

3 Digitale Grundschaltungen in NMOS und CMOS

3.1 Allgemeines zu Schaltkreisfamilien

Digitale Schaltungen dienen der Verarbeitung von digital codierten Nachrichten und Daten. Digital heißt, dass die Variablen - die Komponenten in den Codewörtern - nur zwei Werte annehmen, die üblicherweise mit Null und Eins bezeichnet werden. Die zwei Werte Null und Eins kommen aus der formalen Algebra für Logik, der Bool'schen Algebra, in der Verknüpfungen zwischen Aussagen, die entweder wahr oder falsch sein können, getroffen werden. Solche Verknüpfungen sind Präzisierungen der logischen Begriffe wie „nicht“, „und“, „oder“, „gleich“, „ungleich“ und ähnliche. Zur Entwicklung der Schaltungen dient die Schaltalgebra, die auf der Bool'schen Algebra aufbaut.

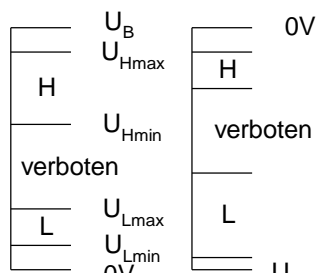


Abbildung 3.1-1: H- und L-Signale in Systemen mit positiver(a) und mit negativer (b) Betriebsspannung

In logischen Schaltungen werden die logischen Werte 0 und 1 durch Signale (Spannungen) in zwei sich nicht überlappenden Bereichen repräsentiert. Es gibt Schaltungstypen (Familien), die mit positiver und solche, die mit negativer Betriebsspannung betrieben werden müssen und demgemäß liegen die Spannungsbereiche für das Ansteuersignal, im positiven oder im negativen Spannungsbereich.

Die sich aus den Ansteuersignalen ergebenden Spannungsbereiche für die Ausgangssignale sind dabei schmaler als diejenigen für die Eingangssignale. Das hat zur Folge, dass bei der Zusammenschaltung mehrerer Grundschaltungen äußere oder innere Störsignale (Rauschen) in gewissen Grenzen nicht zu einem Fehlverhalten der Gesamtschaltung führen.

Die sich nicht überlappenden Bereiche zwischen Eingangs- und Ausgang H bzw. L-Pegel- Signalbereiche werden als Rauschabstände S_H und S_L bezeichnet.

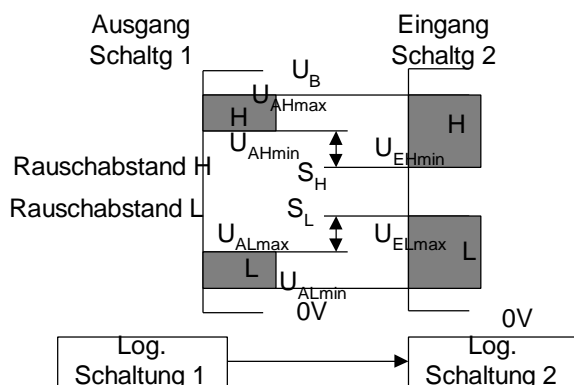


Abbildung 3.1-2: Rauschabstände S_H und S_L

Das Verhalten von logischen Schaltungen wird durch die Arbeitstabelle (Tabelle 3.1-1) beschrieben, in der zu den Kombinationen der Eingangssignale die Ausgangssignale angegeben sind. Je nach Zuordnung des H-Signals zu 1 oder 0 ergibt sich positive oder negative Logik.

Im Beispiel ergeben sich die dualen Funktionen des UNDs bzw. ODERs.

a)															
<table border="1" style="border-collapse: collapse; text-align: center;"> <tr><td>E₁</td><td>L</td><td>H</td><td>L</td><td>H</td></tr> <tr><td>E₂</td><td>L</td><td>L</td><td>H</td><td>H</td></tr> <tr style="border-top: 1px solid black;"><td>A</td><td>1</td><td>1</td><td>1</td><td>0</td></tr> </table>	E ₁	L	H	L	H	E ₂	L	L	H	H	A	1	1	1	0
E ₁	L	H	L	H											
E ₂	L	L	H	H											
A	1	1	1	0											

b)															
<table border="1" style="border-collapse: collapse; text-align: center;"> <tr><td>e₁</td><td>0</td><td>1</td><td>0</td><td>1</td></tr> <tr><td>e₂</td><td>0</td><td>0</td><td>1</td><td>1</td></tr> <tr style="border-top: 1px solid black;"><td>a</td><td>0</td><td>0</td><td>0</td><td>1</td></tr> </table>	e ₁	0	1	0	1	e ₂	0	0	1	1	a	0	0	0	1
e ₁	0	1	0	1											
e ₂	0	0	1	1											
a	0	0	0	1											

c)															
<table border="1" style="border-collapse: collapse; text-align: center;"> <tr><td>e₁</td><td>1</td><td>0</td><td>1</td><td>0</td></tr> <tr><td>e₂</td><td>1</td><td>1</td><td>0</td><td>0</td></tr> <tr style="border-top: 1px solid black;"><td>a</td><td>1</td><td>1</td><td>1</td><td>0</td></tr> </table>	e ₁	1	0	1	0	e ₂	1	1	0	0	a	1	1	1	0
e ₁	1	0	1	0											
e ₂	1	1	0	0											
a	1	1	1	0											

Tabelle 3.1-1: Arbeitstabelle (a) für logische Schaltung mit zwei Eingängen E1 und E2 und einem Ausgang A mit Umwandlung in die Binärtabelle des UNDs bei positiver Logik (b) bzw. des ODERs bei negativer Logik (c).

In der integrierten Schaltungstechnik können Schalterelemente grundsätzlich mit einer Diode oder mit einem Transistor realisiert werden. Praktisch wird jedoch der Transistor eingesetzt, da er als aktives Bauelement im Gegensatz zur Diode bei mehreren Stufen das Signal wieder verstärken und regenerieren kann.

Je nach Transistorart und den damit entwickelten Realisierungen der logischen Grundschaltungen wurden für integrierte digitale Schaltungen verschiedene Technologien (sog. Schaltkreisfamilien) entwickelt. Der Logikteil eines digitalen Systems ist typischerweise aus Bauelementen einer Schaltkreisfamilie aufgebaut. Die Schaltkreisfamilien unterscheiden sich in ihren Eigenschaften und besitzen unterschiedliche Betriebsspannungen. Die Haupttechnologien, die heute in der elektronischen Schaltungsentwicklung eingesetzt werden, sind TTL (Bip-Trans.), CMOS, NMOS (MOS-Transistoren) und Bi-CMOS (Mischtechnologie).

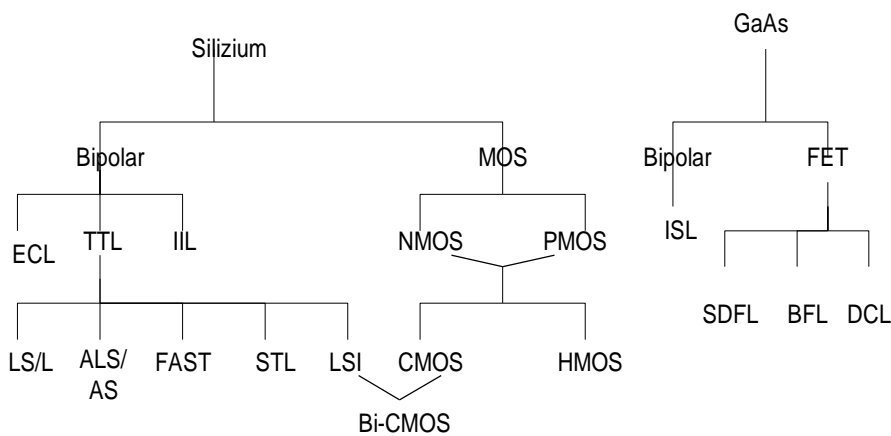


Abbildung 3.1-3: Digitale IC-Technologie-Familien

Die Bipolar-Technologien haben gegenüber den MOS-Technologien den Vorteil, dass die Schalterelemente gute Treiberfähigkeit besitzen und sich gegenüber Störungen (z.B. elektrostatischen Aufladungen) robust verhalten. Allerdings variiert der Innenwiderstand des Transistors mit der angelegten Spannung, so dass sich bei Einsatz als Schalter eine Signalverzerrung ergibt. Dazu kommt, dass der Schalter nur in einer Richtung zu betreiben ist. Der MOS-Transistor ist dagegen symmetrisch aufgebaut und arbeitet somit in beiden Richtungen.

Für einfache Standard-Logik-Bauteile wird die TTL-Technologie (Transistor-Transistor-Logik) eingesetzt. Für schnelle Schaltungen in der Übertragungstechnik (z.B. Laseransteuerungen für Switches und Multiplexer in optischen Übertragungsstrecken) wird die ECL-Technologie eingesetzt.

Mit der MOS-Technologie lassen sich hohe Integrationsdichten (NMOS) und niedriger Leistungsverbrauch (CMOS) realisieren. In NMOS werden hochintegrierte Bausteine wie Mikroprozessoren, signalverarbeitende Bausteine der digitalen Übertragungstechnik und Speicherbausteine gefertigt, während die CMOS-Technologie bei Bausteinen für Mobiltelefone, Uhren, Spiele etc. eingesetzt wird. Der überwiegende Anteil der integrierten digitalen Schaltungen wird in MOS-Technologie gefertigt. Eine Zusammenstellung der verschiedenen Technologien mit ihren Eigenschaften der ist der Tabelle 3.1-2 (aus []) gezeigt.

Familie	Bedeutung	Kommentar
ECL	Emitter coupled logic	Negative Betriebsspannung, deswegen Schwierigkeiten bei der Verbindung mit anderer Logik (interfacing), jedoch sehr schnelle bipolare Schaltungen.
TTL	Transistor-transistor logic	Standard Logikbausteine, z.B. 74er-Serie.
IIL	Integrated injection logic (I ² L)	Schwieriges interfacing, erlaubt jedoch extrem hohe Packungsdichte bei geringem Leistungsverbrauch und kurzen Schaltzeiten
HMOS	High density MOS	Teuer, jedoch hohe Packungsdichte und niedriger Leistungsverbrauch
NMOS	n-Typ MOS	Verwendet n-Kanal MOS-Transistoren mit denen sich hohe Packungsdichte bei skalierbaren Strukturen realisieren lassen. Verglichen mit CMOS hoher Leistungsverbrauch. Einsatz in Speicherbausteinen, Prozessoren etc.
PMOS	p-Typ MOS	Verwendet p-Kanal MOS-Transistoren und ist seit den 80er Jahren durch die NMOS Technik abgelöst worden
CMOS	Complementary MOS	Verwendet n- und p-Kanal MOS-Transistoren und hat sehr geringen statischen Leistungsverbrauch, ist aber andererseits gegenüber NMOS aufwendiger, hat ungünstigere Packungsdichte, langsamere Schaltzeiten und reagiert empfindlicher gegenüber äußeren elektrischen Störungen. Dynamischer Leistungsverbrauch nimmt mit der Taktfrequenz zu. Einsatz: Logische Standardbausteine der Serie 4000 sowie VLSI-Bausteine für ASIC's, FPGA's, Telekommunikationstechnik und Konsumerelektronik.
Bi-CMOS	Bipolar-CMOS	Aufwendige und teure Mischtechnologie mit den positiven Treibereigenschaften wie sie die Bipolartechnik aufweist.
GaAs	Gallium Arsenide	Sehr geringe Schaltzeiten wegen der um der Faktor 10 mal höheren Beweglichkeit der Ladungsträger. Taktfrequenzen bis in den GHz Bereich. Aufwendige und teure Technik mit wesentlich kleineren Scheiben. Nur für Spezialanwendungen in HF-Technik. IC bis einige hundert Transistoren möglich. Durch die bei Si ausgereifere Technologie (kleinere Strukturen) wird der Abstand bezüglich Schnelligkeit gegenüber GaAs stark verringert.
DCL	Direct coupled logic	Ein einfache GaAs-Schaltung, die aber schwierig herzustellen ist, da sie zwei Typen von Transistoren verwendet. Verglichen mit anderen GaAs Logik-Typen rel. geringer Leistungsverbrauch.
BFL	Buffered FET logic	Einfacher zu fertigen als DCL, aber höherer Leistungsverbrauch
SDFL	Schottky diode logic	Fertigungsverfahren einfacher als bei DCL, niedrigerer Leistungsverbrauch als bei BFL

Tabelle 3.1-2: Übersicht über Eigenschaften digitaler IC-Technologien

Die TTL-Technologie wurde verbessert und weiterentwickelt. Die Abkürzungen stehen für:

L= Low-Power TTL

S= Schottky TTL

LS= Low-Power-Schottky,

FAST= Fairchild-Advanced-Schottky -TTL

ALS= Advanced LS

In der Tabelle 3.1-3 (aus []) ist ein Vergleich zwischen den verschiedenen Technologien zu finden, wobei diese jeweils bzgl. der wichtigsten Eigenschaften (nämlich Schaltgeschwindigkeit, Integrationsdichte, und Leistungsverbrauch) der Reihe nach geordnet wurden. Dabei wird allerdings von vergleichbarem Stand der Technologie ausgegangen, also gleichen minimalen Strukturabmessungen.

