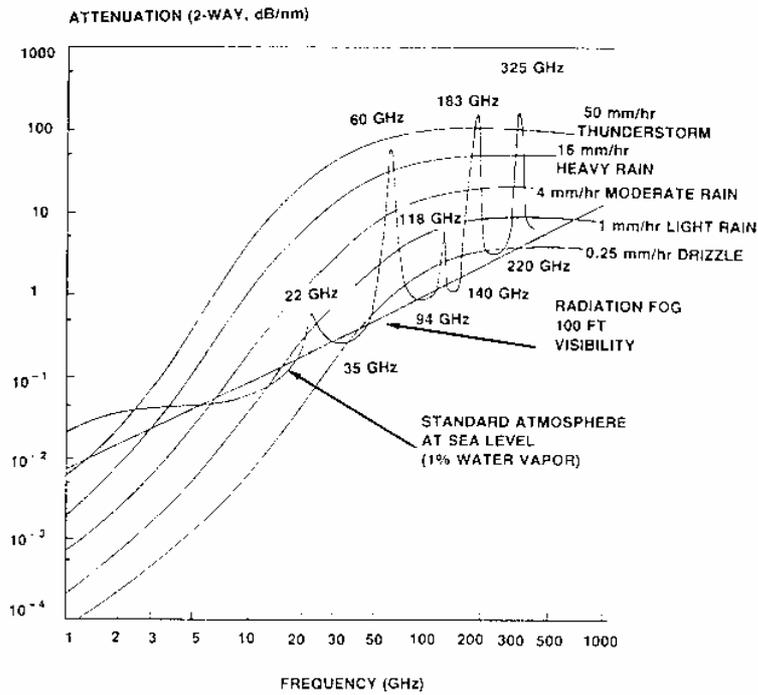
8.3. Freiraumausbreitung

Die Vorteile der Freiraumausbreitung liegen in der Mobilität und Flexibilität der Übertragungsstrecken. Nachteilig sind die geringe Bandbreite (und damit Datenrate), die Störanfälligkeit (Möglichkeit des Abhörens), und die hohen Leistungen (Stichwort Elektrosmog).

Die zur Verfügung stehenden Bandbreiten nehmen naturgemäß mit der Frequenz zu. Gleichzeitig nimmt die atmosphärische Dämpfung zu (siehe Figur 142). Bei hohen Frequenzen sind Sichtverbindungen nötig. Die Figur 143 zeigt exemplarisch die Bedingungen in verschiedenen Frequenzbereichen.

Freiraumübertragung hat sich im Vergleich zur Übertragung per Glasfaser mit wesentlich höheren BER zu arrangieren (Mobiltelefone bis zu $BER = 10^{-3}$). Die sowieso geringe Brutto-Datenrate wird damit zu einem nicht geringen Teil mit Fehlerkorrektur- und Synchronisationscodes belastet.

Die Radiofrequenzen werden national wie international administrativ vergeben und überwacht. In den jeweiligen Frequenzbändern gelten mehr oder weniger strenge technische Vorschriften betreffend Sendeleistung, Modulationsverfahren, erlaubte Niveaus der *out-of-band*-Störungen etc. In Frequenzbändern mit besonders wenigen Vorschriften tummeln sich entsprechend viele Anwendungen. Im Extremfall (2.4...2.5 GHz) arbeiten Mikrowellenöfen und *Bluetooth*-Computerlinks bei derselben Frequenz.



Figur 142: Atmosphärische Dämpfung bei verschiedenen Frequenzen (hin- und zurück, in dB pro nautische Meile [nm]).

Frequenz	1 MHz	100 MHz	10 GHz
Wellenlänge	300 m	3 m	3 cm
Anwendung	AM Radio	UKW, VHF	SAT, WLAN
Bandbreite	10...100 kHz	0.1...6 MHz	... 155 Mb/s
Leistung	0.1...100 kW	0.01...10 kW	0.1...10 W
Richtfaktor	360°	> 60°	> 5°
Distanz	...20'000 km	2...200 km	1...10 km
Sichtverbindung	unnötig, da Reflexion und Beugung	Mehrweg-Ausbreitung unerwünscht	unbedingt; Regen und Nebel stark dämpfend

Figur 143: Einige Kriterien für Radioausbreitung bei verschiedenen Frequenzen

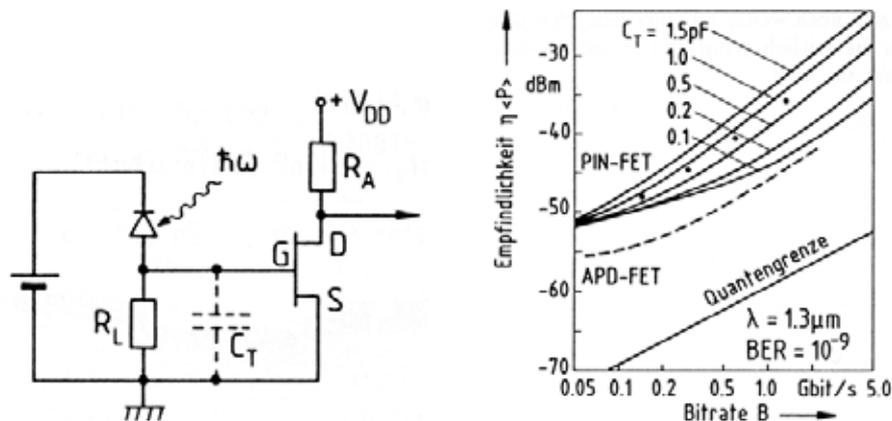
8.4. Kommunikationssysteme

8.4.1. Punkt-zu-Punkt-Verbindung & power budget

Am Beispiel einer Punkt-zu-Punkt-Verbindung soll exemplarisch das Zusammenspiel von Sender, Übertragungsmedium und Empfänger gezeigt werden. Bevor aber die Hardware entworfen wird, müssen die Systemspezifikationen (mindestens Datenrate, BER und Distanz, praktisch auch Kompatibilität, Preis, Zeitrahmen etc.) bekannt sein.

Im Folgenden wird das *power budget* für eine faseroptische Übertragungsstrecke aufgestellt. Gefordert wird eine Datenrate von 1 Gbit/s, $BER = 10^{-9}$, Distanz 80 km. Es stehen Sendemodule mit 1mW, 3mW und 10mW optischer Ausgangsleistung zur Verfügung (deren Kosten steigen überproportional mit der Ausgangsleistung). Die zu verwendende Monomodefaser ist mit 0.3 dB/km spezifiziert und dispersionskompensiert bei 1.3 μm .

Die Figur 8.21 zeigt einen Empfänger mit Photodiode und FET-Verstärkerstufe und das Ergebnis der dazugehörigen Rauschanalyse. Aus dem Datenblatt der Photodiode entnimmt man deren Kapazität von $C_T = 0.5$ pF. Dann zeigt Figur 144: Bei 1 Gbit/s ergibt eine mittlere Empfangsleistung von -39dBm (entspricht 0.125 μW) gerade eine BER von 10^{-9} .



Figur 144 Schaltbild eines Photoempfängers und das dazugehörige, berechnete Rauschverhalten.

Jetzt kann man einfach “zum Sender zurückrechnen”, um zu bestimmen, welche Sendeleistung für $1/8$ μW Empfangsleistung nötig ist (siehe Figur 145). Es stellt sich heraus, dass das 3mW-LD-Modul verwendet werden kann.

Leistung am Empfänger	-39 dBm = 0.13 μW
Empfänger efficiency 80%	1 dB
Splices & connectors	1 dB

Faserdämpfung 80 x 0.3 dB	24 dB
Isolator	1 dB
Einkoppelverluste am Sender	6 dB
Marge Alterung und Reparatur	9 dB
Ausgangsleistung der Sende-LD	+3 dBm = 2.0 mW

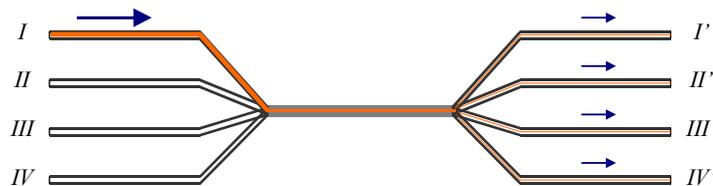
Figur 145: Power budget einer faseroptischen Punkt-zu-Punkt-Verbindung.

8.4.2. Passive Komponenten für faseroptische Netzwerke

Punkt-zu-Punkt-Verbindungen werden recht selten benötigt (etwa für transozeanische Links). Wichtiger sind Netzwerke, die viele Teilnehmer verbinden. Entsprechend der Datenrate und des Datenaufkommens der einzelnen Teilnehmer muss zunächst unter Berücksichtigung von Zuverlässigkeitsaspekten (was darf passieren, wenn ein Teil des Netzes oder/und ein Teilnehmer ausfällt?) die Netztopologie entworfen werden. Das so entstandene Netz wird sicherlich Verzweigungen, möglicherweise Einweg-Verbindungen und gegebenenfalls Links mit stark unterschiedlichen Datenraten enthalten. Die relevanten Komponenten sollen hier kurz vorgestellt werden.

8.4.2.1. Sternkoppler

Ein Sternkoppler (*engl. star coupler*) als passives Element besteht einfach aus mehreren zusammengeschmolzenen (*engl. fused*) Glasfasern. Ein Lichtsignal in Faser I (siehe Fig. 146) wird zu gleichen Teilen auf die Fasern I'...IV' aufgeteilt.



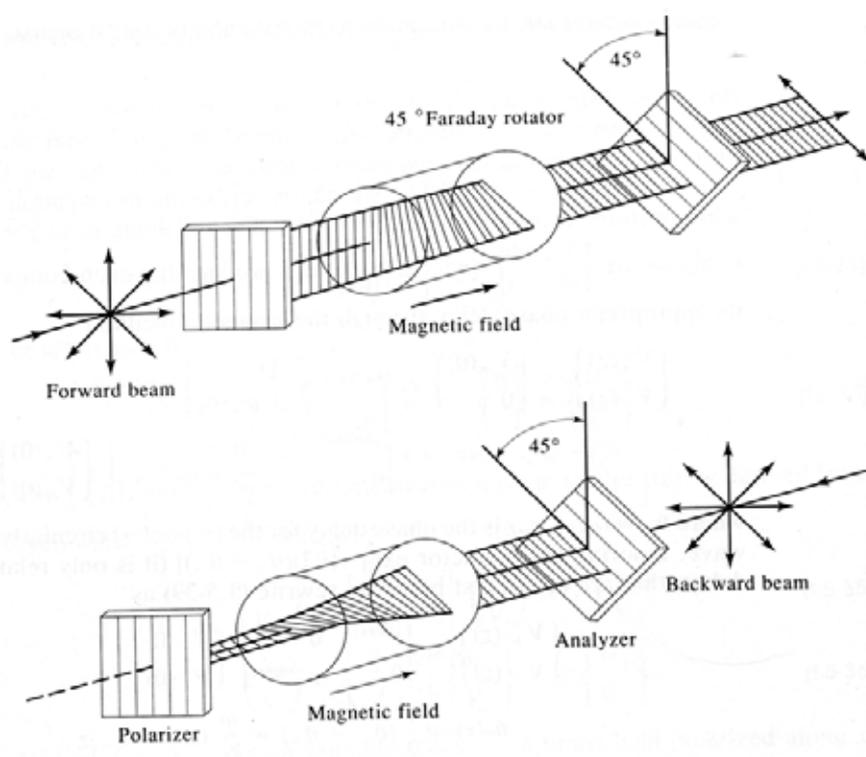
Figur 146: Sternkoppler.

8.4.2.2. Schalter

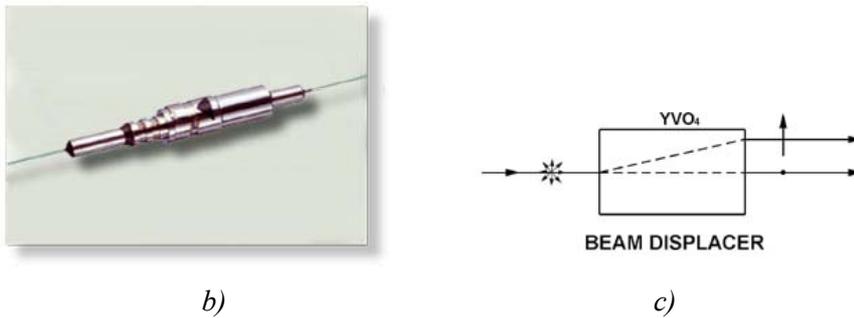
Ein mittels Schaltern aufgebautes Netzwerk hat den Vorteil, dass praktisch wieder Punkt-zu-Punkt-Verbindungen entstehen. Weil aber Schalter meistens elektromechanisch bewegt werden, ist ihre Geschwindigkeit nicht allzu hoch. Durch die zunehmende Miniaturisierung mechanischer Komponenten wird die Schaltzeit aus dem Millisekundenbereich heute in den Mikrosekundenbereich verlagert. Aber auch für langsame Schalter gibt es Anwendungen, etwa zur Überbrückung defekter Komponenten in Repeatern transozeanischer Links.

8.4.2.3. Isolator

Ein optischer Isolator lässt das Licht in einer Richtung hindurch, während er es in der anderen Richtung absorbiert. Insbesondere Laserdioden reagieren sehr empfindlich auf zurück-reflektiertes Licht und werden deshalb oft mit einem Isolator kombiniert. Isolatoren beruhen auf Materialien, die optisch anisotrop sind. Der Faraday-Effekt bewirkt eine Drehung der Polarisations Ebene des Lichts und ist Grundlage für den Einweg-Effekt. Wenn die Drehung der Polarisations Ebene 45° beträgt, dann kann Licht den Isolator und die Polarisationsfilter nur in einer Richtung durchqueren (Figur 147a). Für einen direkt an eine LD gekoppelten Isolator kann sogar auf ein Polarisationsfilter verzichtet werden, weil das Laserlicht stark linear polarisiert ist.



a)



Figur 147: Isolator: a) Prinzip, b) Konstruktive Ausführung, c) Polarisationssteiler

Falls der Isolator unabhängig von der Polarisation des einfallenden Lichtes funktionieren soll, dann muss das Licht zunächst in zwei orthogonale Polarisationen aufgespalten werden. Für jede Polarisation sorgt ein separater Isolator für den gewünschten Einweg-Effekt, und schliesslich werden die beiden Lichtstrahlen wieder kombiniert. Die Figur 147b zeigt die praktische Ausführung eines polarisationsunabhängigen Isolators. Die Figur 147c zeigt das Prinzip der Strahlentrennung mittels eines doppelbrechenden Kristalls.

Der sog. Faraday-Rotator besteht typischerweise aus YIG (Yttrium-Eisen-Granat), das den gewünschten Effekt unter Einfluss eines Permanentmagnetfeldes zeigt. Trennung und Kombination der Polarisationen erfolgen einfach mittels sog. doppelbrechender (*engl. birefringent*) Kristalle, wie etwa CaCO_3 , LiNbO_3 , TiO_2 , YVO_4 .

8.4.2.4. Bragg-Gitter

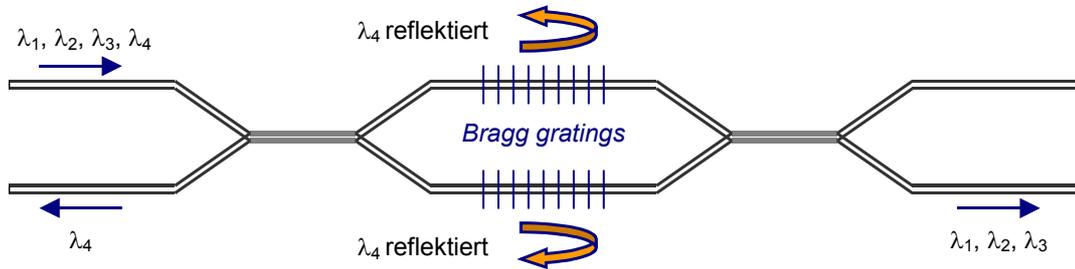
Bragg-Gitter (*engl. Fiber Bragg Gratings*) sind eine periodische Variation des Brechungsindex des Faserkerns in Faserrichtung. Die tatsächliche Änderung der Brechzahl ist sehr klein, dafür sind hunderte oder tausende dieser kleinen Änderungen hintereinander angeordnet. An jeder kleinen Brechzahländerung erfährt eine einfallende Welle eine kleine Reflexion (der allermeiste Teil wird transmittiert). Für eine bestimmte Wellenlänge (definiert durch Brechzahlverhältnis und Periode des Gitters) summieren sich aber all die kleinen reflektierten Anteile auf und Licht genau dieser Wellenlänge wird zu 100% reflektiert (je nach Länge des Gitters kann man auch eine Reflexion von 50% oder einen beliebigen anderen Anteil einstellen). Das Licht aller anderen Wellenlängen wird transmittiert.

Bragg-Gitter lassen sich sehr effizient herstellen. Man nutzt den Effekt, dass z.B. Ge-dotiertes Glas unter UV-Bestrahlung seine Brechzahl dauerhaft ändert. Man braucht also bloss eine Ge-dotierte Faser mit dem Gittermuster UV-bestrahlen. Mittels einer kohärenten, leistungsstarken UV-Lichtquelle (Excimer-Laser) lässt sich das Gittermuster einfach durch Interferenz herstellen.

8.4.2.5. Filter & Multiplexer

Mittels der Faser-Bragg-Gitter lassen sich viele wellenlängenselektive Komponenten einfach herstellen. Die Figur 148 zeigt ein solches Bragg Filter ("Frequenzweiche"). Sie stellt

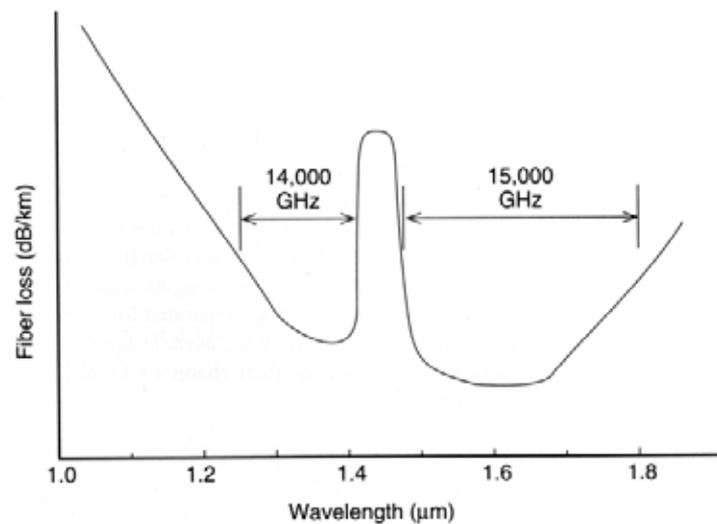
einen sog. Diplexer dar, der zwei Signale selektiv trennt (oder, entgegengesetzt der Pfeilrichtungen, zusammenführt). Multiplexer lassen sich aus kaskadierten Diplexern aufbauen.



Figur 148: Faser Bragg Filter.

8.4.3. WDM - (wavelength division multiplexing) Systeme

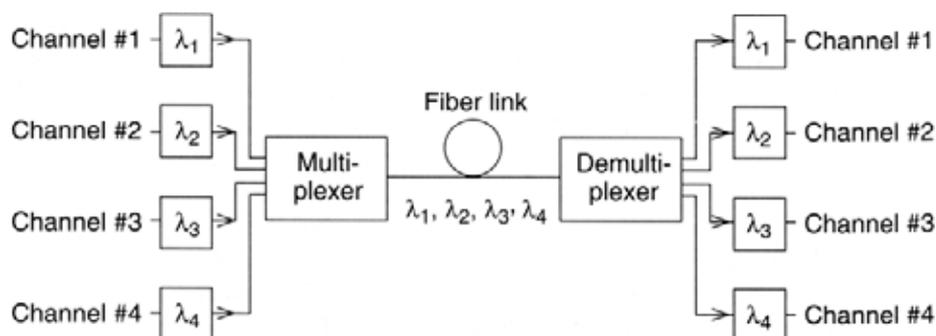
In einer Glasfaser können sich Lichtstrahlen unterschiedlicher Wellenlänge gemeinsam ausbreiten, ohne einander zu beeinflussen (solange die Gesamt-Lichtleistung klein genug ist, um nichtlineare Effekte vernachlässigen zu können). Die sehr grosse optische Bandbreite der Fasern (siehe Figur 149) lässt sich besser ausnutzen, wenn Licht verschiedener Wellenlängen gleichzeitig transportiert wird. Das nennt man Wellenlängen-Multiplex (engl. WDM).



Figur 149: Theoretische Bandbreite der Glasfaser.

Mittels der oben beschriebenen Bragg-Multiplexer lassen sich die Signale von verschiedenen Quellen (jede Quelle mit ihrer spezifischen Wellenlänge) in einer Faser kombinieren,

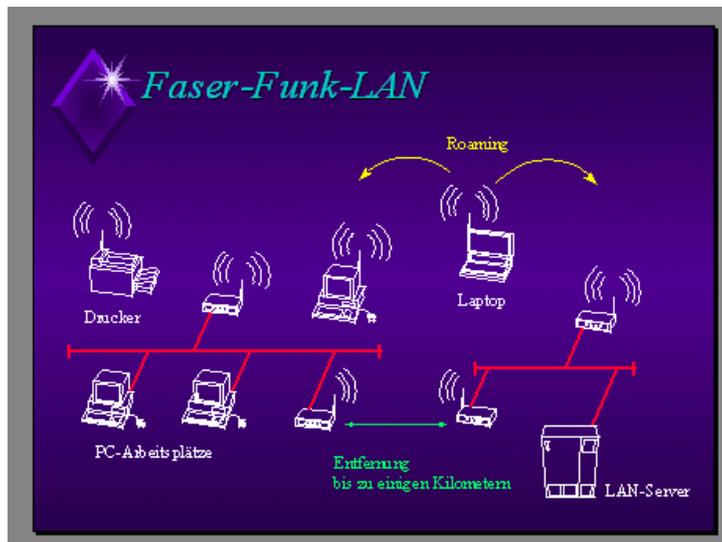
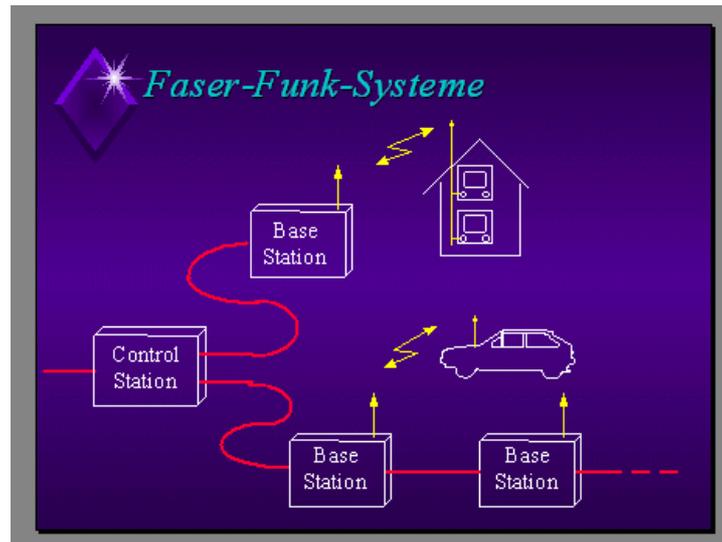
gemeinsam optisch verstärken, und vor den Empfängern wieder selektiv trennen (Figur 150). Die Abbildung 141 (siehe oben) zeigt das optische Spektrum eines WDM-Systems.



Figur 150 WDM System.

8.4.4. Kombination optischer und *wireless*-Netze

Die drei Übertragungsmedien "Koaxial-/Zweidrahtleitung", "Glasfaser" und "Wireless" haben ihre jeweiligen Vor- und Nachteile. In einem praktischen System wird man diese kombinieren, um zu einer optimalen Lösung zu gelangen. Die Glasfaser ist dabei für hohe Datenraten prädestiniert, Funkverbindungen sind für mobile Anwendungen zwingend, aber nicht für hohe Datenraten geeignet, während Kupferkabel niedrige und mittlere Datenraten sehr preiswert übertragen können. Die Figur 151 zeigt zwei heute aktuelle Beispiele für Netze mit mehreren Übertragungsmedien.



Figur 151: Faser-Funk-Systeme.

9. Energietechnik

Mit der Entwicklung der Dampfmaschinen wuchs die Erkenntnis, dass Energie eine physikalische Grösse von besonderem praktischem Wert ist. Insbesondere die praktischen Untersuchungen von Faraday und die theoretischen Arbeiten von Maxwell zeigten, dass der Elektromagnetismus eine besonders interessante und vielfältige Form der Energie ist, welche sich relativ leicht in andere Energieformen umwandeln lässt und ausserdem gut über weite Distanzen übertragen werden kann. Durch Edisons Entwicklung der elektrischen Glühlampe ergab sich die grosse Chance, die gefährlichen Gasbeleuchtungen durch ungefährliche elektrische Beleuchtung zu ersetzen. Dank Edisons Vermarktungstalents gewann die elektrische Energieversorgung rasch eine grosse Bedeutung, obwohl wir heute wissen, dass Elektrizität keineswegs ungefährlich ist.

Edison war ein vehementer Verfechter des Gleichstroms und focht einen langen, erbitterten und teilweise absurden Kampf mit Tesla, welcher erkannte, dass Wechselstrom einfacher und effizienter produziert werden kann. Heute werden die meisten Netze mit Wechselstrom betrieben, wobei sich jedoch Spannung und Frequenz beträchtlich unterscheiden. Eine Ausnahme bilden Gleichstromnetze für den öffentlichen Verkehr.

Da heute in industrialisierten Ländern die Produktion und Verwaltung aller lebenswichtiger Güter von der kontinuierlichen Zufuhr elektrischer Energie abhängt, ist die Energieversorgung von zentraler Bedeutung. In diesem Kapitel werden deshalb die Grundlagen der Energietechnik skizziert.

9.1. Sicherheitsaspekte

Obwohl Edison zunächst aus naheliegenden Gründen die elektrische Energie als harmlos oder gar medizinisch wertvoll anpries, demonstrierte er in seinem Streit mit Tesla auch deren Gefährlichkeit mit der Entwicklung des elektrischen Stuhls. Entgegen Edisons Behauptung, dass Gleichstrom ungefährlich, Wechselstrom hingegen gefährlich sei, wissen wir, dass der elektrische Strom, welcher bei Berührung elektrischer Anlagen durch den Körper fliesst, potentiell gefährlich ist.

Elektrische Ströme können im Körper sehr unterschiedliche Folgen haben. So können Muskeln zur Kontraktion angeregt werden, was zu plötzlichen, unerwünschten Bewegungen mit Unfallfolgen führen kann. Selbstverständlich kann auch der Herzmuskel angeregt werden (insbesondere, wenn der Strompfad von einem Arm zum andern über das Herz führt — Elektriker arbeiten darum gerne einhändig!). Als Folge davon kann Herzstillstand oder Herzflimmern auftreten. Dies ist tödlich, wenn es nicht rasch — zum Beispiel mit gezielten Stromstössen — behoben wird. Eine weitere, bekannte Gefahr sind Verbrennungen, welche in erster Linie an den Kontaktstellen auftreten, weil der Kontaktwiderstand meist wesentlich grösser als der Körperwiderstand ist. Weniger bekannt, aber besonders gefähr-

lich ist die Möglichkeit, dass der elektrische Strom im Körper mittels Elektrolyse Substanzen erzeugt, welche zu Vergiftungen führen können. Nach Stromunfällen wird deshalb empfohlen, womöglich alkalische Getränke in grösseren Mengen aufzunehmen.

