3 FETs in Logik-Schaltungen

Inhalte

3.1 Der MOSFET als Schalter

3.1.1 Generelle Eigenschaften

Gon = 1/Ron = $G_{DS}(U_{DS}=0) = \Delta I_{DS}/\Delta U_{DS}$ @ $U_{DS}=0V$

Tabelle 3.1.1: Vergleich MOSFET – BJT als Mikroschalter

Vorteile als Schalter	Nachteile als Schalter
+ Mikrotechnik ~1/10 der Größe eines Bip. + arbeiten symmetrisch bidirektional + Schalter leckt nicht (kein $I_{G,DC}$), → + leistungslose Steuerung + keine Schaltverzögerungen wg. Übersättig. + gute Stromverteilung unter dem Gate → + keine Neigung zu Hot-Spots → zuverlässig	- Leistungselektronik: phys. groß für große I_D

3.1.2 NMOS-Schalter in programmierbarer Logik (veraltet)

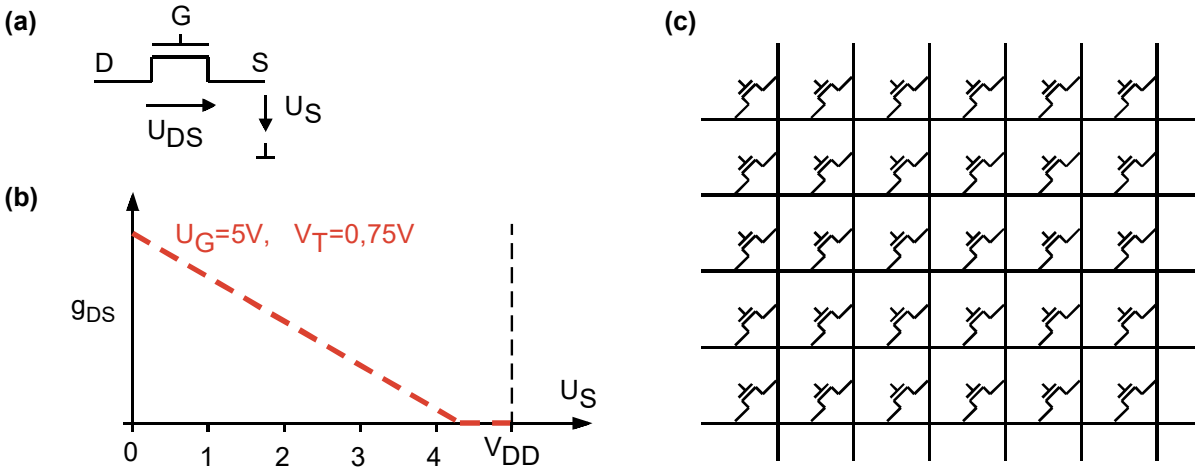


Bild 3.1.2: (a) MOSFET als Schalter, **(b)** $G_{DS}(U_{out})$, **(c)** Anwendung: programmierbare Logik.

Bild 3.1.2(a) zeigt einen NMOSFET als Schalter mit $U_{DS}\geq 0$ (sonst müssten S und D vertauscht werden). Bildteil (b) zeigt die Leitfähigkeit des Schalters $G_{DS}(U_{out})=\Delta I_{DS}/\Delta U_{DS}$ bei $U_{DS}\sim 0V$. Eine typische Anwendung solcher Schalter, die Leitungen in einer programmierbaren Logik verknüpfen können, zeigt Bildteil (c). Gemäß Gl. (F1.a) ist

$$I_D = 2\beta\cdot((U_{GS}-V_T)-U_{DS}/2)U_{DS}\cdot(1+\lambda U_{DS})$$

Für sehr kleine U_{DS} kann man Summenterme wie $\frac{1}{2}U_{DS}$ oder $\lambda\cdot U_{DS}$ (typ.: $\lambda \ll 1/V$) vernachlässigen, eine Multiplikation mit U_{DS} dagegen nicht. Daher folgt

$$\xrightarrow{U_{DS}\to 0} \quad I_D \cong 2\beta\cdot(U_{GS}-V_T)U_{DS}\,.$$

Somit ist für der On-Leitwert eine mit U_{GS}-V_T steuerbare Größe:

$$G_{on}=G_{DS}=\frac{dI_D}{dU_{DS}}=2\beta\cdot(U_{GS}-V_T)=2\beta\cdot((U_G-V_T)-U_S) \qquad \text{für } U_G\geq U_S+V_T.$$