Position	Geschwindigkeit	Packungsdichte	Leistungsverbrauch
1	GaAs (höchste)	I ² L (geringste)	I ² L (niedrigster)
2	ECL	HMOS	CMOS
3	TTL	NMOS	NMOS
4	HMOS	CMOS	HMOS
5	NMOS	TTL	TTL
6	I ² L	GaAs	GaAs
7	CMOS (niedrigste)	ECL (niedrigste)	ECL (höchster)

Tabelle 3.1-3: Reihenfolge der digitalen Technologien bezüglich Geschwindigkeit, Packungsdichte und Leistungsverbrauch

Eine digitale Grundschialtung muß folgende Funktionen leisten: Die Signale müssen in zwei Bereiche quantisiert verarbeitet werden und die Signale müssen logisch verknüpft werden können. Dabei muß die Quantisierung aufrecht erhalten werden. Die beiden Grundfunktion der Quantisierung und Inversion werden von derselben Grundschialtung , nämlich dem Inverter, vorgenommen. Schaltungen mit Verknüpfungsfunktion ergeben sich dann aus den Prinzipien der Schaltalgebra. Die Realisierung des Inverter ist somit das charakteristische und typische für eine spezielle Schaltkreisfamilie und die Eigenschaften der Familie (Schnelligkeit, Leistungsverbrauch, Rauschabstände, Kosten) lassen sich weitgehend auf die Eigenschaften des Inverters zurückführen.

Im folgenden werden die Grundschialtungen nur für die NMOS und CMOS- Technologie besprochen. Für Ausführungen zu den bipolaren Technologien wird auf die Literatur verwiesen (z.B. Weißel- Schubert, Buchanan). Wir behandeln zuerst den Inverter und dann die daraus ableitbaren Schaltungen für die NOR und NAND- Funktion. Eine wichtige Schaltung, die nur in CMOS Technologie möglich ist, ist das Transmission Gate. Dieser Durchgangsschalter arbeitet ohne das Signal zu schwächen und kann in beide Richtungen betrieben werden. Aus den Grundschialtungen lassen sich dann kompliziertere Baugruppen , wie Äquivalenzen, Flip-Flops, Zähler, Register, Speicher, Codeumsetzer, Multiplexer und Rechenschaltungen aufbauen. Neben der Schaltung wird auch der Schichtaufbau der physikalischen Realisierung schematisch dargestellt, der die Grundlage für das Verständnis des Maskenlayouts ist.

3.2 Inverter, NOR und NAND

Inverter

Aufgabe: Überträgt binär quantisiertes Eingangssignal an Ausgang

Qualitätsmerkmale:

- Rauschabstände möglichst groß
- Verzögerungszeit möglichst gering
- Leistungsverbrauch möglichst niedrig

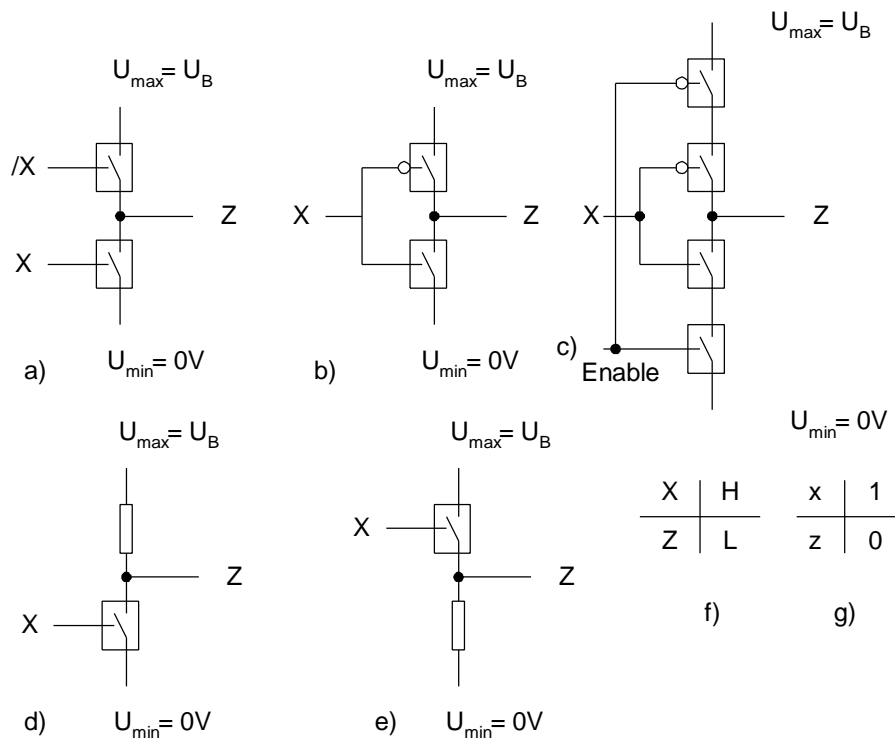


Abbildung 3.2-1: Grundsätzliche Inverter-Strukturen a) Zwei Schalter und zwei Eingänge mit invertierten Signalen. (TTL) b) zwei komplementäre Schalter (CMOS) c) wie b) mit abgetrenntem Ausgang für $\text{Enable} = L$ (Tristate-Inverter) d) aktives Pull-down Element und passives Pull-up Element (TTL NMOS), e) passives Pull-down Element und aktives Pull-up Element (ECL) f) und g) Arbeitstabelle und Wahrheitstabelle für Inverter

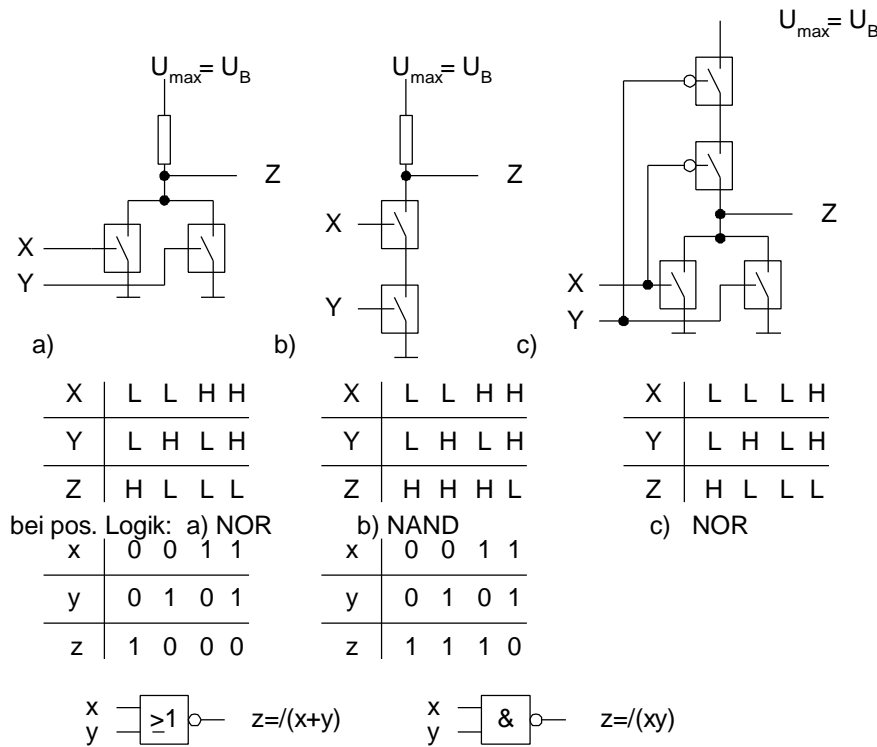


Abbildung 3.2-2: Prinzip der binären logischen Grundsaltungen mit aktiven Pull-down Elementen a) NOR b) NAND und c) NOR mit zwei aktiven Schaltern (CMOS)

Bei mehreren aktiven Pull-down Strukturen können die Ausgänge (fest) zusammen verbunden werden, ohne dass das Signal den quantisierte Bereich verläßt. Die Zusammenfassung ergibt eine AND-Struktur (**wired – AND**). Bei der Verbindung von Ausgängen passiver Pull-down Strukturen ergibt sich eine **wired -OR** Struktur.

3.3 Inverter mit Widerstandslast

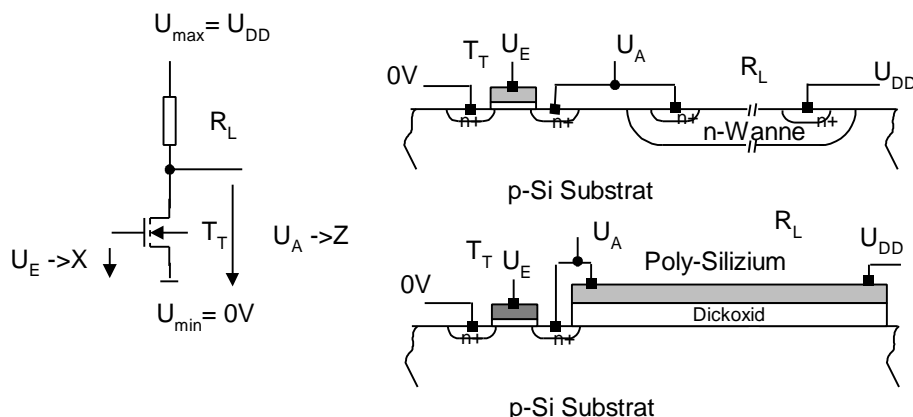


Abbildung 3.3-1: Inverter mit Widerstandslast. Schaltung und prinzipielle Realisierungsmöglichkeiten mit passivem Widerstand

Für $U_E \leq U_{th}$ ist der Transistor geschlossen und die Ausgangsspannung ist U_{DD} . Für $U_E \geq U_{th}$ beginnt der Transistor zu leiten. Die Ausgangsspannung U_A ist gleich der U_{DS} des Transistors und ergibt sich aus dem Spannungsteiler zwischen Lastwiderstand R_L und dem Source-Drain –Widerstand. Bei maximaler Eingangsspannung U_{DD} ist der Widerstand des Transistors am geringsten. Bei geeigneter Dimensionierung kann die Ausgangsspannung im Low Zustand unter die Schwellenspannung gedrückt werden (Abbildung 3.3-3). Aufgrund dieser Restspannung verbraucht der Inverter im Low-Zustand Leistung. Praktisch wird für die Restspannung ein Wert $U_R = 0.2 U_{th}$ bis $U_{th}/2$ gewählt.