Es wird gemeinhin angenommen, die Gefährlichkeit nehme mit der Spannung zu. Dies ist jedoch nicht richtig, obwohl sehr geringe Spannungen sicher harmlos sind. Die im Körper erzeugten physiologischen Erscheinungen hängen zunächst vor allem von der Stromstärke ab. Gemäss dem Ohm'schen Gesetz ist der Strom im Körper proportional zur angelegten Spannung und zum Widerstand. Der Widerstand wiederum setzt sich im wesentlichen aus Kontaktwiderstand und Körperwiderstand zusammen. Insbesondere der Kontaktwiderstand kann sehr unterschiedlich gross sein. Berührt beispielsweise ein Mensch mit feuchten Händen zwei elektrische Drähte, so ist der Kontaktwiderstand relativ gering, so dass bereits bei einer geringen Spannung von unter 100 V ein für den Organismus gefährlicher Strom fließen kann. Berührt hingegen ein Mensch mit einer Hand einen Draht, so fließt der Strom über Hand, Arm, Körper in den Fussboden. In diesem Fall kann der Kontaktwiderstand bei den Füßen sehr hoch sein, wenn gut isolierende Schuhe getragen werden.

Natürlich hängt die physiologische Wirkung des Stroms auch stark von der Dauer ab. Bei eher kleineren Strömen neigt die Muskulatur zu Verkrampfungen, so dass der Mensch sich eventuell von der Stromquelle nicht mehr lösen kann, was zu Elektrolysegefahr führt. Bei höheren Strömen hingegen kann der Mensch durch die Muskelreaktion von der Stromquelle weggeschleudert werden.

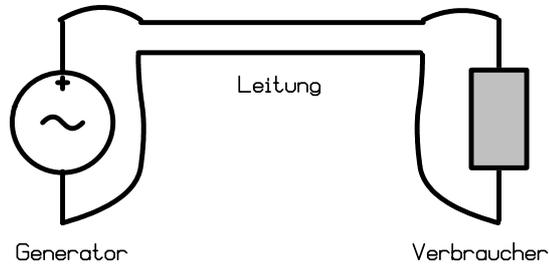
Bereits Tesla hatte bemerkt, dass bei höheren Frequenzen der Strom nur noch wenig in den Körper eindringt und weitgehend auf der Oberfläche bleibt (Skineffekt). Er hatte diesen Effekt in seinem Kampf gegen Edison eindrucksvoll genutzt, indem er selber den Kontakt zu einer Leuchtstoffröhre schaffte und sich so unbeschadet unter Strom setzte. Die heute gängigen Stromnetze werden jedoch niederfrequent bei Spannungen von über 100 V betrieben und sind deshalb prinzipiell gefährlich.

Der Grund für die Verwendung gefährlicher Spannungs- und Frequenzbereiche in der heute üblichen Stromversorgung ist in der Wirtschaftlichkeit zu finden. Um die gleiche Menge Energie bei niedrigeren Spannungen zu transportieren, müsste mit höheren Strömen gearbeitet werden und diese würden grössere Drahtquerschnitte erfordern oder zu grösseren Verlusten in den Drähten führen. Bei höheren Frequenzen würden erhebliche Strahlungsverluste in Freileitungen entstehen. Um diese zu vermeiden, müssten wesentlich teurere, abgeschirmte Kabel entwickelt werden.

9.2. Prinzip der Energieversorgung

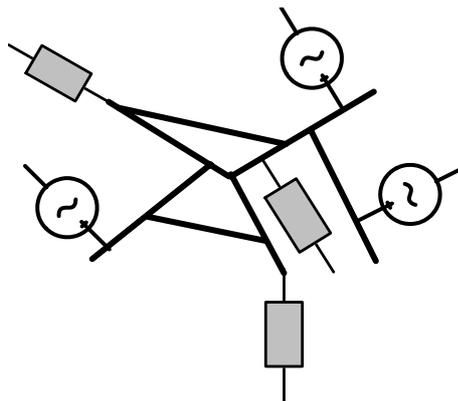
Die notwendige Minimalstruktur eines Energieversorgungssystems ist in Figur 152 wiedergegeben. Es besteht aus einem Generator (Erzeuger), dem Verbraucher und der Verbindungsleitung dazwischen. Man erkennt sofort, dass dieses System unsicher ist: bei jeder

Störung am Generator oder an der Verbindungsleitung bricht die Energieversorgung des Verbrauchers zusammen.



Figur 152: Minimalstruktur der Energieversorgung

In der Praxis sind deshalb viele Kraftwerke (Generatoren) und Verbraucher miteinander verbunden, d.h. vernetzt, wie in Figur 153 schematisch dargestellt. Wichtig ist, dass alle Generatoren in einem Netz genau die gleiche Frequenz aufweisen müssen.



Figur 153: Vermahtes Netz

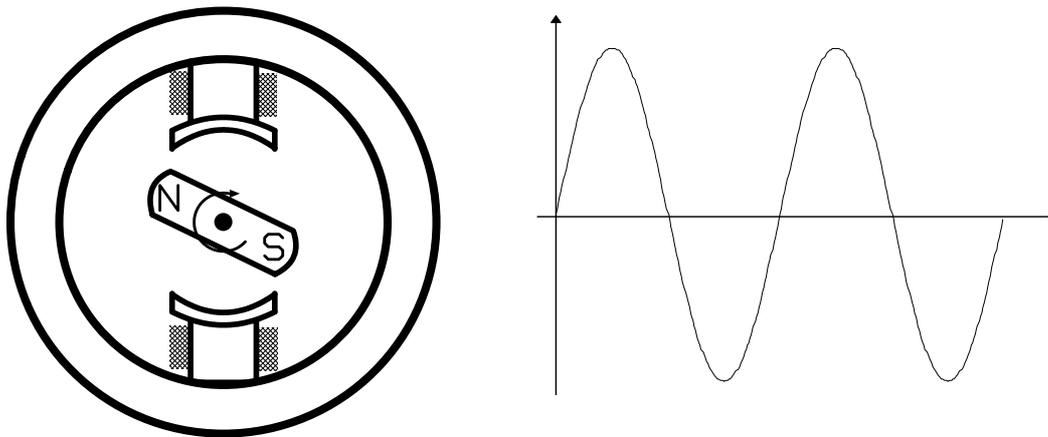
9.3. Erzeugung elektrischer Energie

Energie ist eine physikalische Grösse mit sehr vielen Ausprägungen. Elektrische Energie ist eine davon. Der Energiesatz besagt, dass Energie umgeformt, aber nicht neu entstehen kann. Erzeugung elektrischer Energie meint also Umwandlung nicht-elektrischer in elektrische Energie. So wird beispielsweise in Batterien chemische in elektrische Energie gewandelt. Für die Erzeugung grosser Mengen von elektrischer Energie werden heute ausschliesslich mechanisch-elektrische Wandler verwendet, welche auf dem von Faraday entdeckten Induktionsprinzip beruhen und im wesentlichen dasselbe wie Elektromotoren sind, welche in vielen Geräten des täglichen Lebens zu finden sind.

Das Induktionsprinzip besagt, dass die Veränderung des Flusses des magnetischen Feldes durch eine Schlaufe in der Schlaufe eine elektrische Spannung induziert. Dabei ist die Spannung proportional zur Änderung des Flusses. Durch Hintereinanderschalten mehrerer Schlaufen erhält man eine Spule. Die an der Spule abgegriffene Spannung ist in guter Näherung proportional zur Windungszahl. Zunächst ist eine Spule also ein Wandler, welcher magnetischen Fluss in elektrische Spannung umwandelt. Wird das Magnetfeld zum Beispiel durch die mechanische Bewegung eines Magneten erzeugt, so ergibt sich ein Generator für elektrische Energie. Figur 154 zeigt eine mögliche Realisierung mit einem rotierenden Permanentmagneten. Diese wird allerdings praktisch selten verwendet, da starke Permanentmagnete schwer und spröde sind. Gebräuchlichere Varianten ergeben sich, wenn der Permanentmagnet durch eine stromdurchflossene (extern gespeisene, aktive) Spule ersetzt wird oder wenn Magnet und Spule vertauscht werden, d.h. die (passive) Spule rotiert und der Permanentmagnet fest steht. In jedem Fall wird in der passiven Spule eine Wechselspannung induziert. Diese Spule kann also wie in Figur 152 mit einem elektrischen Verbraucher verbunden werden.

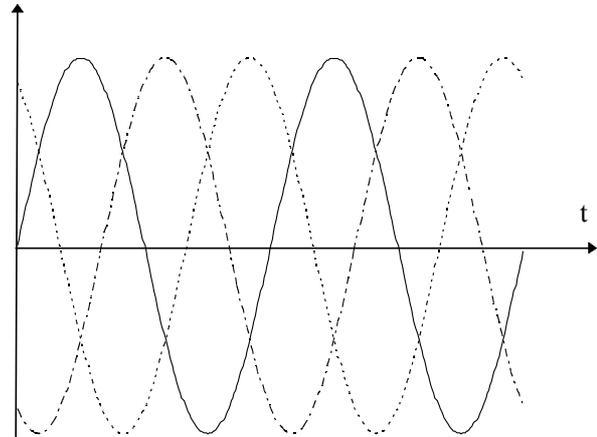
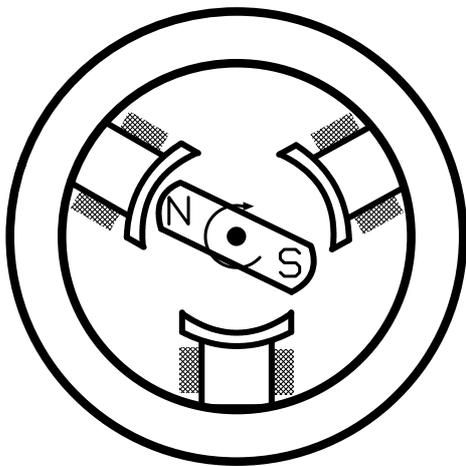
Wird umgekehrt der (passiven) Spule ein Wechselstrom eingeprägt, so erhält man einen elektrisch-mechanischen Wandler, d.h. die zugeführte elektrische Energie führt dann zu einer mechanischen Drehung des Rotors. Dies ist dann ein Wechselstrommotor.

Es wurde in der Einleitung erwähnt, dass Edison Gleichstrom favorisierte. Gleichstrommotoren und Gleichstromgeneratoren basieren ebenfalls auf dem Induktionsprinzip und bestehen ebenfalls aus Spulen und allenfalls Permanentmagneten. Zusätzlich wird ein Kommutator auf der rotierenden Achse angebracht, welcher während jeweils einer halben Umdrehung des Rotors (während der negativen Halbperiode) die Anschlüsse der passiven Spule miteinander vertauscht, so dass die Ausgangsspannung mechanisch gleichgerichtet wird. Edisons Gleichstromgeneratoren erzeugen also im Prinzip keinen Gleichstrom, sondern gleichgerichteten Wechselstrom. Wie aus Wechselstrom Gleichstrom erzeugt werden kann, wird weiter unten ausführlicher dargestellt. Die mechanische Gleichrichtung mit Kommutatoren, d.h. Schaltern, ist technisch wegen Funkenerzeugung und Korrosion der Kontakte problematisch.



Figur 154 Elektrischer Generator, links Aufbau, rechts abgegebene Spannungen
 (2 Perioden entsprechend 2 Umdrehungen dargestellt)
 (N-S): rotierender Magnet, Schraffiert: Wicklungen, in denen Spannung induziert wird.

Wird ein elektrischer Verbraucher an die Klemmen des Generators angeschlossen, so wird der Rotor umso stärker gebremst, je grösser die verbrauchte Energie ist. Im einfachsten Fall eines Ohm'schen Verbrauchers schwankt offensichtlich die Leistung mit der doppelten Frequenz der Wechselspannung. Der Rotor wird also nicht kontinuierlich gebremst, was mechanisch nicht vorteilhaft ist. Um eine kontinuierlichere Belastung des Rotors zu erhalten, kann man mehrere passive Spulen anbringen. Figur 155 zeigt eine besonders einfache Variante mit drei Spulen. Rotiert der Permanentmagnet mit konstanter Geschwindigkeit und sind die Spulen symmetrisch angeordnet, so ergeben sich offensichtlich drei zueinander phasenverschobene Spannungen mit je 120° Phasendifferenz.



Figur155 Dreiphasen-Generator; links Aufbau, rechts abgegebene Spannungen

In der Praxis werden meist derartige Drehstromgeneratoren mit drei symmetrisch angeordneten Spulen (oder drei Paaren von gegenüberliegenden Spulen) verwendet. Für die drei Spannungen eines solchen symmetrischen Dreiphasen-Generators gilt, mit $\omega_1 = 2\pi f_1 = 2\pi / T_1$:

$$\begin{aligned} u_1 &= \hat{u} \sin(\omega_1 t) \\ u_2 &= \hat{u} \sin(\omega_1 t - 120^\circ) \\ u_3 &= \hat{u} \sin(\omega_1 t - 240^\circ) \end{aligned}$$

Werden drei identische Widerstände R_1, R_2, R_3 an die drei Spulen angeschlossen, so ist die verbrauchte Gesamtleistung $p(t)$ zeitlich konstant.

Übungsaufgabe: Berechne $p(t)$ für symmetrische Zwei-, Drei-, und Vierphasensysteme mit symmetrischer Ohm'scher Last (identische Verbraucherwiderstände) und asymmetrischer Ohm'scher Last.

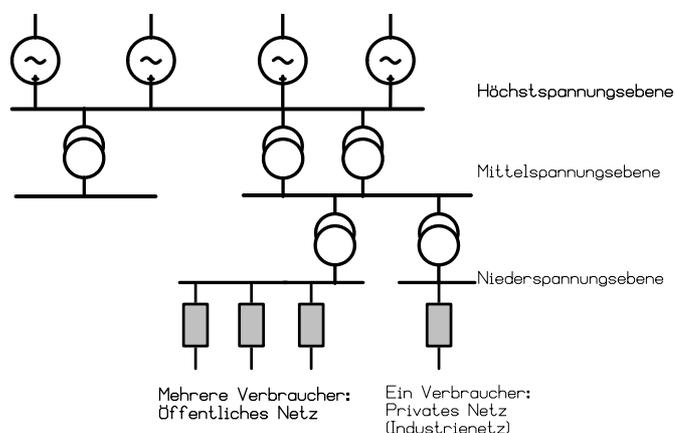
9.4. Übertragung elektrischer Energie

Ein grosser Vorteil der elektrischen Energie für die Energieversorgung liegt in der relative einfachen Übertragung dieser Energieform, wobei im Prinzip sowohl geführte als auch freie Übertragung in Frage kommen. Bei der freien Übertragung werden mittels Sendeannten elektromagnetische Wellen beim Kraftwerk erzeugt und mittels Empfangsantennen beim Verbraucher genutzt. Dieses — von Tesla favorisierte Prinzip — ist aber äusserst

unwirtschaftlich und praktisch höchstens in Ausnahmefällen mit sehr geringem Energiebedarf brauchbar.

Für die geführte Übertragung elektrischer Energie kommen verschiedene Leitungsstrukturen (Mehrdrahtleitungen, Hohlleiter, dielektrische Wellenleiter) in Frage. Will man die Übertragungsverluste bei möglichst geringem Materialaufwand möglichst klein halten, so erweisen sich einfache Kupferdrähte als besonders ökonomisch. Insbesondere für die Übertragung grosser Energiemengen über grosse Distanzen werden deshalb Freileitungen bevorzugt. Ausserdem werden diese bei möglichst hohen Spannungen und relativ kleinen Strömen betrieben um die Ohm'schen Verluste in den Drähten klein zu halten. Da hohe Spannungen schlecht direkt erzeugt und verbraucht werden können, sind sowohl auf Erzeuger- als auch auf Verbraucherseite Transformatoren erforderlich. Die Transformierbarkeit ist übrigens das Hauptargument für die Verwendung von Wechselstrom!

Verbraucher elektrischer Energie haben üblicherweise einen stark schwankenden Energiebedarf. Durch den Anschluss möglichst vieler Verbraucher an ein Kraftwerk, werden diese Schwankungen zwar etwas ausgeglichen, trotzdem ergibt sich ein massiv höherer Verbrauch zu gewissen Tageszeiten. Dies führt zu stark unterschiedlicher Belastung der Kraftwerke. Je nach Kraftwerktyp, kann die Regelung der abgegebenen Energie recht schwierig sein. Dies gilt insbesondere für Kernkraftwerke, aber natürlich auch für Sonnenkraftwerke. Besonders unangenehm macht sich hier bemerkbar, dass bis heute keine brauchbaren Speicher für grosse Mengen elektrischer Energie bekannt sind. Als Ausweg bietet sich deshalb die möglichst grossflächige Vernetzung der unterschiedlichsten Kraftwerke an. Figur 156 zeigt schematisch die Struktur eines solchen Netzes.



Figur 1565 Struktur (Europäisches) Verbundnetz

Die Generatoren (Kraftwerke) sind an eine gemeinsame (vermaschte) Hochspannungsleitung (Höchstspannungsebene 110 kV, 220 kV, 380 kV) angeschlossen, welche mit vielen Parallelpfaden ganz Europa wie ein Rückgrat durchzieht.

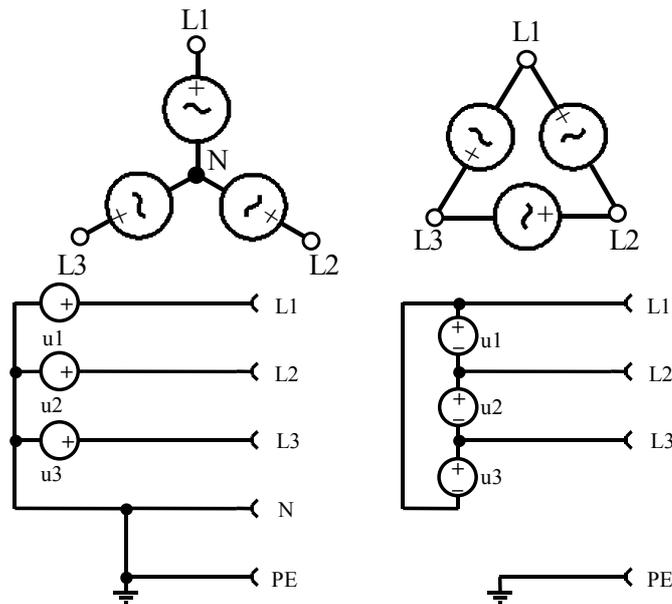
An dieses Netz wird dann, über Transformatoren an ausgesuchten Punkten, die Spannung auf einen kleineren Wert transformiert (Bereich 11 kV bis 75 kV). Diese Spannungsebene wird für die Mittelverteilung verwendet (Längen von einigen 10 km). Nahe beim Verbraucher wird dann diese Mittelspannung auf die eigentliche Verbraucherspannung (bei uns 230/400 V), transformiert und verteilt.

Für die Verteilungen wird aus technischen Gründen je nach Spannungsebene verschiedene Techniken verwendet. Im Niederspannungsgebiet sind es fast ausschliesslich Kabel mit einem dem Nennstrom entsprechenden Querschnitt. Auf der Höchstspannungsebene können praktisch nur Freileitungen verwendet werden. Auf der Mittelspannungsebene werden beide Verfahren verwendet.

Betrachtet man das ganze Versorgungsnetz vom Anschlusspunkt des Verbrauchers her, so kann man es in sehr guter Näherung durch eine ideale Spannungsquelle und eine Quellenimpedanz Z_N ersetzen. Im europäischen Netz kann man dabei davon ausgehen, dass bei 50 Hz diese Impedanz den Wert von $(0,4 + j 0,25) \Omega$ nicht überschreitet; in städtischen Gebieten ist sie im allgemeinen viel kleiner. Die Frequenz wird dabei so genau gehalten, dass früher Uhren gebaut wurden, welche anstelle einer Quarzfrequenz die Netzfrequenz verwendeten.

9.5. Dreiphasensysteme

Wie im Abschnitt 9.3 erwähnt, ist die gleichzeitige Erzeugung von drei Spannungen mit unterschiedlicher Phase (Drehstrom) vorteilhaft und deshalb üblich. Für die Übertragung bedeutet dies zunächst, dass dreimal so viele Leitungen benötigt werden, wenn jede Phase getrennt weitergeleitet werden kann. Viel ökonomischer ist es jedoch, die drei Phasen miteinander zu verbinden. Wegen der Symmetrie sind im Prinzip zwei Schaltungen interessant, wie in Figur 157 angedeutet. Die Sternschaltung verwendet einen Nulleiter, welcher mit allen drei Spulen verbunden ist, während die nicht mit dem Nulleiter verbundenen Spulenden je mit einem Polleiter verbunden werden.



Figur 157: Mögliche Zusammenschaltungen der drei Quellen.
 Oben: semi-graphischer Darstellung, unten elektrisches Schema.
 Links Sternschaltung mit gemeinsamem ("neutralem") Leiter N, rechts Serie- oder Dreieckschaltung.
 Zusätzlich angegeben: Schutzterde PE ("Protective Earth")

Aufgabe: Zeige dass der Strom im Nulleiter bei symmetrischer Last verschwindet.

Bei nahezu symmetrischer Last ist der Strom im Nulleiter viel kleiner als der Strom in den drei Polleitern. Wird dafür gesorgt, dass die Last ausgeglichen ist, so kann der Nulleiter relativ dünn dimensioniert werden oder im Extremfall wegfallen.

Offensichtlich kommt die Dreiecksschaltung mit lediglich drei Polleitern aus, da hier jeder Phasenleiter mit je zwei Spulenenden kontaktiert ist. Der Nulleiter erübrigt sich.

Die Sternschaltung hat den Vorteil, dass einerseits die Originalspannungen u_1 , u_2 und u_3 zur Verfügung stehen, andererseits aber auch die Spannungen zwischen zwei Polleitern ("verkettete Spannung"), welche der Differenz der beiden Quellenspannungen entspricht, also z.B.

$$\begin{aligned}
 u_{12} &= u_1 - u_2 \\
 &= \hat{u} \sin(\omega_1 t) - \hat{u} \sin(\omega_1 t - 120^\circ) \\
 &= \sqrt{3} \hat{u} \sin(\omega_1 t + 30^\circ)
 \end{aligned}$$

Analog lassen sich die beiden andern Spannungen berechnen.

In der Niederspannungsverteilung wird praktisch ausschliesslich die Sternschaltung verwendet. Einphasige Verbraucher können dabei entweder zwischen einem Polleiter ($L_{1..3}$) und dem gemeinsamen Leiter (Neutralleiter) angeschlossen werden oder aber zwischen zwei Polleitern.

Damit stehen wie erwähnt zwei verschiedene Spannungen zur Verfügung:

- Polleiter-Neutralleiter die Spannung U (= 230V in Europa);
- Polleiter-Polleiter die Spannung $U\sqrt{3}$ (= 400V in Europa).

Kleine Lasten werden in der Regel einphasig angeschlossen, währenddem grössere Verbraucher häufig dreiphasig angeschlossen werden.

Beim Anschluss von einphasigen Lasten werden diese soweit möglich gleichmässig über die Phasen verteilt, so dass jede etwa gleich stark belastet ist.

Aus Sicherheitsgründen werden metallische Gehäuse zusätzlich über die Schutz Erde in Figur 157 mit dem Nullpotential der Erde verbunden. Damit ist gewährleistet, dass im Falle einer Beschädigung das Gehäuse immer auf dem gleichen Potential wie die Erde bleibt und somit eine Personengefährdung vermieden wird.

9.6. Eigenschaften des Produktes "Elektrische Energie" EN 50160:1994

HINWEIS: Der Inhalt dieses Abschnittes ist kein Prüfungstoff, aber eine wichtige Ergänzung.

Die hier noch beigefügte Norm ist eine Regelung betr. die Qualität der elektrischen Energie.

VOLTAGE CHARACTERISTICS OF ELECTRICITY SUPPLIED BY PUBLIC DISTRIBUTION SYSTEMS

Die Vorgeschichte: Im Rahmen der Liberalisierungsbestrebungen der EG wird elektrische Energie als ein Produkt wie viele andere betrachtet, so dass eine Produktnorm notwendig wird, in welcher die relevanten Eigenschaften beschrieben sind und die Grenzwerte für die Abweichungen von den Nenndaten fixiert werden¹. Diese Umschreibung wird vorgenommen, trotzdem bekannt ist, dass "Elektrizität" das wohl einzige Produkt ist, dessen Eigenschaften weitgehend allein durch die Kunden bzw. Verbraucher negativ beeinflusst werden. Die Norm betrifft sowohl öffentliche Niederspannungsnetze (Nennspannung 230/400V) wie auch Mittelspannungsnetze (bis Nennspannung < 35 kV).

Mit Rücksicht auf das Produkthaftpflichtgesetz, in welchen die Elektrizität explizit als Produkt aufgeführt ist, gewinnt allerdings diese Norm eine vollständig neue Bedeutung.

Die wichtigsten Daten im Niederspannungsnetz unter normalen Betriebsbedingungen sind:

Netzfrequenz:

50 Hz \pm 1% während 95% einer Beobachtungszeit von einer Woche, maximal +4 / -6% während 100% einer Woche

Netzspannung

¹ Diese Norm wurde Ende März 1994 angenommen, Sie ist ab Mitte 1995 in Kraft gesetzt.

Nennwert $\pm 10\%$, gemessen als 10-Minuten-Effektivwert, während 95% einer Woche (+10% und -15% über gesamte Messzeit). In Europa ist generell der Wert 230V (bzw. 400V Polleiter-Polleiter) normiert.

(Schnelle) Spannungsschwankungen

Meist kleiner als 5%, unter besonderen Umständen auch 10%

Flickerintensität (Mass für die optische Empfindung von Helligkeitsschwankungen) $P_{Lt} = 1$

Spannungseinbrüche

Definiert als Absenkungen unterhalb 90% der Nennspannung bis zu ca. 1% der Nennspannung während weniger als 3 Minuten (meistens weniger als Absenkung auf 40% während weniger als 1 Sekunde). Man muss je nach Netz mit bis zu 1000 derartigen kurzzeitigen Spannungseinbrüchen rechnen

Kurze Versorgungsunterbrechungen

Absenkungen bis nahe 0% (physikalisch gesehen: Schalteröffnung) und einer Dauer von meistens < 1 Sekunde (maximale Dauer < 3 Minuten) treten einige zehn bis einige hundert Mal im Jahre auf

Lange Versorgungsunterbrüche

Es muss mit 10 bis 30 ungeplanten Unterbrüchen mit einer Dauer über 3 Minuten gerechnet werden

Überspannungen

Netzfrequente Überspannungen können vereinzelt während Schaltoperationen oder bei Erdschlüssen auftreten, übersteigen aber in der Regel 1,5 kV nicht. Transiente (nicht netzfrequente) Überspannungen mit Anstiegszeiten unterhalb Millisekunden überschreiten in der Regel 6 kV nicht; ihr Energieinhalt kann stark variieren

Spannungsunsymmetrie

Jeder 10-Minuten-Effektivwert der Gegenkomponente weist während 95% des Wochenintervalles in der Regel einen Wert auf, der kleiner ist als 2% der Mitkomponente

Oberschwingungen

Die 10-Minuten-Effektivwerte der Spannungsüberschwingungen der ersten 25 Harmonischen (die Komponenten oberhalb der Ordnungszahl 25 sind nicht definiert) überschreiten in der Regel während 95% der Zeit eines Wochenintervalles die Kompatibilitätspegel gemäss IEC 1000-2-2 (siehe Tabelle 2) nicht. Der gesamte Klirrfaktor erreicht unter Berücksichtigung der ersten 40 Harmonischen maximal den Wert von 8%. Individuelle 10-Minuten-Effektivwerte können bis zum 1,5-fachen dieser Werte gehen.