Der Schalter soll leiten bei $U_G=V_{DD}$. Bei größer werdendem U_S nimmt der On-Leitwert ab, denn er ist proportional $(U_G-V_T)-U_S$. Daher kann er U_S nicht auf Spannungen höher als $U_S\geq U_G-V_T$ treiben, weil $U_{GS}=\geq V_T$ mit $U_{GS}=U_G-U_S$ sein muss.

3.1.3 Transmission-Gate: komplementäre (=ergänzender) MOSFETs

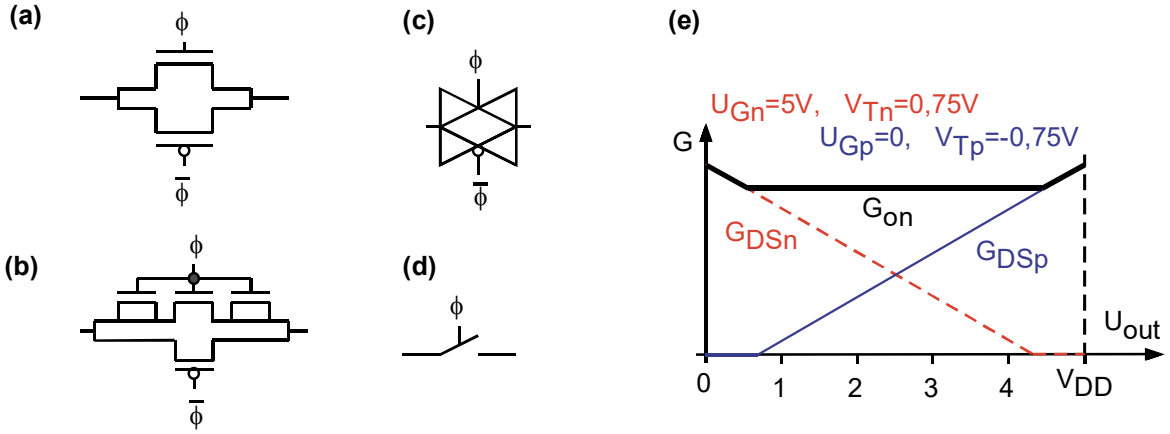


Bild 3.1.3: (a) Transmission-Gate (TG), **(b)** kapazitiv ausgewogenes TG, **(c)** Symbol für TG, **(d)** Schaltersymbol für TG, **(e)** Leitwert G_{on} des TG für U_{DS}~0V.

Bildteil (a) oben zeigt zwei komplementäre FETs. Deren Leitwerte mit ßn=ßp so eingestellt sind, dass ihre Summe Gon gemäß Bildteil (e) über einen weiten Spannungsbereich konstant ist. Bildteil (c) zeigt das Symbol für ein Transmission-Gate (TG) und Bildteil (d) einen Schalter, der als TG realisiert sein kann.

Bildteil(b) zeigt ein kapazitiv ausgewogenes TG. Da ßn=ßp verlangt, dass die Gate-Fläche und somit die Kapazität des PMOSFETs ca. 2,7 mal größer ist, als die des NMOSFETs, wird in Bildteil (a) der PMOSFET während der Taktflanken von Φ entsprechend mehr Ladung einstreuen (bekannt als „clock feed through“). Damit die durch den PMOSFET eingestreute Ladung vom NMOSFET kompensiert wird, muss der NMOSFET eine gleich große Gate-Fläche haben. Damit die Symmetrie der Leitwerte gemäß Bildteil (e) erhalten bleibt, verwendet man zwei kurzgeschlossene NMOSFETs so, dass die gesamte Kapazität der drei N-FETs genauso groß ist wie die des P-FETs.

