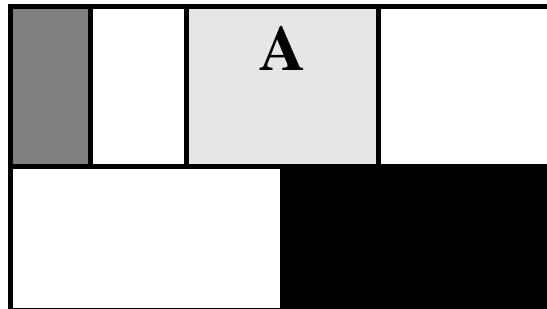


ADDENDUM



BIPOLAR DIGITAL DESIGN

Qualitative understanding of the bipolar device

π

Bipolar device models

π

The ECL Gate

π

Advanced bipolar digital design

π

- A.1 Introduction
- A.2 The Bipolar Transistor
 - A.2.1 A First Glance at the Device
 - A.2.2 Static Behavior
 - A.2.3 Dynamic Behavior
 - A.2.4 The Actual Bipolar Transistor—Secondary Effects
 - A.2.5 SPICE Models for the Bipolar Transistor
- A.3 The Bipolar ECL Inverter
 - A.3.1 Issues in Bipolar Digital Design: A Case Study
 - A.3.2 The Emitter-Coupled Logic (ECL) Gate at a Glance
 - A.3.3 Robustness and Noise Immunity: The Steady-State Characteristics
 - A.3.4 ECL Switching Speed: The Transient Behavior
 - A.3.5 Power Consumption

A.3.6 Looking Ahead: Scaling the Technology

A.4 Bipolar Gate Design

A.4.1 Logic Design in ECL

A.4.2 Differential ECL

A.4.3 Current Mode Logic

A.4.4 ECL with Active Pull-Downs

A.4.5 Alternative Bipolar Logic Styles

A.5 Exercises and Design Problems

A.1 Introduction

MOS transistors took over the digital integrated circuit market in the 1970s, mainly as a consequence of their high integration density. Before that time, most digital gates were implemented in the bipolar technology. The dominance of the bipolar approach to digital design was exemplified in the wildly and widely successful TTL (Transistor-Transistor Logic) logic series, which persisted until the late 1980s. Bipolar digital designs occupy only a small portion of the digital market today. They still are the technology of choice when very high performance is required, yet even there CMOS is becoming highly competitive. This trend will continue in the future, as the reduced supply voltages of the deep-submicron technologies make bipolar design exceedingly hard.

Because of this reduced importance, we decided to remove bipolar design from the 2nd Edition of “Digital Integrated Circuits — A Design Perspective”, and to make the material of the 1st edition freely available at the web-site as a set of addenda. We hope that this helps to address the concerns of those designers for whom bipolar design is still a necessity.

In this addendum, we first present a brief overview of the bipolar device and its models. This is followed by an extensive description and analysis of the Emitter-Coupled Logic (ECL) gate, the dominant bipolar digital gate at present. After a discussion on how to build complex logic gates in ECL, the chapter is concluded with an overview of the BICMOS approach to digital design that combines MOS and bipolar devices into a single gate.

A.2 The Bipolar Transistor

A.2.1 A First Glance at the Device

Figure A.1a shows a cross section of a typical *npn* bipolar (junction) transistor structure. The heart of the transistor is the region between the dashed lines and consists of two *np* junctions, connected back to back. In the following analysis, we will consider the idealized transistor structure of Figure A.1b. The transistor is a three-terminal device, where the two *n*-regions, called the *emitter* and the *collector*, sandwich the *p*-type *base* region. In contrast to the source and drain regions of the MOSFET, the emitter and collector regions are not interchangeable, as the emitter is much more heavily doped than the collector.

Depending upon the voltages applied over the device terminals, the emitter-base and collector-base junctions are in the *forward*- or *reverse-biased* condition. Enumeration of all possible combinations results in Table A.1, which summarizes the operation modes of the bipolar device. In digital circuits, the transistor is operated by preference in the cut-off or forward-active mode. Operation in the saturation or reverse regions is, in general, avoided as the circuit performance in those regions tends to deteriorate.

In a superficial way, the operation of the device can be summarized as follows:

- As both junctions are reverse biased in the *cut-off* mode, no current flows into or out of the device. It can be considered *off*.

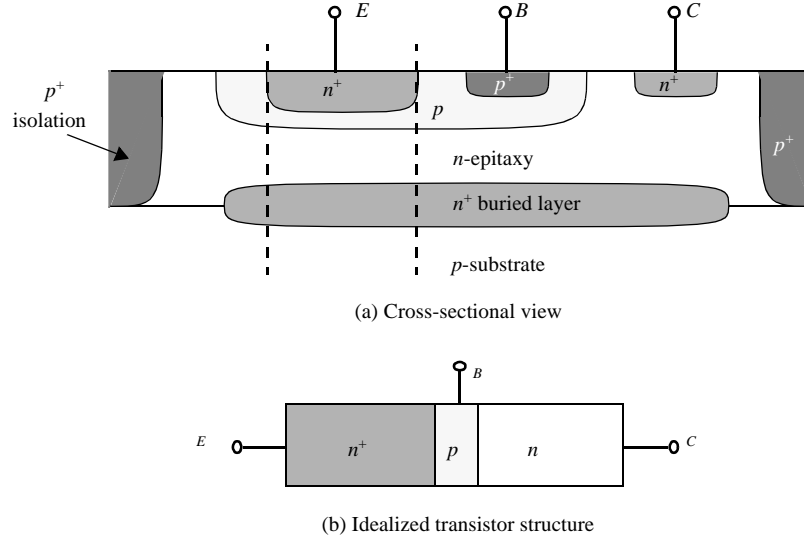


Figure A.1 *npn* bipolar transistor.

Table A.1 Modes of operation of the bipolar transistor.

Mode	Emitter Junction	Collector Junction
Cut-off	Reverse	Reverse
Forward-active	Forward	Reverse
Reverse-active	Reverse	Forward
Saturation	Forward	Forward

- In the *active* mode, the transistor acts as a current amplifier. The current flowing into the base results in a collector current that is β times larger. Furthermore, there exists an exponential relationship between the emitter-base voltage and the collector current. This relation is similar to the forward-bias condition of the junction diode. In the reverse active condition, this current gain is small and virtually nonexistent ($\beta \approx 1$) in contrast to the forward-active mode, where values of over 100 can be observed.
- Finally, when the device saturates, a substantial drop in current gain occurs. Typical for the saturation mode is the low value of the emitter-collector voltage.

Some important differences with the MOS transistor jump immediately into view. First of all, the *exponential relationship* between input voltage and output current makes it possible to drive large currents with small voltage excursions. This has a beneficial impact on the performance. On the other hand, the control terminal (i.e., the base) of the bipolar transistor carries an input current when the device is in active or forward mode. This

means that the *input resistance* of the device is small compared to the MOS transistor. As will become apparent later, this feature makes the device not as amenable to high-density digital design.