Zwischenharmonische Spannungen

Sie sind generell wesentlich kleiner als die benachbarten Oberschwingungen. Der Wert ist auf jeden Fall unterhalb 0,5%.

Signalspannungen auf der Versorgungsspannung

Für die 3-Sekunden-Effektivwerte sind für 95% des Wochenintervalles Grenzwerte vorgeschlagen, angegeben in % der Nennspannung (9% für 100 - 500 Hz, 5% für 900Hz bis 9 kHz, noch 1.1% bei 95 kHz (die nicht konstanten Bereiche sinken linear in einer doppeltlogarithmischen Darstellung)).

Im Mittelspannungsnetz gelten ähnliche Gesichtspunkte mit allerdings leicht verschiedenen Zahlenwerten.

Eigenschaft	Nennwert	Toleranz/Wertebereich
Frequenz	50 Hz	1%
Spannung	Nennspannung	$\pm 10\%$ (+10 und -15%)
Spannungsänderungen	kleiner 5%, evtl. 10% Flickerintensität $P_{L,t} = 1$	
Spannungseinbrüche	Spannung kleiner 90%	bis 1000
Unterbrüche	Spannung etwa null	bis 30
Symmetrie		besser 2%
Oberschwingungen		wie IEC 1000-2-2
Interharmonische		wenige % bis % der Nennspannung

Tabelle 1: Übersicht über die Grenzwerte gem. EN 50160

Odd harmonics non-multiple of 3		Odd harmonics multiple of 3		Even harmonics	
Harmonic Order h	Harmonic Voltage %	Harmonic Order h	Harmonic Voltage %	Harmonic Order h	Harmonic Voltage %
5	6	3	5	2	2
7	5	9	1.5	4	1
11	3.5	15	0.4	6	0.5
13	3	21	0.3	8	0.5
17	2	$21 < h \leq 45$	0.2	10	0.5
$17 < h \leq 49$	$0.5 + 1.5 \times (17/h)$			$10 < h \leq 50$	0.2

Tabelle 2: Kompatibilitätspegel gem. IEC 1000-2-2

9.7. Elektrische Verbraucher

Elektrische Verbraucher sind im Prinzip Wandler, welche die zugeführte elektrische Energie in andere Energieformen umwandeln oder aber der zugeführten elektrische Energie eine andere Form geben (elektrisch-elektrische Wandler). Zur ersten Gruppe gehören Elektromotoren, elektrische Heizungen und Lampen, während Transformatoren, Gleichrichter und viele andere elektrische Geräte zur zweiten Gruppe gehören. Die Vielfalt der Verbraucher führt auch zu einer Vielfalt von Ansprüchen, welche an das elektrische Netz gestellt werden. Einige davon sollen kurz aufgelistet werden:

- Niedrige Kosten
- Hohe Zuverlässigkeit (keine Unterbrüche der Versorgung)
- Konstante Spannung (möglichst unabhängig von der Last)
- Konstante Frequenz

Das elektrische Netz kann allerdings nicht alle Wünsche der Verbraucher gleichzeitig befriedigen. So ist die Zuverlässigkeit des Netzes in kritischen Bereichen (Spitäler, Banken, etc.) nicht ausreichend und für viele elektronische Geräte wären Gleichspannungen von

wenigen Volt wünschenswert. Es ist deshalb oft eine Umformung der vom Netz zur Verfügung gestellten Energie nötig.

9.8. Netzgeräte

Elektronische Geräte beziehen zwar in der Mehrzahl die zum Betrieb notwendige Energie aus dem öffentlichen Netz, intern werden sie aber mit Gleichspannungen verschiedener Grösse betrieben. Logisch arbeitende Systeme (z.B. solche mit Rechnern) benutzen in der Regel eine Spannung von 5 V, mit Tendenz zu kleineren Werten. Für die analoge Signalverarbeitung wird mit bipolaren Speisespannungen gearbeitet, normalerweise $\pm 15\text{V}$, in Sonderfällen auch mit höheren Spannungen.

Damit stellt sich die Frage, wie aus dem Wechselstrom des Netzes eine Gleichspannung der gewünschten Höhe mit genügender Stabilität gewonnen werden kann.

Zweck des Netzgerätes ist es, aus der Eingangswchselspannung eine geregelte Gleichspannung zu erzeugen. Hierzu werden im wesentlichen folgende drei Komponenten verwendet:

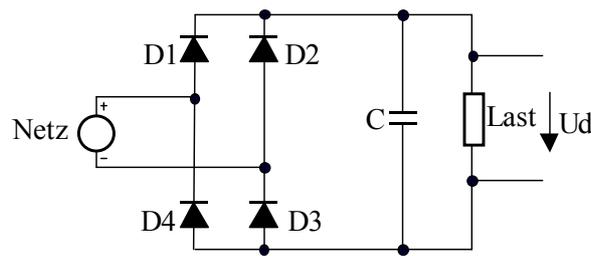
- Transformatoren (zwei magnetisch gekoppelte Spulen) zur Transformation von Wechselspannungen $u_1(t)$ in $u_2(t) = c \cdot u_1(t)$, wobei c konstant
- Gleichrichter, welche im wesentlichen $u(t)$ in $|u(t)|$ umformen
- Elektronische Schaltungen zur Glättung und Regelung erzeugen aus $|u(t)|$ eine nahezu konstante Gleichspannung

Transformatoren werden hier nicht weiter behandelt.

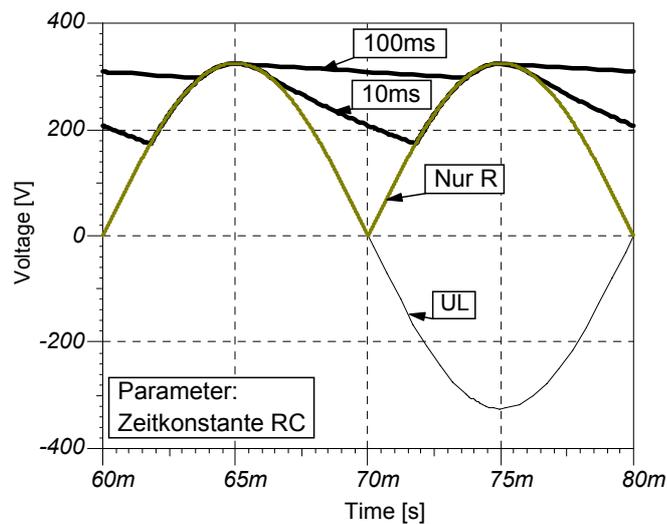
Gleichrichter bestehen im Prinzip aus Schaltern. Heute bestehen Gleichrichter meist einer oder mehreren Dioden. Das Prinzip des besonders häufig verwendeten Brückengleichrichters ist in Figur 158 dargestellt. Für die Erklärung des Funktionsprinzips sei zuerst der Kondensator C nicht anwesend. Ist die (sinusförmig angenommene) Netzspannung positiv, so kann Strom nur durch die Diode D1, die Last und die Diode D3 fließen; D2 und D4 sperren. Ist die Netzspannung negativ, so kann Strom nur über den Weg D2-Last-D4 fließen. über der Last liegt somit immer eine positive Spannung, wie in Figur 11.8 (Kurve "Nur R") angegeben. Man erhält zwar eine stets positive Spannung, die aber zeitlich keineswegs konstant ist. Für den Betrieb von einfachen Geräten (z.B. Gleichstrommotoren) kann dies ausreichend sein, nicht aber für elektronische Geräte.

Um eine zeitkonstante Spannung zu erhalten, wird die gleichgerichtete Wechselspannung geglättet. Im Prinzip werden dabei alle Frequenzen mit einem Tiefpassfilter herausgefiltert. Die einfachste (aber nicht beste) Variante eines Tiefpassfilters besteht aus einem RC-Glied, d.h. einem Widerstand und einem Kondensator. Ist die Last ein Ohm'scher Widerstand, so braucht nur ein Kondensator C eingefügt zu werden (siehe Figur 158). Beim ersten Ein-

schalten ist der Kondensator noch ungeladen; seine Spannung folgt also in der ersten Viertelperiode der anliegenden Netzspannung. Sobald dann nach 90° elektrisch die Netzspannung anfängt abzusinken, tritt der Fall ein, dass sie kleiner wird als die Ladespannung des Kondensators. Damit sperren die Dioden und es fließt vom Netz her kein Ladestrom mehr in den Kondensator; er entlädt sich entsprechend der Zeitkonstante RC über die Last. Dies gilt solange, bis die Netzspannung wieder grösser wird als die Kondensatorspannung, dann setzt wieder Leitung der Dioden ein. Diese Vorgänge sind für zwei verschiedene Zeitkonstanten RC in Figur 159 wiedergegeben.



Figur 158: Prinzipschaltung Einphasiger Zweiweg-Gleichrichter mit kapazitiver Glättung (siehe Text)



Figur 159: Einfluss der Parameter der Glättung auf die Gleichspannung

Aufgabe: Berechne den Mittelwert U_d der Ausgangsspannung für die Extremfälle 1) rein Ohm'sche Last (ohne Glättungskondensator C) und 2) ideale Glättung ($RC \rightarrow \infty$)

Da grosse Kapazitäten teuer und gross sind, wird mit einem RC -Tiefpass normalerweise keine ausreichende Glättung erreicht. Es wird deshalb meist eine elektronische Regelung nachgeschaltet.

Da Transformatoren bei Netzfrequenz (50-60Hz) relativ gross werden, wird oft ein etwas komplizierteres Vorgehen gewählt, welches sich folgendermassen skizzieren lässt:

- Gleichrichtung der Netzspannung
- Glättung der gleichgerichteten Spannung
- Elektronisches Zerhacken (erzeugt eine Rechteckspannung hoher Frequenz (bis ca. 100 kHz) durch wiederholtes Ein/Ausschalten)
- Transformation auf gewünschte Spannung
- Gleichrichtung
- Glättung (hier wegen der hohen Frequenz viel einfacher!)

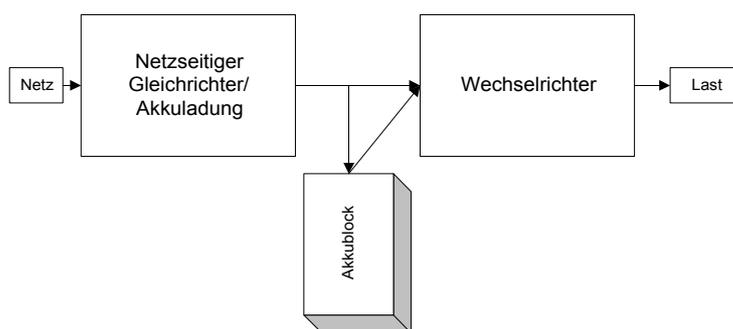
9.9. USV

Auch gut ausgelegte Netzgeräte können nur kurze Spannungsunterbrüche überbrücken. Für länger dauernde muss zu einer besseren Speicherung von elektrischer Energie gegriffen werden, welche mit Akkumulatoren realisierbar ist (sieht man von mitlaufenden Notstromgeneratoren ab). Die notwendige Energie wird also chemisch gespeichert. Nur bei Netzausfall wird darauf zurückgegriffen. Diese Gruppe von Speisungen heisst "Unterbrechungsfreie Stromversorgung (USV)".

Es gibt hier zwei verschiedene Betriebsmodi:

- Normalbetrieb: Das Netz ist vorhanden, die Akkumulatoren werden geladen gehalten;
- Störfall: Das Netz fehlt, die notwendige Energie wird den Akkumulatoren entnommen.

Da allerdings die Geräte für normale Netzspeisung konzipiert sind muss die von den Akkumulatoren erzeugte Gleichspannung in eine ausreichend sinusförmige Wechselspannung der gewünschten Frequenz umgeformt werden.

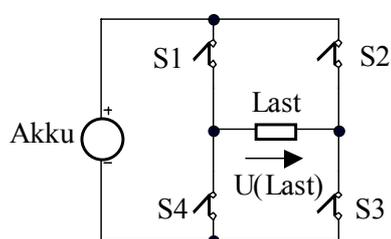


Figur 160: Eine mögliche Struktur einer USV

Der netzseitige Gleichrichter muss, neben der Speisung der Last auch die Ladung des Akkus gewährleisten. Dies erfordert Steuerbarkeit, da je nach Ladezustand des Akkus der La-

destrom (entladener Akku) oder aber die Ladespannung (Halteladung bei geladenem Akku) geregelt werden müssen.

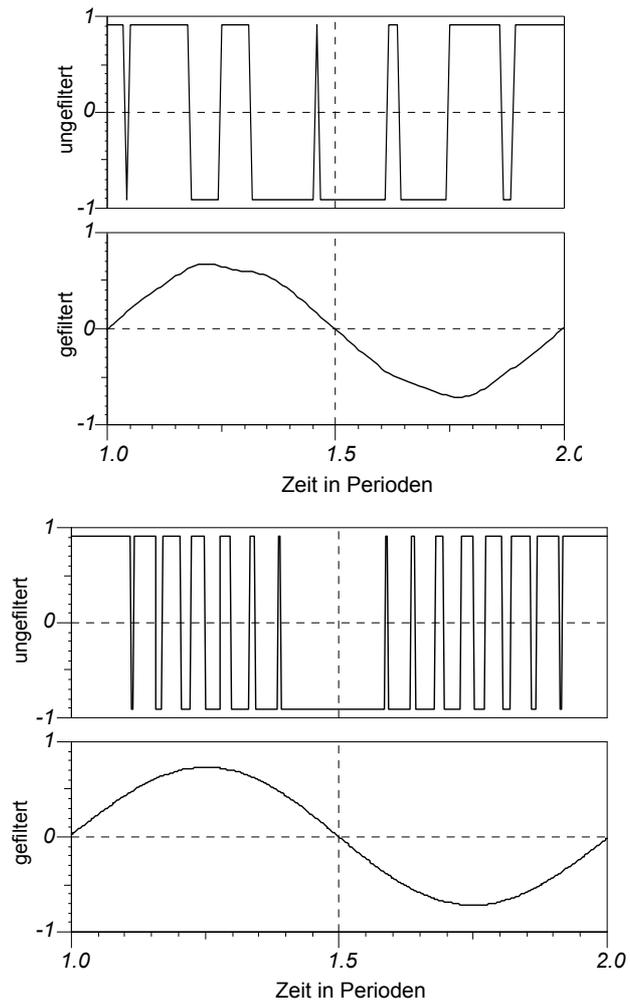
Der ausgangsseitige Wechselrichter formt die Gleichspannung im Zwischenkreis (gleich der Spannung des Akkus) in eine Wechselspannung von z.B. einphasig 230 V / 50 Hz um, mit welcher dann die Lasten gespeist werden. Es wäre naheliegend und billigst, aus der Gleichspannung einfach eine rechteckförmige Wechselspannung zu erzeugen; dies kann aber je nach angeschlossenem Gerät zu Funktionsstörungen führen. Bei moderneren USV wird daher ein angenähert sinusförmiger Strom mittels Pulsweitenmodulation (PWM) der Spannung und anschließende Filterung erzeugt. Die hierzu verwendete Schaltung ist in Figur 161 prinzipiell wiedergegeben.



Figur 161: Prinzipschaltung eines Wechselrichters

Sind die Schalter S1 und S3 geschlossen, so liegt eine positive Spannung über der Last; beim Schliessen von S2 und S4 eine negative.

Das Ergebnis ist in Figur 162 angedeutet. Die dort jeweils oben angegebene Spannung ist direkt die Ausgangsspannung des Wechselrichters, durch eine relativ wenig aufwendige Filterung wird daraus die darunter dargestellte sinusförmige Spannung erzeugt.



Figur 162: Ausgangsspannung (ungefiltert und gefiltert): Oben nur Prinzip, unten realitätsnäher: eine Periode wird durch wesentlich mehr Pulse angenähert und damit die Spannung wesentlich "besser". Durch die Filterung tritt eine Phasenverschiebung auf.

10. Messtechnik

Messtechnik ist ein sehr weites, anspruchsvolles Gebiet, welches hier nur andeutungsweise behandelt werden kann. Informatiker werden meist an drei Orten mit der Messtechnik in Berührung kommen:

- bei der Inbetriebnahme von neuen Rechnersystemen oder bei der Fehlersuche in solchen;
- beim Einsatz von Rechnern für die Verarbeitung von Messdaten;
- beim Einsatz von Rechnern für die Steuerung und Regelung von Prozessen.

In allen Fällen ist es notwendig, die Prinzipien und grundlegenden Begriffe der verwendeten Werkzeuge zu kennen. In diesem Kapitel sollen die dazu notwendigen Grundlagen erarbeitet werden.

10.1 Einleitung und Grundlagen

Beim Messen werden physikalische Grössen quantifiziert, d.h. Zahlenwerte zugewiesen. Messung ergibt damit automatisch einen direkten Zugang von der Physik und andern Disziplinen der Naturwissenschaften zur Informatik. Da bei Messungen riesige Datenmengen gehören, besteht ein Teil der Messtechnik in der Kunst der Datenverarbeitung und Datenkompression.

Spätestens seit Heisenberg weiss man, dass der Messvorgang einen schwerwiegenden Eingriff in das beobachtete physikalische System bedeutet. Messtechnik beinhaltet auch die Kunst, die Folgen dieses Eingriffs auf das System entweder vernachlässigbar klein zu halten oder durch Nachbearbeitung der Messdaten zu eliminieren.

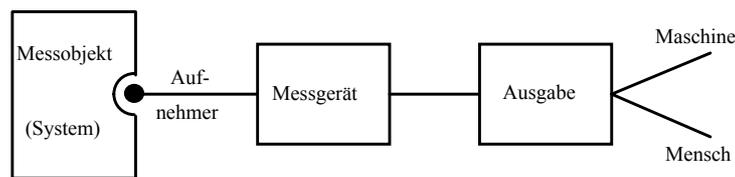
Messungen werden aus verschiedenartigsten Gründen vorgenommen und diese haben schliesslich einen beträchtlichen Einfluss auf die Anforderungen, welche an die Messung gestellt werden.

- Höchste Anforderungen an die Messgenauigkeit werden in der Physik gestellt, wenn durch Messungen ein physikalisches Gesetz verifiziert werden soll. Dabei können der apparative Messaufwand und die Messdauer riesig gross werden. Vielfach wird dabei nicht das Gesetz selber, sondern eine Konsequenz des Gesetzes beobachtet. Ein typisches Beispiel ist die Verifikation des Coulomb'schen Gesetzes mit dem Cavendish Experiment.
- Bei der Entwicklung, Optimierung und Fabrikation technischer Geräte wird mit Messungen die Funktionsweise und Qualität geprüft. Die Anforderungen an die Messgenauigkeit sind dabei beträchtlich geringer, ebenso aber auch die Zeit, welche für diese Messungen zur Verfügung steht. Insbesondere bei der Qualitätskontrolle sollte der

Messvorgang relativ kurz und vollautomatisch vonstatten gehen. Dies gilt insbesondere für die Herstellung von Massenprodukten.

- Bei der Regelung von Prozessen ist die Messgenauigkeit oft unkritisch, die Messzeit kann aber bei schnell laufenden Prozessen extrem kurz sein, wodurch an die Messsensoren und die anschliessende Datenverarbeitung hohe Anforderungen gestellt werden. Ein Beispiel ist die elektronische Regelung von Verbrennungsmotoren.

Die Durchführung einer Messung erfordert im Allgemeinen eine Wirkkette gemäss folgendem Bild:



Figur 163: Wirkkette einer Messeinrichtung

- Messobjekt mit zu messender physikalischer Eigenschaft
- Aufnehmer, welcher die Verbindung zwischen Messobjekt und Messgerät mit Transformation der physikalischen Grössen, vornimmt ("Sensor")
- Messgerät, das die eigentliche Umwandlung der zu messenden Grössen in einen Zahlenwert vornimmt
- Ausgabeeinheit, welche den Zahlenwert in einer für den Benutzer sinnvollen Art darstellt

Sensoren sind im Prinzip physikalische Wandler, welche dem Messsystem Energie entnehmen und diese in eine dem Messgerät adäquate Form umwandeln. Für elektronische Messgeräte sind die Ausgabegrössen des Sensors elektrische Ströme und Spannungen. Um das System möglichst wenig zu stören, wird — im Gegensatz zu den Wandlern in der Energietechnik — versucht, mit möglichst geringen Energien auszukommen. Die Ströme und Spannungen, welche der Sensor dem Messgerät liefert sind deshalb meist sehr klein. Als Folge davon muss die Messleitung zwischen Sensor und Messgerät oft sehr sorgfältig ausgelegt werden, da diese sonst leicht gestört werden kann. Als weitere Folge werden oft Messverstärker erforderlich, an welche hohen Anforderungen gestellt werden.

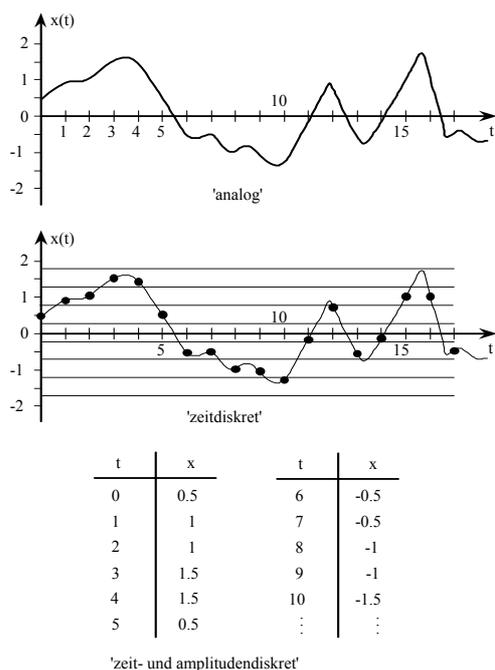
Sensoren können höchst komplexe Objekte sein, deren Entwicklung viel physikalisches Wissen erfordert; sie können aber auch so primitiv und einfach sein, dass sie kaum wahrgenommen werden. Letzteres gilt insbesondere bei der elektronischen Messung von Strömen und Spannungen bei nicht sehr hohen Frequenzen, wo die Messleitung aus einfachen Drähten besteht, welche mit den Messpunkten im zu messenden System kontaktiert werden. Die Messkontakte sind dann „Sensoren“.

Im eigentlichen Messgerät wird die zu messende Grösse mit einer bekannten Grösse verglichen. Das Messgerät ist also im Prinzip ein Komparator. Ein typisches Beispiel ist die Balkenwaage zur Messung von Gewichten. Das Ergebnis dieses Vergleichs muss schliesslich in eine Form gebracht werden, die dem Benutzer angepasst ist, zum Beispiel in einen Zeigerausschlag oder in eine dezimale oder binäre Zahl (beispielsweise für Rechner). Dabei kommen wieder Wandler zum Einsatz. Messgeräte bestehen deshalb aus Verstärkern, Komparatoren und Wandlern, wobei die Verstärker in einfacheren Fällen fehlen.

10.2 Zeitabhängigkeit der Messgrössen

Zunächst erfordert jede Messung Zeit. Insbesondere wenn kleine Messgrössen mit hoher Genauigkeit gemessen werden sollen, so kann die Messung beträchtlich lange Zeit erfordern. Dies ist wohl der Grund warum sich Physiker in ihren Untersuchungen gerne mit statischen Phänomenen beschäftigen. Obwohl mit zunehmender Messdauer und Messaufwand die Messgenauigkeit einer statischen Grösse f erhöht werden kann, ist eine exakte Messung einer reellwertigen Grösse prinzipiell ausgeschlossen, da eine reelle Zahl einen unendlichen Informationsgehalt hat. Der Messwert f_0 ist also immer eine Näherung des physikalischen Wertes f . Der absolute Messfehler $e_a = f_0 - f$ und der relative Messfehler $e_r = (f_0 - f)/f$ lassen sich nur aus der Konstruktion der Messanordnung abschätzen. Genauigkeitsangaben auf Messgeräten sind mit Vorsicht zu geniessen, da meist durch die Messleitung, den Sensor und den Eingriff in das zu messenden System weitere Fehler gemacht werden.

In der Praxis sind Messgrössen stets zeitveränderlich, d.h. Funktionen der Zeit $f(t)$, wobei die Zeit selber eine zu messende Grösse ist. Da jede Messung eine minimale Zeit Δt in Anspruch nimmt und nur endliche Informationsmengen verarbeitet werden können, führt eine Messung von $f(t)$ stets auf ein endliches, diskretes Set von Messwerten $f_i(t_i)$, $i=1, 2, \dots, N$, wobei die Zeitwerte ebenso wie die Funktionswerte fehlerbehaftet, d.h. ungenau sind.



Figur 164: Signaldarstellung

Die Unmöglichkeit, eine Zeitfunktion exakt zu messen, ist zwar für Physiker unangenehm, kommt aber dem Informatiker gelegen. Allerdings ist der Informationsgehalt von gemessenen Zeitfunktionen $f(t)$ oft so gross, dass eine Datenreduktion erforderlich wird. Man sucht deshalb gerne nach aussagekräftigen, charakteristischen Grössen und speziellen Eigenschaften von $f(t)$, welche die Beschreibung vereinfachen.

Gleichmässige Drehbewegungen und ähnliche, häufig anzutreffende physikalische Erscheinungen führen zu Zeitfunktionen, welche in guter Näherung periodisch sind. Der Informationsgehalt einer periodischen Funktion $f_p(t)$ ist zwar im Prinzip immer noch unendlich, jedoch reicht zu deren Beschreibung eine einzige Periode T , weil $f_p(t+k \cdot T) = f_p(t)$ für alle ganzzahligen k gilt. Das erforderliche Beobachtungsintervall T ist also endlich.

Anstelle der Periodendauer T misst man oft die Anzahl Perioden pro Zeiteinheit, d.h. die Frequenz $f = 1/T$. Dies ist in der Praxis ein Zählvorgang, der insbesondere bei kurzen Perioden einfacher und genauer ist als die direkte Messung von T . Oszillatoren erzeugen häufig nahezu sinusförmige Signale der Form $f(t) = A \cdot \sin(\omega t + \varphi)$, wobei $\omega = 2\pi/T$ die Kreisfrequenz und φ die Phase bezeichnet. Die Charakterisierung der Periodizität mit der Kreisfrequenz ω vereinfacht vor allem die mathematische Beschreibung.