Graphische Ermittlung der Inverterkennlinie $U_A=f(U_E)$

Der Strom durch den Transistor kann aus dem Kennlinienfeld ermittelt werden, in welches die Lastgerade für den Widerstand eingezeichnet wird. Durchläuft die Eingangsspannung den Spannungshub von $0V$ bis zu U_{DD} ($=5V$), so wird die Lastgerade von rechts nach links durchlaufen. Aus den Schnittpunkten der Lastgerade mit den Kurven des Kennlinienfeldes kann man zu jeder Eingangsspannung die Ausgangsspannung ablesen, siehe Abbildung 3.3-2 und Abbildung 3.3-3 .

$$U_{DD} - U_A = R_L \cdot I_{DS}$$

und $I_{DS} = f(U_E, U_A)$

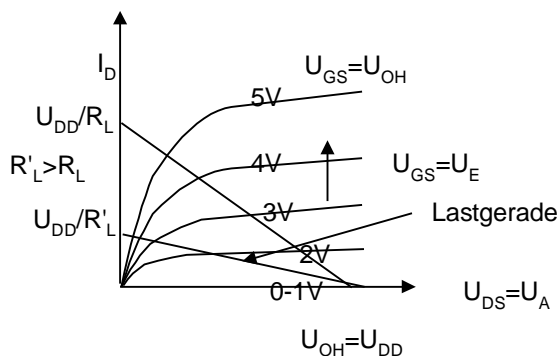


Abbildung 3.3-2: Zur graphischen Ermittlung der Inverter-Übertragungskennlinie

Statische Verlustleistung im Low-Zustand:

$$P_{stat} = U_{DD} I_R \quad I_R \approx \frac{U_{DD}}{R_L}$$

$$P_{stat} \approx \frac{U_{DD}^2}{R_L}$$

High- und Low-Bereiche

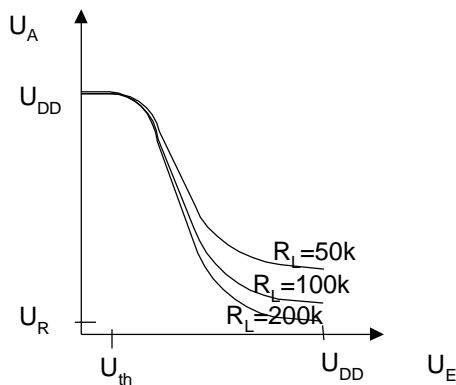


Abbildung 3.3-3: Übertragungskennlinie des Inverters mit Widerstandslast für drei verschiedene Widerstandswerte

Die Spannungsbereiche für Low und High sind als die Abschnitte festgelegt, in denen die Kennlinie eine Steigung $|dU_A/dU_E| = |V| < 1$ aufweist. Der Bereich, in dem $|V| > 1$, ist verboten, da es hier bei Zusammenschaltung mehrerer Gatter durch Störspannungen zu Fehlfunktionen kommen kann.

Aufgrund der nichtlinearen Kennlinie werden die erlaubten Bereiche für High und Low in schmalere Bereiche am Ausgang transformiert. Bei Hintereinanderschaltung von zwei Invertiern wird ein High-Signal angehoben und ein Low Signal abgesenkt. Dieses Verhalten führt dazu, daß

Störspannungen, die in einer logischen Schaltung auf die Signale einwirken, nach Durchlaufen mehrerer Glieder der Schaltung wieder entfernt werden. Das Logiksignal wird regeneriert.

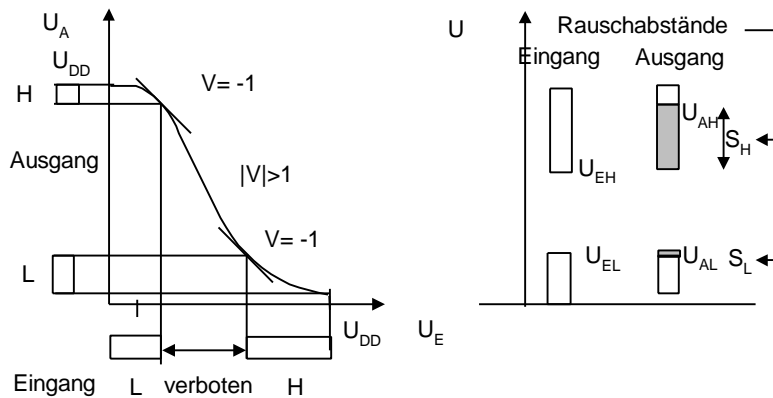


Abbildung 3.3-4: Inverterkennlinie, High und Low Bereiche, sowie Rauschabstände

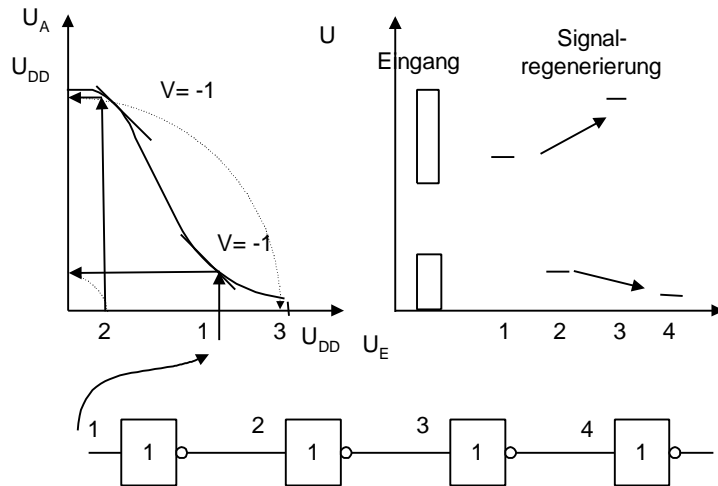


Abbildung 3.3-5: Regenerierung eines gestörten Signals beim Durchlaufen einer Inverterkette

Die Signalentwicklung in einer Inverterkette läßt sich auch in einem Diagramm verfolgen, das die Inverterkennlinie und ihre Inverse enthält. Man erkennt, daß es zwei stabile Zustände gibt, je nachdem, ob das Signal $> U_{DD}/2$ oder $< U_{DD}/2$ ist.

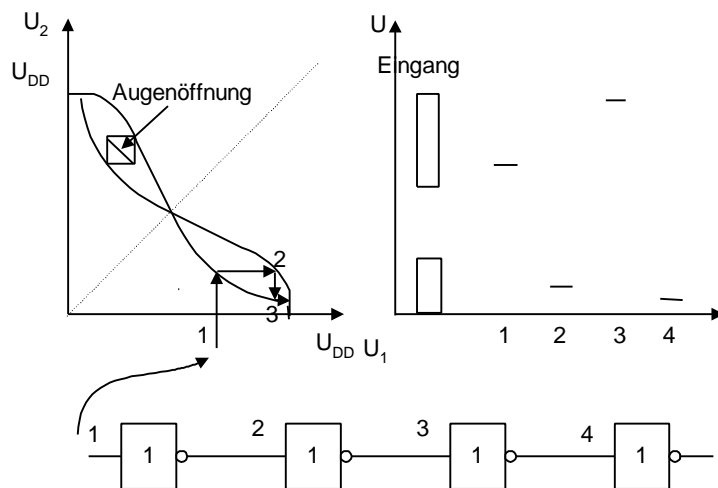


Abbildung 3.3-6: Überlagerung der Kennlinien in einer Inverterkette

Störabstand

Es werden zwei rückgekoppelte Inverter (Flip-Flop) betrachtet, an deren Eingänge additiv (durch kapazitive Einkopplung) die Störspannungen $\Delta U_1 = \Delta U_2 = \Delta U$ gelegt werden. Der Störabstand entspricht dem Störsignal ΔU , das aufgewendet werden muß, damit das Flip-Flop aus einem stabilen Punkt in den anderen kippt. Durch die Störspannung wird die Inverter-Kennlinie um ΔU in U_2 Richtung (nach oben) verschoben $U_2=f(U_1)$, während die Inverse um ΔU in U_1 Richtung (nach links) verschoben wird. Dadurch wird der Bereich zwischen

dem stabilen Punkt A' und dem Umschlagpunkt B' schmaler. Durch weiteres Erhöhen von ΔU verschwinden die Schnittpunkte B' und A'. Daraus folgt: der stabile Bereich existiert nicht mehr.

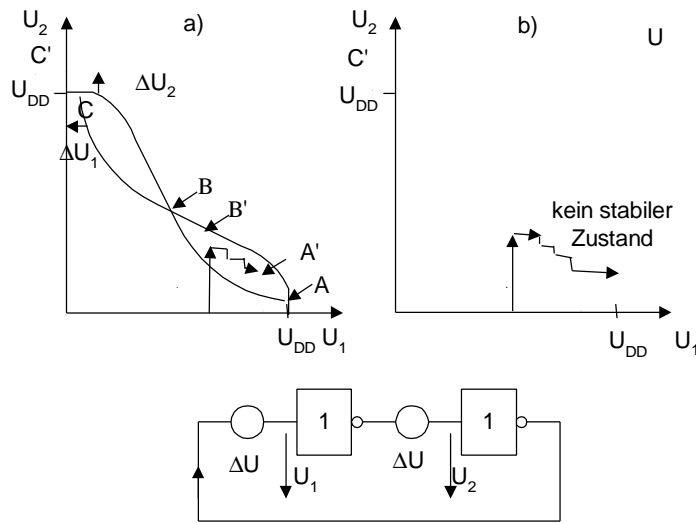


Abbildung 3.3-7: Rückgekoppelte Inverter mit zwei stabilen Zuständen A und C und einem labilen Punkt. Mit zunehmender Störspannung wird das Augenfenster für den zweiten stabilen Zustand A' kleiner und verschwindet für den Störabstand

Schreibt man in die „Augenöffnung“ der überlagerten Kennlinien ein Rechteck ein, so ergibt die Diagonale ungefähr den Störabstand.

Schaltverhalten

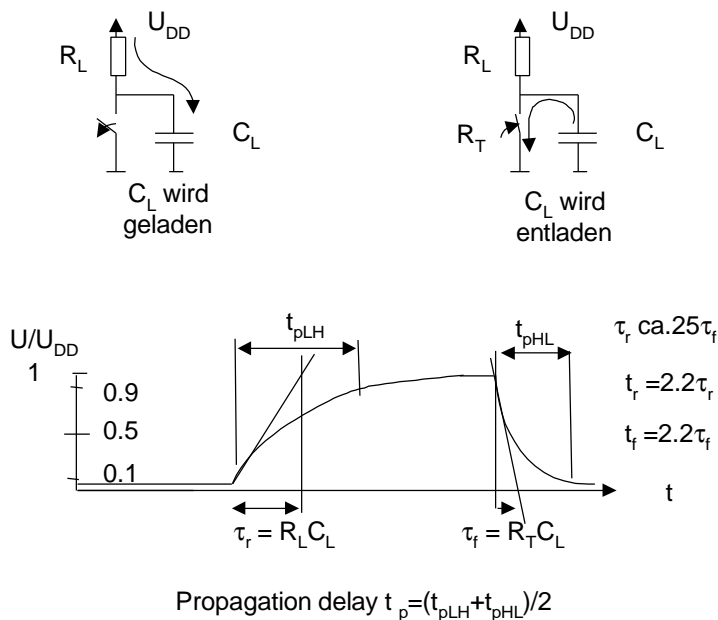


Abbildung 3.3-8: Schaltverhalten des Inverters mit Widerstandslast.