It should come as no surprise that a complementary device, called the *pnp* bipolar junction transistor, can be conceived as well. They have only been used sparingly in digital design—adding a high-quality *pnp* to a bipolar process tends to raise the cost substantially—and are therefore treated superficially in the rest of this book. Schematic symbols for both *npn* and *pnp* devices, as well as sign conventions for voltage and currents are pictured in Figure A.2.

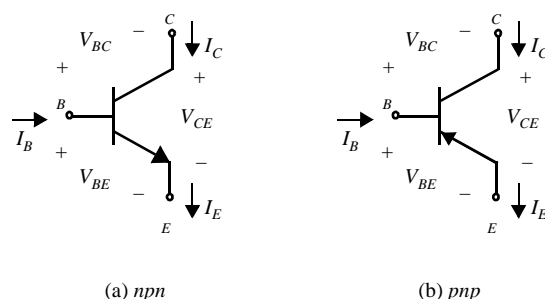


Figure A.2 Bipolar transistor schematic symbols and sign conventions.

A.2.2 Static Behavior

Forward-Active Region

Figure A.3 shows a cross section of the idealized transistor structure of Figure A.1b as well as the minority carrier concentrations in the emitter, base, and collector regions. The concentrations are plotted for the forward-active operation mode. That is, the base-emitter (*be*) junction is forward-biased, while the base-collector (*bc*) junction is in reverse-bias condition. The subscripts *e*, *b*, and *c* are used to denote the various regions. As we know from our diode study, the forward bias causes excess minority carriers on the *be* side, while the reverse bias at the *bc* end causes the minority concentration to approach zero. We assume (without loss of generality) that the short-base diode model is valid for all junctions.

The *law of the junction*, described in Chapter 3 of the textbook, is still valid and can be used to evaluate the minority carrier concentrations at the boundaries of the base region

$$\begin{aligned} n_b(0) &= n_{b0} e^{V_{BE}/\phi_T} \\ n_b(W) &= n_{b0} e^{V_{BC}/\phi_T} \approx 0 \end{aligned} \tag{A.1}$$

Obviously, a concentration gradient exists within the base region causing minority electrons to diffuse from the emitter to the collector end. The width of the base region, which is substantially below 1 μm in contemporary technologies, is sufficiently smaller than diffusion length L_n . Consequently, the minority carriers in the base region display a linear gradient, similar to the case of the short-base diode. The value of this current is readily computed

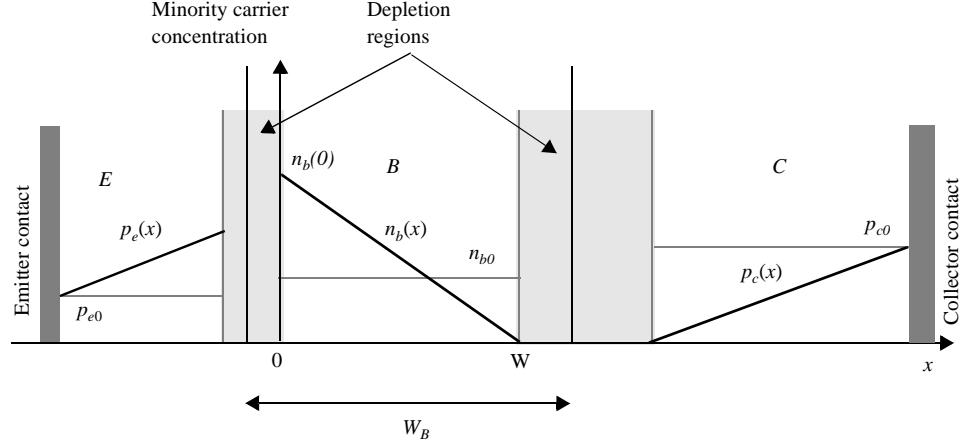


Figure A.3 Minority carrier concentration profile in an *npn* transistor under forward-active conditions.

$$\begin{aligned}
 I_{diff} &= qA_E D_b \frac{dn_b}{dx} = qA_E D_b \left(\frac{n_b(0) - n_b(W)}{W} \right) = \frac{qA_E D_b}{W} n_b(0) \\
 &= \frac{qA_E D_b n_{b0}}{W} e^{V_{BE}/\phi_T}
 \end{aligned} \tag{A.2}$$

where A_E is the area of the of the transistor under the emitter, and D_b is the diffusion coefficient for minority carriers in the base region.

The following picture now emerges. The forward bias of the *be* junction reduces the potential barrier. Electrons, the majority carriers in the emitter, diffuse from the emitter to the base, where they become minority carriers. As the emitter is more heavily doped than the base, we assume for the moment that the *be* junction is a one-sided junction and that the hole current can be ignored. The concentration gradient in the base causes the injected (or emitted) electrons to diffuse towards the collector. From Eq. (A.2), it follows that this diffusion current is an exponential function of the applied base-emitter bias-voltage. Once arrived at the reverse-biased *bc* junction, these electrons are swept towards the collector by the local electrical field (the collector potential is positive with respect to the base). The current across the junction (or the collector current) is, therefore, a drift current. This yields the following expression for the collector current

$$I_C = \frac{qA_E D_b n_{b0}}{W} e^{V_{BE}/\phi_T} = I_S e^{V_{BE}/\phi_T} \tag{A.3}$$

where I_S is the saturation current. Since the thickness of the depletion regions is, generally, negligible with respect to the base region W_b , it is safe to assume that $W \approx W_b$. Be aware also that the derivation of Eq. (A.3) is based on some extra assumptions:

- The minority carrier concentration in the base region is substantially below the majority carrier concentration. This is called the *low-level injection* criterion.

- All voltage drops occur over the depletion regions. This assumes that the neutral regions are perfect conductors. This simplification will be accounted for later by adding series resistances.

The main assumption until now, however, was that all electrons make it safely from the emitter to the collector, or that $I_C = I_E$. In reality, this is clearly not the case. A more accurate picture of the currents flowing in the bipolar *npn* transistor is pictured in Figure A.4.

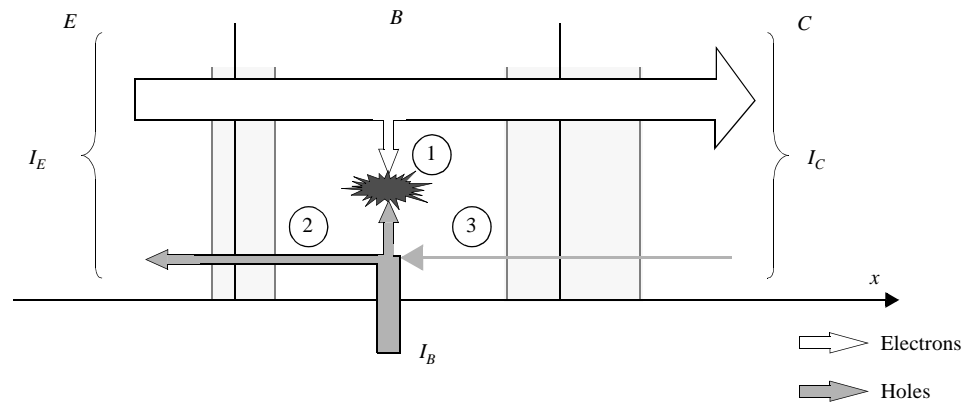


Figure A.4 Components of the base current.