In der Messtechnik wird übrigens die Periodizität der zu messenden Funktion $f_p(t)$ üblicherweise nicht so ausgenutzt werden, dass nur während einer einzigen Periode gemessen wird, da dadurch keine hohe Genauigkeit erzielt wird und ausserdem bei hohen Frequenzen Probleme entstehen. So wird beispielsweise während jeder Periode nur ein einziger Zeitwert $f_i(t_i)$ gemessen („abgetastet“) und so die gesamte Information über mehrere Perioden gesammelt — das ist insbesondere bei hohen Frequenzen hilfreich. Dies ist das Prinzip von

Sampling-Oszillographen und der Sichtbarmachung schneller, periodischer Bewegungen mit Stroboskopen.

Beliebige, periodische Funktionen $f^p(t)$ können bekanntlich in Fourier-Reihen der Form

$$f_p(t) = a_0/2 + \sum a_n \cos n\omega_0 t + \sum b_n \sin n\omega_0 t$$

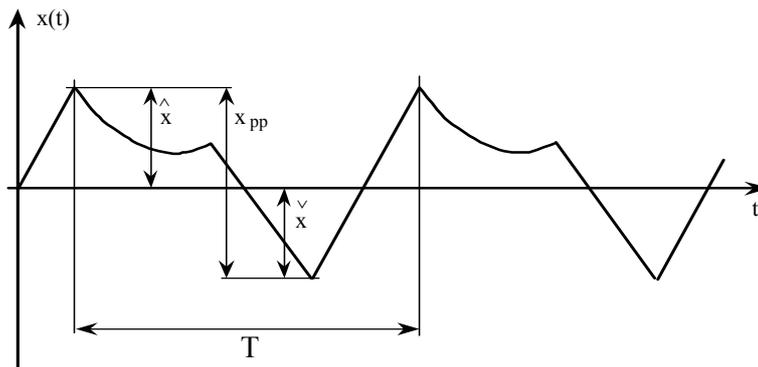
entwickelt werden, wobei für die Grundfrequenz $\omega_0 = 2/T$ und für die Fourierkoeffizienten a_n und b_n

$$a_n = \frac{2}{T} \int_0^T f_p(t) \cos(n\omega_0 t) dt, \quad b_n = \frac{2}{T} \int_0^T f_p(t) \sin(n\omega_0 t) dt$$

gilt. In vielen Fällen ergibt sich eine ausreichend genaue Beschreibung der gegebenen, periodischen Funktion durch Angabe von relativ wenigen Fourierkoeffizienten, so dass die Charakterisierung durch das Fourierspektrum — d.h. der Werte der Fourierkoeffizienten — einer Datenreduktion entspricht. Das Spektrum ist übrigens diskret, d.h. es treten nur diskrete Frequenzen $\omega = n\omega_0$ auf.

Mit dem Grenzübergang $T \rightarrow \infty$ werden auch nichtperiodische Funktionen der Fourieranalyse zugänglich gemacht. Dabei geht die Grundfrequenz ω_0 bei den Fourierreihen gegen 0. Die Summen müssen dann durch Integrale und die Koeffizienten a_n, b_n durch Funktionen $a(\omega), b(\omega)$ ersetzt werden. D.h. man erhält anstelle des diskreten Fourierspektrums ein kontinuierliches Spektrum. Es mag zunächst der Eindruck entstehen, die Beschreibung beliebiger Zeitfunktionen mit Hilfe von Fourierspektren sei eher kompliziert als hilfreich. In der Praxis erweisen sich viele Signale jedoch als relativ schmalbandig, d.h., das betreffende Fourierspektrum ist nur in einem relativ engen Frequenzbereich $\omega_1 < \omega < \omega_2$ wesentlich von 0 verschieden, womit wieder eine Datenreduktion erreicht werden kann. Darüber hinaus sind spektrale Darstellungen wichtig, weil viele Bauelemente durch ihren Frequenzgang charakterisiert werden. Dies gilt insbesondere für Filter. Man erinnere sich, dass Filter vornehmlich aus Kapazitäten und Induktivitäten aufgebaut werden, welche im Zeitbereich durch Differentiale und Integrale beschrieben werden. Die Fouriertransformation ersetzt diese im Frequenzbereich durch einfache Multiplikationen und Divisionen. Filter sind übrigens auch die passenden Bauelemente für die messtechnische Erfassung von Fourierspektren (Spektrum-Analysator).

Insbesondere in der Digitaltechnik und in der Energietechnik ist die detaillierte Erfassung der Zeitfunktionen oft überflüssig und es reicht die Angabe einiger weniger charakteristischer Daten, insbesondere Spitzen- und Mittelwerte. Diese charakteristischen Grössen sind im folgenden Bild am Beispiel eines periodischen Signals angegeben.



Figur165 Charakteristische Werte eines periodischen Signals

- **Spitzenwerte**

\hat{x} : Gipfelwert, positiver Spitzenwert, pos. peak-Wert

\check{x} : Talwert, negativer Spitzenwert, neg. peak-Wert (Anmerkung: bei sinusförmigen Größen spricht man auch von Scheitelwert bzw. Sohlenwert)

x_{pp} Spitze-Spitze-Wert, peak-to-peak-Wert, Schwingungsbreite

- **Mittelwerte:**

Von den vielen möglichen Mittelwertbildungen sind folgende üblich:

1) Der lineare Mittelwert:

$$F_0 = \frac{1}{T} \int_0^T f(t) dt$$

2) Der Betragsmittelwert (aus messtechnischen Gründen häufig benutzt):

$$F_r = \frac{1}{T} \int_0^T |f(t)| dt$$

3) Der quadratische Mittelwert:

$$F^2 = \overline{f^2} = \frac{1}{T} \int_0^T f^2(t) dt$$

Dabei ist T bei periodischen Funktionen die Periodendauer (oder auch ein ganzzahliges Vielfaches davon) und bei beliebigen Zeitfunktionen das Messintervall.

Im Falle der Fourier-Reihen ist offenbar der Koeffizient $a_0 = 2F_0$.

Für die in einem elektrischen Widerstand R umgesetzte Leistung $p(t)$ gilt beispielsweise

$$p(t) = R \cdot i^2(t)$$

Der lineare Mittelwert der Leistung ist damit verknüpft mit dem quadratischen Mittelwert des Stromes $i(t)$:

$$P = \frac{1}{T} \int_0^T p(t) dt = \frac{1}{T} \int_0^T Ri^2 dt = \overline{Ri^2}.$$

Aufgabe: Zeige dasselbe sinngemäss für die Spannung $u(t)$.

Üblicherweise gibt man allerdings nicht den quadratischen Mittelwert an, sondern seine positive Wurzel, genannt Effektivwert:

$$F = \sqrt{\frac{1}{T} \int_0^T f^2(t) dt}.$$

Im englischen Sprachgebrauch wird übrigens der Effektivwert "root mean square value" (abgekürzt: rms-value) genannt.

Aufgabe: Zeige, dass bei sinusförmigen Grössen für den Effektivwert $F = \hat{f} / \sqrt{2}$ gilt.

10.3 Messverfahren und Messgeräte

Angesichts der enormen Vielfalt von Verfahren und Geräten, welche im Laufe der letzten Jahrhunderte entwickelt wurden, ist selbst eine kurze Übersicht hier unmöglich. Wir beschränken uns deshalb auf einen kurzen Abriss der in der elektrischen Schaltungstechnik wichtigsten Verfahren und Geräte. Das heisst, dass wir uns im Folgenden auf die Messung von elektrischen Spannungen und Strömen in leitenden Drähten konzentrieren.

10.3.1 Drehspulinstrumente

Drehspulinstrumente werden zwar immer mehr von elektronischen Geräten verdrängt, verdeutlichen wegen ihrer Einfachheit jedoch die verschiedenen Probleme der Messtechnik besonders gut. Grundlage der Drehspulinstrumente ist die Tatsache, dass auf einen stromdurchflossenen Draht in einem Magnetfeld eine Kraft ausgeübt wird, welche proportional zur Stromstärke ist. Diese Kraft kann mit einem Kraftmessgerät gemessen werden, z.B. mit einer Federwaage. Um die Kraft zu verstärken, kann man das Magnetfeld vergrössern — was z.B. grosse, d.h. schwere Permanentmagnete erfordern würde — oder den Strom durch eine Spule mit vielen Windungen führen, was auch beim Elektromotor gemacht wird. Ein Drehspulinstrument wird tatsächlich ähnlich wie ein Elektromotor aufgebaut. Üblicherweise wird eine drehbare Spule verwendet, an deren Achse ein Zeiger und eine Spiralfeder

montiert werden. Letztere dient als einfache Federwaage. Ist die Auslenkung des Zeigers proportional zur Federkraft, so wird diese auch proportional zum Strom in der Spule, womit der Zeiger auf einer linearen Skala die Stromstärke angibt.

Offenbar kann die Ablesegenauigkeit durch Verlängerung der Zeiger relativ leicht gesteigert werden, dies führt aber nicht nur zu grösseren Instrumenten, sondern auch zu physikalischen Problemen, wie Trägheit der Anzeige, mechanische Stabilität, Reibung in den Lagern der Achse usw. Zur Verbesserung kann beispielsweise der Zeiger durch einen kleinen Spiegel ersetzt werden, welcher von einem Laserstrahl beleuchtet wird, welcher wiederum auf die Anzeige gelenkt wird. Obwohl sich dadurch die Genauigkeit verbessern und die Trägheit der Anzeige verringern lässt, so bleibt das Drehspulinstrument doch ein relativ träges Instrument, welches nur für die Messung von langsam veränderlichen Strömen in Frage kommt.

Ganz andere Probleme entstehen durch die endliche Auslenkung, welche den Messbereich einschränkt, Nichtlinearität der Federkraft, Alterung der Feder usw. Schliesslich kann ein Drehspulinstrument durch mechanische Beanspruchung, ebenso wie durch zu grosse Ströme irreversiblen Schaden nehmen.

Nehmen wir an, der maximal zulässige Ausschlag eines Drehspulinstrumentes sei 1 mA, so muss der Benutzer dafür sorgen, dass dem Instrument keine deutlich höheren Ströme zugeführt werden. Wie kann er dies tun, wenn er die Grösse des zu messenden Stromes nicht kennt? Natürlich könnte er mit Messgeräten für höhere Ströme beginnen und diese solange durch empfindlichere Geräte ersetzen bis eine Anzeige sichtbar wird. Dies ist aber umständlich und teuer. Betrachtet man das elektrische Ersatzschaltbild eines Drehspulinstrumentes, so findet man einfache Möglichkeiten der Messbereichserweiterung. Wie jede reale Spule kann auch die Messspule des Drehspulinstrumentes durch eine Serienschaltung eines Widerstand R_S mit einer Induktivität L_S in guter Näherung dargestellt werden. Bei langsam veränderlichen Strömen kann L_S vernachlässigt werden. Damit ist es möglich, durch Zuschalten zusätzlicher, externer Widerstände den Messbereich beliebig zu erweitern.

Aufgabe: Finde verschiedene Möglichkeiten um den Messbereich eines Drehspulinstrumentes (Innenwiderstand R_S , maximaler Strom I_{max}) mit 1) einem und 2) zwei zusätzlichen Widerständen um das Zehnfache zu erweitern.

Durch die Messbereichserweiterung mit Widerständen wird üblicherweise die Messgenauigkeit vermindert, da präzise Widerstände sehr teuer sind.

Aufgabe: Schätze die Messgenauigkeit der in der vorhergehenden Aufgabe vorgeschlagenen Instrumente, wenn die Messgenauigkeit des Drehspulinstrumentes 1% und die Genauigkeit der Widerstände 2% beträgt.

Wegen des Ohm'schen Gesetzes lassen sich mit dem Drehspulinstrument nicht nur Ströme, sondern auch Spannungen messen. Dazu muss lediglich die Anzeige ausgewechselt werden, wobei in guter Näherung $U_{max} = R_S \cdot I_{max}$ gilt. Was misst denn nun das Drehspulinstrument eigentlich? Offenbar wird im Drehspulinstrument elektrische Energie verbraucht,

bzw. in Wärme umgeformt. Es handelt sich also um einen Wandler, der nicht ohne Energie auskommt. Diese Energie wird der elektrischen Schaltung entzogen, welche ausgemessen wird. Damit ist klar, dass die Messung auf die Schaltung zurück wirkt und die Ströme und Spannungen in dieser verändert. Damit diese Beeinflussung möglichst klein ist, sollte der Widerstand R_S für eine Strommessung möglichst klein und für eine Spannungsmessung möglichst gross sein. Bei der Messbereichserweiterung ist es deshalb sinnvoll, bei Strommessungen Parallelwiderstände zu R_S und bei Spannungsmessungen Seriewiderstände zu R_S hinzuzuschalten.

Aufgabe: Schalte zu einem Drehspulinstrument mit gegebenen R_S und I_{max} einen Widerstand R so hinzu, dass ein möglichst gutes Messgerät für Spannungen bis zu U_{max} entsteht.

Da $R = U/I$ gilt und mit einem Drehspulinstrument U und I gemessen werden können, lässt sich indirekt auch der Widerstand R messen. Als Alternative dazu wird allerdings gerne eine Widerstandsmessbrücke verwendet. Dabei wird im Prinzip der unbekannte Widerstand mit bekannten Widerständen verglichen (Komparator). Das Drehspulinstrument dient dann dem Nullabgleich und dessen Innenwiderstand R_S hat keinen Einfluss auf die Messung.

Ist der Innenwiderstand R_S eines Drehspulinstrumentes bekannt, so lässt sich der dadurch entstehende Messfehler rechnerisch eliminieren.

Aufgabe: Zur Widerstandsmessung wird ein unbekannter Widerstand R mit einer Spannungsquelle U_q mit (unbekanntem) Innenwiderstand R_q verbunden. Zur Verfügung stehen ein Strommessgerät mit Innenwiderstand R_I und ein Spannungsmessgerät R_U , deren Messbereich für die Messung angepasst sei. Welche Widerstandswerte misst man, wenn man die Innenwiderstände der Messgeräte nicht berücksichtigt und wie folgt misst:

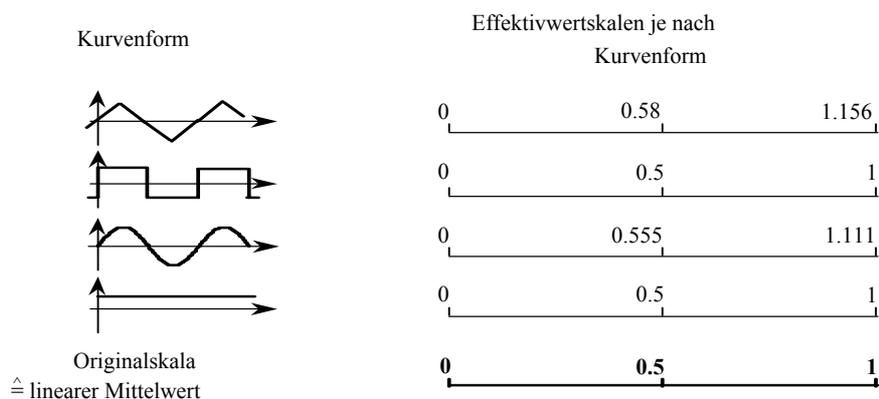
- 1) Messe den Strom durch R ohne gleichzeitig die Spannung über R zu messen und messe anschliessend nur die Spannung über R .
- 2) Messe Strom und Spannung gleichzeitig, wobei
 - a) das Strommessgerät direkt an R angeschlossen und das Spannungsmessgerät an der Spannungsquelle angeschlossen wird und
 - b) das Strommessgerät an der Spannungsquelle und das Spannungsmessgerät über dem Widerstand kontaktiert wird. Welches Verfahren ist das genaueste? Wie ist R — unter Berücksichtigung der Innenwiderstände der Messgeräte — für die drei Varianten zu berechnen?

Bei periodischen Vorgängen mit nicht zu tiefer Frequenz misst das Drehspulinstrument wegen seiner mechanischen Trägheit den linearen Mittelwert. Durch Zuschalten eines Gleichrichters (Diode) lässt sich der Betragsmittelwert messen. Dabei ist allerdings zu beachten, dass Dioden keine idealen Gleichrichter sondern nichtlineare Bauelemente sind.

Aufgabe: Eine Diode wird zur Gleichrichtung zu einem Strommessgerät mit Innenwiderstand R_S in Serie geschaltet. Im Durchlassbereich kann die Diode in guter Näherung durch eine Spannungsquelle U_D approximiert werden, im Sperrbereich sperrt sie in guter Näherung ideal. Nun wird der Strom durch einen Widerstand R , gemessen, welcher an eine

Spannungsquelle mit $u_q(t) = U_q \sin(\omega t)$ und Innenwiderstand R_q angeschlossen wird. Was zeigt das Messinstrument an, wenn ω so hoch ist, dass der Messzeiger nicht oszilliert und so tief, dass die Induktivität des Messinstrumentes vernachlässigbar ist?

Drehspulinstrumente werden oft mit zuschaltbaren Dioden und Widerständen versehen und damit zu „Universalinstrumenten“ mit umschaltbarem Messbereich für verschiedene Ströme oder Spannungen. Ist eine Diode oder ein Gleichrichter zugeschaltet, so werden üblicherweise Effektivwerte angezeigt. Diese gelten aber nur bei sinusförmigen Messgrößen! Figur 166 zeigt Effektivwertskalen für einige einfache Kurvenformen, wenn ein idealer Brückengleichrichter verwendet wird.



Figur 166: Umrechnungsfaktoren Betragsmittelwert zu Effektivwert

Aufgabe: Berechne die Effektivwertskala für die Kurven in Figur 166 wenn eine einzige ideale Diode zur Gleichrichtung (Einweggleichrichter) verwendet wird.

10.3.2 Digitalmessgeräte

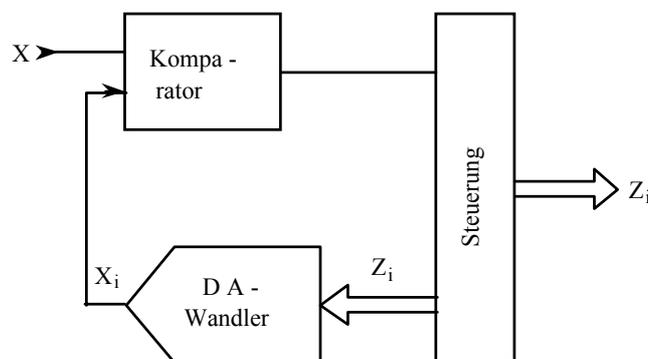
Um eine (analoge) Messgröße digital auszugeben, werden Analog/Digital Wandler, abgekürzt A/D-Wandler, eingesetzt. Diese sind im Prinzip immer elektronische Komparatoren. Ist beispielsweise die Messgröße eine Spannung U_{in} , so wird im A/D-Wandler eine Spannung U_k (z.B. $U_k = k \cdot U_0$ mit k ganzzahlig — üblicherweise intern binär dargestellt — und U_0 ausreichend klein) erzeugt, welche so lange verändert wird, bis sie praktisch gleich U_{in} ist (z.B. gilt dann $k \cdot U_0 < U_{in} < (k+1) \cdot U_0$).

Ein besonders einfaches Verfahren mit $U_k = k \cdot U_0$ besteht darin, bei $k = 0$ zu beginnen und anschliessend k solange zu erhöhen, bis $U_{in} < U_k$ gilt. Für Informatiker ist es sicher sofort klar, dass dies kein effizienter Suchalgorithmus ist.

Aufgabe: Suche nach effizienteren Versionen, mit welchen sich k in möglichst wenig Schritten finden lässt und zwar wenn $k = 1, 2, \dots, 2^N$ ist und die Wahrscheinlichkeit für das Auftreten für einen beliebigen Wert $k = 1/2^N$ ist. Wie sieht ein effizientes Verfahren aus,

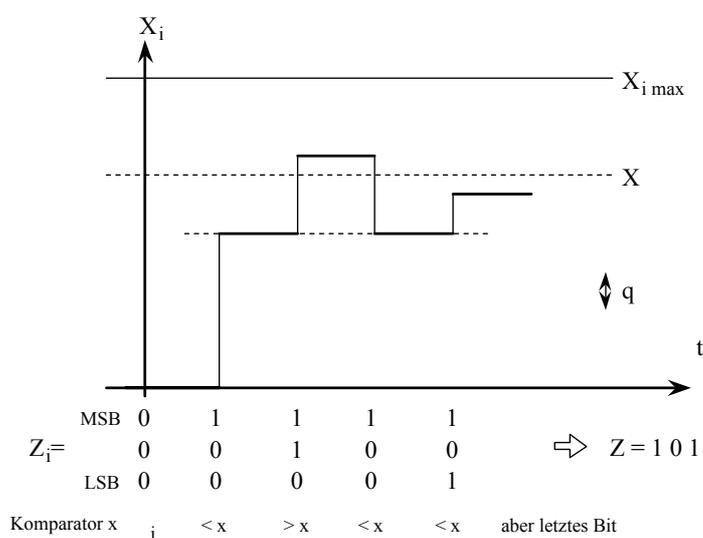
wenn mehrere Messungen $k_1, k_2, k_3, \dots, k_i$ hintereinander ausgeführt werden und bekannt ist, dass sich die Eingabegrösse nur langsam ändert?

Der derzeit am häufigsten verwendete A/D-Wandler-Typ ist der successive-approximation Wandler, dessen Prinzip in Figur 12.5 bzw. Figur 12.6 wiedergegeben ist.



Figur 167: Struktur eines successive-approximation Wandlers

Das Steuerwerk setzt sukzessive alle Bits von Z_i auf 1 und erhält dann vom Komparator die Information, ob $X_i > X$ oder $X_i < X$: im Fall $X_i > X$ wird das zuletzt "1" gesetzte Bit wieder gelöscht.



Figur 168: Verfahren der successive approximation

Heute übliche A/D-Wandler haben Auflösungen zwischen 8 und 10 bit ($\hat{=} 2^8 = 256$ bzw. $2^{10} = 1024$ Stufen), nur in Spezialfällen werden Auflösungen von 12 bit (4096 Stufen) oder sogar 16 bit (65536 Stufen) verlangt. In der elektronischen Messtechnik sind so hohe Auflösungen wenig gebräuchlich, da dann auch sämtliche analoge Vorstufen entsprechende Genauigkeit aufweisen müssten und so hohe Genauigkeiten meist nicht erforderlich sind.

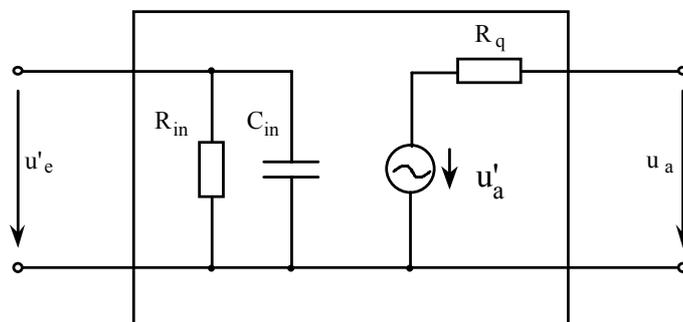
Der successive approximation Wandler enthält übrigens einen D/A-Wandler, das Gegenstück zum A/D-Wandler. Dieser bewirkt die Rückkonversion einer Zahl in eine elektrische Spannung. Im Prinzip werden dabei die Spannungswerte $U_0 \cdot 2^n$, welche zu einem Bit gehören, welches gleich eins ist, summiert. Ist beispielsweise $k = 5 = 1 \cdot 2^2 + 0 \cdot 2^1 + 1 \cdot 2^0$, so wird $U_k = U_0 \cdot 2^2 + 0 \cdot 2^1 + U_0 \cdot 2^0$. Wichtig ist dabei, dass die Spannungen $U_0 \cdot 2^n$ mit hoher Genauigkeit und Konstanz (z.B. geringer Temperaturabhängigkeit) erzeugt werden.

10.3.3 Messverstärker

Genauso wie Drehspulinstrumente haben A/D-Wandler einen beschränkten Messbereich. Wir haben gesehen, dass sich dieser mit einfachen Massnahmen (Widerständen) nach oben erweitern lässt. Bei A/D-Wandlern ist oft eine höhere Empfindlichkeit erwünscht, da der Standardmessbereich (Grössenordnung 10 V) relativ gross ist. Drehspulinstrumente haben oft einen wesentlich empfindlicheren Eingang, jedoch für Spannungsmessungen einen zu geringen Innenwiderstand. Mit Messverstärkern kann in beiden Fällen Abhilfe geschafft werden.

Messverstärker sind meist sehr zuverlässige und genaue Operationsverstärker mit hohem Eingangswiderstand. Durch zusätzliche Beschaltung mit Präzisionswiderständen kann der Messbereich — ähnlich wie bei Drehspulinstrumenten — eingestellt werden. Wichtig ist, dass bei hohen Frequenzen eine Impedanzanpassung von Messsonde, Messleitung und Eingang des Messverstärkers erforderlich ist, da es sonst zu unerwünschten Reflexionen des Messsignals kommt. Da die Messleitungen meist auf 50 Ohm Wellenwiderstand ausgelegt werden, muss dann auch der Eingangswiderstand des Messverstärkers auf 50 Ohm gelegt werden.

Reale (Spannungs-) Verstärker entnehmen dem Messobjekt immer eine gewisse Leistung. Diese Belastung kann genügend genau durch einen ohmschen Widerstand R_{in} mit dazu Parallelgeschalteter Kapazität C_{in} dargestellt werden. Desgleichen ist auch ausgangsseitig der Verstärker nicht ideal, sondern in erster Näherung als reale gesteuerte Quelle mit Innenwiderstand R_q aufzufassen. Damit ergibt sich folgendes Ersatzschaltbild für den realen Verstärker:



Figur169: Ersatzschema eines realen Verstärkers

Typische Werte sind:

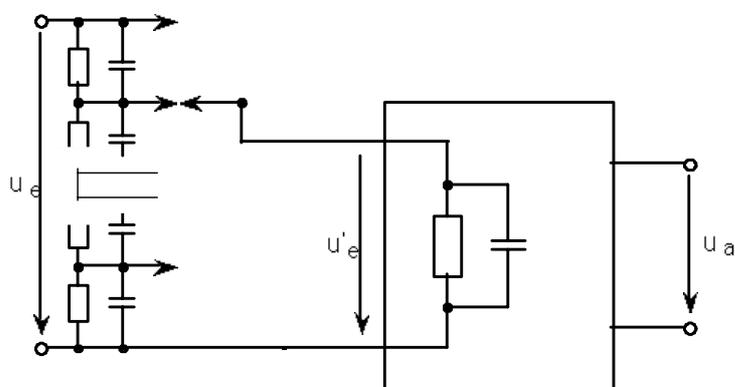
R_{in} : 1 ... 100 M Ω (universelle Anwendung); 50 Ω ("HF": Hochfrequenztechnik)

C_{in} : 20 ... 50 pF

R_q : einige Ω ... 50 Ω ... 600 Ω

Wichtig für Messverstärker sind neben Linearität und Genauigkeit auch Temperaturunabhängigkeit, Störungssicherheit, dynamisches Verhalten und Frequenzgang.