Damit die Restspannung niedrig genug ist, um einen ausreichenden Störabstand zu erreichen, muß der Widerstand des Schalttransistors im eingeschalteten Zustand (über das W/L Verhältnis) etwa 25 mal kleiner als der Lastwiderstand gewählt werden (dann ist $U_R = 0.2 U_{th}$). Für die Schwellenspannung wird bei $U_{DD} = 5 \text{ V}$ ein Wert von ca. 1 V gewählt. Dann sperrt einerseits der Transistor für $U_{GS} = 0 \text{ V}$ sicher, andererseits bleibt aber die Schaltgeschwindigkeit hoch.

Praxis : $R_L > 100 \text{ k}\Omega$. Dies bereitet in Standard MOS-Technologien (diffundierte Widerstände) große Schwierigkeiten. Bei Verwendung eines Widerstands bestehend aus einer meanderförmigen Bahn aus undotiertem polykristallinem Silizium würde sich bei einer Standarddicke und einer Standardbreite von je 1 μm eine Länge von 1000 μ ergeben , was gegenüber dem Transistor zu einer unverhältnismäßig großen Fläche führen würde. Daher werden für den Lastwiderstand aktive Bauelemente eingesetzt: ein passiver selbstleitender n-MOSFET (sog. depletion load transistor) bei NMOS oder ein aktiv geschalteter p-MOSFET in der CMOS-Technik.

Ausnahme SRAM: über dem Transistor wird in einer speziellen Technik ein hochohmiger Polysilizium – Lastwiderstand angebracht.

3.4 NMOS-Inverter

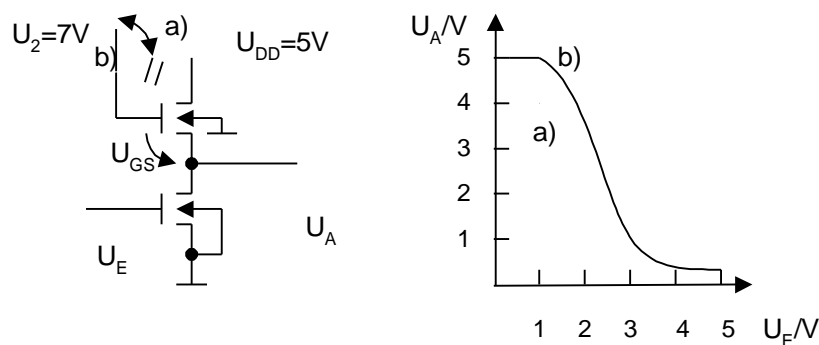


Abbildung 3.4-1: NMOS-Inverter mit Anreicherungslasttransistor. T_L a) Gate von T_L mit U_{DD} verbunden ergibt $U_{OH} < U_{DD}$, b) Zusätzlicher Spannungsquelle U_2 für $U_{GD,L}$

In Abbildung 3.4-1 sind Schaltungsmöglichkeiten gezeigt, in denen der Lasttransistor vom selben Typ wie der Schalttransistor ist. Diese Schaltungsvarianten haben jedoch gravierende Nachteile : Abbildung 3.4-1 a) Gate des Lasttransistors ist mit Spannungsversorgung verbunden $\rightarrow U_{OH} = U_{DD} - U_{th}$, Abbildung 3.4-1 b) zusätzliche Spannungsquelle mit $U_2 > U_{DD}$

Praktisch wird der Inverter mit einem Lasttransistor vom Verarmungstyp (depletion-load-transistor, selbstleitend), realisiert (Abbildung 3.4-2).

Vorteile zu Widerstandslast:

- Geringerer Platzbedarf
- Bessere Störabstände

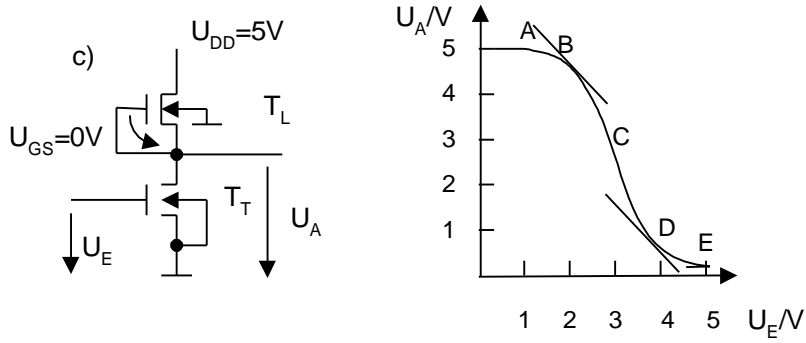


Abbildung 3.4-2: NMOS –Inverter mit depletion- Lasttransistor T_L . Konstruktion der Kennlinie siehe Abbildung 3.4-3.

Graphische Erläuterung zum Zustandekommen der Kennlinie siehe Abbildung 3.4-3

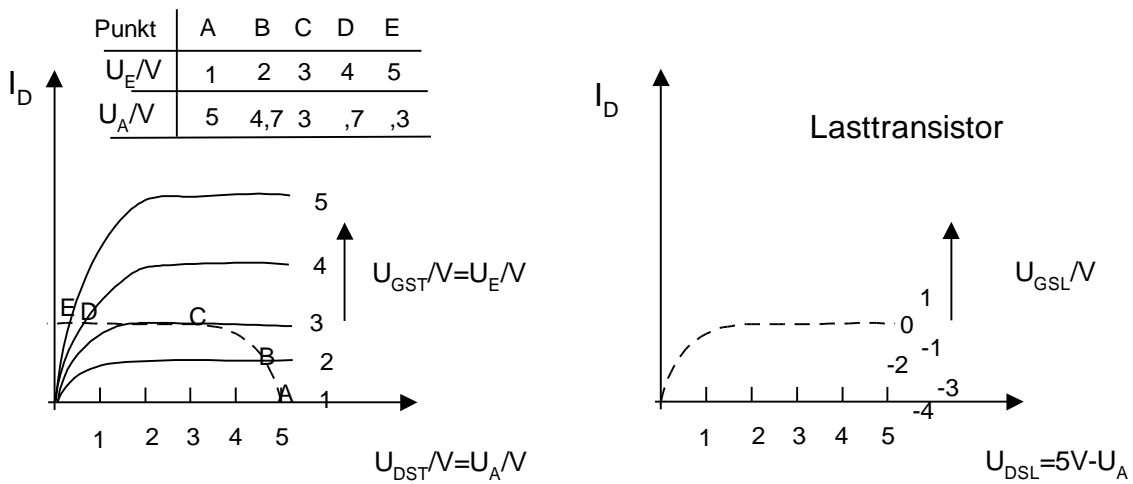


Abbildung 3.4-3: Graphische Konstruktion der Kennlinie des NMOS-Inverters mit (passivem) n-Kanal Transistor vom Avarmungstyp (normally on).

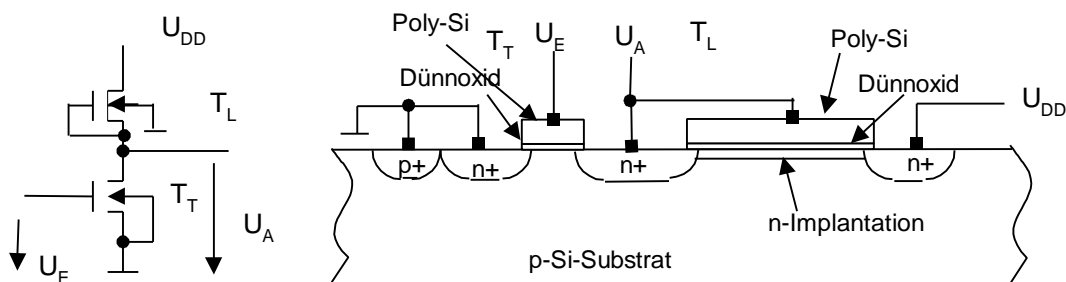


Abbildung 3.4-4: Schichtaufbau NMOS-Inverter

Beispiel: Dimensionierung der Transistoren im NMOS-Inverter

Wir berechnen den Strom durch beide Transistoren für den Fall, dass sich der Transistor im Low Zustand befindet und wählen für die Restspannung einen Wert der kleiner als die Schwellenspannung ist (Wahl: $U_R = 0.2U_{th}$). Daraus ergibt sich das Verhältnis der Verstärkungsfaktoren und damit die Geometrieverhältnisse. Zahlenwerte: $U_{DD} = 5\text{ V}$, $U_{thT} = 1\text{ V}$, $U_{thL} = -4\text{ V}$.

Aus der Zeichnung entnimmt man, daß im Low-Zustand des Inverters der Treibertransistor im Widerstandsgebiet und der Lasttransistor im Sättigungsgebiet arbeitet. Aus dem Gleichsetzen der Ausdrücke für den Strom durch den Treibertransistor und den Lasttransistor folgt der Ausdruck für die Restspannung: also

$$I = b_T \left[(U_{GSL} - U_{thT}) U_{DST} - \frac{U_{DST}^2}{2} \right] \approx b_T [(U_{DD} - U_{thT}) U_R]$$

$$I = \frac{b_L}{2} [U_{GSL} - U_{thL}]^2 = \frac{b_L}{2} [-U_{thL}]^2$$

$$\frac{b_L}{2} [-U_{thL}]^2 = b_T [(U_{DD} - U_{thT}) U_R] \Rightarrow U_R = \frac{b_L U_{thL}^2}{2 b_T (U_{DD} - U_{thT})} = 0.2 U_{thT}$$

$$\Rightarrow \frac{b_L}{b_T} = \frac{2 U_R (U_{DD} - U_{thT})}{U_{thL}^2} = \frac{2 \cdot 0.2 \text{V} (5\text{V} - 1\text{V})}{4^2 \text{V}^2} = 0.1$$

$$b = k_p \frac{w}{L} \Rightarrow \left(\frac{w}{L} \right)_L = 0.1 = \frac{1/3}{3.3/1}$$

Bei gleichen Kanallängen sollte also die Weite des n-Kanal-Transistors etwa 10 mal so groß sein wie die Weite des n-Kanal-depletion load-Transistors („ratioed logic“). Falls die Kanallänge verändert werden kann ist das Verhältnis w/L für den Treibertransistor 3.3/1 und den Lasttransistor 1/3 zu wählen (Abbildung 3.4-5) („ratioed logic“).

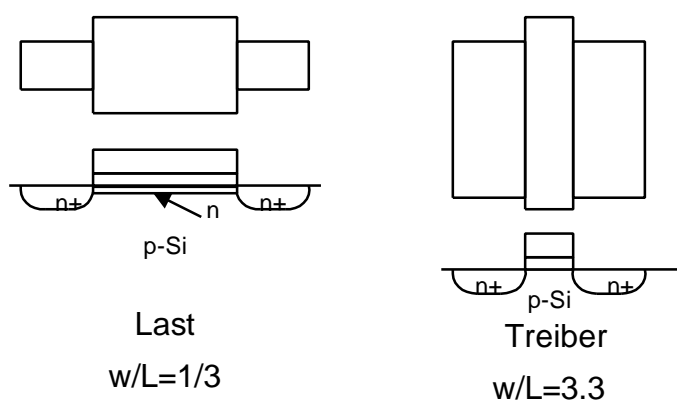


Abbildung 3.4-5: Dimensionierung des NMOS Inverters

Schaltzeiten: Es gelten die Formeln

$$t_r = \frac{4C_L}{b_L U_{DD}} \quad t_f = \frac{5C_L}{b_T U_{DD}} \quad t_p = \frac{C_L}{U_{DD}} \left(\frac{3}{b_L} + \frac{1.5}{b_T} \right)$$

Beispiel:

$$U_{DD} = 5V, C_L = 0,1 \text{ pF}, (W/L)_L = 1/3, (W/L)_T = 2, k_p = 40 \mu\text{F/V}^2 \rightarrow t_p = 4,9 \text{ ns}$$

Pseudo-CMOS-Logik

In der CMOS-Technologie wird anstelle des passiven (nicht vom Steuersignal beeinflussten Last bzw. Pull-Up Transistors) ein aktiver p-Kanal Transistor eingesetzt. Dies führt zu entscheidenden Vorteilen , insbesondere, daß der Inverter im Ruhezustand keinen Strom verbraucht. Der Nachteil ist allerdings, daß beim Aufbau von logischen Gattern im Pull-Up Zweig mehr Transistoren nötig sind als bei der NMOS-Technologie. Dies führt dazu, daß auch in der CMOS-Technologie gelegentlich von einer Inverterstruktur mit passivem p-Kanal Lasttransistor (vom Anreicherungstyp) gebraucht gemacht wird, der Pseudo-NMOS-Logik.