As shown, holes are flowing into the base to supply the following currents:

- *Recombination current* i_r —Occasionally, minority electrons diffusing through the base recombine with majority hole carriers. To maintain charge neutrality in the base region, the vanished holes have to be replaced.
- *be-junction hole current* i_{bh} —The forward-biased base-emitter junction carries an electron as well as a hole current. Although the latter is substantially smaller due to the one-sided nature of the *be* junction, a small base current is still required to supply the carriers.
- *bc-junction hole current* i_{bch} —This current actually flows from the collector to the base and equals the saturation current of a reverse-biased junction. From the discussion on *pn*-junction diodes, we know that this component is negligible.

This analysis demonstrates that the base current I_B is relatively small; the smaller the better, actually. It relates to the collector current I_C by a constant ratio, called the *forward current gain* β_F :

$$\beta_F = \frac{I_C}{I_B} \quad (\text{A.4})$$

For typical digital bipolar processes, β_F varies between 50 and 100. This means that a small hole current into the base sustains a large electron current at the collector; hence the current gain.

The relationship between I_E and I_B is obtained by enforcing the current conservation law

$$I_E = I_B + I_C = I_B(\beta_F + 1) \quad (\text{A.5})$$

Finally, by combining Eq. (A.4) and Eq. (A.5), we can relate I_C and I_E .

$$\frac{I_C}{I_E} = \frac{\beta_F}{\beta_F + 1} = \alpha_F \leq 1 \quad (\text{A.6})$$

where α_F is called the *forward common-base current gain*.

Reverse-Active Region

In this operation region, the situation is reversed—the base-collector junction is forward biased, while the base-emitter junction is in reverse mode. The picture is essentially the same as in the forward active mode, as shown in Figure A.5. The major difference is that

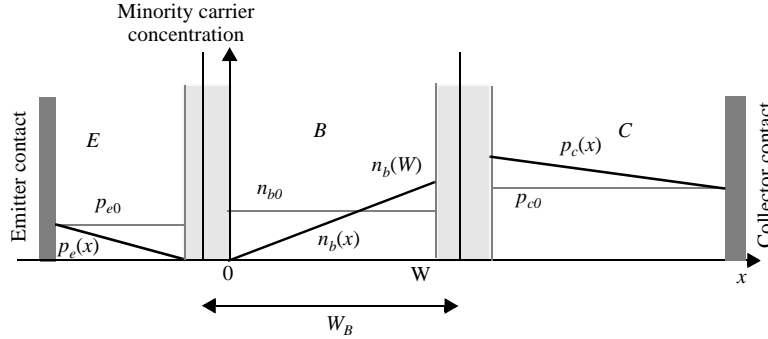


Figure A.5 Minority carrier concentration profile in reverse-active operation mode.

the gradient of the base minority carrier distribution is reversed, so that the diffusion current now is directed from collector to emitter. We use a similar approach to derive the current expression

$$I_E = -\frac{qA_C D_b n_{b0}}{W_B} e^{V_{BC}/\phi_T} \quad (\text{A.7})$$

where A_C is the collector area. Notice that we observe the current conventions of Figure A.2, which explains the negative sign of I_E . While this expression looks similar to Eq. (A.3), some major differences come to light when examining the components of the base current. While the recombination and base-emitter fractions stay at roughly the same level, the hole current from the base to the collector is substantial and actually exceeds the flow of electrons over the junction. The latter is explained by the fact that the base has a higher doping level than the collector, so that its current in forward bias is dominated by the flow of holes (Eq. (2.10)). These holes have to be provided by the base current, which

becomes comparable to the emitter current. The *reverse current gain* β_R is, therefore, small

$$\beta_R = -\frac{I_E}{I_B} \approx 1 \quad (\text{A.8})$$

and so is the *reverse common-base current gain*

$$\alpha_R = \frac{I_E}{I_C} = \frac{\beta_R}{\beta_R + 1} \ll 1 \quad (\text{A.9})$$

Saturation Region

In this mode of operation, both junctions are forward-biased (V_{BE} and $V_{BC} > 4\phi_T$). We only consider the case where the emitter junction has a stronger bias (or $V_{BE} > V_{BC}$). The opposite case is called the *reverse-saturation condition* and occurs rarely in digital circuits.

Under this condition, excess minority carriers are present at both the emitter and collector boundaries of the base region, although a somewhat higher concentration is present at the emitter end. The resulting minority carrier concentration is plotted in Figure A.6.

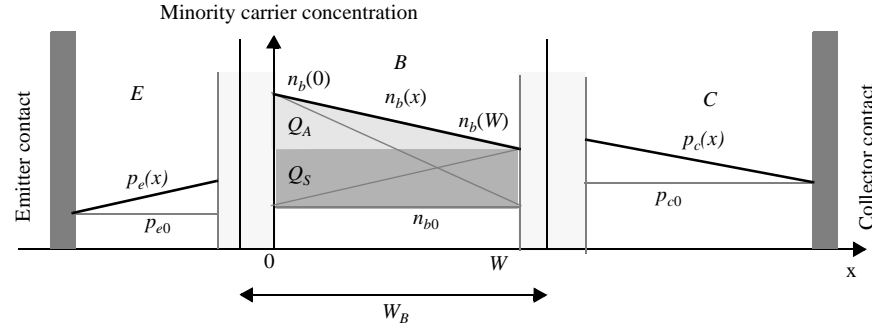


Figure A.6 Minority carrier concentration in saturation region ($V_{BE} > V_{BC} > 4\phi_T$).

Due to the short base width, a linear carrier gradient is still appropriate, and a diffusion current flows from emitter to collector, albeit of a smaller value than in the forward-active region. An important increase in the base charge is furthermore apparent. This gives rise to an increase in the recombination component of the base current. The combination of both factors indicates that the current gain is reduced in the saturation region, and is substantially smaller than β_F . A simple application of Kirchhoff's voltage law also shows that saturation condition corresponds to a small value of V_{CE} .

$$V_{CE} = V_{BE} - V_{BC} \quad (\text{A.10})$$

with $V_{BE} > V_{BC} > 4\phi_T$. When the device is in deep saturation, $V_{CE(sat)}$ generally ranges between 0.1 to 0.2V.

Cut-off Mode

Finally, in cut-off mode, both diodes are reverse-biased. The corresponding concentration profile is shown in Figure A.7. No excess base charge is present. The currents into the terminals are limited to the saturation currents of the reverse-biased diodes and are extremely small. The transistor is considered to be in the off-state.