Zur Messbereichserweiterung können — ähnlich wie bei Drehspulinstrumenten — umschaltbare Widerstandsketten verwendet werden. Sollen die Eingangskapazitäten berücksichtigt werden, ergibt sich beispielsweise:

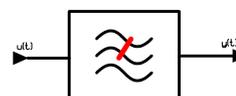


Figur 170; Verstärker für weiten Eingangsspannungsbereich, links kompensierter sprunghafter Teiler vor dem eigentlichen Verstärker

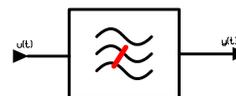
10.3.4 Filter

Filter sind ein wichtiges Hilfsmittel in der Signalverarbeitung. Von speziellem Interesse sind:

Tiefpass Nur Frequenzkomponenten zwischen 0 und einer oberen Grenzfrequenz werden durchgelassen, die übrigen Komponenten werden unterdrückt.



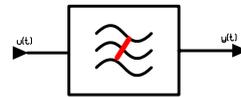
Hochpass Nur Frequenzkomponenten oberhalb einer unteren Grenzfrequenz werden durchgelassen, der untere Teil des Frequenzbereiches wird unterdrückt.



Bandpass Nur ein bestimmtes Frequenzband wird durchgelassen, alle andern Komponenten werden unterdrückt.



Bandsperre Ein bestimmtes Frequenzband wird unterdrückt, alle andern Komponenten werden durchgelassen.



Für die Messung von Frequenzgängen mit Spektrum-Analysatoren sind mehrere hochwertige Bandpass-Filter erforderlich. Dabei werden diese so ausgelegt, dass sich die Frequenzbänder möglichst lückenlos aneinander anschliessen ohne sich jedoch zu überlappen. Am Ausgang jedes Bandpass-Filters lässt sich dann messen, wie gross der spektrale Anteil im betreffenden Frequenzband ist. Um eine hohe spektrale Auflösung zu erhalten, sind viele Bandpass-Filter mit jeweils engem Frequenzband erforderlich.

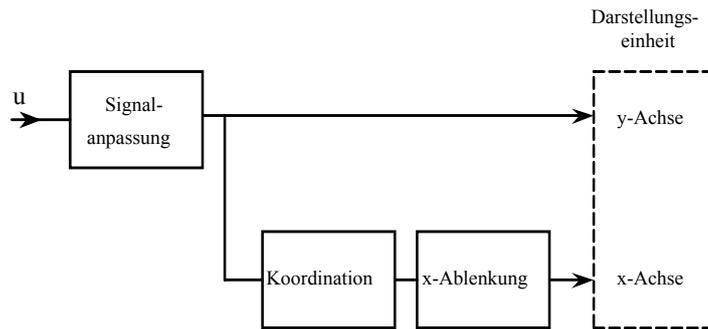
10.4 Messgeräte für Momentanwerte

Messgeräte für Momentanwerte dienen der Darstellung des zeitlichen Verlaufs einer oder mehrerer Grössen, welche in der Regel als Spannungen verfügbar sein müssen. Bei ausreichend langsam veränderlichen Messgrössen können einfache Drehspulinstrumente zu Messschreibern umgebaut werden, indem z.B. an der Zeigerspitze ein Schreiber angebracht wird, welcher den Wert auf einem bewegten Papierband aufzeichnet. Für die meisten Momentanwertmessungen in der Elektrotechnik sind solche Messschreiber aber viel zu langsam und man ersetzt deshalb den Zeiger durch einen Elektronenstrahl. Je nach Zweck unterscheidet man konventionelle Oszilloskope oder Logic Analyser (nur logische Werte "1" oder "0").

10.4.1 Elektronenstrahloszilloskop

Im Elektronenstrahloszilloskop wird ein Momentanwert als Ablenkung eines Elektronenstrahls in einer Elektronenstrahlröhre dargestellt. In der vertikalen Achse (y-Achse) wird dabei der Momentanwert aufgetragen, in der horizontalen Achse (x-Achse) eine zeitproportionale Grösse. Diese Ablenkung wird auf einem phosphoreszierenden Schirm (Emission von Licht beim Auftreffen von Elektronen) sichtbar gemacht. (Andere Namen sind: Kathodenstrahloszilloskop, kurz "KO", englisch: Cathode Ray Oscilloscope, kurz "CRO").

Man kommt damit zu folgendem, sehr einfachen Blockschaltbild (u: zu messender Momentanwert):



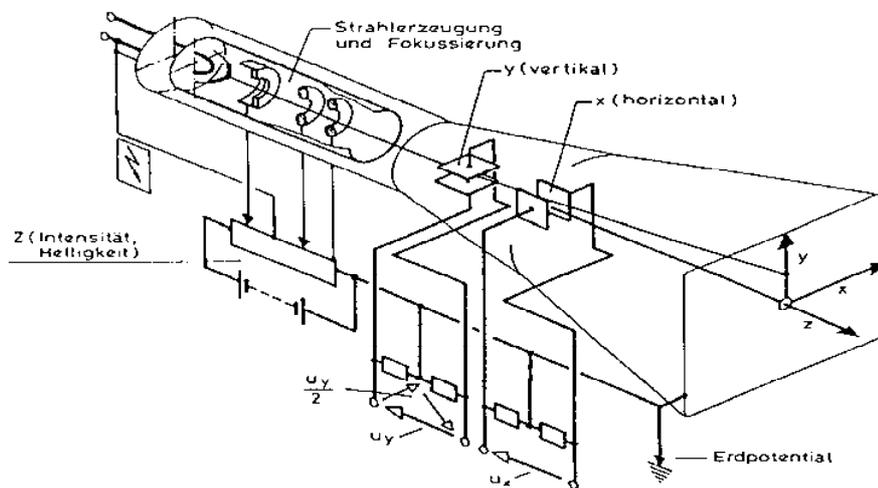
Figur 171: Prinzipieller Aufbau eines KO

10.4.1.1 Grundstruktur

Elektronenstrahlröhre

Die Elektronenstrahlröhre (Kathodenstrahlröhre, cathode ray tube, CRT) dient zur Darstellung einer Spannung als Auslenkung eines Elektronenstrahles auf einen Bildschirm.

Der prinzipielle Aufbau einer Elektronenstrahlröhre ist in Figur 12.10 skizziert: Der im Strahlerzeugungs- und Fokussierungssystem erzeugte Elektronenstrahl hoher Geschwindigkeit wird im elektrostatischen Feld zwischen den Ablenkplatten in horizontaler und vertikaler Richtung abgelenkt und prallt am Ende der Röhre auf einen Leuchtblag auf, in dem die kinetische Energie der Elektronen zum Teil in sichtbares Licht umgesetzt wird.



Figur 172: Elektronenstrahlröhre: Vereinfachter Aufbau

Eine Berechnung der Vorgänge in der Elektronenstrahlröhre soll hier nicht durchgeführt werden. Aus dieser Berechnung folgt, dass die Ablenkung in x - bzw. y -Richtung auf dem Leuchtschirm proportional zur Spannung u_x bzw. u_y zwischen den entsprechenden Ablenkplatten ist:

$$y = k_y \cdot u_y \quad \text{und} \quad x = k_x \cdot u_x.$$

Damit lässt sich also jeder Punkt auf dem Leuchtschirm mit Hilfe entsprechender Spannungen ansteuern. Dies gilt bis zu sehr hohen Frequenzen der Ablenkspannungen in der Größenordnung einiger 100 MHz.

HINWEIS: Neben der hier behandelten elektrostatischen Ablenkung kann eine magnetische mittels entsprechender Spulen vorgenommen werden. Allerdings ist die Linearität weniger gut und es sind nur relativ langsame Ablenkungen bis ca. 100 kHz möglich. Sie werden dort eingesetzt, wo das Bild nicht in Echt-Zeit aufgebaut werden muss. Hauptanwendungsbereich ist im Fernsehen und bei den davon abgeleiteten Monitoren für Rechner.

y-Verstärker

Die üblichen Ablenkempfindlichkeiten $1/k_y$ bzw. $1/k_x$ von der Größenordnung einige V/cm auf dem Bildschirm verunmöglichen in den meisten Fällen den direkten Anschluss der zu messenden Spannung an die Platten, da sie je nach Art der ausgeführten Messung im Bereich μV , mV, V (zum Beispiel Elektronik) bis kV (Hochspannungstechnik) liegen. In den meisten Fällen muss deshalb den Platten noch ein Verstärker vorgeschaltet werden, der die Anpassung der Messspannung an die Ablenkempfindlichkeit gestattet. Diese Verstärker sind am Ausgang immer als Differentialausgang geschaltet, damit das Potential zwischen den Platten gleich null bleibt.

x-Ablenkung

Kathodenstrahloszillographen werden üblicherweise zur Betrachtung des zeitlichen Verlaufs einer Spannung $u_y(t)$ verwendet, so dass die x-Achse als Zeitachse verwendet wird. Damit muss im Bereich linker Bildrand bis rechter Bildrand gelten:

$$x = c \cdot t u_x = a \cdot t$$

(Rampenförmige Spannung)

Für periodische Vorgänge muss auch die Rampenspannung wiederkehrend erzeugt werden, und zwar koordiniert mit dem darzustellenden Signal. Üblich ist dabei die Triggerung: Überschreitet, bzw. unterschreitet u_y einen vorwählbaren Wert U_T , so entstehen in einer entsprechenden elektronischen Schaltung Triggerimpulse. Entweder der negative oder der positive Impuls, je nach gewünschter Flankenpolarität, löst dann den Sägezahn aus, welcher mit einer gegebenen wählbaren Geschwindigkeit ansteigt. Damit lässt sich der Bildschirm horizontal in Zeiteinheiten eichen, zum Beispiel 1 s/cm, 1 ns/cm, 50 μs /cm, usw. Der Rücklauf erfolgt bei einer festen Maximalspannung, welche dem rechten Bildrand entspricht. Gleichzeitig mit dem Strahlrücklauf wird die Strahlhelligkeit auf Null reduziert und solange auf Null gehalten, bis der nächste Triggerimpuls eintritt.

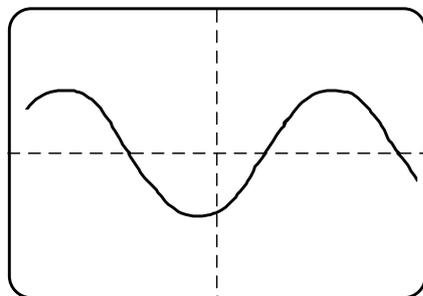
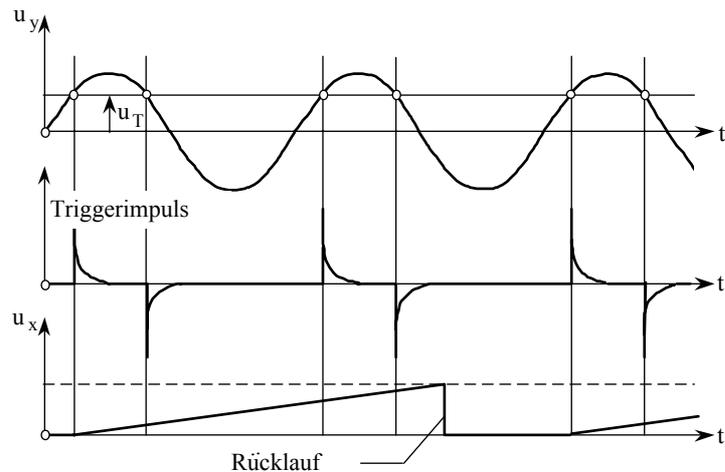


Bild auf dem Schirm

Figur 173: Zum Prinzip der Triggerung
 (Hier: Verwendung der positiven Flanke):
 Der max. Wert von u_x entspricht dabei dem rechten Bildrand, die Spannung $u_x = 0$ dem linken.

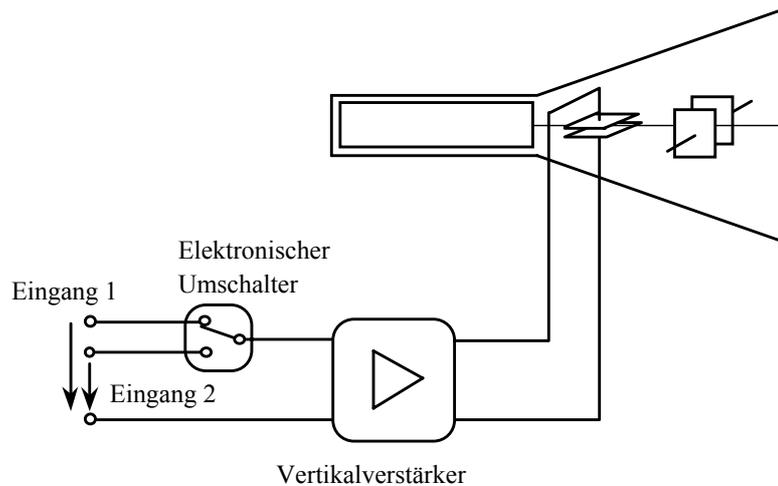
Trigger-Einstellmöglichkeiten:

- ◇ + oder – (die positiven oder negativen Trigger-Impulse (Bild 173) lösen die Ablenkspannung aus
- ◇ interne, externe oder Netz: intern von u_y ; extern von beliebiger Quelle; Netz von 50 Hz-Netzfrequenz

10.4.1.2 Zweistrahl-Oszillograph

Zur Darstellung von 2 oder mehr Signalen gleichzeitig auf dem Bildschirm sind prinzipiell zwei Wege gangbar:

- ◇ Echter Zweistrahl (Dual Beam): 2 Strahlen, 2 Systeme in 1 Strahlröhre (heute nur noch selten verwendet für neue Geräte)
- ◇ Unechter Zweistrahl (auch Vierstrahl) mit Einstrahlröhre



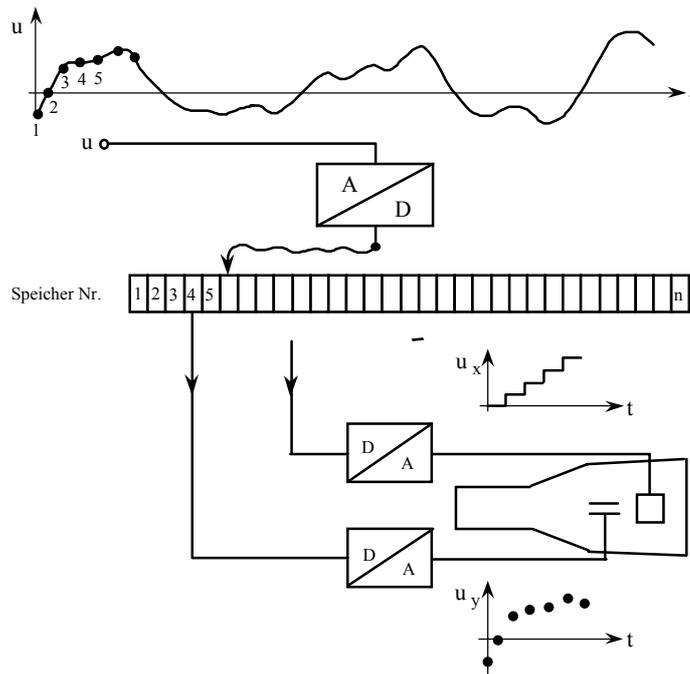
Figur 174: Prinzip des unechten Zweistrahlzilloskops

Beim elektronischen Umschalter sind 2 Betriebsarten möglich: rasch umgeschaltet, bzw. zerhackt (chopped) oder nach jedem Horizontaldurchlauf umgeschaltet (alternate). Alternate ist dabei die bestgeeignete Möglichkeit bei schnellen Ablenkgeschwindigkeiten.

Vorsicht Für eine phasenrichtige Darstellung muss dafür gesorgt werden, dass die zur Triggerung verwendete Spannung immer dem gleichen Kanal (Eingang 1 oder Eingang 2 oder ext. Triggerung) entstammt.

10.4.1.3 Digitaler Speicheroszillograph

Beim digitalen Speicheroszillographen wird das Bild nicht als durchgehender Kurvenzug auf dem Bildschirm dargestellt, sondern die zu speichernde Spannung wird in regelmäßigen Abständen abgetastet und in einen Zahlenwert umgewandelt. Diese Zahlenwerte werden sukzessive in einen elektronischen Speicher eingeschrieben (Figur 175), bis der ganze zur Verfügung stehende Speicherplatz gefüllt ist. Anschliessend werden diese Werte fortlaufend (1, 2, ..., $n-1$, n , 1, 2, ..., $n-1$, n , 1, ...) wieder abgefragt und über einen Digital/Analog-Wandler wieder in eine analoge periodische Spannung zurückgewandelt. Die "Adressen" der Speicherzellen (Nr. der Zelle) geben dabei das Signal für die x-Ablenkung.



Figur 175: Prinzip des digitalen Speicheroszillographen

Weitere Betriebsarten, bei denen fortlaufend eingelesen wird (roll mode) sind möglich. Da die Information gespeichert ist, muss der Triggerpunkt zeitlich gesehen nicht am Anfang des Speichers, sondern kann auch später liegen (Pretrigger), so dass auch Vorgänge vor dem Erreichen des eigentlichen Triggerpegels erfassbar sind (wichtig z.B. bei einmaligen Vorgängen).

Dieses Verfahren wird bei den meisten modernen Oszilloskopen angewendet. Die zulässigen Abtastraten liegen in der Größenordnung 100 MHz, in Sonderausführungen bis 1 GHz bei Speichertiefen von 1024 (= 1K) Punkten bis zu 50 und mehr K. Die Vertikalauflösung beträgt meist 8 bit, was entsprechend der Genauigkeit der Vorverstärker auch genügend ist. Sehr häufig können auch weitere Operationen, wie Dehnung (vertikal und horizontal), numerisches Auslesen der gespeicherten Werte und Bestimmung charakteristischer Mittelwerte vorgenommen werden. Zusätzlich ist es auch möglich, mittels spezieller Interfaces (General Purpose Interface Bus GPIB, siehe 10.5) die Oszilloskope als Teil eines Rechnergesteuerten Mess-Systems einzusetzen.

10.4.2 Logic Analyser

Logische Signale sind dadurch gekennzeichnet, dass sie nur zwei signifikante Werte annehmen können, welche konventionell mit "logisch null" und "logisch eins" bezeichnet werden. Welche physikalischen Größen diesen logischen Werten zugeordnet sind, ist systemabhängig. In einer viel verbreiteten Familie von Logikbausteinen ist die Speisespannung 5 V (selten sind auch Familien mit 15V anzutreffen). Man bezeichnet einen Wert

zwischen 0 V und 0,7 V als "tief (LOW)" und zwischen 2,4 V und 5 V als "hoch (HIGH)". Dies sagt nichts über die Zuordnung zu 0 und 1 aus.

Bei der Analyse logischer Zustände ist man somit (ausser bei gewissen Fällen der Störungssuche) lediglich an einer Aussage über LOW oder HIGH bzw. an der entsprechenden graphischen Darstellung interessiert. Aus diesem Grunde werden die Eingangssignale nicht als solche abgebildet, sondern mittels entsprechender Schwellwertkreise in die beiden Zustände umgesetzt. Dies ist eine erste Operation, welche in den Eingangskreisen vorgenommen wird. Da zudem die Pegel in der Regel klein sind (5 V, evtl. 15 V) können die Eingangsverstärker einfach gehalten werden; sie bestehen meist nur aus einem hochohmigen Eingangsverstärker und dem Pegelbildner. Als grösstes Problem kommt hinzu, dass häufig sehr viele (16, 24, ...) Signale gleichzeitig zu beobachten sind, man braucht also entsprechend viele Eingangsstufen. Sehr häufig sind sie als kleinvolumige Sonde ausgelegt, welche direkt in die Schaltung eingeklemmt werden.

Für die Triggerung wird nicht ein Pegel verwendet, wie dies bei den Oszilloskopen üblich, sondern eine vorwählbare Kombination von logischen Zuständen ("Triggerwort"). Häufig wird auch noch von der untersuchenden Schaltung ein CLOCK-Signal (innere "Uhr" des Systems) zugeführt.

Ansonsten sind Logic Analyser ähnlich strukturiert wie Digitalspeicheroszilloskope, mit dem Unterschied dass für die Signale nur Ein-Bit Speicher vorgesehen werden müssen.

Komfortable Logic Analyser haben, für die Störungssuche (z.B. Schwingungen beim Schalten der Zustände durch schlecht angepasste Kreise oder Leitungen), die Möglichkeit in einen normalen Oszilloskopmodus umzuschalten, evtl. sogar automatisch.

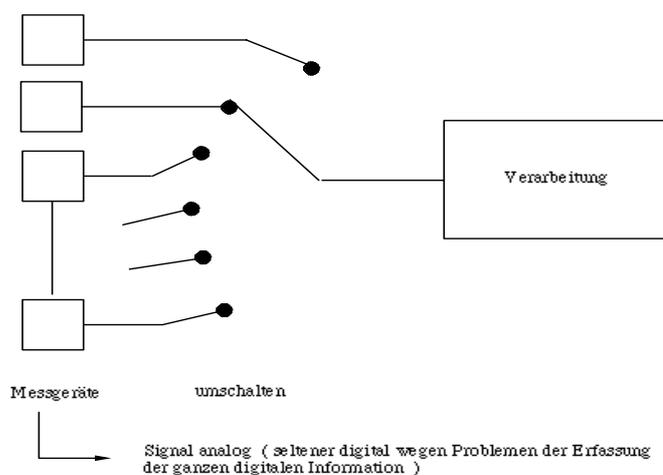
10.5 Messsysteme

10.5.1 Strukturen

Problemstellung :

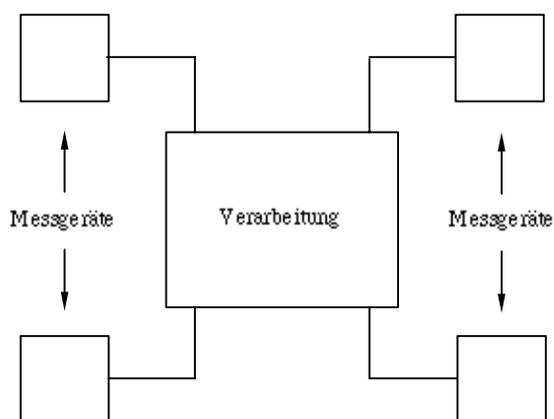
In sehr vielen Fällen kann die notwendige Information über den Zustand eines Prozesses nicht aus einer Messgrösse gewonnen werden, sondern aus einer Kombination mehrerer Messgrössen, d. h. der Ausgänge mehrerer Messgeräte. Wegen der notwendigen Signalverarbeitung werden heute meist digitale Systeme verwendet. Verschiedene Philosophien sind denkbar:

- Messstellen umschalten (langsame Prozesse)



Figur 1178: Messsystem-Struktur mit Umschalter

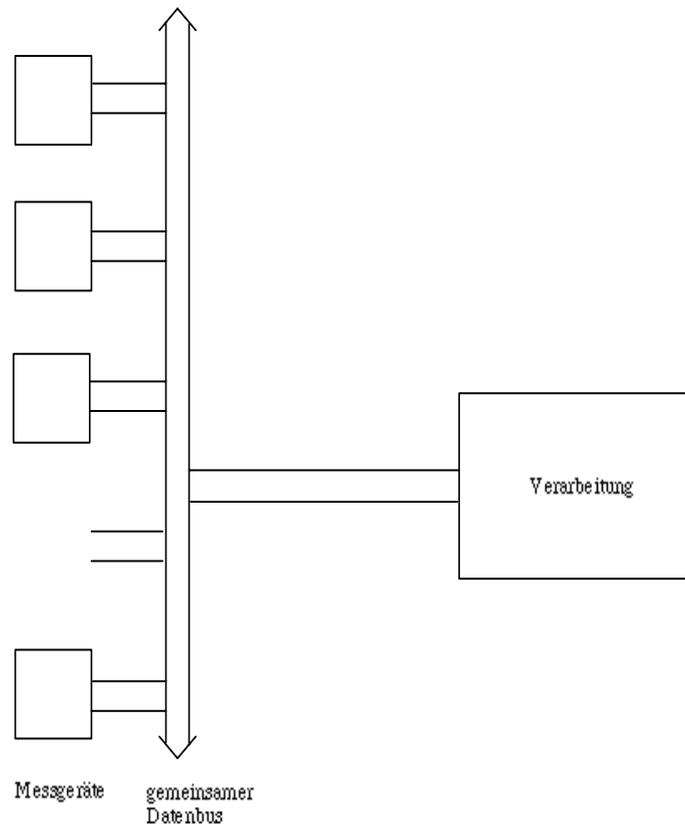
- Sternanschluss



Figur 179: Messsystem-Struktur mit Sternanschluss

Die Ausgangswerte der Messgeräte liegen kontinuierlich an und können verschiedener Struktur (analog, digital) sein. Die Verarbeitung entscheidet über die Auswahl der Messgrößen.

- Parallel-Anschluss (Bussystem), nur digital



Figur 180: Messsystem-Struktur mit gemeinsamem Bus

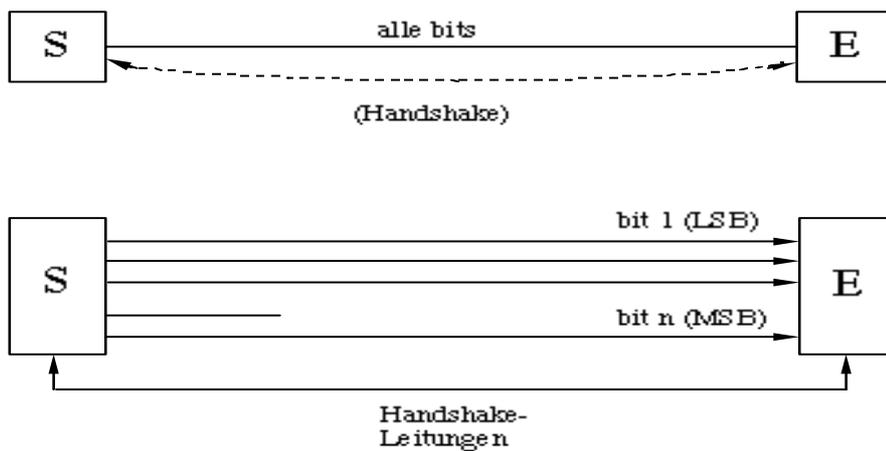
Als Signalarten werden verwendet:

- analog :
 - Spannungssignale (1V, 5V, 10V), eher störempfindlich
 - Stromsignale (0...20mA, 4...20mA) vor allem für längere Leitungen und relativ langsam veränderliche Signale, 4...20mA bevorzugt.
- frequenzanalog: Die Information liegt in der Frequenz, meist als Form von Impulsen. Die Auswertung benötigt einen Frequenzzähler mit analogem oder digitalem Ausgang.
- digital: meist TTL-Niveau mit serieller oder paralleler Übertragung der einzelnen Bits.

Für die digitale Datentübertragung werden üblicherweise normierte Verfahren angewendet, z.B.:

- seriell: RS 232 bzw. RS 485
- parallel: IEC 625 (auch GP-IB oder HP-IB genannt) (neben herstellerspezifischen Systemen wie INTEL Multibus, VME Bus, ...)

In jedem Fall muss sichergestellt werden, dass Sender und Empfänger wissen, wann ein Wort ausgesendet wird. Dies wird meist mit dem Begriff "Handshake" umschrieben.