Bei VLSI-Bausteinen in CMOS-Technologie wird somit neben der leistungssparenden CMOS-Technik für einzelne Module auch die platzsparende Pseudo-NMOS Technik eingesetzt, wenn das insgesamt Vorteile bringt [Klar].

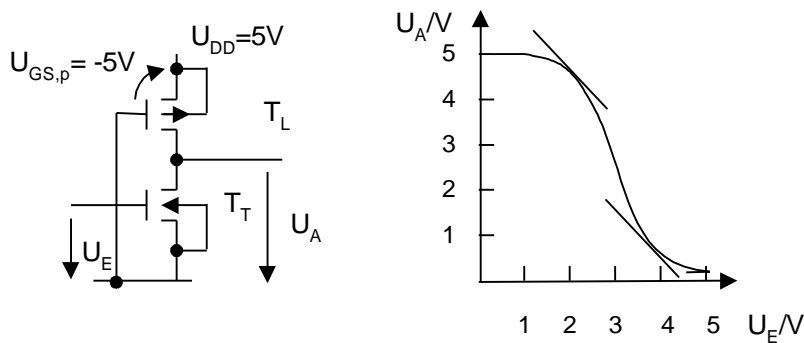


Abbildung 3.4-6: Pseudo-NMOS-Inverter in der CMOS-Technologie

Der Pseudo-NMOS Inverter (Abbildung 3.4-6) verhält sich wie der NMOS Inverter, hat also insbesondere eine Restspannung im Low- Zustand, $U_{OL} = U_R$, und deswegen einen statischen Leistungsverbrauch.

Die Dimensionierung eines Pseudo-NMOS-Inverters wird aus der Bedingung abgeleitet, daß die Schaltschwelle bei $U_{DD}/2$ liegen soll. Dann ist automatisch die Bedingung

$$U_{OL} = (1/2) U_{th} \text{ erfüllt.}$$

$$\frac{b_n}{2} [U_E - U_{thn}]^2 \approx b_p [(U_{DD} - |U_{thp}|) U_{DSp}]$$

$$U_E = U_{DS,p} = 2.5V, \quad U_{thn} = |U_{thp}|, \quad \frac{m_n}{m_p} = 2 \quad \Rightarrow \quad \left(\frac{W}{L} \right)_n = 4$$

Bei gleichen Kanallängen muß die Weite des n-Kanal-Transistors etwa viermal so groß wie die Weite des p-Kanal-Transistors („ratioed logic“) gewählt werden , vgl. NMOS Inverter : etwa 10 mal .

3.5 CMOS-Inverter

In der CMOS Technologie wird der Inverter aus zwei aktiven Schaltern, einem n-Kanal und dem dazu komplementären p-Kanal Transistor, aufgebaut. Für das Eingangspotential $U_E < U_{DD}/2$ wird über die Transistoren der Ausgang des Inverters an $U_A = U_{DD}$ gelegt für $U_E > U_{DD}/2$ an $U_A = 0V$. Dabei ist jeweils nur ein Transistor geöffnet während der andere geschlossen ist, so daß im Ruhezustand kein Strom fließt. Lediglich beim Schalten fließt ein Strom durch beide Transistoren.

Vorteile:

- $P_{stat} = 0$
- $t_r = t_f$
- $S_H = S_L$

Nachteile

- etwas mehr Fläche als NMOS
- kompliziertere Technologie
- latch-up Problem

Für die Abbildungen wird $U_{DD} = 5V$, $U_{thn} = 1V$ und $U_{thp} = -1V$ angenommen.

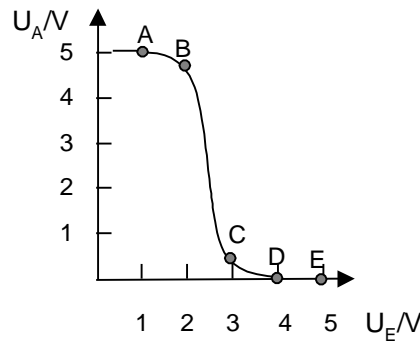
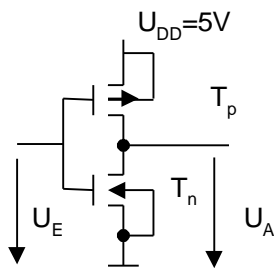
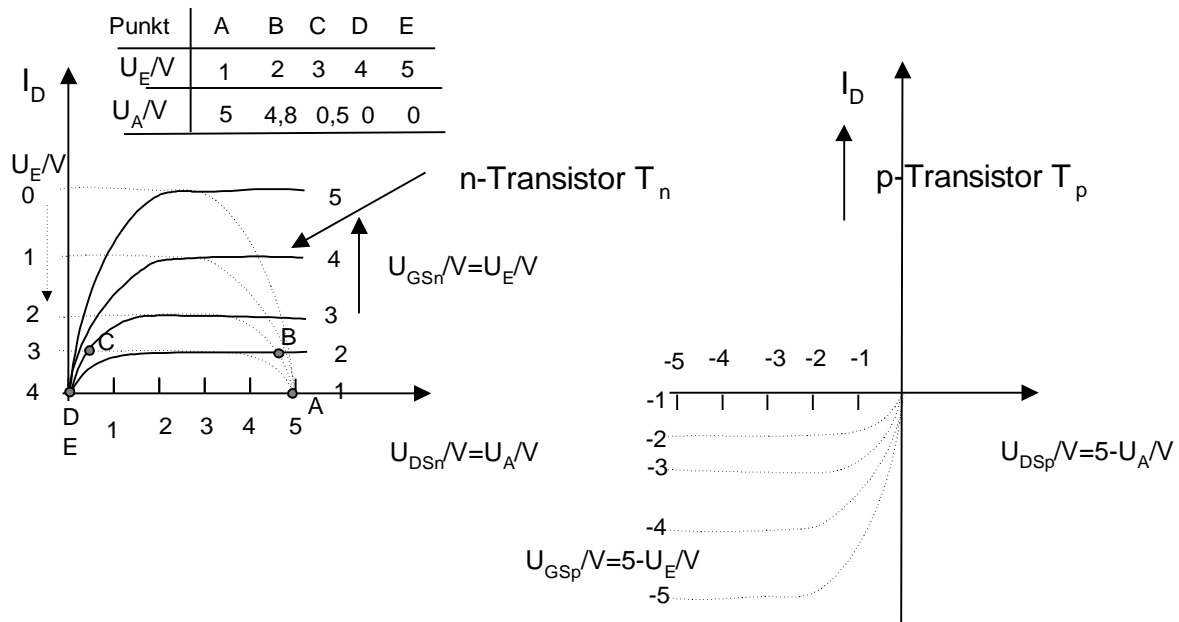


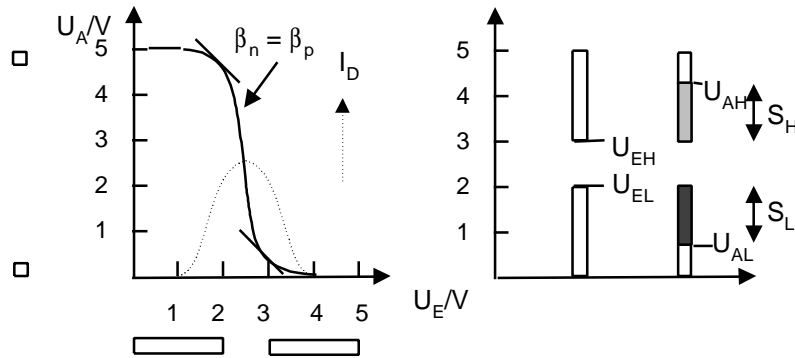
Abbildung 3.5-1: CMOS-Inverter und Kennlinie

Abbildung 3.5-2: (unten) Graphische Konstruktion der CMOS-Inverter-Kennlinie

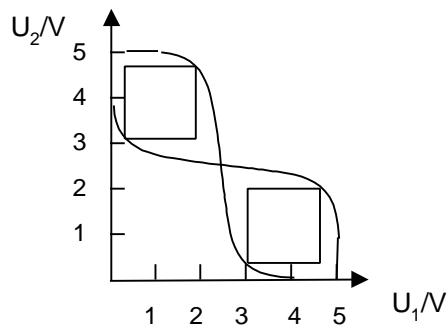


Aus der graphischen Konstruktion der Kennlinie erkennt man,

- bei gleicher Verstärkung der Transistoren ist die Kennlinie symmetrisch bzgl. des halben Betriebsspannungswertes
- es gibt keine Restspannung
- die Rauschabstände sind größer als bei NMOS
- für $\beta_n = \beta_p$ ist $S_L = S_H$ und $t_r = t_f$



Ein Strom fließt nur während des Schaltens. Für $\beta_n < \beta_p$ ist die Kurve nach links verschoben und umgekehrt.



Beispiel:

Dimensionierung eines optimalen CMOS Inverters mit symmetrischer Kennlinie:

Aus $\beta_n = \beta_p$ folgt $(W/L)_p / (W/L)_n = \mu_n / \mu_p$

Das Verhältnis μ_n / μ_p ist für Gatelängen $> 1 \mu\text{m}$ ungefähr 2 und nimmt im Sub-Mu-Bereich weiter ab.

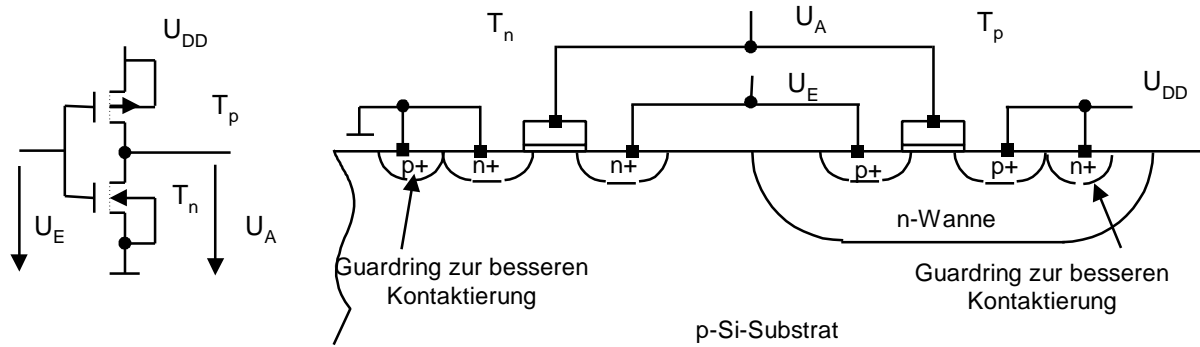


Abbildung 3.5-3: CMOS-Inverter : Schichtaufbau . Der Anschluß der Leiterbahnen und der dazu notwendigen Oxidisationsschichten ist nicht gezeigt.

In der Abb. 3.5-4 sind Layout Varianten gezeigt [Klar]. Dabei sind die geometrischen Verhältnisse nur schematisch angegeben und die unterschiedlichen Kanalweiten zur Einstellung gleicher Verstärkung beim n und p-Transistor nicht berücksichtigt. A) und b) unterscheiden sich durch die Plazierung der Diffusionsgebiete, einmal senkrecht zu den Signalleitungen, das andere Mal parallel dazu. c) bis e) zeigen Varianten, bei denen zusätzliche Leitungen eingefügt sind..