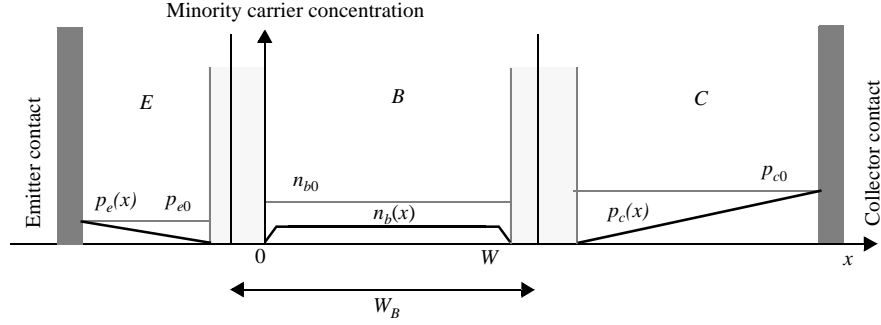


Figure A.7 Minority carrier profile in cut-off mode.

A Global View

The different operation modes of the bipolar transistors can be unified into a single set of equations, called the *Ebers-Moll model*. The model combines both forward and reverse currents, and in its most general form is expressed by the following set of equations:

$$I_C = I_F - \frac{I_R}{\alpha_R} \quad I_E = \frac{I_F}{\alpha_F} - I_R \quad I_B = I_E - I_C \quad (\text{A.11})$$

where

$$I_F = I_S(e^{V_{BE}/\phi_T} - 1) \quad I_R = I_S(e^{V_{BC}/\phi_T} - 1) \quad (\text{A.12})$$

and where I_S , α_F , and α_R are defined by Eqs. (A.3), (A.6), and (A.9) respectively. This form is often used for the computer representation of transistor large-signal behavior.

We now have all the information needed to plot the I - V characteristics of the bipolar transistor. The behavior of an MOS transistor was completely described by the variable set (I_D , V_{GS} , and V_{DS}), two of which could be chosen independently. In the bipolar case, the presence of the base current introduces one extra independent parameter. Therefore, a dual set of plots is needed to completely characterize the device. This is illustrated in the following pair of graphs. Figure A.8 plots I_C as a function of V_{CE} with I_B as a parameter for both the reverse and the forward operation modes. Observe the reduced current gain in reverse operation. Figure A.9 shows I_C as a function of V_{BE} , plotted on a logarithmic scale. The reduced slope of the currents at higher current levels is due to a number of secondary effects, which will be introduced later in the chapter.

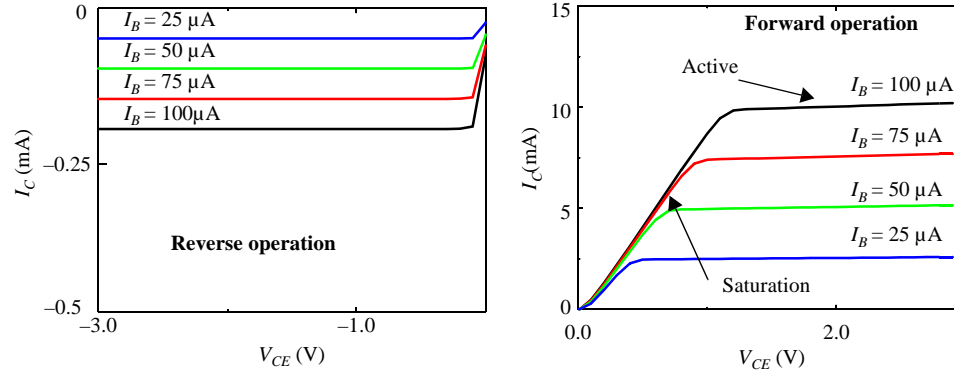


Figure A.8 Collector current as a function of V_{CE} with I_B as a parameter.

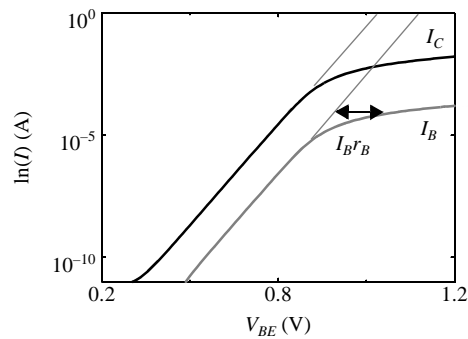


Figure A.9 Collector and base current as a function of V_{BE} (for a constant V_{CE} of 3 V).

Problem A.1 Current Gain of Bipolar Transistor.

Determine β_F and β_R of the device plotted in Figure A.8.

Problem A.2 *pnp* Transistor Characteristics.

Plot, approximately, the I - V characteristics of a *pnp* transistor.

Model for Manual Analysis

The Ebers-Moll model, while being accurate, tends to be too complex to be useful in the first-order analysis of a bipolar circuit. A simplified model of the bipolar transistor, intended for manual analysis, is therefore at hand. A summary of such a model, categorized per operation region, is given below (similar models can be established for the reverse operation modes):

- **Cut-off**— $V_{BE} < V_{BE(on)}$, $V_{BC} < V_{BC(on)}$

All currents are zero: $I_C = I_B = 0$

- **Forward-active**— $V_{BE} \geq V_{BE(on)}$, $V_{BC} < V_{BC(on)}$

A first-order model for large signal computations consists of a base-emitter diode and a current source (Figure A.10a). An even simpler version of this model, which is often useful, replaces the input diode by a voltage source with a fixed voltage $V_{BE(on)}$, which ranges between 0.6 and 0.75 V. This represents the fact that the base-emitter voltage varies only a little in the forward-active region due to the steep exponential characteristic (Figure A.10b).

- **Forward-saturation**— $V_{BE} \geq V_{BE(sat)}$, $V_{BC} \geq V_{BC(on)}$, $V_{BE} > V_{BC}$

During saturation, the collector-emitter voltage is assumed to be fixed at a voltage $V_{CE(sat)}$, which produces the model of Figure A.10c.