3.2 Bipolar / CMOS (BiCMOS)Technologiebeispiel

3.2.1 Vergleich der Vor- und Nachteile verschiedener Logiktechnologien

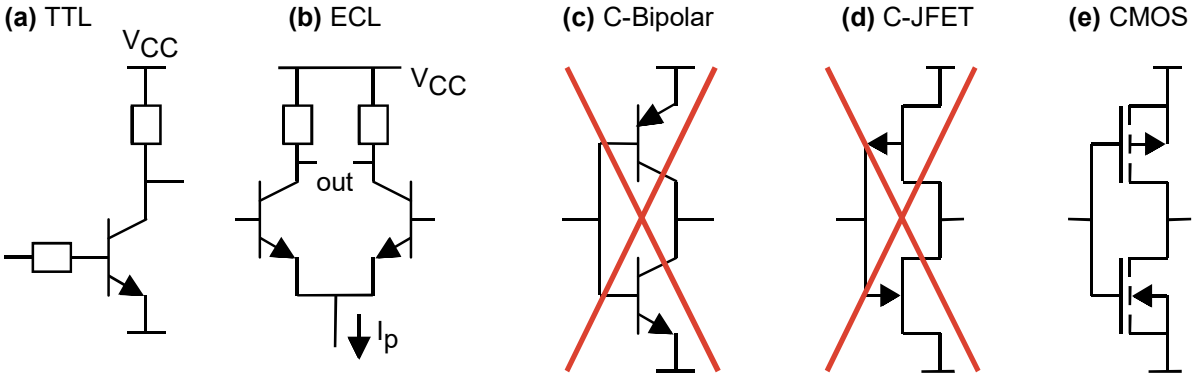
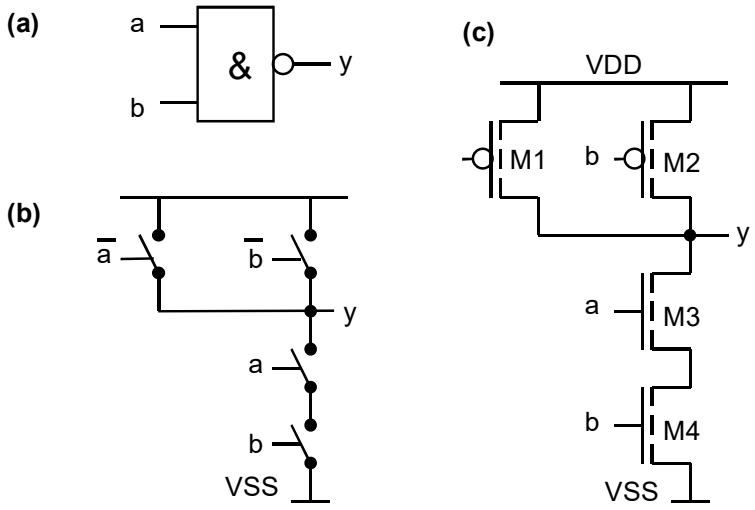


Bild 3.2.1.1: **(a)** TTL, **(b)** ECL, **(c,d)** Komplementäre Bipolar-/JFET-Technologie nicht möglich mit Versorgungsspannungen >1,2V, **(d)** CMOS

ECL: hält Transistoren sättigungsfrei unter Strom => sehr schnell + hoher Stromverbrauch.
MOSFETs nur in Si möglich, alle andere Technologien (GaAs, InPh,...) können nur JFETs.

Bild 3.2.1.2:
(a) NAND-Gatter Symbol
(b) NAND-Gatter realisiert mit idealen Schaltern.
(c) Die Schalter in (b) wurden durch MOS-FETs ersetzt, wobei N-Kanal MOSFETs prinzipiell gegen VSS und P-Kanal MOSFETs prinzipiell gegen V_{DD} schalten müssen.