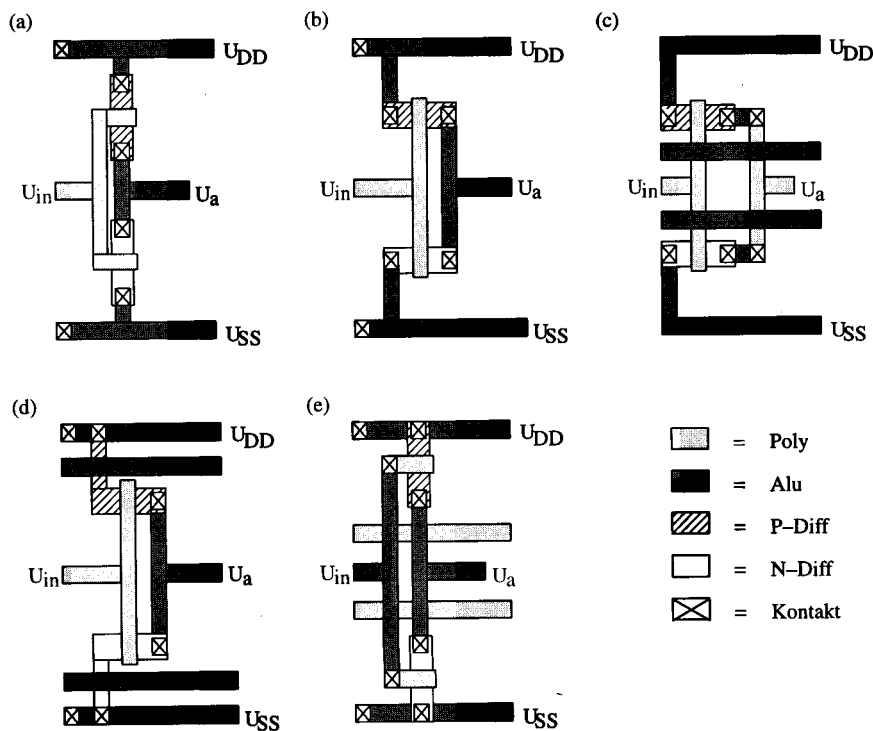


Abbildung 3.5-4: Layout-Varaianten für den CMOS-Inverter ohne Berücksichtigung der Dimensionierungsvorschriften für w/L a) und b) . In Abbildung c) bis e) ist gezeigt wie zusätzliche Signalleitungen eingefügt werden können.

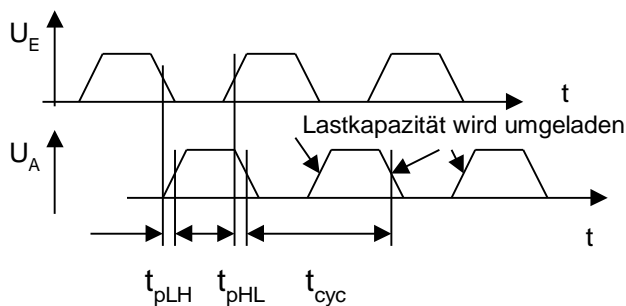
Schaltzeiten: Es gelten die Formeln

$$t_r = \frac{4C_L}{b_p U_{DD}} \quad t_f = \frac{4C_L}{b_p U_{DD}}$$

Dynamische Verlustleistung

Da der CMOS-Inverter nur während des Schaltens Strom zieht ist die Verlustleistung von in CMOS gefertigten IC's geringer als die von NMOS-IC's. Die Verlustleistung steigt proportional der Taktfrequenz an.

Verlustleistung eines Inverters, der mit dem Takt $f_{cyc} = 1/t_{cyc}$ geschaltet wird:



Während einer Periode des Taktes wird die Lastkapazität einmal aufgeladen und einmal entladen. Dabei muß die Energie $W = \frac{1}{2} C_L U_{DD}^2$ zu und wieder abgeführt werden. Damit ergibt sich die Verlustleistung zu

$$P = \frac{dW}{dt} = \frac{2W}{t_{cyc}} = C_L U_{DD}^2 f_{cyc}$$

In einem IC sind während jeder Periode des Taktes nicht alle Schaltelemente aktiv. Für eine Abschätzung geht man davon aus, daß die Hälfte der Gatter (der Anzahl N) aktiv ist; dann ist die Verlustleistung des IC's gegeben durch:

$$P_{dyn} = \frac{N}{2} C_L U_{DD}^2 f_{cyc}$$

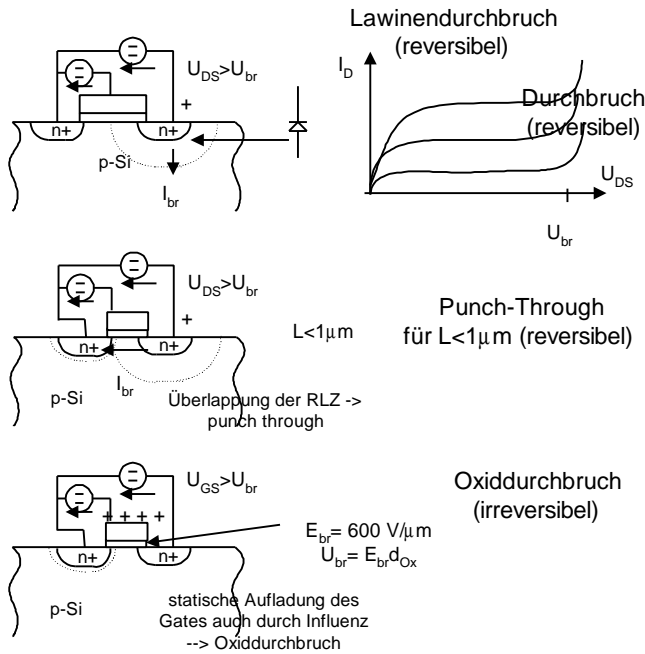
Bei hohen Taktfrequenzen kann der Leistungsverbrauch einer CMOS-Schaltung den statischen Leistungsverbrauch einer NMOS-Schaltung (oder auch TTL) übersteigen. Maßnahmen zur Verringerung der Verlustleistung bei VLSI-Bausteinen:

- C_L erniedrigen
- U_{DD} erniedrigen

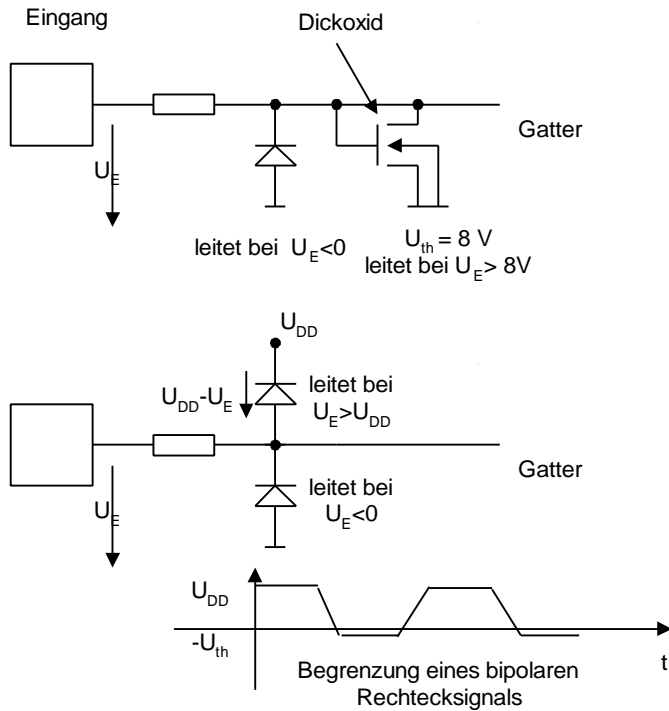
Zusätzlich werden zur effektiven Wärmeabfuhr Kühlkörper und Ventilatoren angebracht.

Beispiel: $C_L = 0,1 \text{ pF}$, $U_{DD} = 5 \text{ V}$, $N = 10^6$ Gatter, $f_{cyc} = 10 \text{ MHz} \rightarrow$
 $P_{dyn} = (10^6/2) \cdot 0,1 \cdot 10^{-12} \cdot \text{F} \cdot 25 \cdot \text{V}^2 \cdot 10^7 \text{ Hz} = 12,5 \text{ W}$

3.6 Durchbruchmechanismen



Schutzschaltungen



Latch-up

In der **CMOS-Technologie** wird aufgrund des Aufbaus des Inverters eine parasitäre Thyristor-Struktur (zwei verkoppelte bipolare Transistoren npn und pnp) erzeugt, die bei Normalbetrieb und Spannungen geschlossen ist, aber bei bestimmten Signalspannungsübergängen gezündet werden kann und zur Zerstörung des Bauelements oder zu schwerwiegenden Fehlern in der Funktion der Schaltung führen kann.

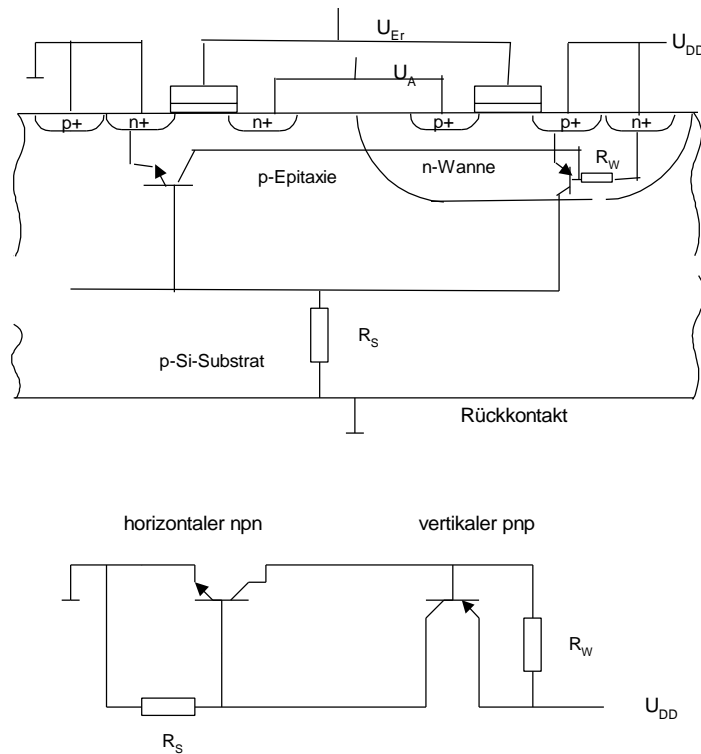


Abbildung 3.6-1: CMOS-Inverter mit parasitärer Thyristor-Struktur, gebildet aus einem horizontalen bipolaren npn und einem lateralen bipolaren pnp-Transistor. Aufgrund dieser Struktur kann es zum latch-up-Effekt kommen

Wie in der Abbildung 3.6-1 zu sehen ist enthält der CMOS-Inverter zwei rückgekoppelte Bipolar-Transistoren, sowie zwei Widerstände R_S und R_W . Der Latch-up Effekt kann ausgelöst werden durch

- Überschreiten des Nennwertes der Versorgungsspannung und des Massepotentials
- Schnelle Schwankungen der Versorgungsspannung

In diesen Fällen entsteht am Widerstand R_W ein Spannungsabfall der den vertikalen pnp-Transistor zum Leiten bringt. Dann fließt auch ein Strom in die Basis des npn-Transistors und bringt diesen zum Leiten. Nunmehr besteht zwischen Versorgungsspannung und Masse des Substrates eine leitende Verbindung. Dadurch kann es im ungünstigsten Fall zu hohen Strömen kommen, die das Bauteil zerstören. In logischen Schaltungen kann auch ein Gatter in einem Zustand „hängen“ bleiben. Der Effekt kann erst durch Abschalten der Versorgungsspannung wieder entfernt werden.