Example A.1 Static *npn* Model for Manual Analysis

The *npn* transistor of Example A.3 is to a first degree modeled by the following parameters:

$$\beta_F = 100, I_S = 1.E-17 \text{ A}, V_{BE(on)} = 0.7 \text{ V}, V_{BE(sat)} = 0.8 \text{ V}, V_{CE(sat)} = 0.1 \text{ V}.$$

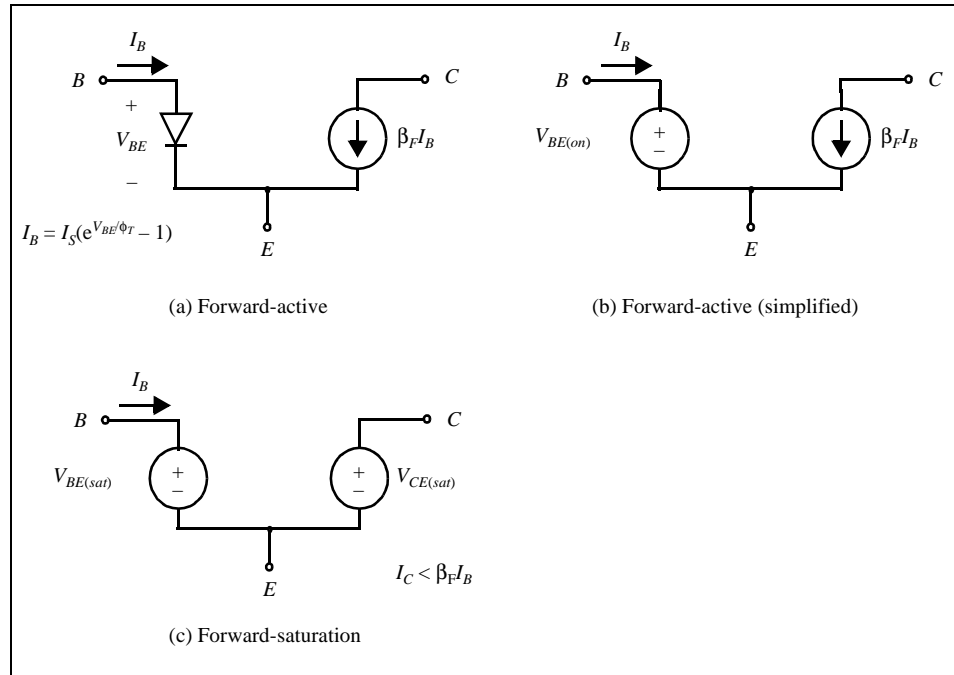


Figure A.10 First-order large signal models for *npn* bipolar transistor

A.2.3 Dynamic Behavior

As we have done for diodes and MOS devices, we model the dynamic behavior of the bipolar transistor by a set of capacitances. In the bipolar device, these capacitive effects originate from three sources: the base-emitter and base-collector depletion regions, the collector-substrate diode, and the excess minority carrier charge in the base. We discuss each of them individually.

The Base-Emitter and Base-Collector Depletion Capacitances

Depending upon the operation region, the *be* and *bc* junctions are either forward- or reverse-biased. This has an important impact on charge storage in the depletion region and, hence, on the equivalent capacitance. The nature of the junction capacitance and its model should be familiar by now, as it has been discussed extensively in previous sections. To combat the nonlinearity of the junction capacitance, the linearized large-scale model of Eqs. (2.17)–(2.18) (repeated below for clarity) is often used. In determining the value of the K_{eq} multiplier factor, one should be careful in delineating exactly the range of the voltage swing over the diode, since this affects the accuracy.

$$C_{eq} = K_{eq}C_{j0} \quad (\text{A.13})$$

where

$$K_{eq} = \frac{-\phi_0^m}{(V_{high} - V_{low})(1 - m)} [(\phi_0 - V_{high})^{1-m} - (\phi_0 - V_{low})^{1-m}] \quad (\text{A.14})$$

Both junction capacitances are thus adequately modeled by the following parameters:

- Their zero-bias value (C_{be0} and C_{bc0} , respectively). Although the value of the collector capacitance is typically a trifle larger than its counterpart at the emitter side, both values are comparable. For instance, for an *n*pn device in an advanced bipolar process with emitter size of $0.6 \mu\text{m} \times 2.4 \mu\text{m}$, $C_{be0} = 6.7 \text{ fF}$, and $C_{bc0} = 7.5 \text{ fF}$ [Yamaguchi88].
- The junction grading coefficients m_{be} and m_{bc} .
- The built-in potentials.

The Collector-Substrate Capacitance

As can be observed from the cross-sectional view of the bipolar transistor (Figure A.1), the *n*-type collector is isolated from the substrate by an inverse-biased diode. The model is complicated by the presence of the n^+ buried layer and the p^+ isolation regions. This results in a parasitic capacitance, which is, once again, a nonlinear junction capacitance. Since the area of the collector-substrate junction is considerably larger than the emitter area, the zero-bias value of this parasitic capacitance can be substantial and might dominate the performance of the device. The value of C_{cs0} typically ranges between 1 and 2.5 times the base-emitter capacitance. Advanced processes reduce its impact by providing

trench isolation between devices (i.e., a vertical plug of SiO₂ inserted between transistors), or even by using isolating substrates.

The Base Charge—An Introduction

When deriving the diode current model, we were able to relate the diode current to the excess minority charge in the neutral regions. This proved to be beneficial, especially when studying the dynamic characteristics of the device. A similar approach works for the bipolar transistor as well. The important parameter here is the *excess minority carrier charge in the base region*, which is a strong function of the operation mode.

1. Forward-active

Assuming that the minority carriers in the base region display a linear gradient, as shown in Figure A.3, the total excess minority charge in the base is readily computed.

$$Q_F = qA_E W_B \left(\frac{n_b(0) + n_b(W) - 2n_{b0}}{2} \right) \approx qA_E W_B \frac{n_b(0)}{2} \quad (\text{A.15})$$

this yields the following relationship between the collector current of Eq. (A.3) and the excess base charge,

$$I_C = \frac{Q_F}{\tau_F} \quad (\text{A.16})$$

with τ_F the *mean forward-transit time*, which can be interpreted as the mean time for the minority carriers to diffuse from the emitter to the collector.

$$\tau_F = \frac{W_B^2}{2D_B} \quad (\text{A.17})$$

Observe the similarity between the forward transit time and the mean transit time of the short-base diode, Eq. (2.20). Keeping this transit time as short as possible is clearly a good idea, as it results in a larger current for a smaller base charge. This can be accomplished by reducing the base width to a maximal extent. This trend is apparent in contemporary bipolar processes, where W_B varies between 50 and 100 nm. Values of τ_F range between 5 and 30 psec.

In a similar fashion, we can relate the base current I_B to the base charge

$$I_B = \frac{Q_F}{\tau_{BF}} \quad (\text{A.18})$$

with τ_{BF} the *minority carrier lifetime* in the base region during forward-active operation. It accounts for the three factors shown in Figure A.4. Taking into account that the base and collector currents are related by the forward current gain, a relationship between the base transit and lifetimes is easily derived.

$$\beta_F = \frac{\tau_{BF}}{\tau_F} \quad (\text{A.19})$$

2. Reverse-active

A similar analysis yields the following charge-driven model for the reverse currents.

$$I_E = -\frac{Q_R}{\tau_R} \quad I_C = -\frac{Q_R}{\tau_R} - \frac{Q_R}{\tau_{BR}} \quad \text{and} \quad \beta_R = \frac{\tau_{BR}}{\tau_R} \quad (\text{A.20})$$

3. Forward-saturation

This excess base charge in the saturation region can be divided into a forward and a reverse component, as shown by the dotted lines in Figure A.6. This results in the following expression for I_C ,

$$I_C = \frac{Q_F}{\tau_F} - \frac{Q_R}{\tau_R} \frac{1}{\alpha_R} = \frac{Q_F}{\tau_F} - Q_R \left(\frac{1}{\tau_R} + \frac{1}{\tau_{BR}} \right) \quad (\text{A.21})$$

It is conceptually more interesting to consider a different subdivision as illustrated by the shadings of grey in Figure A.6.