3.2.2 Beispiel einer BiCMOS-Technologie und -Schaltung

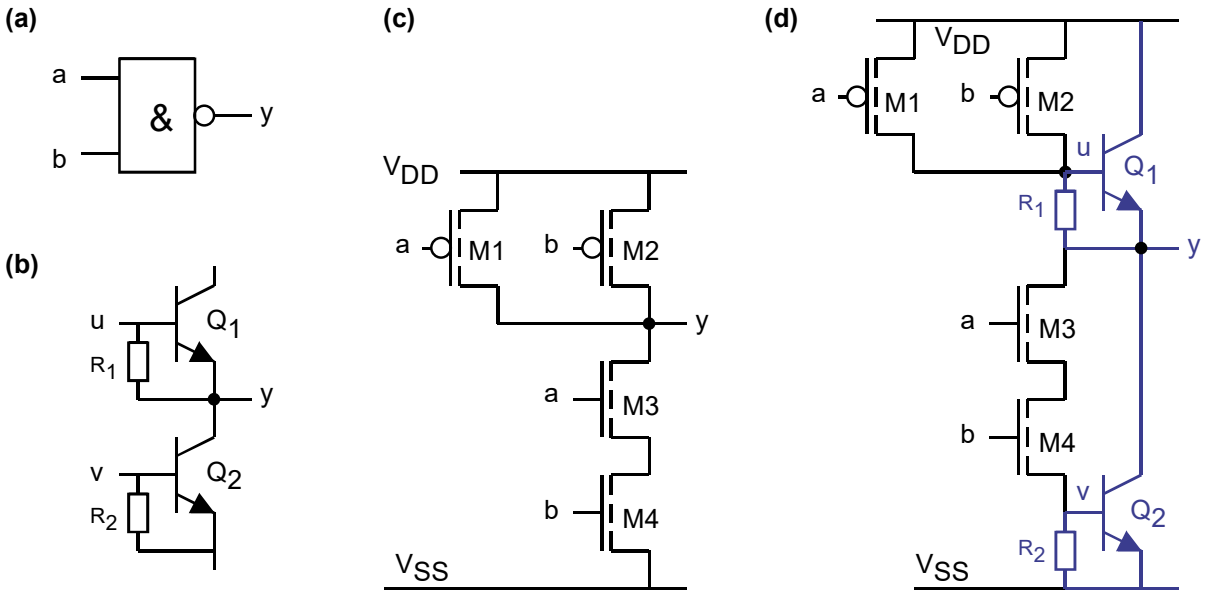


Bild 3.2.2: BiCMOS Gatter (aus [1]): **(a)** Realisiertes NAND-Gatter, **(b)** „full swing totem-pole unit“, **(c)** CMOS-Realisierung, **(d)** BiCMOS-Realisierung, mit npn-Transistoren.

Bild 3.2.2 zeigt eine Möglichkeit ein CMOS-Gatter in BiCMOS-Technologie zu realisieren. Die guten Eigenschaften der CMOS und der bipolaren Technologie werden kombiniert:

- CMOS-Vorteil: Preiswert (kleine Flächenverbrauch, relativ einfache Technologie).
- CMOS-Vorteil: Kein Stromverbrauch im stationären Zustand bei y='0' oder y='1'.
- Bipolar-Vorteil: Starke Treibereigenschaften -> schnell

Weitere positive Eigenschaften des BiCMOS-Gatters

- Ausgangsspannung mit „full swing“, also von V_{SS} ... V_{DD} möglich
- Beliebig viele Eingänge möglich
- Eingangsspannung mit „full swing“, also von V_{SS} ... V_{DD} möglich
- Es wird kein pnp-Transistor benötigt (würde zusätzliche, sehr tiefe n-Wanne benötigen)

Nachteile des BiCMOS-Gatters:

- Es werden pipolare Transistoren benötigt
- Es werden Widerstände benötigt

Wegen der nicht zu vernachlässigenden Nachteile wird man nur solche Gatter nur dort verwenden, wo große Lasten mit steiler Flanke getrieben werden müssen (z.B. Takt-Puffer).