Figur 181: Serielle und parallele Datenübertragung

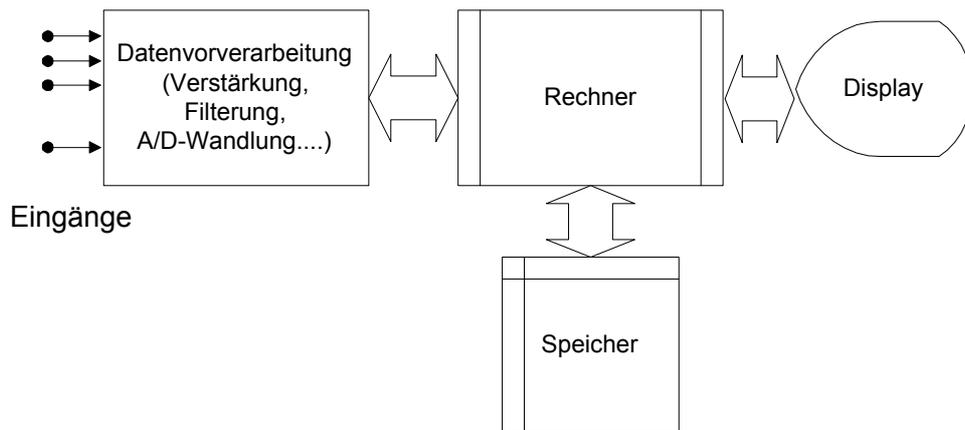
Bei der seriellen Übertragung werden die einzelnen Bits des digitalen Wortes nacheinander übertragen (eine einzige Leitung!), bei der parallelen Übertragung werden alle bit simultan über eine Leitung pro bit übertragen

10.5.2 Virtuelle Instrumente

Jede Messwertverarbeitung weist immer eine ähnliche Struktur bzw. ein ähnliches Pflichtenheft auf:

1. Erfassung der interessierenden Grössen
2. Umformung auf gewünschten Signalpegel
3. Evtl. Umformung in ein digitales Signal
4. Eigentliche Signalverarbeitung (analog oder digital) und Merkmalsextraktion
5. Ausgabe und/oder Speicherung des Resultates.

Diese Operationen sind graphisch in Figur 12.18 dargestellt.

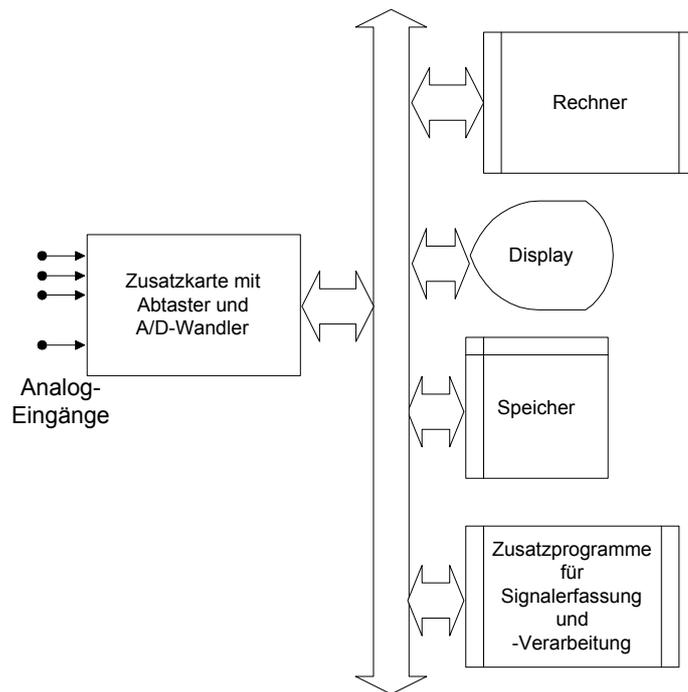


Figur 182: Generelle Struktur einer Messdatenverarbeitungsanlage

Rechner sind für die Punkte 4 und 5 bestens geeignet und haben den Vorteil, dass die Signalverarbeitung einfach mittels Softwareänderung an die Bedürfnisse angepasst werden kann. Auch für die graphische Darstellung der Resultate stehen mächtige Standardpakete zur Verfügung.

Was noch fehlt sind ausreichend flexible Eingangssysteme insbesondere für die Punkte 2 und natürlich 3. Hier hat nun die Philosophie der "Virtuellen Instrumente" eingehakt. Sie geht von zwei Grundideen aus:

1. Datenvorverarbeitung: Es wird angenommen, dass die einzelnen Eingangsgrößen bereits mit normierten Pegeln (also z.B. ± 100 mV, ± 1 V, ± 10 V) vorliegen. In der Eingangsstufe werden diese Signale abgetastet und in eine Zahl umgewandelt, welche lokal gespeichert wird.
2. Die eigentliche Messoperation (also z.B. Mittelwertbestimmung, Frequenzanalyse, Vergleich mit Grenzwerten etc.) wird rein softwaremässig, also durch Programmteile, vorgenommen.



Figur 183: Struktur für virtuelles Messgerät

Damit wird, verglichen mit einem Universalrechner (PC), lediglich eine Zusatzkarte mit der Signalverarbeitung gemäss Pkt. 1 vorstehend gebraucht, der Rest der Operationen (Befehl zum Abtasten, Datentransfer aus der Zusatzkarte, Verarbeitung etc.) ist reine Software. Da es sich in der Regel ja immer um die gleichen Operationen handelt, werden diese einzeln programmiert und dann durch den Anwender zu einem Paket gebündelt. Entsprechend der üblichen Philosophie wird in der Regel eine graphische Programmieroberfläche angeboten, bei welcher die einzelnen Module als Funktionsblöcke zusammengeschaltet werden können.

Im Allgemeinen sind die Abläufe nicht auf maximale Geschwindigkeit getrimmt; allerdings ist mit den heutigen Rechnern die Verarbeitungsgeschwindigkeit ausreichend hoch, um übliche technische Prozesse damit zu erfassen und sogar zu regeln. Hierzu muss auf den erwähnten Zusatzkarten auch noch die Möglichkeit der Datenausgabe vorgesehen werden.

Anhang

Tabellen und Übersichten

Masseinheiten

Größen		Basiseinheit	
Name	Symbol	Name	Symbol
Länge	l	Meter	m
Masse	m	Kilogramm	kg
Zeit	t	Sekunde	s
elektr. Stromstärke	I	Ampère	A
Temperatur	T	Kelvin	K
Stoffmenge	N	Mol	mol
Lichtstärke	I_v	Candela	cd

Elektrische Grössen

Name	Symbol	Name	Symbol	
Frequenz	f, ν	Hertz	Hz	$1 \text{ Hz} = 1 \text{ s}^{-1}$
Kraft	F	Newton	N	$1 \text{ N} = 1 \text{ kg m/s}^2$
Druck	p	Pascal	Pa	$1 \text{ Pa} = 1 \text{ N/m}^2$
mechanische Spannung	σ			
mechanische Festigkeit	R			
Arbeit	W	Joule	J	$1 \text{ J} = 1 \text{ N m}$
Energie	E			
Wärmemenge	Q			
Leistung	P	Watt	W	$1 \text{ W} = 1 \text{ J/s}$
Wärmestrom	Φ			
elektrische Ladung, Elektrizitätsmenge	Q	Coulomb	C	$1 \text{ C} = 1 \text{ A s}$
elektrisches Potential	V	Volt	V	$1 \text{ V} = 1 \text{ W/A} = 1 \text{ J/C}$
Potentialdifferenz, Spannung	U			
elektromotorische Kraft	E			
elektrische Kapazität	C	Farad	F	$1 \text{ F} = 1 \text{ C/V}$
elektrischer Widerstand	R	Ohm	Ω	$1 \Omega = 1 \text{ V/A}$
elektrischer Leitwert	G	Siemens	S	$1 \text{ S} = 1 \Omega^{-1}$
magnetischer Fluss	Φ	Weber	Wb	$1 \text{ Wb} = 1 \text{ V s}$
Induktion, magnetische Flussdichte	B	Tesla	T	$1 \text{ T} = 1 \text{ Wb/m}^2$
Induktivität	L	Henry	H	$1 \text{ H} = 1 \text{ Wb/A}$
Lichtstrom	Φ, Φ_v	Lumen	lm	$1 \text{ lm} = 1 \text{ cd sr}$
Beleuchtungsstärke	E, E_v	Lux	lx	$1 \text{ lx} = 1 \text{ lm/m}^2$

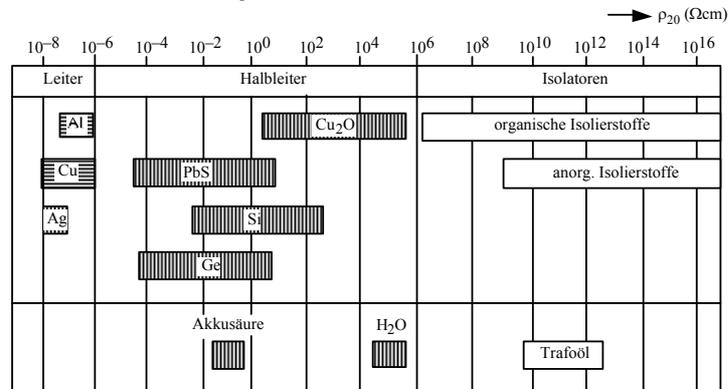
Mechanische Grössen

Bezeichnung	Zusammenhang	Einheiten
Geschwindigkeit	$v = ds/dt$	m/s
Beschleunigung	$a = dv/dt$	m/s^2
Kraft	$F = m \cdot a$	$kgm/s^2 = \text{Newton (N)}$
Arbeit	$W = F \cdot s$	$kgm^2/s^2 = Nm = \text{Joule (J)}$
Energie	$W = 1/2 mv^2$	J
Leistung	$P = F \cdot v = \frac{dW}{dt}$	$kgm^2/s^3 = J/s = \text{Watt (W)}$

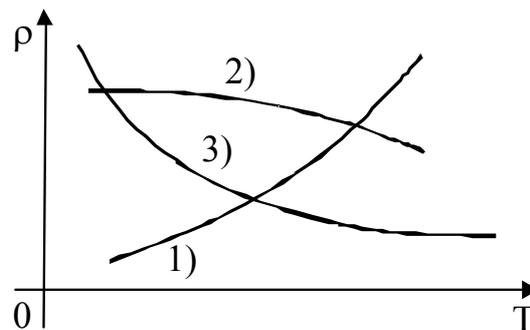
Multiplikationsfaktoren

Faktor mit welchem die Einheit zu multiplizieren ist	Vorsatz		Faktor mit welchem die Einheit zu multiplizieren ist	Vorsatz	
	Symbol	Name		Symbol	Name
10^{-18}	a	Atto	10^1	da	Deka
10^{-15}	f	Femto	10^2	h	Hekto
10^{-12}	p	Pico	10^3	k	Kilo
10^{-9}	n	Nano	10^6	M	Mega
10^{-6}	μ	Mikro	10^9	G	Giga
10^{-3}	m	Milli	10^{12}	T	Tera
10^{-2}	c	Zenti	10^{15}	P	Peta
10^{-1}	d	Dezi	10^{18}	E	Exa

Spezifischer Widerstand ρ verschiedener Stoffe bei 20 °C



Prinzipielle Temperaturabhängigkeit des spezifischen Widerstands ρ :



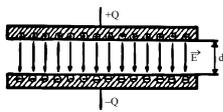
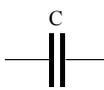
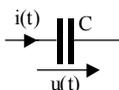
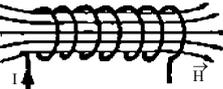
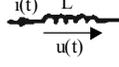
1) Metalle, 2) Metall-Legierungen, 3) Halbleiter

T ist die absolute Temperatur in Kelvin (K) dar, $T = \vartheta + 273,2$ mit ϑ in °C. In der Umgebung der Zimmertemperatur ($\vartheta = 20^\circ\text{C}$) kann folgende Näherungsformel für den Widerstand verwendet werden:

$$R(\vartheta) = R_{20} [1 + \alpha_{20} (\vartheta - 20^\circ)].$$

α ist der Temperaturkoeffizient des spezifischen Widerstandes. Er ist positiv für Metalle (etwa 10^{-3} bis $10^{-4} \text{ }^\circ\text{C}^{-1}$), fast Null für spezielle Metall-Legierungen und negativ für Halbleiter. Bei der Abkühlung auf sehr tiefe Temperaturen springt bei manchen Stoffen der spezifische Widerstand plötzlich auf Null. Dieser Effekt wird mit Supraleitung bezeichnet.

Darstellung (Definition) von Kapazität und Induktivität

	geometrisch	physikalisch	mathematische		
			Schaltzeichen	Einheit	Beziehung
Kapazität	<p>Anordnung von zwei Elektroden, zwischen denen sich ein elektrisches Feld (E) bilden kann</p> 	<p>Energiespeicher für elektrische Energie</p> $W = \frac{1}{2} C U^2$		<p>Farad (F)</p> $1F = \frac{1As}{1V}$	 $i(t) = C \frac{du(t)}{dt}$
Induktivität	<p>Anordnung eines spiralförmig gewickelten Drahtes, in dem sich ein magnetisches Feld (H) bilden kann</p> 	<p>Energiespeicher für magnetische Energie</p> $W = \frac{1}{2} L I^2$		<p>Henry (H)</p> $1H = \frac{1Vs}{1A}$	 $u(t) = L \frac{di(t)}{dt}$

Dielektrizitätskonstanten gebräuchlicher Materialien

Anstelle der Dielektrizitätskonstanten ε schreibt man üblicherweise

$$\varepsilon = \varepsilon_0 \varepsilon_r.$$

Dabei ist

$$\varepsilon_0 = 8,86 \cdot 10^{-12} \text{ As/Vm}$$

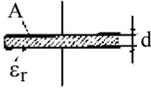
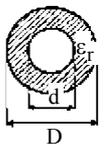
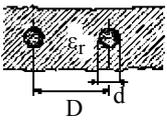
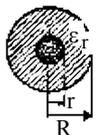
die Vakuum-Dielektrizitätskonstante.

Die relative Dielektrizitätskonstante ε_r ist eine dimensionslose Zahl. Die folgende Tabelle gibt die Werte von ε_r für einige wichtige Stoffe an.

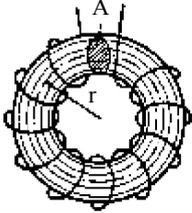
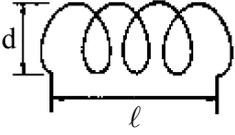
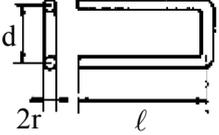
Bariumtitanat (ferroelektrisch)	1000 - 4000	Polystyren	2.5
Glas	2 - 16	Porzellan	4.5 - 6
Glimmer	4 - 8	Quarz	4.3 - 4.7
Luft	≈ 1	Teflon	2.1
Papier (trocken)	1.8 - 2.6	Wasser (rein)	81

Tabelle: Relative Dielektrizitätskonstanten wichtiger Materialien

Näherungsformeln zur Berechnung von Kapazitäten

	<p>Plattenkondensator (Randeffekte vernachlässigt)</p> <p>Beispiel: $A = 10 \times 10 \text{ cm}$ $\varepsilon = \varepsilon_0 \varepsilon_r$ $d = 1 \text{ mm}$ $C \approx 88 \text{ pF}$ $\varepsilon_r = 1$ (Luft)</p>	$C = \varepsilon \frac{A}{d}$
	<p>Zylinderkondensator (zwei konzentrische Zylinder)</p> <p>Beispiel: (koaxiales Kabel) $D = 7 \text{ mm}$ $d = 2 \text{ mm}$ $C/\ell \approx 100 \text{ pF/m}$ $\varepsilon_r = 1$ $\ell = \text{Länge}$</p>	$C = 2\pi\varepsilon \frac{\ell}{\ln\left(\frac{D}{d}\right)}$
	<p>Zwei parallele Drähte</p> <p>Beispiel: (Telefon-Freileitung) $D = 0.5 \text{ m}$ $d = 3 \text{ mm}$ $C/\ell \approx 7 \text{ pF/m}$ $\varepsilon_r = 1$ $\ell = \text{Länge}$</p>	$C = \frac{\pi\varepsilon\ell}{\ln\left[\frac{D}{d} + \sqrt{\left(\frac{D}{d}\right)^2 - 1}\right]}$ <p>für $D \gg d$</p> $C = \frac{\pi\varepsilon\ell}{\ln\left[2\frac{D}{d}\right]}$
	<p>Zwei konzentrische Kugel</p> <p>Beispiel: (Erde gegenüber des Weltalls) $r = 16380 \text{ km}$ $C \approx 700 \mu\text{F}$ $\varepsilon_r = 1$</p>	$C = 4\pi\varepsilon \frac{Rr}{R-r}$ <p>für $R \gg r$</p> $C = 4\pi\varepsilon r$

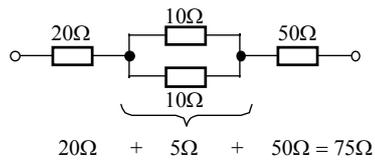
Näherungsformeln zur Berechnung von Induktivitäten

	<p>Ringspule</p> <p>(Das Feld verläuft ausschliesslich innerhalb des Toroids) A = Spulenquerschnitt r = mittlerer Radius des Spulenringes N = Windungszahl</p>	$L = \mu_0 \mu_r \frac{A}{2\pi r} N^2$								
	<p>Lange Zylinderspule</p> <p>(lang: Länge $\ell \gg$ Durchmesser D)</p> <p>Es gilt die gleiche Formel für wie die Ringspule</p> <p>Wird ein Kernmaterial verwendet, so muss dieses einen geschlossenen magnetischen Kreis bilden.</p>	$L = \mu_0 \mu_r \frac{A}{\ell} N^2$								
	<p>Kurze Zylinderspule (einlagig)</p> <p>F = Formfaktor (Korrektur für die Streufelder an den Spulenden)</p> <table border="1" data-bbox="646 1081 933 1171"> <tbody> <tr> <td>ℓ/D</td> <td>0.1</td> <td>1</td> <td>10</td> </tr> <tr> <td>F</td> <td>0.2</td> <td>0.68</td> <td>0.95</td> </tr> </tbody> </table>	ℓ/D	0.1	1	10	F	0.2	0.68	0.95	$L = \mu_0 \mu_r \frac{A}{\ell} N^2 F$
ℓ/D	0.1	1	10							
F	0.2	0.68	0.95							
	<p>Paralleldrahtleitung mit kurzgeschlossenem Ende</p> <p>Die Formel gilt nur solange die Wellenlänge \gg Länge ℓ oder</p> $f \ll \frac{c}{\ell}$ <p>c = Lichtgeschwindigkeit f = Frequenz</p> <p>Bei der Berechnung der Formel wurde das Feld im Draht vernachlässigt</p>	$L = \mu_0 \mu_r \frac{\ell}{\pi} \ln\left(\frac{d-r}{r}\right)$								

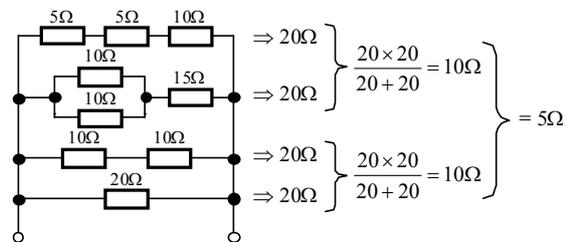
Übungsbeispiele

Beispiele zur Serien- und Parallelschaltung von Widerständen

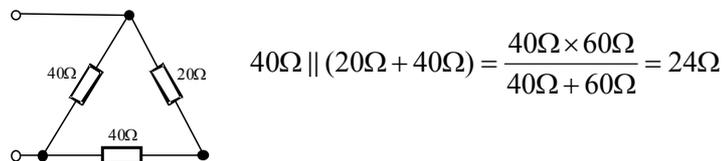
1)



2)

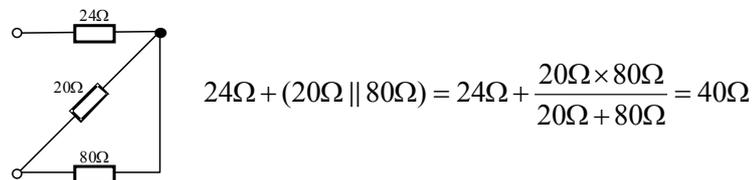


3)



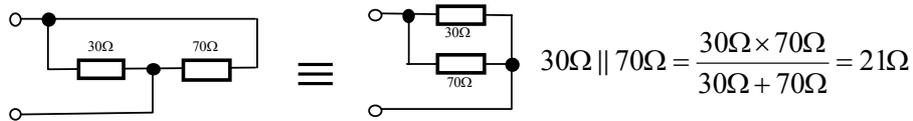
Diese ist keine Dreieckschaltung!

4)



Geometrisch parallel ist nicht immer elektrisch parallel!

5)

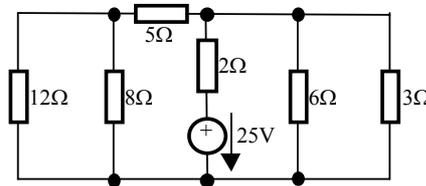


Geometrisch in Reihe ist nicht immer elektrisch in Serie!

Bei einer Parallelschaltung ist der Gesamtwiderstand immer kleiner als der kleinste Widerstand.

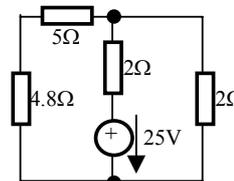
Beispiele von Gleichstromschaltungen

6) Wie gross ist die von der Quelle abgegebene Leistung?

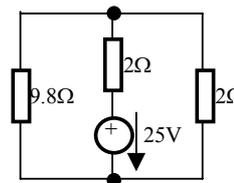


Lösung: Man muss zuerst die die Quelle umgebenden Widerstände als einen "Ersatz"-Widerstand angeben.

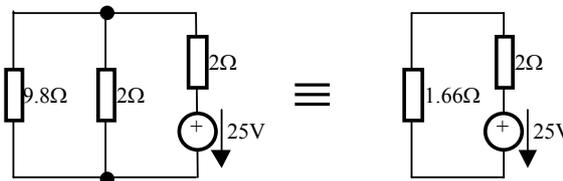
a) $12\Omega \parallel 8\Omega = 4.8\Omega$ und $6\Omega \parallel 3\Omega = 2\Omega$



b) $4.8\Omega + 5\Omega = 9.8\Omega$

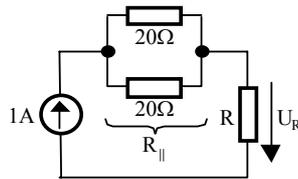


c) $9.8\Omega \parallel 2\Omega = 1.66\Omega$



d) $I = \frac{25V}{2\Omega + 1.66\Omega} = \frac{25V}{3.66\Omega} = 6.8A, \quad P = 25V \times 6.8A = 171W$

7)



- a) Wie gross ist R für $U_R = 25V$?
 b) Wie gross ist dann die Spannung der Quelle?

$$b) \quad R_{||} = \frac{20\Omega \times 20\Omega}{20\Omega + 20\Omega} = 10\Omega$$

$$I_{ges} = 1A$$

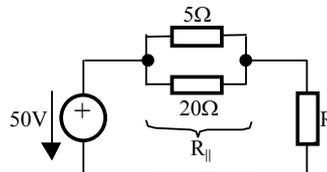
$$U_{R_{||}} = 1A \times 10\Omega = 10V$$

Somit $U_{ges} = U_R + U_{R_{||}} = 35V$ an der Klemmen der Quelle.

$$a) \quad R = \frac{U_R}{I_{ges}} = \frac{25V}{1A} = 25\Omega$$

$$\text{oder auch} \quad \frac{U_R}{U_{R_{||}}} = \frac{R}{R_{||}} \Rightarrow R = R_{||} \frac{U_R}{U_{R_{||}}} = 10\Omega \frac{25V}{10V} = 25\Omega$$

8) Berechne den Wert von R , wenn die Leistung im 5Ω -Widerstand $20W$ beträgt.



Lösung: Wir wollen also zunächst den Spannungsabfall über $R_{||}$ finden, um den über R zu kennen. Dann können wir auch R bestimmen.

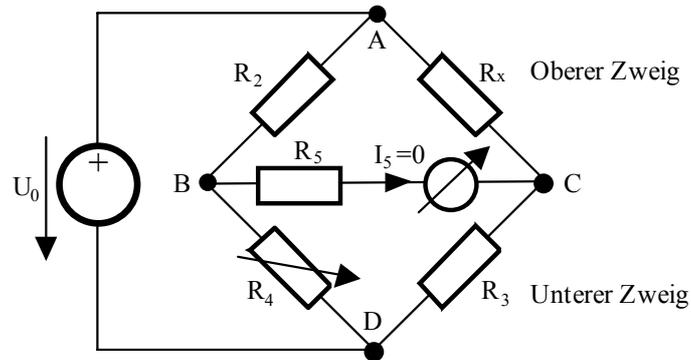
$$a) \quad 20W = P = \frac{U_{R_0}^2}{R = 5\Omega}$$

$$U_{R_0} = \sqrt{P \cdot R} = \sqrt{20W \times 5\Omega} = 10V$$

$$U_R = U_0 - U_{R_0} = 40V$$

$$b) \quad R_0 = \frac{5\Omega \times 20\Omega}{5\Omega + 20\Omega} = 4\Omega \quad \frac{U_{R_0}}{U_R} = \frac{4\Omega}{R} \quad R = 4\Omega \frac{40V}{10V} = 16\Omega$$

- 9) Ein etwas komplizierteres Beispiel stellt die *Wheatstone Brücke* dar. Diese dient zur hochgenauen Messung von Widerständen (R_x). Diese können aus dem Verhältnis $R_x = R_2 \cdot (R_3 / R_4)$ bestimmt werden (R_2, R_3, R_4 bekannt), wenn sich die Brücke im abgestimmten Zustand befindet, d.h. der Strom durch den Querzweig null ist. Folgende Überlegung wird das verdeutlichen.