The first component, called the *critical base charge* Q_A , has a triangular shape. It represents the charge needed to bring the transistor to the *edge-of-saturation (eos)*, which is the boundary condition between forward-active and saturation. At the *eos*, the equations governing the forward-active mode are still valid.

$$I_{C(eos)} = \frac{Q_A}{\tau_F} = \beta_F I_{B(eos)} \quad (\text{A.22})$$

where $I_{B(eos)}$ is the corresponding base current.

The second component, Q_S , is called the *overdrive charge* and is the result of both junctions being in forward bias. Being rectangular, it does not contribute to the gradient of the excess minority carriers, nor to the collector current. It arises from the pushing of more current into the base than is required to reach the saturation condition. The current in excess of the $I_{B(eos)}$ is called the *excess base current* I_{BS} and is related to the overdrive charge by a time constant τ_S , called the *saturation time constant*,

$$I_{BS} = \frac{Q_S}{\tau_S} \quad (\text{A.23})$$

which is a weighted average of τ_F and τ_R [Hodges88].

$$\tau_S = \frac{\alpha_F(\tau_F + \alpha_R\tau_R)}{1 - \alpha_F\alpha_R} \approx \frac{\alpha_R}{1 - \alpha_R}\tau_R \quad (\text{A.24})$$

(assuming that $\tau_F \ll \tau_R$ and $\alpha_F \approx 1$).

With the transistor in saturation, an expression for the (static) base current is now readily derived

$$I_B = I_{B(eos)} + I_{BS} = \frac{Q_A}{\tau_{BF}} + \frac{Q_S}{\tau_S} \quad (\text{A.25})$$

It is, especially, the overdrive charge that makes the saturation region an operation mode to avoid in digital design. Recall from the discussion of diode dynamics that it takes time to remove or provide this minority excess charge. Bringing a transistor in and out of saturation is a slow operation that is inconsistent with the high performance requirements we impose on digital circuits.

Changing the Base Charge—The Diffusion Capacitance

From the above, it becomes apparent that a considerable amount of excess minority carrier charge is stored in the base region of an active or saturated bipolar transistor. Changing the operation mode of the device requires the addition or the removal of this charge.

As with the diode, this minority excess charge can be modeled by capacitances in parallel with the *be* junction (representing the forward charge Q_F) and the *bc* junction (representing the reverse charge Q_R). The value of these capacitors is a strong function of the transistor operation mode and the current levels. A large-signal approach, in the sense of Eqs. (2.23)–(2.24), is therefore appropriate. Observe that the base charge is an exponential function of the voltage over the junctions.

As an example, during forward-active operation Q_R is ignorable. Only a single capacitor between base and emitter, representing $Q_F = I_C \tau_F$, has to be considered. This capacitor is often called *the diffusion capacitance* C_d , as it represents the carriers diffusing through the base region. Similar to the approach used for the diode, the **small-signal value** of C_d is readily computed.

$$C_d = \tau_F \frac{I_C}{\phi_T} \quad (\text{A.26})$$

The large-signal value (called C_D) is comparable to Eq. (2.24). During saturation, the base charge can be represented as the sum of the forward and reverse charges. For manual analysis, it is often more convenient to divide the base charge into the *overdrive* and the *edge-of-saturation* (eos) charge in correspondence with the model discussed in the previous section. Bringing a transistor out of saturation simply requires the removal of the overdrive charge.

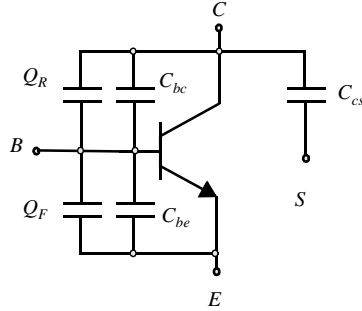
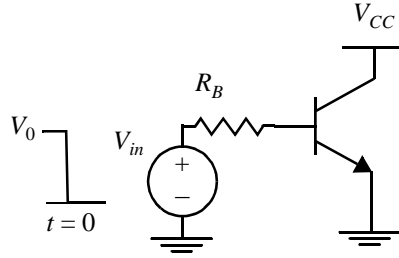
Bipolar Transistor Capacitive Model

The effects of all the above contributions are combined into a single capacitive model for the bipolar transistor, as shown in Figure A.11. Its components can be identified based on the preceding discussion. Having an intuitive feeling for what components are important for each operation condition is what makes a good bipolar designer.

Bipolar Transistor Switching Time—A Case Study

Consider the simple circuit of Figure A.12. We analyze the case where the transistor is initially on, or $V_0 > V_{BE(on)}$, and in forward-active operation mode. At time $t = 0$, the input voltage drops to 0, and the device is turned off.

Similar to the diode case, the turn-off process proceeds in two phases. In a first step, the base charge is removed, turning the transistor off. During this phase, the base-emitter

Figure A.11 Capacitive model of an *npn* bipolar transistor.Figure A.12 Example of *npn*-transistor-based circuit.

voltage remains relatively constant. In the second phase, the base-emitter voltage drops to zero. This requires the (dis)charging of the junction capacitances. We analyze each of these phases individually. Again, be aware that this break-up is a simplification.

1. Removal of the base charge

Before time $t = 0$, the excess minority-carrier charge in the base equals

$$Q_F(t = 0) = Q_{F0} = I_C(t = 0) \times \tau_F$$

where $I_C(t = 0)$ is expressed by the following equation:

$$I_C(t = 0) = \beta_F \times I_B(t = 0) = \beta_F(V_0 - V_{BE(on)})/R_B$$

We may assume that during the base-charge removal the base-emitter voltage remains relatively constant and is approximately equal to $V_{BE(on)}$. This means the space charge of the *be* and *bc* junctions is constant also, and not a factor at this point. The length of this interval is thus uniquely defined by the time it takes to remove the base charge. Under the assumption of a fixed base-emitter voltage, the base current that accomplishes this is constant and equals

$$I_B = (0 - V_{BE(on)})/R_B$$

This current actually serves two purposes: (1) to sustain the recombination in the base and (2) to remove the base charge. This dual purpose is expressed by the following equation,

$$I_B = \frac{Q_F}{\tau_{BF}} + \frac{dQ_F}{dt} \quad (\text{A.27})$$

Solving the differential equation, taking the initial condition for Q_F into account, yields $Q_F(t)$,

$$Q_F(t) = \tau_{BF}[I_B - (I_B(t=0) - I_B)e^{-t/\tau_{BF}}] \quad (\text{A.28})$$

and the time to remove the excess charge ($Q_F = 0$).

$$t_{base} = \tau_{BF} \ln \left[\frac{I_B(t=0) - I_B}{-I_B} \right] \quad (\text{A.29})$$

The recombination factor in Eq. (A.27) is often small compared to I_B and actually drops off very rapidly. It can be ignored for a first-order analysis, which means that all base current is assumed to be used for base-charge removal. This yields a simplified expression for t_{base} .

$$t_{base} = \frac{\Delta Q_F}{I_B} = -\frac{Q_{F0}}{I_B} = -\tau_{BF} \frac{I_B(t=0)}{I_B} = \frac{C_D \Delta V}{I_B} \quad (\text{A.30})$$

Observe especially the last component of Eq. (A.30). C_D stands for the large-signal equivalent capacitance (defined over the range ΔV). This equation states that, if the recombination can be ignored, the base charge removal can be modeled as the discharging of a capacitor.