Aufgaben der Widerstände R_1, R_2:

- Schnelligkeit: Entladung der Basen von Q_1 und Q_2, wenn diese die Q's übersättigt sind,
- „Full swing“ am Ausgang: R_1, R_2 ermöglichen maximale Aufladung der kap. Last gegen V_{DD} über M_1, M_2, R_1 u. Restentladung einer kapazitiven Last gegen V_{SS} über M_3, M_4, R_2.

3.3 CMOS-Technologie: Eigenschaften und technische Bedeutung

Die CMOS-Technologie ist heutzutage die alles dominierende Technologie in der Digitaltechnik. Diese Tatsache beruht auf der Kombination zweier Eigenschaften des MOSFETs:

1. Leistungslose Steuerung
2. Gate im gesamten Spannungsbereich 0V...V_{DD} isoliert
3. Verfügbarkeit komplementärer Transistoren

Damit lässt sich erreichen, dass ein logisches CMOS-Gatter Leistung ausschließlich für Schaltvorgänge benötigt. Die Schaltungsgröße ist daher wärmetechnisch begrenzt auf die kühlbare Anzahl an Schaltvorgängen pro Sekunde und Chip. (Die in einer Kapazität gespeicherte Energie ist $\mathbf{E_C = \frac{1}{2}CU^2}$, der Leistungsbedarf bei f Umladungen pro Sekunde $P = f \cdot E_C = \frac{1}{2} f C U^2$. Die Leistungsaufnahme einer CMOS-Technologie ist also proportional der Taktfrequenz f, der Gate-Kapazitäten C_G und dem Quadrat der Versorgungsspannung V_{DD}.)

Alle anderen Technologien benötigen einen Ruhestrom proportional zur Anzahl der Gatter. Ihre Schaltungsgröße ist daher wärmetechnisch begrenzt auf die kühlbare Anzahl an Gattern pro Chip.

Betrachtet man die drei Inverter in den drei verschiedene Technologien in Bild 3.3.1, dann kann erst einmal nur der CMOS-Inverter einen Ausgangs-Low-Pegel von 0V erreichen.