Damit $I_5 = 0$ (Brückenstrom) muss auch $U_{BC} = 0$ sein (da $R_5 < \infty$). Daraus folgt

$$U_{AB} = U_{AC}, \quad \text{bzw.} \quad U_{BD} = U_{CD}$$

Weiterhin kann man für den oberen Zweig mit der Spannungsteilerregel arbeiten

$$\frac{U_{AB}}{U_0} = \frac{R_2}{R_2 + R_4} \Rightarrow U_{AB} = U_0 \frac{R_2}{R_2 + R_4}$$

und

$$\frac{U_{AC}}{U_0} = \frac{R_x}{R_x + R_3} \Rightarrow U_{AC} = U_0 \frac{R_x}{R_x + R_3}$$

Aus $U_{AB} = U_{AC}$ folgt

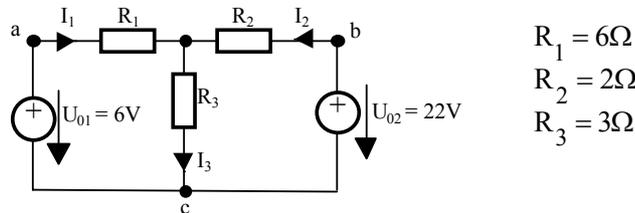
$$\frac{R_2}{R_2 + R_4} = \frac{R_x}{R_x + R_3} \Rightarrow R_x R_2 + R_x R_4 = R_2 R_x + R_2 R_3$$

und somit

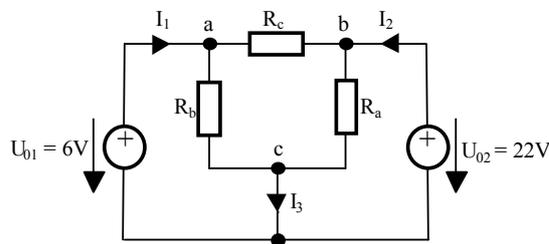
$$\boxed{R_x = R_2 \frac{R_3}{R_4}}$$

Stern-Dreieckumwandlung

Beispiel:



Stern-Dreieck Umwandlung



$$R_a = \frac{R_1 R_2 + R_2 R_3 + R_3 R_1}{R_1} = \frac{6 \times 2 + 2 \times 3 + 3 \times 6}{6} = \frac{36\Omega^2}{6\Omega} = 6\Omega$$

$$R_b = \frac{R_1 R_2 + R_2 R_3 + R_3 R_1}{R_2} = \frac{36\Omega^2}{2\Omega} = 18\Omega$$

$$R_c = \frac{R_1 R_2 + R_2 R_3 + R_3 R_1}{R_3} = \frac{36\Omega^2}{3\Omega} = 12\Omega$$

Ströme durch die drei Dreieck-Widerstände

$$I_a = \frac{U_{02}}{R_a} = \frac{22V}{6\Omega} = 3.66A, \quad I_b = \frac{U_{01}}{R_b} = \frac{6V}{18\Omega} = 0.33A, \quad I_c = \frac{U_{02} - U_{01}}{R_c} = \frac{22V - 6V}{12\Omega} = 1.33A$$

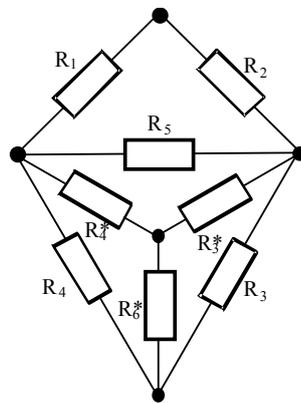
Ströme durch die drei ursprünglichen Widerstände

$$I_1 = I_b - I_c = -1A, \quad I_2 = I_a + I_c = 5A, \quad I_3 = I_a + I_b = 4A$$

Beispiel: Berechne mit Hilfe der Dreieck-Stern Umwandlung den Gesamtwiderstand der Wheatstone Brücke an den Klemmen a-b.

Lösung: Man wählt entweder das obere oder untere Dreieck zur Umwandlung aus. Nach der Umwandlung (hier des unteren Dreiecks) ersetzt die Sternschaltung mit den Widerständen R_n^* die ursprüngliche Dreieckschaltung. Die

Werte dieser neuen Widerstände ergeben sich direkt aus den oben hergeleiteten Umformungen

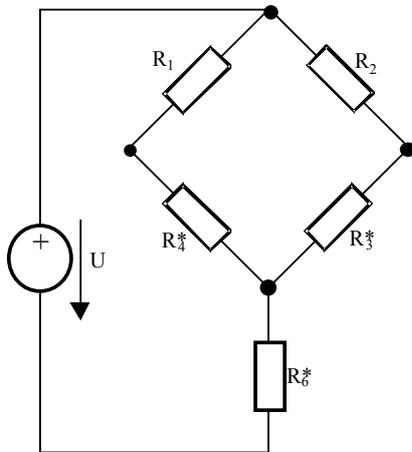


$$R_3^* = \frac{R_5 R_3}{R_3 + R_4 + R_5}$$

$$R_4^* = \frac{R_4 R_5}{R_3 + R_4 + R_5}$$

$$R_6^* = \frac{R_3 R_4}{R_3 + R_4 + R_5}$$

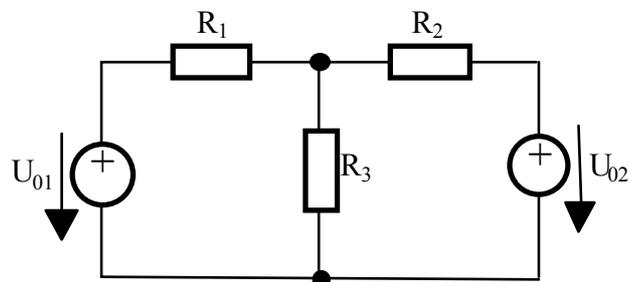
Somit kann die Schaltung wesentlich einfacher dargestellt werden und die Berechnung des Gesamtwiderstands gestaltet sich einfach:



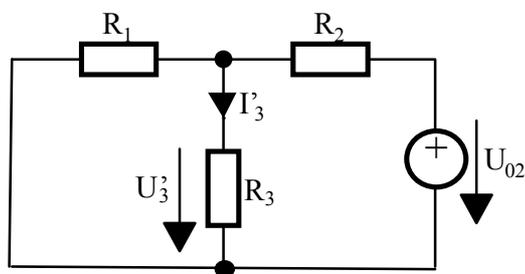
$$R_{\text{ges}} = R_6^* + \frac{(R_1 + R_4^*)(R_2 + R_3^*)}{R_1 + R_4^* + R_2 + R_3^*}$$

Superpositionsprinzip

Beispiel: Bestimme den Strom durch den Widerstand R_3

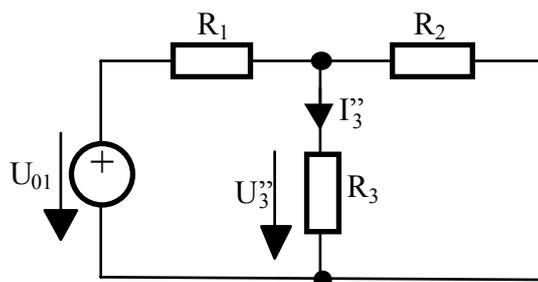


Lösung: Wir setzen zunächst die Quelle $U_{01} = 0$ und berechnen der durch die Quelle U_{02} verursachte Strom I_3'



$$I_3' = \frac{U_{R_3}'}{R_3} = \frac{U_{02}R_1}{R_1R_2 + R_2R_3 + R_3R_1}$$

Im 2. Schritt wird die Quelle $U_{02} = 0$ gesetzt und der Strom I_3'' verursacht durch die Quelle U_{01} berechnet



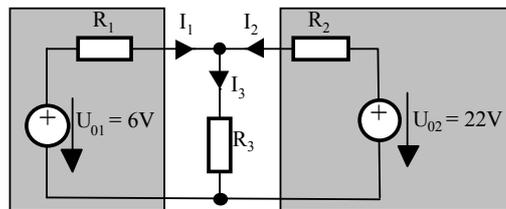
$$I_3'' = \frac{U_{R_3}''}{R_3} = \frac{U_{01}R_2}{R_1R_2 + R_2R_3 + R_3R_1}$$

Schliesslich werden beide Ergebnisse addiert

$$I_3 = I_3' + I_3'' = \frac{U_{01}R_2}{R_1R_2 + R_2R_3 + R_3R_1} + \frac{U_{02}R_1}{R_1R_2 + R_2R_3 + R_3R_1}$$

Quellenumwandlung

Beispiel 1: Wandle im folgenden Schaltungsbeispiel die realen Spannungsquellen in reale Stromquellen um. Widerstände R_1 und R_2 werden als die Innenwiderstände der Quellen angesehen



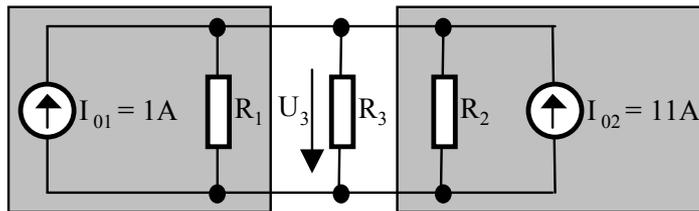
$$R_1 = 6\Omega$$

$$R_2 = 2\Omega$$

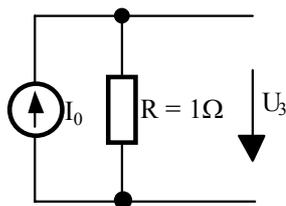
$$R_3 = 3\Omega$$

Lösung: Die Umwandlung der Spannungsquellen in äquivalente Stromquellen führt zu einer Parallelschaltung der Widerstände R_1 bis R_3 . Die äquivalenten Ströme der Stromquellen sind

$$I_{01} = \frac{U_{01}}{R_1} = \frac{6\text{V}}{6\Omega} = 1\text{A}, \quad I_{02} = \frac{U_{02}}{R_2} = \frac{22\text{V}}{2\Omega} = 11\text{A}$$



Die parallelen Stromquellen und Widerstände können deshalb zu einer Quelle und einem Widerstand zusammengefasst werden.



Gesamtwiderstand R :

$$R = 6\Omega // 3\Omega // 2\Omega = \frac{1}{\frac{1}{6\Omega} + \frac{1}{3\Omega} + \frac{1}{2\Omega}} = \frac{1}{\frac{1+2+3}{6}} = 1\Omega$$

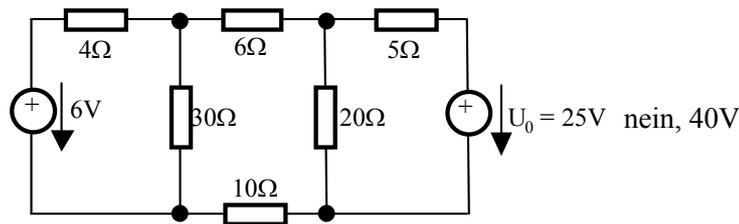
Gesamtstrom und Spannung I_0 , U_3 :

$$I_0 = I_{01} + I_{02} = 1\text{A} + 11\text{A} = 12\text{A}, \quad U_3 = I_0 \cdot R = 12\text{A} \times 1\Omega = 12\text{V}$$

Die Ströme durch die Einzelwiderstände ergeben sich dann leicht über das Ohmsche Gesetz

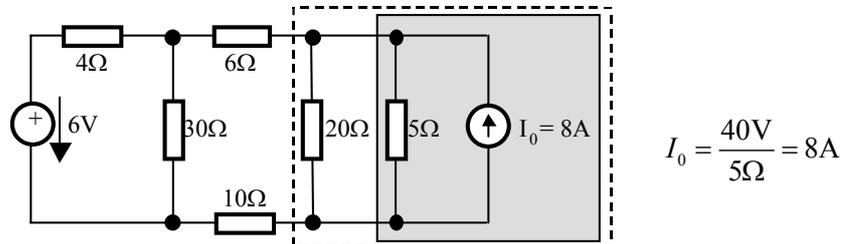
$$I_3 = \frac{U_3}{R_3} = \frac{12\text{V}}{3\Omega} = 4\text{A}, \quad I_1 = \frac{U_{01} - U_3}{R_1} = \frac{6\text{V} - 12\text{V}}{6\Omega} = -1\text{A}, \quad I_2 = \frac{U_{02} - U_3}{R_2} = \frac{22\text{V} - 12\text{V}}{2\Omega} = 5\text{A}$$

Beispiel 2: In der folgenden Schaltung möchte man wissen, wie viel Leistung die 6V-Quelle Leistung abgibt oder aufnimmt.



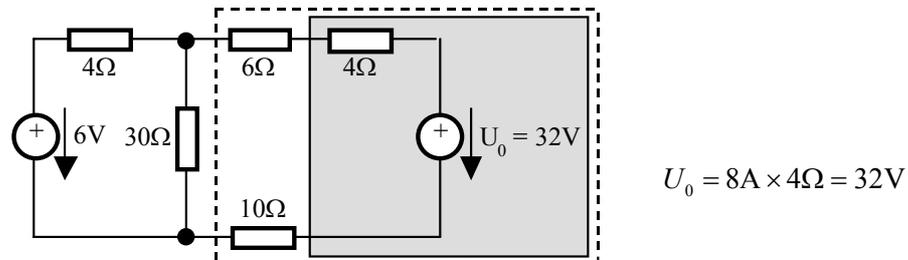
Lösung: Zunächst vereinfachen wir die Schaltung für die 6V-Quelle

1. Schritt: Umwandeln der 25V-Quelle in eine Stromquelle



Zusammenfassen der beiden Widerstände $20\Omega // 5\Omega = \frac{20\Omega \times 5\Omega}{20\Omega + 5\Omega} = 4\Omega$

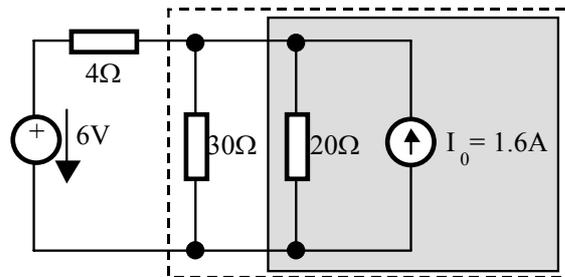
2. Schritt: Umwandeln der resultierenden Stromquelle in eine Spannungsquelle



Die in Serie liegenden Widerstände werden ebenfalls zusammengefasst

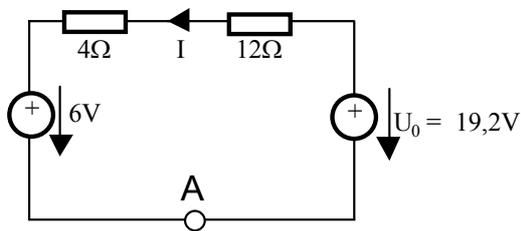
$$6\Omega + 4\Omega + 10\Omega = 20\Omega$$

3. Schritt: Wiederum wird die Spannungsquelle in eine Stromquelle umgewandelt und die resultierende Parallelschaltung der Widerstände zusammengefasst



$$\frac{30\Omega \times 20\Omega}{30\Omega + 20\Omega} = 12\Omega$$

4. Schritt: Im letzten Schritt wird die vorige Stromquelle wieder in eine Spannungsquelle verwandelt und man erhält eine einfache Serienschaltung zweier Widerstände mit zwei Spannungsquellen.



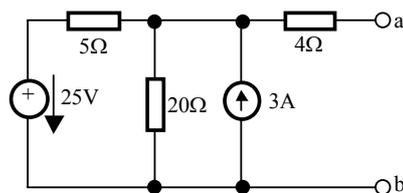
$$I = \frac{19.2\text{V} - 6\text{V}}{12\Omega + 4\Omega} = 0.85\text{A}$$

$$P_{6\text{V}} = -0.85\text{A} \times 6\text{V} = -4.95\text{W}$$

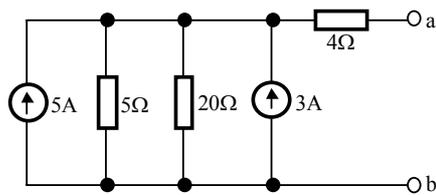
Die 6V-Quelle nimmt Leistung auf!

Thévenin und Norton

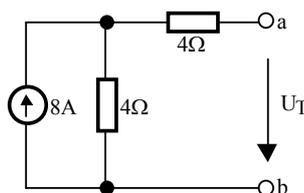
Beispiel: Die folgende Schaltung soll in eine Thévenin Äquivalenz umgewandelt werden. Bestimme die Spannung und den Widerstand an den Klemmen a-b.



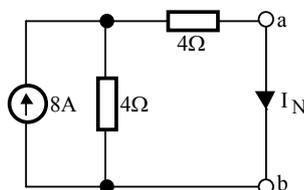
Eine Möglichkeit bestünde darin, die Leerlaufspannung U_T und den Kurzschlussstrom I_K zu bestimmen. Dazu benutzen wir zunächst die Quellenumwandlung und wandeln die Spannungsquelle in eine Stromquelle um



Jetzt ist es leicht, die parallelgeschalteten Stromquellen und Widerstände zusammenzufassen



und die Leerlaufspannung zu bestimmen $U_T = 8A \times 4\Omega = 32V$



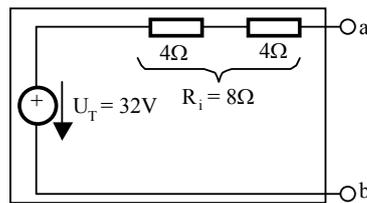
Der Kurzschlussstrom I_N teilt sich in diesem Fall zur Hälfte auf beide Widerstände auf:

$$I_N = 8A \frac{2\Omega}{4\Omega} = 4A$$

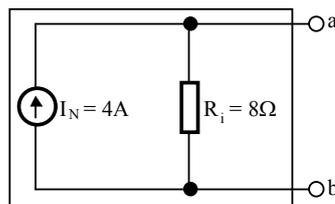
und der Innenwiderstand ergibt sich aus

$$R_i = \frac{U_T}{I_N} = \frac{32V}{4A} = 8\Omega$$

Durch weitere Quellenumwandlung kann daraus leicht eine Thévenin Äquivalenz geschaffen werden.



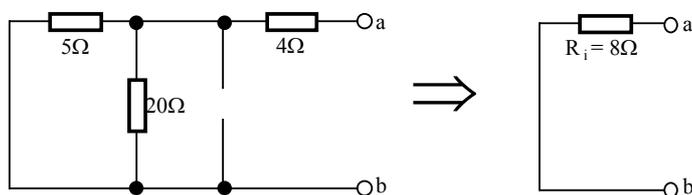
Eine weitere Quellenumwandlung führt zur Norton Äquivalenz



Eine weitere Möglichkeit, den Innenwiderstand R_i zu bestimmen ist folgende:

- a) Spannungsquelle als Kurzschluss
- b) Stromquelle als offener Kreis

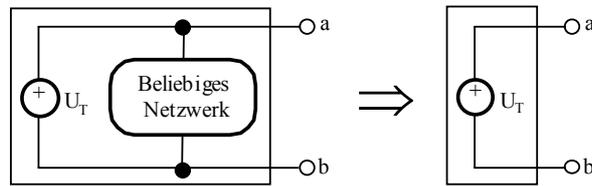
Die so entstandene Widerstandsschaltung kann zusammengefasst werden:



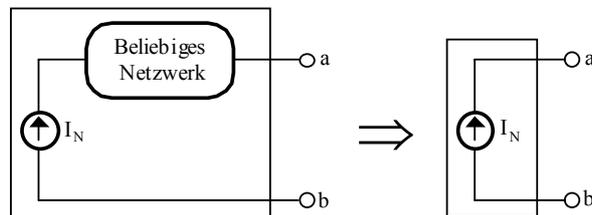
Wichtige Spezialfälle

Eine ideale Spannungsquelle in Parallel mit einem beliebigen Netzwerk (der aber keine andere Spannungsquelle, insbesondere kein Kurzschluss, darstellen darf) wirkt von aussen

wie die Spannungsquelle allein ($R_i = 0$). Das Netzwerk spielt dabei keine Rolle. Das geht aus der Definition der idealen Spannungsquelle, vgl. Abschnitt 2.1.4



Eine ideale Stromquelle in Serie mit einem beliebigen Netzwerk (der aber keine andere Stromquelle, insbesondere kein offener Kreis, darstellen darf) wirkt von aussen wie die Stromquelle allein ($R_i = \infty$). Das Netzwerk spielt dabei keine Rolle. Das geht aus der Definition der idealen Stromquelle, vgl. Abschnitt 2.1.4

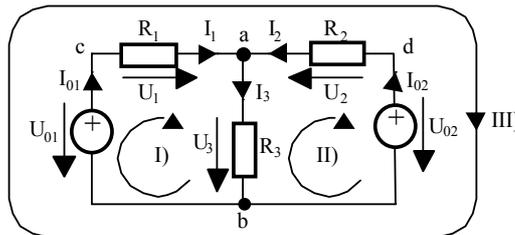


Beide Ersatzschaltungen können zur Vereinfachung von Netzwerken sehr nützlich sein.

Anwendung der Kirchhoff'schen Gleichungen

In den Fällen, wo eine Schaltung sich nach allen Vereinfachungen nicht lösen lässt, muss man zu den Kirchhoff'schen Gleichungen greifen. Wir werden zuerst ihre heuristische Anwendung anhand ein Paar einfacher Beispiele zeigen.

Beispiel 1



$$R_1 = 6\Omega$$

$$R_2 = 2\Omega$$

$$R_3 = 3\Omega$$

Finde Gleichung und Wert für Strom I_3 durch R_3

Lösung: Wir haben 8 Unbekannten, neben I_3 noch $I_1, I_2, I_{01}, I_{02}, U_1, U_2$ und U_3 . Wir brauchen deshalb 8 unabhängige Gleichungen.

Kirchhoff'scher Stromgesetz bei den Knoten a und b: $\sum_k I_k = 0$

In den Knoten einlaufende Ströme positiv

- a) $I_1 + I_2 - I_3 = 0$
- b) $-I_{01} - I_{02} + I_3 = 0$
- c) $I_{01} - I_1 = 0$
- d) $I_{02} - I_2 = 0$

Die Gleichung für Knoten b ist (unter Berücksichtigung von c) und d)) gleich derjenigen für Knoten a und deshalb überflüssig.

Kirchhoff'scher Spannungsgesetz in Kreis I), II) und III): $\sum_k U_k = 0$

- I) $-U_{01} + U_1 + U_3 = 0$
- II) $U_{02} - U_3 - U_2 = 0$
- III) $-U_{01} + U_1 - U_2 + U_{02} = 0$

Auch diese Gleichungen sind nicht unabhängig, da III) = I) + II). Deshalb sind diese Gleichungen nicht unabhängig, eine davon muss weggelassen werden, Wir lassen III) aus, da komplizierter.

Ohm'scher Gesetz für die Widerstände

$$U_1 = R_1 I_1$$

$$U_2 = R_2 I_2$$

$$U_3 = R_3 I_3$$

Somit haben wir 8 Gleichungen für 8 Unbekannte, die jetzt aufgelöst werden. Die Gleichungen c) und d) sind trivial, somit werden I_{01} und I_{02} eliminiert.

Wenn wir die Widerstandsgleichungen in I) und II einsetzen:

$$\text{I) } -U_{01} + R_1 I_1 + R_3 I_3 = 0$$

$$\text{II) } U_{02} - R_3 I_3 - R_2 I_2 = 0$$

$$\text{Aus I) finde } I_1 = \frac{U_{01} - R_3 I_3}{R_1}$$

$$\text{Aus II) finde } I_2 = \frac{U_{02} - R_3 I_3}{R_2}$$

Einsetzen in der Kirchhoff'scher Stromgleichung für Knoten a

$$\frac{U_{01} - R_3 I_3}{R_1} + \frac{U_{02} - R_3 I_3}{R_2} - I_3 = 0$$

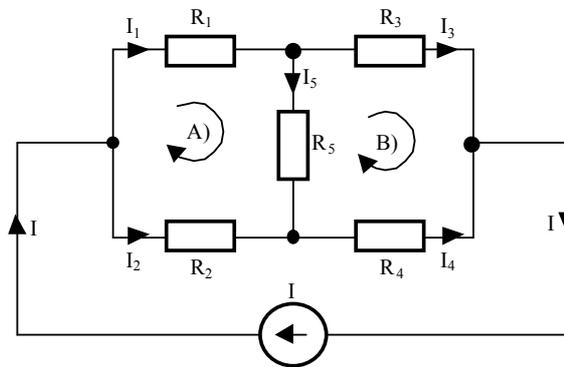
$$I_3 \left(1 + \frac{R_2 R_3 + R_3 R_1}{R_1 R_2}\right) = \frac{U_{01}}{R_1} + \frac{U_{02}}{R_2}$$

$$I_3 = \frac{U_{01}}{R_1} \frac{R_1 R_2}{R_1 R_2 + R_2 R_3 + R_3 R_1} + \frac{U_{02}}{R_2} \frac{R_1 R_2}{R_1 R_2 + R_2 R_3 + R_3 R_1}$$

$$I_3 = \frac{U_{01} R_2}{R_1 R_2 + R_2 R_3 + R_3 R_1} + \frac{U_{02} R_2}{R_1 R_2 + R_2 R_3 + R_3 R_1}$$

Setzt man die Werte der Widerständen und den Quellenspannungen ein, erhält man $I_3 = \underline{4\text{A}}$

Beispiel 2: *Wheatstone Brücke*, dient zur hochgenauen Messung von Widerständen



Lösung: Die Schaltung enthält 5 Widerstände und (für $I_5 \neq 0$) 5 voneinander verschiedene unbekannte Ströme.

Knotenströme

$$I = I_1 + I_2, \quad I_1 = I_5 + I_3, \quad I_4 = I_5 + I_2$$

Die Gleichung $I = I_3 + I_4$ ist überflüssig, da nicht unabhängig: sie kann aus den anderen drei hergeleitet werden, z.B.

$$\begin{aligned} I_3 &= I_1 - I_5 \\ +I_4 &= I_5 + I_2 \\ \hline I_3 + I_4 &= I_5 + I_2 = I \end{aligned}$$

Kreisspannungen, zusammen mit den Widerstandsgleichungen

$$\begin{aligned} \text{A)} \quad I_1 R_1 + I_5 R_5 - I_2 R_2 &= 0 \\ \text{B)} \quad I_3 R_3 - I_4 R_4 - I_5 R_5 &= 0 \end{aligned}$$

Man könnte auch eine den Kirchhoff'sche Spannungsgesetz auf den Kreis $R_1-R_3-R_4-R_2$ anwenden. Die daraus gewonnene Gleichung wäre aber auch überflüssig, da sie aus den Gleichungen A) und B) hergeleitet werden kann.

Wir wollen die Unbekannten auf I_5 reduzieren, um die Bedingung $I_5 = 0$ (abgegliche Wheatstone Brücke) einsetzen zu können

Diese Ströme werden in A) und B) eingesetzt

$$\begin{aligned} \text{A)} \quad I_1 R_1 + I_5 R_5 - (I - I_1) R_2 &= 0 \\ I_1 R_1 + I_5 R_5 - I R_2 + I_1 R_2 &= 0 \end{aligned}$$

$$I_1(R_1 + R_2) + I_5 R_5 - I R_2 = 0$$

$$I_1 = \frac{I R_2 - I_5 R_5}{R_1 + R_2}$$

$$\text{B) } (I_1 - I_5)R_3 - (I_5 + I - I_1)R_4 - I_5 R_5 = 0$$

$$(I_1 - I_5)R_3 - (I_5 + I - I_1)R_4 - I_5 R_5 = 0$$

$$I_1(R_3 + R_4) - I_5(R_3 + R_4 + R_5) - I R_4 = 0$$

$$(I R_2 - I_5 R_5)(R_3 + R_4) = I_5(R_3 + R_4 + R_5)(R_1 + R_2) - I R_4(R_1 + R_2)$$

$$\begin{aligned} I_5 &= I \frac{R_2(R_3 + R_4) - R_4(R_1 + R_2)}{R_5(R_1 + R_2 + R_3 + R_4) + (R_1 + R_3)(R_2 + R_4)} \\ &= I \frac{R_2 R_3 - R_1 R_4}{R_5(R_1 + R_2 + R_3 + R_4) + (R_1 + R_3)(R_2 + R_4)} \end{aligned}$$

Im abgeglichenen Zustand ($I_5 = 0$) folgt

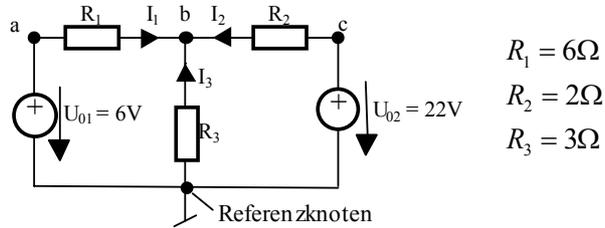
$$R_2 R_3 = R_1 R_4$$

$$\text{bzw. } \frac{R_1}{R_2} = \frac{R_3}{R_4}$$

Wenn die Kirchhoff'schen Gleichungen einfach wahllos aufgestellt werden, riskiert man, zu viele oder zu wenige Unbekannte zu bekommen und einen unter- oder überbestimmten Gleichungssystem zu erhalten, vor allem wenn das Netzwerk nicht einfach ist. Daher muss man systematisch vorgehen.