2. Changing the space charge of the *be* and *bc* junctions

Once the transistor is off, the rest of the transient is devoted to discharging the junction capacitances to 0 volt. During this time span, the circuit is modeled as a first-order RC-circuit. The time to reach 90% of the final value is a simple function of the time constant of the network

$$t_{space} = 2.2R_B(C_{be} + C_{bc}) \quad (\text{A.31})$$

Example A.2 Dynamic Behavior of an *npn* Transistor

Assume that the bipolar transistor in Figure A.12 is characterized by the following parameters: $\beta_F = 100$, $V_{BE(on)} = 0.75$ V, $\tau_F = 10$ psec, $C_{be0} = 20$ fF, $C_{bc0} = 22$ fF, $C_{cs0} = 47$ fF, $m = 0.33$, $\phi_0 = 0.75$ V. The following circuit parameters are further assumed: $V_{CC} = 2$ V, $V_0 = 1$ V, and $R_B = 5$ k Ω .

First we compute the initial values of the currents. Since the V_{CE} of the transistor is equal to 2 V, the transistor is, obviously, in the forward-active mode. The base current is expressed as

$$I_B(t=0) = (1 - 0.75) / (5 \times 10^3) = 50 \mu\text{A}$$

After the switching of the input, the base current changes direction, or

$$I_B = (-0.75) / (5 \times 10^3) = -150 \mu\text{A}$$

The time to remove the base charge can now be computed with the aid of Eq. (A.29).

$$t_{base} = (100 \times 10 \text{ psec}) \ln (200 \mu\text{A} / 150 \mu\text{A}) = 288 \text{ psec}$$

The simplified expression of Eq. (A.30), on the other hand, yields 333 psec. This is sufficiently close to be useful as a first-order approximation.

As mentioned, the circuit behaves as an RC -equivalent network during the space-charge removal. To approximate the time constant of the network, it is necessary first to linearize the junction capacitances over the voltage range of interest. Observe that the collector voltage remains fixed during the transient, so that C_{bc} can be considered as a grounded capacitor. The voltage over the base-emitter junctions decreases from 0.75 V to 0 V, while the voltage over the bc junction increases from 1.25 V to 2 V. Injecting those values into Eq. (2.18) yields the following values of K_{eq} : $K_{eq(be)} = 1.5$ and $K_{eq(bc)} = 0.68$. The value of t_{depl} can now be computed

$$t_{depl} = 2.2 \times 5 \text{ k}\Omega \times (1.5 \times 20 \text{ fF} + 0.68 \times 22 \text{ fF}) = 495 \text{ psec}$$

which is considerably larger than the base-charge removal period.

These results are validated by the SPICE simulations shown in Figure A.13. During the removal of the base charge, a fast decrease in the collector current can be observed. At the end of the base-charge removal phase, the collector current drops to 0, the transistor goes into the cut-off mode, and the base-emitter voltage decays to 0 V in an exponential manner. The simulated values of t_{base} and t_{depl} equal 302 psec and 429 psec (the transition between both regions is somewhat fuzzy and is determined here by the point where the extrapolated linear collector current crosses the zero-axis). The discrepancy between the predicted and simulated values of t_{depl} is due to the linearization, which is more accurate in the reverse- than the forward-biased regions of the diodes.

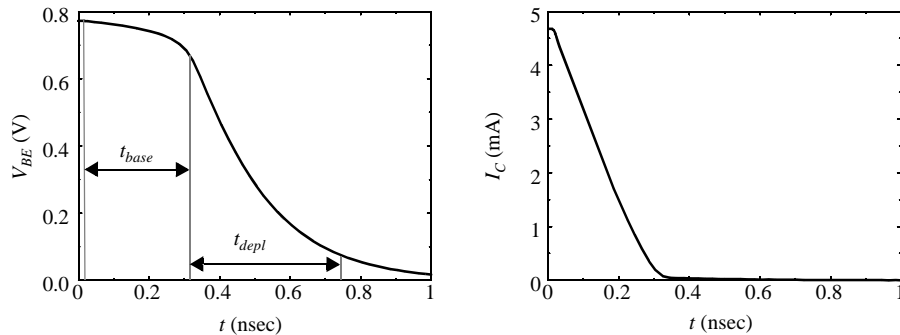


Figure A.13 Simulated transient response of the circuit of Figure A.12.

Problem A.3 Turning on the Bipolar Transistor

Derive the time it takes to turn on the nnp transistor in the example of Figure A.12, assuming the same device and circuit parameters and assuming a step from 0 V to 1 V at the input.

A.2.4 The Actual Bipolar Transistor—Secondary Effects

In the previous sections, we have sketched a model of an ideal nnp device. As always, the actual transistor is somewhat more complex, and some second-order parameters have to be considered.

The Early Voltage

In the analysis of the forward-active operation mode, we assumed the base-collector junction to be reverse-biased and to have no effect on the value of the collector current (Eq. (A.3)). This is a useful approximation for first-order analysis, but is not strictly true in practice. The reason for this is easily understood. Until now, we ignored the width of the depletion regions with respect to the width of the base region W_B . In other words, we assumed that the width of the neutral base region W is equal to the base width W_B . For thin base regions, this is not really the case. Furthermore, an increase in the collector-base voltage increases the reverse bias and, consequently, the width of the depletion region. Since this decreases the effective base width, it reduces the forward transit time τ_F and raises the collector current I_C .

The collector current therefore increases as a function of the collector voltage, as is shown in Figure A.14, which plots the collector voltage as a function of V_{CE} with V_{BE} as a parameter. It is interesting to observe that all characteristics extrapolate to a single point on the V_{CE} -axis, called the *early voltage* V_A . For integrated circuit transistors, V_A varies from 15 V to 100 V. The inverse of the early voltage is analogous to the channel-length modulation parameter λ in the MOS transistor.