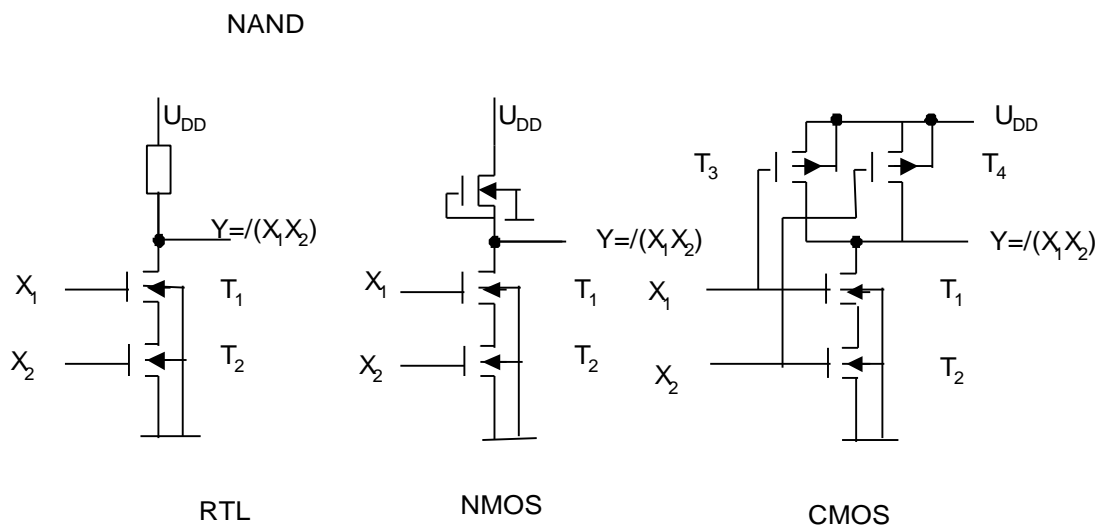
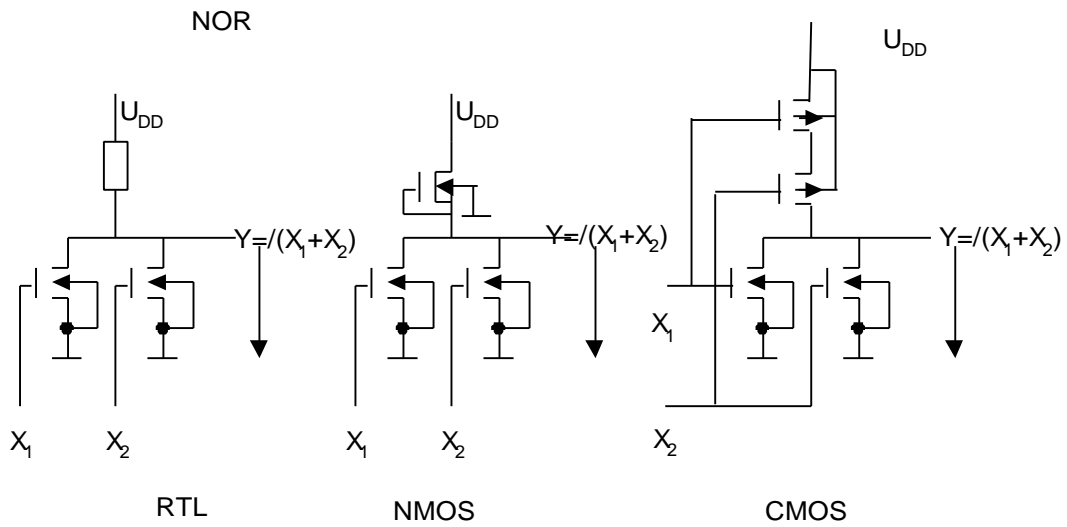
Der latch-up Effekt erweist sich als grundsätzliches Problem bei der CMOS-Technologie. Durch folgende technologische Maßnahmen läßt sich jedoch die latch-up Gefahr verringern:

- Spezielles Design z.B. große Abstände der Source-Drain-Gebiete zu den Wannenträndern
- Niedrige Widerstände R_S durch niederohmiges Substrat und p+-Guardring neben der n+-Wanne.
- Niederohmiger n+-Guardring für den Versorgungsspannungsanschluß

Vergleich NMOS- mit der CMOS-Technologie

Eigenschaft	NMOS	CMOS
Packungsdichte	groß	kleiner
Komplexität	Einfache Technologie	Kompliziertere Technologie
Minimale Zahl der Diffusionen/Implantationen	3	4
Leistungsverbrauch	$P_{\text{stat}} > 0$	$P_{\text{stat}} = 0,$ $P_{\text{dyn}} = (N/2)C_L U_{\text{DD}}^2 f_{\text{cyc}}$
Schaltzeiten	$t_r \gg t_f$	$t_r = t_f$, langsamer
Rauschabstände	$S_H \gg S_L$	$S_H = S_L$
Sonstiges		Latch-up gefährdet

3.7 Gatter und Speicher in MOS-Technik



L = Low-Pegel Bereich, H = High-Pegelbereich, s = sperrt, l = leitet, n. d. = nicht definiert

RTL und NMOS-NAND

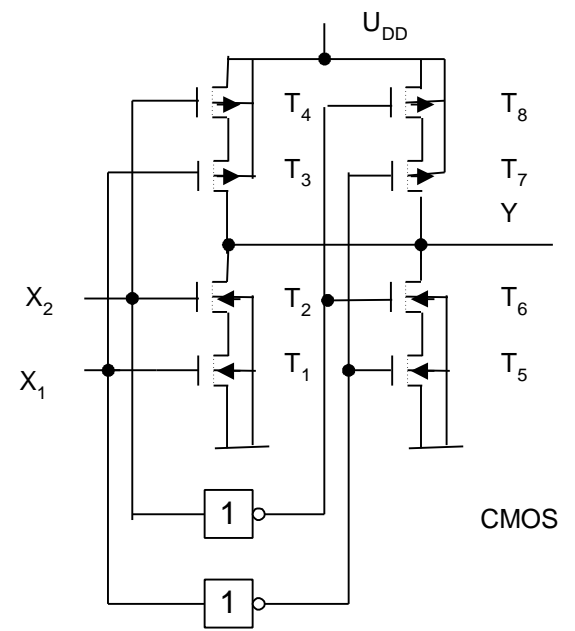
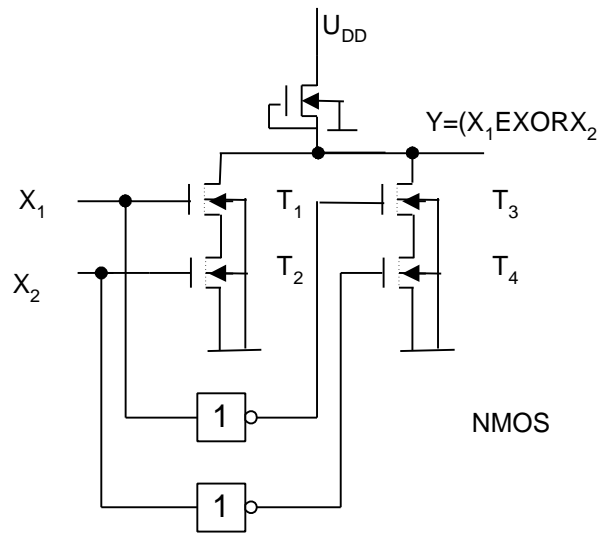
X ₁	X ₂	T ₁	T ₂	Y
L	L	s	s	H
L	H	s	n.d.	H
H	L	l	s	H
H	H	l	l	L

CMOS-NAND

X ₁	X ₂	T ₁	T ₂	T ₃	T ₄	Y
L	L	s	s	l	l	H
L	H	l	s	s	l	H
H	L	n.d.	s	l	s	H
H	H	l	l	s	s	L

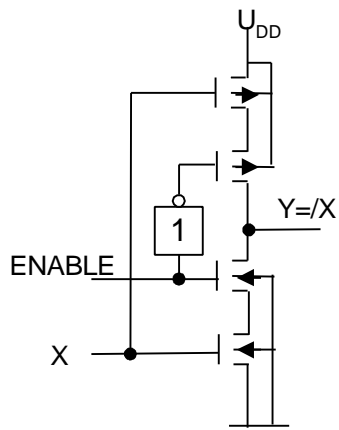


EXOR

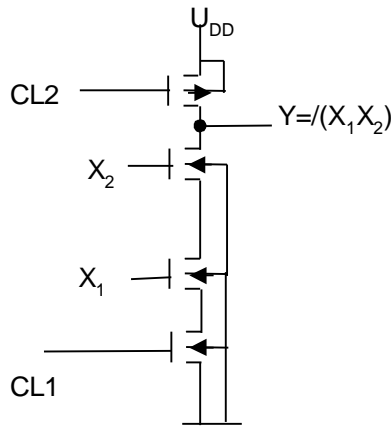


CMOS-EXOR

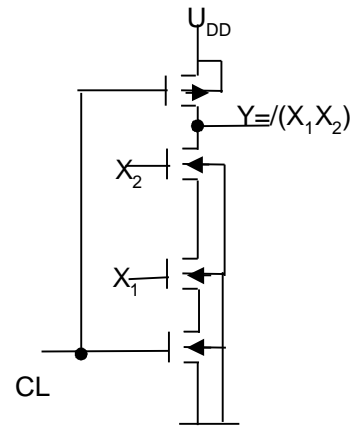
X ₁	X ₂	T ₁	T ₂	T ₃	T ₄	T ₅	T ₆	T ₇	T ₈	Y
L	L	s	s	l	l	l	l	n.d.	s	L
L	H	s	l	l	s	l	s	l	s	H
H	L	l	n.d.	l	s	l	s	s	l	H
H	H	l	l	n.d.	s	l	s	s	l	L



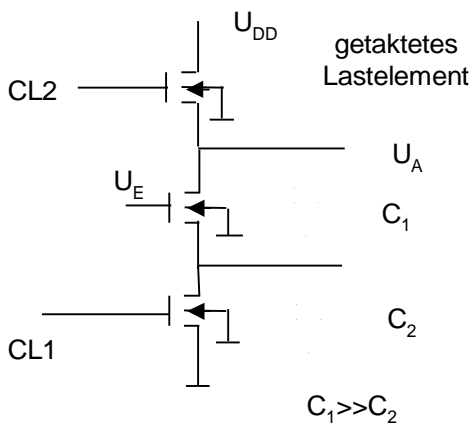
CMOS-Tristate-Inverter



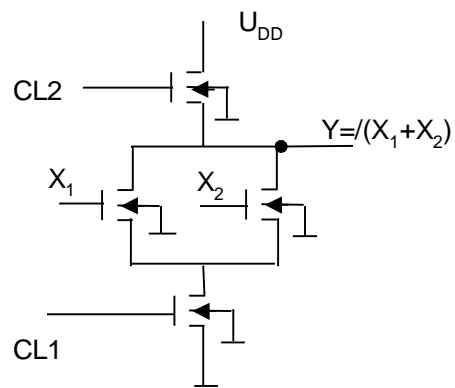
CMOS-NOR dynamische Logik



CMOS-NOR dynamische Logik ratio-Technik, Pseudo-CMOS



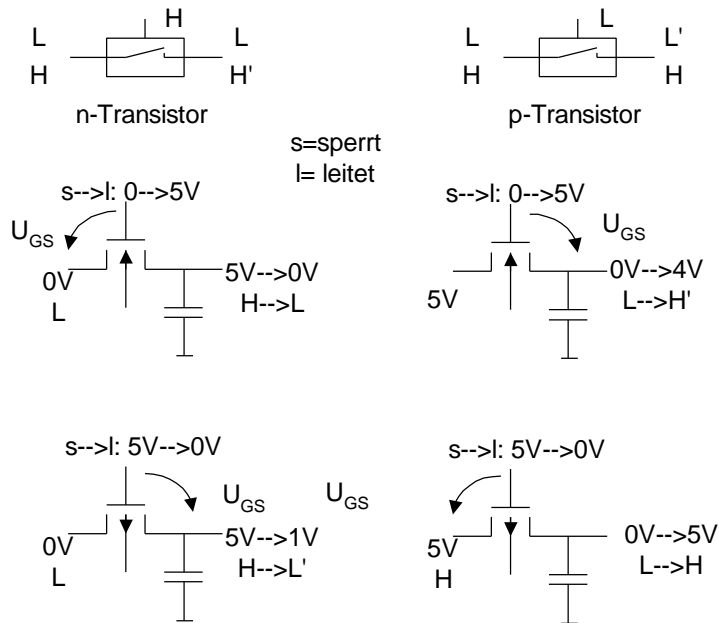
NMOS-Inverter in dynamischer ratio-logic