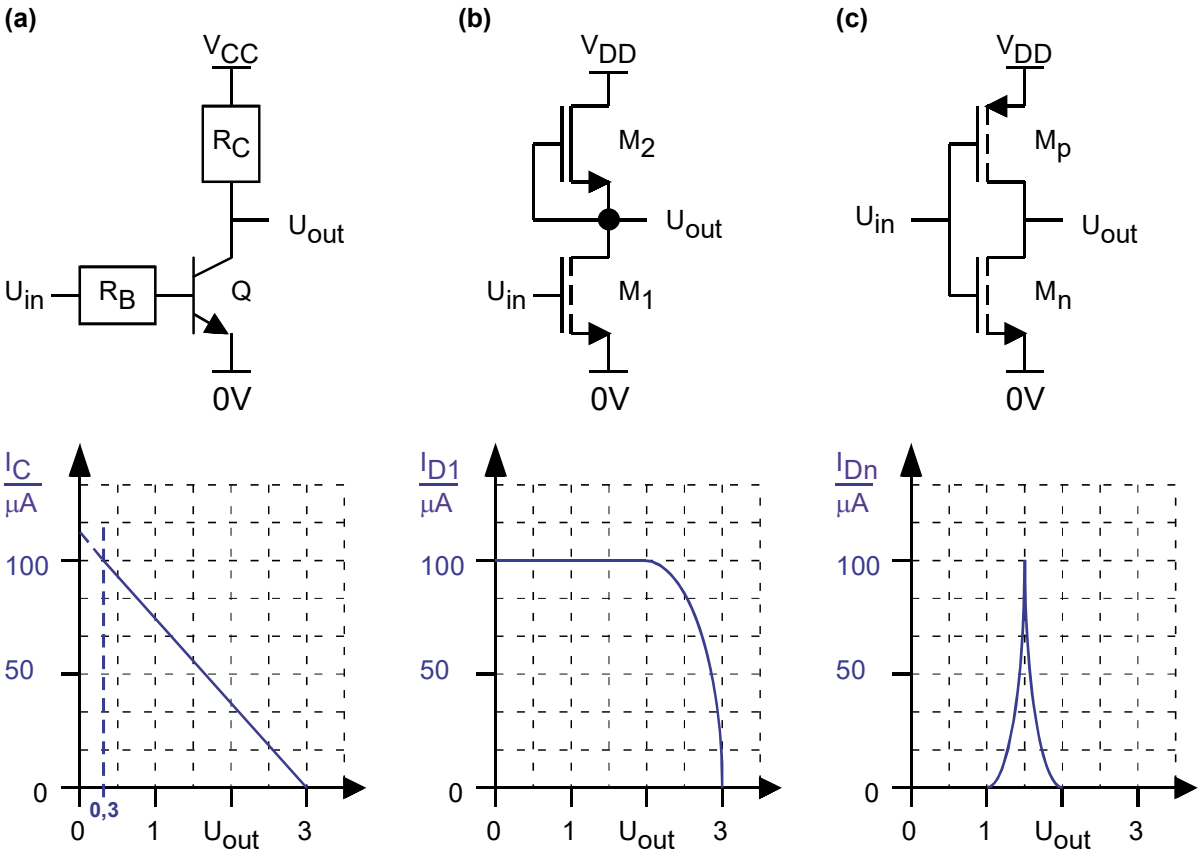


Bild 3.3.1: Inverter in drei Technologien **(a)** bipolar, (b) NMOS, (c) CMOS.

Beispiel zu Bildteil (a): Der bipolare Inverter in Bildteil (a) habe einen Ausgangs-Low-Pegel von 0,3V bei V_{CC}=3V und R_C=27KΩ?

Dann ist der Kollektorstrom I_C bei Ausgangs-High-Pegel (U_{in}=0,3V) µA und bei Ausgangs-Low-Pegel I_C = (3V-0,3V)/27KΩ = 100µA.

Beispiel zu Bildteil (b): Der NMOS-Inverter in Bildteil (b) besteht aus zwei n-Kanal-MOSFETs, nämlich dem selbstsperrenden M_1 (z.B. mit der Schwellenspannung V_{T1}=1V) und mit dem selbstleitenden M_2 (z.B. mit V_{T2}=-1V und I_{DSS}=100µA).

Nimmt man für beide λ=0 an, dann ist der Drainstrom I_D bei Ausgangs-High-Pegel 0µA bei Ausgangs-Low-Pegel ist (willkürlich angenommen) $I_{D2} = I_{DSS}$ = 100µA Wegen U_{GS2}=0V.

Der selbstleitende n-Kanal-MOSFET M_2 sättigt bei $U_{DS2,sat} = U_{GS2} - V_{T2}$ = 0V–(-1V) = 1V. Dies entspricht einer Ausgangsspannung von $U_{out,sat2} = V_{DD} - U_{DS2,sat}$ = 3V - 1V = 2V. Für kleinere U_{out} arbeitet M_2 als Stromquelle.

Beispiel zu Bildteil(c): Ein CMOS-Gatter führt nur während des Schaltvorganges einen statischen Strom, wie es in (c) für einen maximalen Ausgangsstrom von 100µA skizziert ist.

Folgerungen: Bipolar und NMOS ziehen nur bei Ausgangs-Low-Pegel einen statischen Strom, CMOS in keinem der Pegel.

Andere logische Gatter, z.B. NOR, in diesen Technologien erhält man, indem man dem Biopolartransistor in Bildteil (a) oder dem FET M_1 in Bildteil (b) gleichartige Transistoren parallel schaltet. Bei CMOS werden weitere komplementäre Transistorpaare eingefügt. Der Stromverbrauch ändert sich durch solche Modifikationen nicht gegenüber dem oben berechneten Stromverbrauch eines Inverters.

In einer großen Digitalschaltung kann man davon ausgehen, dass ca. 50% aller Gatterausgänge eine logische '0' und der Rest eine logische '1' treiben. Daher ist die Größe einer digitalen CMOS-Schaltung wärmetechnisch begrenzt durch Schaltvorgänge pro Sekunde und Chip. In allen anderen Technologien ist sie begrenzt durch Gatter pro Chip.