Knotenanalyse

Beispiel 1



Knotenspannungen

$$U_a = U_{01} = 6\text{V}, \quad U_c = U_{02} = 22\text{V}, \quad U_b \text{ unbekannt}$$

Lösung: Spannungen in den Widerständen

$$U_1 = U_a - U_b = 6\text{V} - U_b$$

$$U_2 = U_c - U_b = 22\text{V} - U_b$$

$$U_3 = -U_b$$

Kirchhoffschen Knotengleichungen

Knoten a und c: trivial

$$\text{Knoten b:} \quad I_1 + I_2 + I_3 = 0$$

$$\frac{U_1}{R_1} + \frac{U_2}{R_2} + \frac{U_3}{R_3} = 0$$

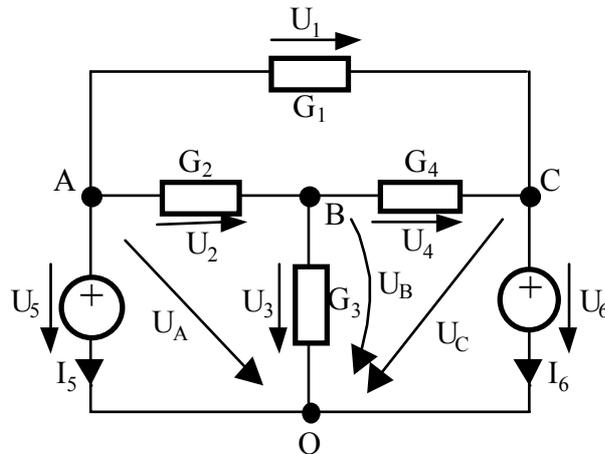
$$\frac{6\text{V} - U_b}{R_1} + \frac{22\text{V} - U_b}{R_2} + \frac{-U_b}{R_3} = 0$$

Das Einsetzen der Werte für R_1 , R_2 und R_3 ergibt $U_b = 12\text{V}$

Für die Ströme erhalten wir

$$I_1 = \frac{6\text{V} - U_b}{6\Omega} = -1\text{A}, \quad I_2 = \frac{22\text{V} - U_b}{2\Omega} = 5\text{A}, \quad I_3 = \frac{-U_b}{3\Omega} = -4\text{A}$$

Beispiel 2



Lösung: Dazu wird zunächst ein Knoten des Netzwerks als Bezugsknoten gewählt. Den übrigen Knoten werden Knotenspannungen zugeordnet, deren Richtungen vom jeweiligen Knoten zum Bezugsknoten zeigen. Die Spannungen können gemäss den Umlaufgleichungen durch die Knotenspannungen ausgedrückt werden, womit die Umlaufgleichungen automatisch erfüllt sind (z.B. Umlauf b)

$$\text{Umlauf a: } \underbrace{(U_A - U_C)}_{U_1} - \underbrace{(U_A - U_B)}_{U_2} - \underbrace{(U_B - U_C)}_{U_4} = 0$$

$$\text{Umlauf b: } \underbrace{(U_A - U_B)}_{U_2} + \underbrace{U_B}_{U_3} - \underbrace{U_A}_{U_5} = 0$$

$$\text{Umlauf c: } \underbrace{-U_B}_{U_3} + \underbrace{(U_B - U_C)}_{U_4} + \underbrace{U_C}_{U_6} = 0$$

Also muss man nur noch die Knotenspannungen aufschreiben. Unter Berücksichtigung der Ströme in den Widerständen erhält man für den Knoten A

$$G_1 U_1 + G_2 U_2 + I_5 = 0$$

Oder ausgedrückt durch die Knotenspannungen

$$G_1(U_A - U_C) + G_2(U_A - U_B) + I_5 = 0$$

bzw.

$$(G_1 + G_2)U_A - G_2 U_B - G_1 U_C = -I_5$$

Daran erkennt man bereits die Ansätze für das folgende Bildungsgesetz: Die Knotengleichung für Knoten A enthält als Koeffizient der Knotenspannung U_A die Summe der im Knoten A zusammengeführten Leitwerte. Man nennt diesen den Knotenleitwert. Der Koeffizient der Knotenspannung U_B ist der Leitwert, der die Knoten A und B verbindet und der Koeffizient der Knotenspannung U_C der Leitwert, der die Knoten A und C verbindet. Beide Leitwerte heissen Kopplungsleitwerte und sie sind stets negativ.

In ähnlicher Weise erhält man für Knoten B

$$-G_2U_A + (G_2 + G_3 + G_4)U_B - G_4U_C = 0$$

Und für Knoten C

$$-G_1U_A - G_4U_B + (G_1 + G_4)U_C = -I_6$$

Der Quellenstrom wird allgemein auf die rechte Seite geschrieben (auf den Knoten zufließende Ströme erscheinen auf der rechten Gleichungsseite mit positivem Vorzeichen, abfließende mit negativem). Obige Gleichungen kann man in Tabellenform zusammenfassen

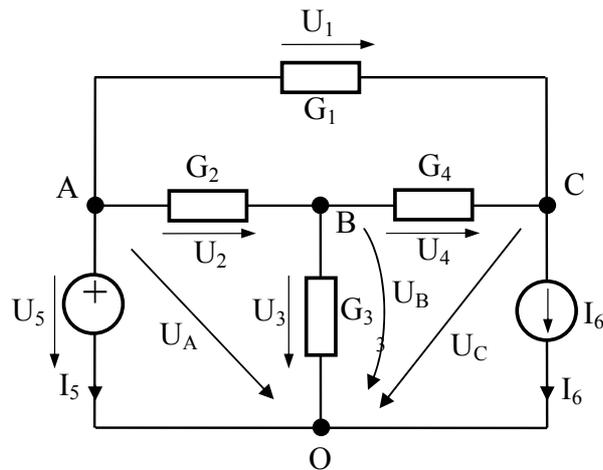
	U_A	U_B	U_C	
A:	G_1+G_2	$-G_2$	$-G_1$	$-I_5$
B:	$-G_2$	$G_2+G_3+G_4$	$-G_4$	0
C:	$-G_1$	$-G_4$	G_1+G_4	$-I_6$

Oder auch in Matrizenschreibweise darstellen

$$\begin{bmatrix} G_1 + G_2 & -G_2 & -G_1 \\ -G_2 & G_2 + G_3 + G_4 & -G_4 \\ -G_1 & -G_4 & G_1 + G_4 \end{bmatrix} \begin{bmatrix} U_A \\ U_B \\ U_C \end{bmatrix} = \begin{bmatrix} -I_5 \\ 0 \\ -I_6 \end{bmatrix}$$

Diese Matrizenschemata mit den drei Knotenleitwerten auf der Hauptdiagonalen und den zu diesen symmetrisch angeordneten Kopplungsleitwerten nennt man auch Leitwertmatrix. Das Gleichungssystem beschreibt die Eigenschaften des Netzwerks, und zwar unabhängig davon, ob etwa die Quellen als Strom- oder Spannungsquellen vorgegeben sind.

Beispiel 3: In der Schaltung ist die linke Quelle eine bekannte Spannungsquelle, die rechte eine bekannte Stromquelle, d.h. U_5 und I_6 sind bekannt. Gesucht ist die Spannung U_2 .



Lösung: Zur Lösung der Aufgabe benötigen wir die Zeilen B und C aus (3.2.1). Man erhält also für die unbekannt Knotenspannungen U_B und U_C das Gleichungssystem:

U_B	U_C	
$G_2+G_3+G_4$	$-G_4$	G_2U_5
$-G_4$	G_1+G_4	$G_1U_5-I_6$

Nach U_B aufgelöst erhält man:

$$U_B = \frac{[(G_1 + G_4)G_2 + G_1G_4]U_5 - G_4I_6}{(G_1 + G_4)(G_2 + G_3) + G_1G_4}$$

Die Spannung am Widerstand R_2 ist

$$U_2 = U_A - U_B = U_5 - U_B$$

U_B eingesetzt und einige Umformungen durchgeführt ergibt

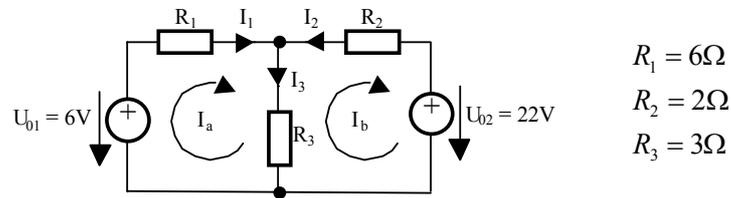
$$U_2 = \frac{G_3(G_1 + G_4)U_5 + G_4I_6}{(G_1 + G_4)(G_2 + G_3) + G_1G_4}$$

Bzw. nach Erweiterung mit den Widerständen R_1 bis R_4

$$U_2 = \frac{R_2(R_1 + R_4)U_5 + R_1R_2R_3I_6}{(R_1 + R_4)(R_2 + R_3) + R_2R_3}$$

Maschenanalyse

Beispiel



Für jede Masche schreiben wir eine Kirchhoff'sche Spannungsgleichung

$$\begin{cases} -U_{01} + I_a R_1 + I_a R_3 - I_b R_3 = 0 \\ U_{02} + I_b R_2 + I_b R_3 - I_a R_3 = 0 \end{cases}$$

Werte eingesetzt:

$$\begin{cases} -6\text{V} + I_a \cdot 6\Omega + I_a \cdot 3\Omega - I_b \cdot 3\Omega = 0 \\ 22\text{V} + I_b \cdot 2\Omega + I_b \cdot 3\Omega - I_a \cdot 3\Omega = 0 \end{cases}$$

Aus der 2. Gleichung:

$$I_b = -\frac{22\text{V}}{5\Omega} + \frac{3\Omega}{5\Omega} I_a$$

und in der ersten eingesetzt:

$$-30\text{V} + I_a \cdot 45\Omega + 66\text{V} - I_a \cdot 9\Omega = 0$$

$$\underline{I_a = -1\text{A}}$$

Schaltverhalten einer Kapazität

Beispiel: Ein Spannungsimpuls

$$\begin{aligned}u(t) &= 0 && \text{bei } t \leq 0 \\u(t) &= 4t \text{ V} && \text{bei } 0 \leq t \leq 1\text{s} \\u(t) &= 4e^{-(t-1\text{s})} \text{ V} && \text{bei } 1\text{s} \leq t \leq \infty\end{aligned}$$

wird an eine Kapazität $C = 0.5 \mu\text{F}$ angelegt.

a) Finde Strom, Leistung und Energie für die Kapazität

b) Skizziere die Spannung, Strom und Leistung, sowie die Energie als Funktion der Zeit

Lösung: Strom $i = C \frac{du(t)}{dt}$

$$\begin{aligned}i(t) &= (0.5\mu\text{F}) \cdot 0 = 0 && t \leq 0 \\i(t) &= (0.5\mu\text{F}) \cdot 4\text{V/s} = 2\mu\text{A} && 0 \leq t \leq 1\text{s} \\i(t) &= (0.5\mu\text{F}) \cdot (-4e^{-(t-1\text{s})} \text{V/s}) = -2e^{-(t-1\text{s})} \mu\text{A} && 1\text{s} \leq t \leq \infty\end{aligned}$$

Leistung: $p(t) = u \cdot i = C u \frac{du}{dt}$

$$\begin{aligned}p(t) &= 0 && t \leq 0 \\p(t) &= 4t\text{V} \cdot 2\mu\text{A} = 8t \mu\text{W} && 0 \leq t \leq 1\text{s} \\p(t) &= (4e^{-(t-1\text{s})} \text{V}) \cdot (-2e^{-(t-1\text{s})} \mu\text{A}) = -8e^{-2(t-1\text{s})} \mu\text{W} && 1\text{s} \leq t \leq \infty\end{aligned}$$

Energie: $W = \frac{1}{2} C U^2$

$$\begin{aligned}w(t) &= 0 && t \leq 0 \\w(t) &= \frac{1}{2} (0.5\mu\text{F}) \cdot (16t^2 \text{V}^2) = 8t \mu\text{J} && 0 \leq t \leq 1\text{s} \\w(t) &= \frac{1}{2} (0.5\mu\text{F}) \cdot (16e^{-2(t-1\text{s})} \text{V}^2) = 4e^{-2(t-1\text{s})} \mu\text{J} && 1\text{s} \leq t \leq \infty\end{aligned}$$

- Energie wird im Kondensator gespeichert, wenn die Leistung positiv ist ($0 \rightarrow 1\text{s}$)
- Energie wird vom Kondensator abgegeben, wenn die Leistung negativ ist ($t > 1\text{s}$)

Energie im Intervall $0 \leq t \leq 1\text{s}$:

$$\int_0^1 p dt = \int_0^1 8t dt = 4t^2 \Big|_0^1 = 4\mu\text{J} \quad \text{Energie aufgenommen von C}$$

Energie im Intervall $1s \leq t \leq \infty$:

$$\int_0^1 p dt = \int_{1s}^{\infty} -8e^{-2(t-1)} dt = -8 \frac{e^{-2(t-1)}}{-2} \Big|_{1s}^{\infty} = -4 \mu J$$

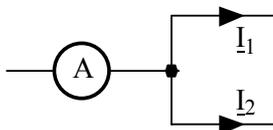
Zuvor aufgenommene Energie wird nun wieder abgegeben.

Analyse linearer Netzwerke bei stationären sinusförmigen Vorgängen mittels der komplexen Darstellung

Beispiel 1: Bestimme die Gesamtspannung \underline{U} an den Klemmen zweier in Serie geschalteter Spannungsquellen $\underline{U}_1 = 20V \angle 0^\circ$ und $\underline{U}_2 = 30V \angle 45^\circ$.

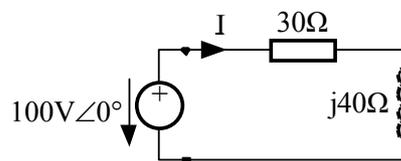
$$\begin{aligned} \underline{U}_1 &= 20V \angle 0^\circ = 20V + j0 \\ \underline{U}_2 &= 30V \angle 45^\circ = 21.2V + j21.2V \\ \underline{U}_1 + \underline{U}_2 &= 41.2V + j21.2V = 46.35V \angle 27.2^\circ \end{aligned}$$

Beispiel 2: Bestimme die Anzeige des Amperemeters (A), wenn die Teilströme $\underline{I}_1 = 15 A \angle -20^\circ$ und $\underline{I}_2 = 8 A \angle 25^\circ$ betragen. Wie gross ist der Phasenwinkel des Gesamtstromes?

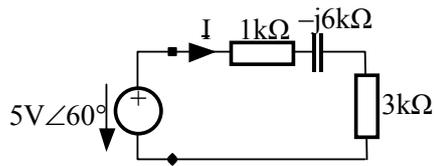


$$\begin{aligned} \underline{I}_1 &= 15A \angle -20^\circ = 14.1A - j5.13A \\ \underline{I}_2 &= 8A \angle 25^\circ = 7.25 + j3.38A \\ \underline{I}_1 + \underline{I}_2 &= 21.35A - j1.75A = 21.42A \angle -4.7^\circ \end{aligned}$$

Beispiel 3: Bestimme die von den Spannungsquellen gelieferten Ströme in folgenden Schaltungen.



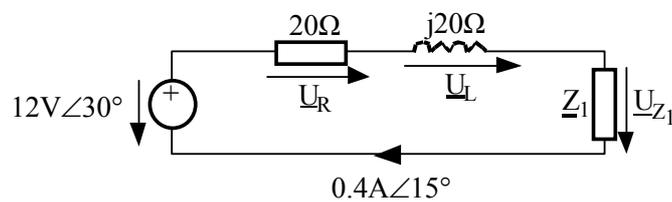
$$\begin{aligned} \underline{Z} &= 30\Omega + j40\Omega = 50\Omega \angle 53.13^\circ \\ \underline{I} &= \frac{\underline{U}}{\underline{Z}} = \frac{100V \angle 0^\circ}{50\Omega \angle 53.13^\circ} = 2A \angle -53.13^\circ \end{aligned}$$



$$\underline{Z} = 4\text{k}\Omega - j6\text{k}\Omega = 7.21\text{k}\Omega \angle -56.3^\circ$$

$$\underline{I} = \frac{\underline{U}}{\underline{Z}} = \frac{5\text{V} \angle 60^\circ}{7.21\text{k}\Omega \angle -56.3^\circ} = 0.7\text{mA} \angle 116.3^\circ$$

Beispiel 4: Bestimme die Impedanz \underline{Z}_1 und die Teilspannungen \underline{U}_R , \underline{U}_L und \underline{U}_{Z_1} in folgender Schaltung.



$$\underline{Z}_{\text{tot}} = \frac{12\text{V} \angle 30^\circ}{0.4\text{A} \angle -15^\circ} = 30\Omega \angle (30^\circ - (-15^\circ)) = 30\Omega \angle 45^\circ = 21.213\Omega + j21.213\Omega$$

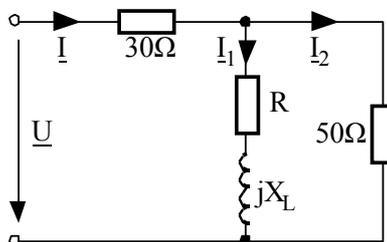
$$\begin{aligned} \underline{Z}_1 &= \underline{Z}_{\text{tot}} - (R - jX_L) = 21.213\Omega + j21.213\Omega - 20\Omega - j20\Omega = 1.213\Omega + j1.213\Omega \\ &= 1.715\Omega \angle 45^\circ \end{aligned}$$

$$\underline{U}_R = 0.4\text{A} \angle -15^\circ \cdot 20\Omega = 8\text{V} \angle -15^\circ$$

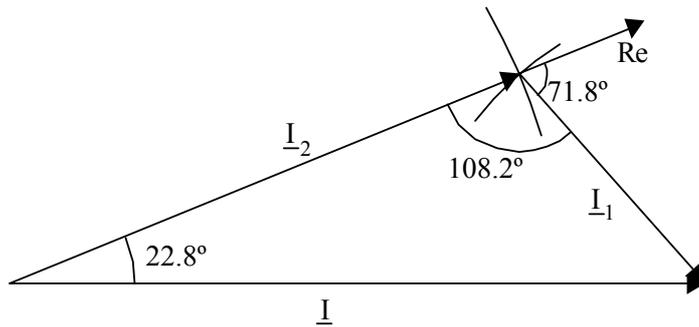
$$\underline{U}_L = 0.4\text{A} \angle -15^\circ \cdot 20\Omega \angle 90^\circ = 8\text{V} \angle 75^\circ$$

$$\underline{U}_{Z_1} = 0.4\text{A} \angle -15^\circ \cdot 1.715\Omega \angle 45^\circ = 0.686\text{V} \angle 30^\circ$$

Beispiel 5: In der folgenden Schaltung sind $\underline{I} = 1\text{A}$, $\underline{I}_1 = 0.4\text{A}$ und $\underline{I}_2 = 0.8\text{A} \angle 0^\circ$ bekannt. Bestimme R , X_L und \underline{U} .



Zuerst muss man die Gleichung $\underline{I} = \underline{I}_1 + \underline{I}_2$ auflösen, entweder graphisch



oder mit Hilfe des Cosinussatzes

$$I^2 = 0.8^2 + 0.4^2 - 2 \cdot 0.8 \cdot 0.4 \cos x$$

$$\cos x = \frac{1 - 0.64 - 0.32}{2 \cdot 0.32} = -0.3125 \Rightarrow x = \pm 108.2^\circ$$

Es hat zwei Lösungen (zwei Schnittpunkte zwischen den Kreisen). Man muss diejenige nehmen, bei welcher der Strom eine negative Phase aufweist, damit am Schluss die Impedanz induktiv resultiert. Die andere Lösung würde eine kapazitive Impedanz ergeben.

$$\underline{I}_1 = 0,4\text{A} \angle (108.2^\circ - 180^\circ) = 0,4\text{A} \angle -71.8^\circ$$

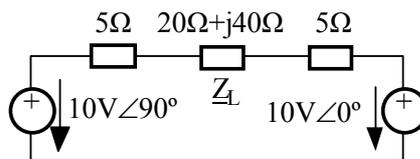
$$\underline{U}_2 = 50\Omega \times 0.8\text{A} = 40\text{V} \angle 0^\circ$$

$$\underline{Z} = \frac{\underline{U}_2}{\underline{I}_1} = \frac{40\text{V} \angle 0^\circ}{0.4\text{A} \angle -71.8^\circ} = 100\Omega \angle 71.8^\circ = R + jX_L \text{ (da } R \text{ und } X_L \text{ in Serie!)}$$

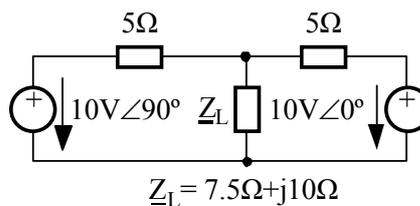
$$R = 100\Omega \cos 71.8^\circ = 31.2\Omega$$

$$X_L = 100\Omega \sin 71.8^\circ = 95.0\Omega$$

Beispiel 6: Berechne den Strom durch \underline{Z}_L in folgenden Schaltungen.



$$\underline{I}_{Z_L} = \frac{10\text{V} \angle 0^\circ - 10\text{V} \angle 90^\circ}{5\Omega + \underline{Z}_L + 5\Omega} = \frac{10\text{V} - j10\text{V}}{30\Omega + j40\Omega} = \frac{14.4\text{V} \angle -45^\circ}{50\Omega \angle 53^\circ} = 0.28\text{A} \angle 98^\circ$$



Quellenumwandlung

$$\underline{I}_1 = \frac{10\text{V}\angle 90^\circ}{5\Omega} = 2\text{A}\angle 90^\circ; \quad \underline{I}_2 = 2\text{A}\angle 0^\circ$$

Gesamtstrom

$$\underline{I}_{\text{tot}} = \underline{I}_1 + \underline{I}_2 = 2\text{A}\angle 90^\circ + 2\text{A}\angle 0^\circ = 2\text{A} + j2\text{A} = 2.83\text{A}\angle 45^\circ$$

Gesamtimpedanz

$$\begin{aligned} \underline{Z}_{\text{tot}} &= (5\Omega \parallel 5\Omega) \parallel \underline{Z}_L = \frac{2.5\Omega (7.5\Omega + j10\Omega)}{2.5\Omega + (7.5\Omega + j10\Omega)} = \frac{(18.75 + j25)[\Omega^2]}{10\Omega + j10\Omega} = \\ &= \frac{31.25\Omega^2 \angle 53.1^\circ}{14.14\Omega \angle 45^\circ} = 2.21\Omega \angle 8.1^\circ \end{aligned}$$

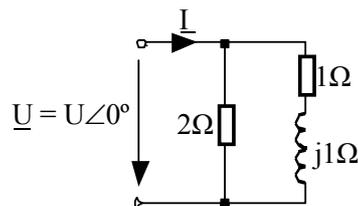
Spannung an \underline{Z}_L

$$\underline{U} = \underline{I}_{\text{tot}} \cdot \underline{Z}_{\text{tot}} = 2.83\text{A}\angle 45^\circ \cdot 2.21\Omega \angle 8.1^\circ = 6.25\text{V}\angle 53.1^\circ$$

Strom durch \underline{Z}_L

$$\underline{I}_{\underline{Z}_L} = \frac{\underline{U}}{\underline{Z}_L} = \frac{6.25\text{V}\angle 53.1^\circ}{12.5\Omega \angle 53.1^\circ} = 0.5\text{A}\angle 0^\circ$$

Beispiel 7: Bestimme für folgende Schaltung das Leistungsdreieck unter der Annahme, dass die im 2Ω -Widerstand umgesetzte Leistung 200 W beträgt.



$$I_R = \sqrt{\frac{P}{R}} = \sqrt{\frac{200\text{W}}{2\Omega}} = 10\text{A}$$

$$U = I \cdot R = 2\text{A} \cdot 10\Omega = 20\text{V}; \quad \underline{U} = 20\text{V}\angle 0^\circ$$

$$\underline{I}_Z = \frac{\underline{U}}{1\Omega + j1\Omega} = \frac{20\text{V}\angle 0^\circ}{1.414\Omega \angle 45^\circ} = 14.14\text{A}\angle -45^\circ = 10\text{A} - j10\text{A}$$

$$\underline{I} = \underline{I}_Z + \underline{I}_R = 10\text{A} + 10\text{A} - j10\text{A} = 20\text{A} - j10\text{A} = 22.36\text{A}\angle -26.53^\circ$$

$$\underline{S} = \underline{U} \cdot \underline{I}^* = 20\text{V}\angle 0^\circ \cdot 22.36\text{A}\angle 26.57^\circ = 447.2\text{VA}\angle 26.57^\circ$$

$$P = S \cos \varphi = 447.2\text{VA} \cos 26.57^\circ = 400\text{W}$$

$$Q = S \sin \varphi = 447.2\text{VA} \sin 26.57^\circ = 200\text{Var}$$

Beispiel 8: Ein einphasiger Wechselstrommotor nimmt von einem $220\text{ V}/50\text{ Hz}$ Netz 3 A auf und liefert bei einem Wirkungsgrad von 85% eine Leistung von 450 W . Berechne den zum Motor parallel zu schaltenden Kondensator, damit der Leistungsfaktor auf 1 verbessert wird. Wie gross ist in diesem Fall der vom Netz aufgenommene Strom?

$$\text{Scheinleistung } S = 220\text{V} \times 3\text{A} = 660\text{VA}$$

Elektrische Wirkleistung

$$P_{\text{el}} = \frac{P_{\text{mech}}}{\text{Wirkungsgrad}} = \frac{450\text{W}}{0.85} = 529.4\text{W}$$

$$\cos \varphi = \frac{S}{P_{\text{el}}} = \frac{529.4\text{W}}{660\text{VA}} = 0.8$$

$$\sin \varphi = \sqrt{1 - \cos^2 \varphi} = \sqrt{1 - 0.8^2} = 0.6$$

Blindleistung: $Q = S \sin \varphi = 660 \times 0.6 = 396\text{Var}$

Damit der Leistungsfaktor $\cos \varphi = 1$ wird, muss die kapazitive Blindleistung des Kondensators gleich der induktiven Blindleistung des Motors sein

$$Q = \frac{U^2}{X_C} = U^2 \omega C$$

$$C = \frac{Q}{U^2 \omega} = \frac{396\text{Var}}{(220\text{V})^2 2\pi 50\text{s}^{-1}} = 26\mu\text{F}$$