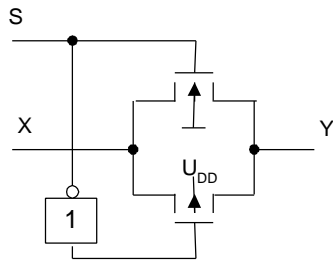
NOR in dynamischer ratio-logic

Pass-Transistor

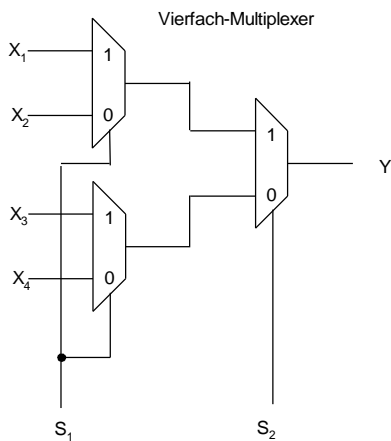
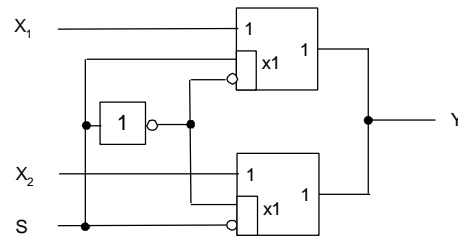
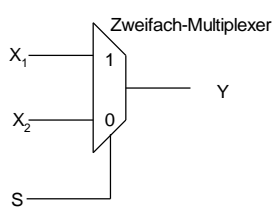
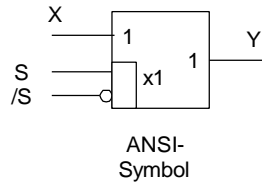
Um logische Signale auf einer Signalleitung je nach Wunsch zu unterbrechen oder durch-zu-lassen wird ein Schalter benötigt. Verwendet man als Schalter einen n-Kanal Transistor, so zeigt es sich, daß bei geschlossenem Schalter (leitendem Transistor) das Low-Signal unverändert weitergegeben wird, während das High-Signal am Ausgang des Transistors um die Schwellenspannung verringert weitergeleitet wird. Es tritt beim High-Signal also eine Signaldegradation ein. Bei Verwendung eines p-Kanal-Transistor als Schalter sind die Verhältnisse analog: Das Low-Signal wird um die Schwellenspannung degradiert (erhöht) weitergegeben; während das High-Signal unverändert bleibt.



Transmission Gate



S	X	Y
L	L	*
L	H	*
H	L	L
H	H	H

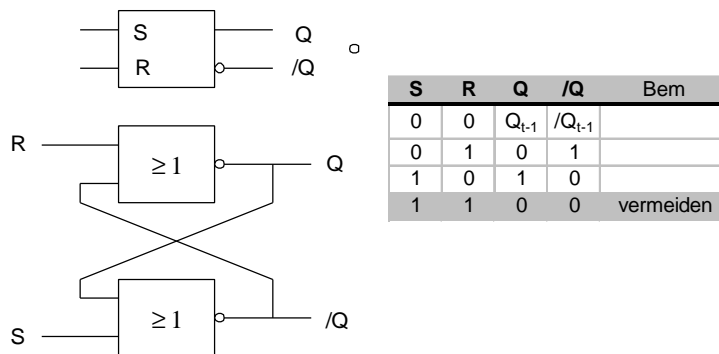


S ₁	S ₂	Y
1	1	X ₁
0	1	X ₂
1	0	X ₃
0	0	X ₄

Speicherschaltungen:

Die einfachste Speicherschaltung ist ein RS-Flip-Flop.

Geht das Signal am Set-Eingang auf 1 und liegt am Reset-Eingang eine 0 an, so wird das Flip-Flop gesetzt, d.h. der Ausgang Q geht auf 1 und der invertierte Ausgang /Q auf 0. Wenn das Signal am Set-Eingang wieder auf Null geht bleibt der Ausgangszustand erhalten

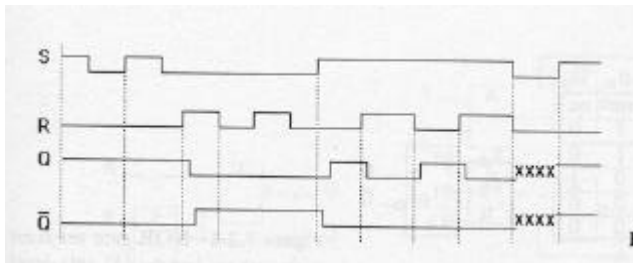


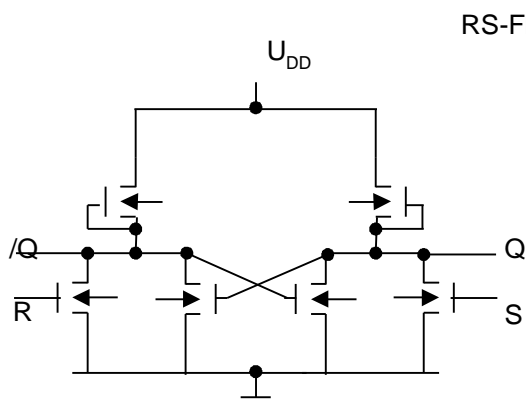
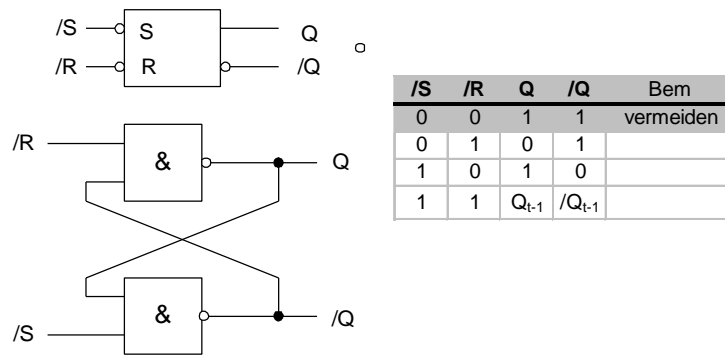
Liegt am beiden Eingängen eine Null an und geht das Rest-Signal auf 1, so wird das Ausgang Q auf Null gesetzt und der Ausgang /Q geht auf 1, das Flip-Flop wird zurückgesetzt. Dieser Zustand bleibt auch erhalten, wenn der Reset-Eingang wieder auf Null geht.

Führt aus einem der betrachteten Eingangszustände ein Wechsel des Eingangssignals (oder beider) zu zwei Einsen am Eingang, so gehen beide Ausgänge auf 0. Werden nun

beide Eingänge gleichzeitig auf Null gesetzt, so ist der Ausgangszustand von den zufälligen Gatterlaufzeiten abhängig, also undefiniert.

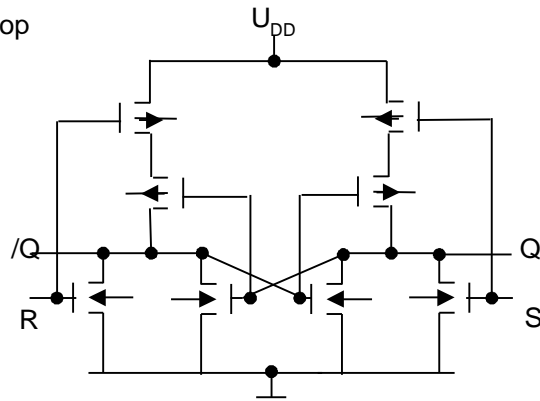
Deshalb muß darauf geachtet werden, daß entweder nicht beide Eingänge gleichzeitig 1 werden können, oder der Übergang von 1 nach 0 nicht gleichzeitig an beiden Eingängen erfolgen kann





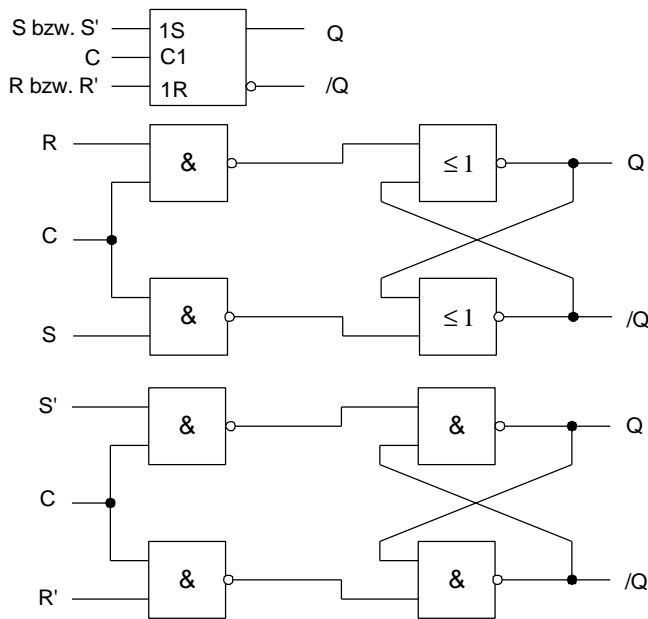
NMOS

RS-Flip-Flop



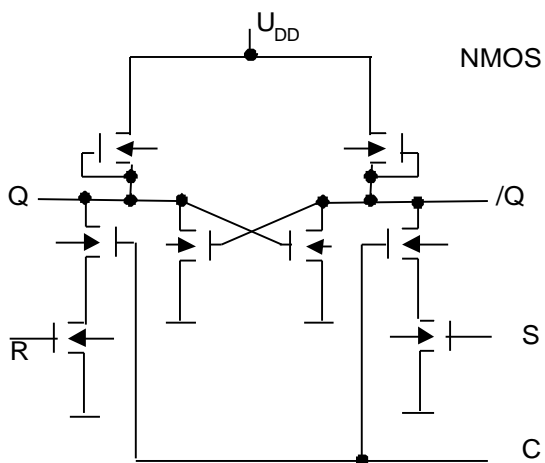
CMOS

Taktzustandgesteuertes RS-Flip-Flop



C	S S'	R R'	Q	/Q	Bem
0	X	X	Q_{t-1}	$/Q_{t-1}$	
1	0	0	Q_{t-1}	$/Q_{t-1}$	
1	0	1	0	1	
1	1	0	1	0	
1	1	1	1	1	vermeiden

Taktgesteuertes RS-Flip-Flop



D-Flip-Flop: Für die Dauer des H-Pegels des Taktes erscheint die Eingangsinformation am Ausgang. Am Ende des Taktes bleibt die zuletzt gespeicherte Information erhalten (transparentes Flipflop)

Verschiedene Realisierungsmöglichkeiten:

mit NAND Gates in NMOS : 12 Transistoren

mit NAND Gates und Inverter in NMOS : 11 Transistoren

mit Transmission Gates und Inverter in CMOS : 10 Transistoren