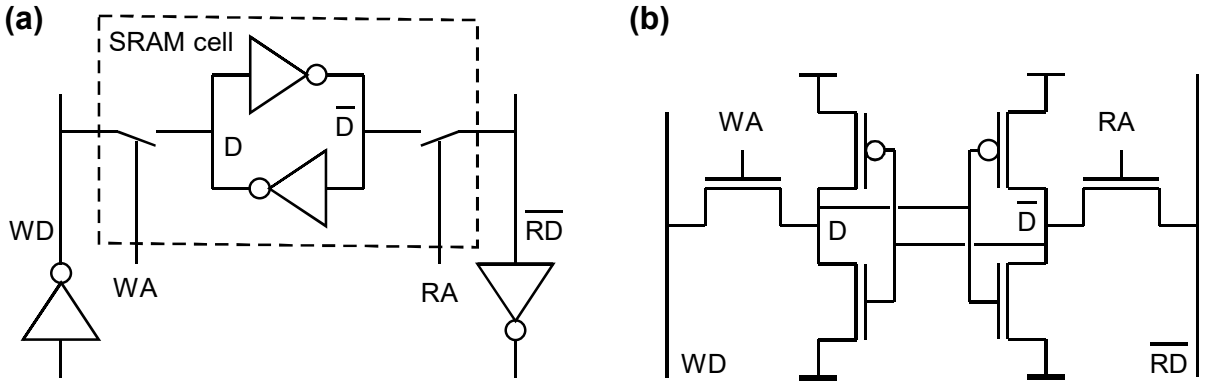


Bild 3.3.2: "Static Random Access Memory" (SRAM) Zelle **(a)** Gatterebene mit den Leitungen "Write Data" (WD), "Write Address" (WA), "Read Address" (RA) und "Read Data" (RD), **(b)** Standardrealisierung in CMOS als 6T(ransistor)-Zelle.

Bild 3.3.2 zeigt einen Vorschlag für eine SRAM-Zelle (gestrichelt umrandet), die aus zwei gleichen Invertern mit einer der oben diskutierten Technologien aufgebaut ist. Tab. 1-1 zeigt den Stromverbrauch als Funktion der Technologie und des logischen Zustands.

Tabelle 3.3.1: Stromverbrauch einer statischen SRAM-Zelle.

Technologie:	bipolar	NMOS	CMOS
D='1'	10 nA	10 nA	0 µA
D='0'	10 nA	10 nA	0 µA

Ein 32 GB SRAM enthält $32 \cdot 10^9$ x 8 Bits $\approx 256 \cdot 10^9$ Bits. Tabelle 2.3.3-2 zeigt den Stromverbrauch eines solchen 32 GB SRAMs für die verschiedenen Technologien. Benötigt jedes Bit 100nA Strom, dann ergibt sich ein Gesamtstromverbrauch von über 2560 A für die bipolare und die NMOS-Technologie. Um diesen Strom in das Chip hereinzubringen müsste dessen Durchgangswiderstand bei 3V Versorgungsspannung <0,5mΩ sein, was erhebliche Ansprüche an die Leiterbahnen und Stecker stellen würde.

Tabelle 3.3.2: Stromverbrauch eines 8 GB SRAMs.

Technologie:	bipolar	NMOS	CMOS
Stromverbrauch:	2560 A	2560 A	0 A

Es sei angemerkt, dass dynamische RAM (DRAM) Zellen ihren Zustand nicht dauerhaft halten können, sondern einen internen "Refresh" - Mechanismus benötigen, der die Information von Zeit zu Zeit ausliest und wieder zurückschreibt. Die veraltete Bauweise gemäß Bild 3.3.3(a) nutzt die Gate-Kapazität eines MOSFETs, um Ladung zu speichern und kann ihre Information zerstörungsfrei auslesen. Sie verbraucht jedoch mehr Chip-Fläche, als eine Kapazität, die man mit Hilfe eines geätzten Grabens in die Tiefe bauen kann. Eine moderne DRAM-Zelle besteht im wesentlichen aus einer Kapazität, die so klein ist im Vergleich zur Kapazität der Schreib-/Leseleitungen, dass man ihre Information gerade noch auslesen kann.

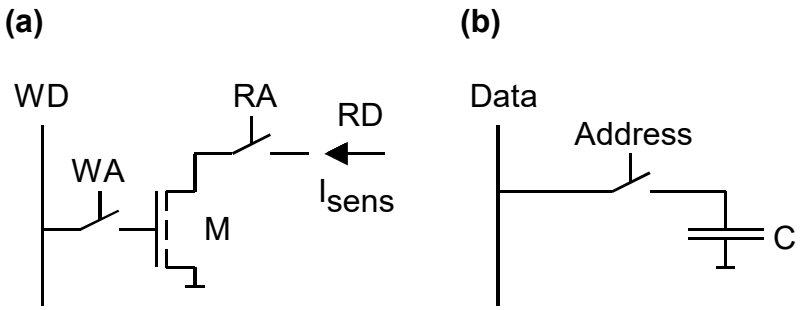


Bild 3.3.3:
Dynamische RAM (DRAM) Zellen:
(a) veraltete Version,
(b) neue Version.

3.4 *Miller*-Effekt in der Digitaltechnik

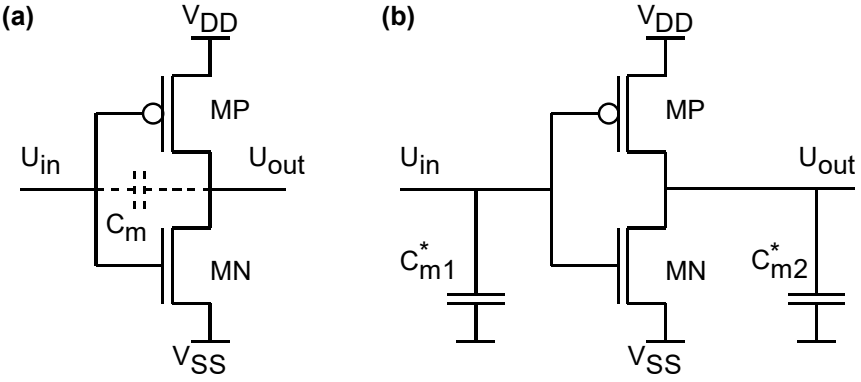


Bild 3.4:
Digitaler CMOS-Inverter
(a) Miller-Kapazität C_m.
(b) mit äquivalenten Kapazitäten C^*_{m1} und C^*_{m2}.

Die Miller-Kapazität in Bild 3.4 ist gestrichelt eingezeichnet, weil sie nicht absichtlich eingebaut wird, sondern aus den unvermeidlichen Kapazitäten Gate-Drain-Kapazitäten besteht: $C_m = C_{GD}(M_N) + C_{GD}(M_P)$.

Die äquivalenten Kapazitäten C^*_{m1} und C^*_{m2} ergeben sich damit zu $C^*_{m1} = C_m\ (1-A_V)$ und $C^*_{m2} = C_m\ (1-1/A_V)$. Tabelle 2 liefert $A_V = -1$ für die stationären Endwerte.

Tabelle 3.4: Zusammenhang zwischen U_{in} und U_{out}

U_{in} =	V_{SS}	V_{DD}
U_{out} =	V_{DD}	V_{SS}

Für die äquivalenten Kapazitäten ergibt sich damit $C^*_{m1} = C^*_{m2} = 2C_m$. Der doppelte Spannungshub an C_m wird umgerechnet in einen einfache Spannungshub an C^*_{mx}.

Anmerkung zu Praxis.
Parasitäre Kapazitäten von Leitungen, Transistoren und anderen Bauelementen werden von Designprogrammen (z.B. für FPGAs) automatisch aus dem Layout extrahiert und bei der sogenannten „Backannotation“ (dt. Rücknotation [der extrahierten der Kapazitäten]) berücksichtigt, um eine unvorteilhafte Designs zu vermeiden und eine möglichst exakte Simulation zu ermöglichen.

3.5 Referenzen

[1]