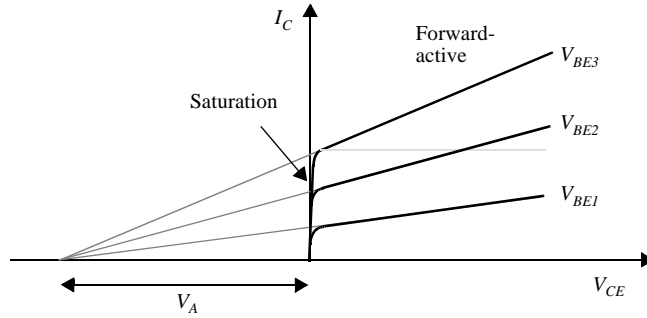


Figure A.14 Bipolar transistor characteristics showing the early voltage effect.

The collector-current Eq. (A.3) can be adapted to compensate for the early effect. Due to additional computational complexity this model is rarely used in manual calculations and is reserved for computer models.

$$I_C = I_S \left(1 + \frac{V_{CE}}{V_A} \right) e^{V_{BE}/\phi_T} \quad (\text{A.32})$$

Transistor Breakdown Voltages

The mechanism of avalanche breakdown in a pn -junction was described in section 2.3.4. Similar effects occur at the be and bc junctions of the bipolar transistor. This limits the voltages that can be applied to the device. The breakdown voltages depend upon the transistor configuration and are commonly expressed as BV_{CBO} (common base configuration breakdown voltage), V_{CEO} (common emitter configuration) and BV_{EBO} (base-emitter breakdown). The values of the breakdown voltages vary from 6–8 V (BV_{EBO}) to several tens of volts for the others. Since breakdown creates excessive power dissipation and might lead to device destruction when sustained, it is avoided in digital bipolar design.

Parasitic Resistances

Parasitic resistances are produced by the finite resistance of the neutral regions of the transistor, as shown in Figure A.15. While r_E is normally very small (1–5 Ω), r_B and r_C can be substantial and have a significant impact on the device performance due to the high doping level of the emitter region.

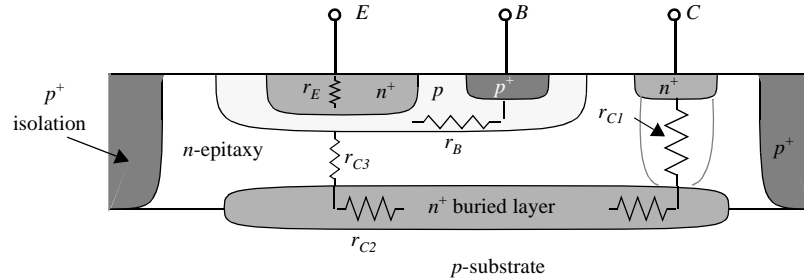


Figure A.15 npn transistor structure, showing parasitic resistances.

The collector resistance consists of three components, labeled r_{C1} , r_{C2} , and r_{C3} . Of these, the latter is in general the dominant component. In advanced processes, r_{C3} is reduced by adding a low-resistance n^+ region below the collector (called a *deep collector*), shown in dotted lines in Figure A.15. The value of r_C varies between 20 Ω (with deep collector) to almost 1 k Ω (without). The value of r_B ranges from 50 Ω to 500 Ω and unfortunately varies with the collector current due to an effect called *current crowding*. This effect causes the transistor action to occur at the edges of the emitter area, instead of the central portion, which results in a varying distance between the active base region and the base contact.

The presence of these parasitic resistances is particularly noticed under high collector-current conditions. The voltage drop over the resistances reduces the voltage differences at the internal device terminals and, as a result, the transistor currents. Simulation accuracy is strongly contingent upon a careful modeling of the resistances. The designer can, if necessary, reduce the parasitic effects by modifying the transistor structure. Available options include increasing the emitter area, and providing multiple base and emitter terminals.

β_F Variations

The ideal model states that in the forward-active region, the current gain β_F is a constant. This parameter, in fact, does vary with the operating conditions of the device as shown in Figure A.16, which plots the values of $\ln(I_C/I_B)$ as a function of V_{BE} . It can be observed that for intermediate values, a constant value of β_F can indeed be observed. At low current values, a degradation of the current gain occurs. This is attributed to an increase in the base current, caused by the recombination of carriers in the *be* depletion region. This effect is present at all current levels, but only has an impact under very low current conditions. At the other end of the spectrum, the collector current drops below the ideal current as a result of *high-level injection* effects—as the injected minority carrier density in the

base approaches the majority carrier density, the hole current from base to emitter becomes substantial leading to a decline in the collector current.

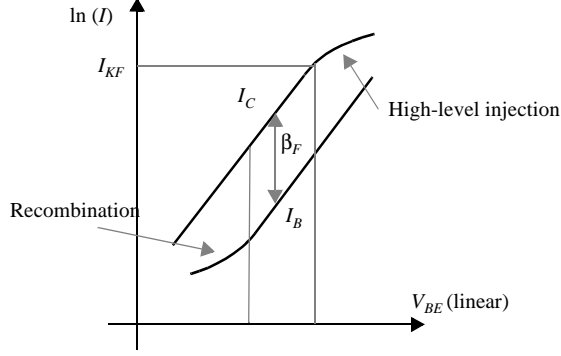


Figure A.16 Deterioration of the current gain at low and high current levels.

In addition to high-level injection, the β_F at high currents is also affected by the onset of the *Kirk effect* or *base stretching*. This occurs when the level of injected electrons into the collector becomes comparable to the donor-atom doping density of the collector region. This causes a collapse of the space-charge region, and the base stretches out into the collector regions, resulting once again in a current-gain degradation.

The effects of the low- and high-level injection can be modeled by modulating the emission coefficient n in the base and collector currents Eq. (A.33). For instance, the high-level injection effect is modeled by modifying n from the standard value of 1 to 2. The point of onset of the high-level injection effects is called the *knee current* (I_{KF}).

$$I_C = I_S e^{V_{BE}/n\phi_T} \quad (\text{A.33})$$

The impact of some of the high-current and parasitic resistance effects is observed in the I - V plots of Figure A.9. Notice especially how the voltage drop over the base resistance causes both the collector and base currents to degrade for high values of V_{BE} .

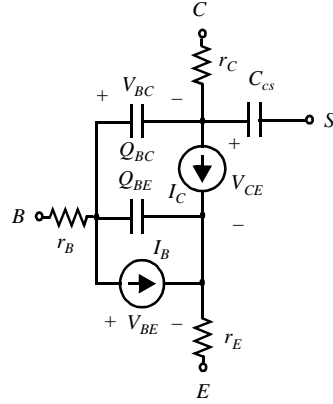
A.2.5 SPICE Models for the Bipolar Transistor

Earlier versions of SPICE supported two separate models for the bipolar transistor: the *Ebers-Moll (E-M)* and the charge-based *Gummel-Poon (G-P)* model. Both were merged in later versions into the *modified Gummel-Poon* model which incorporates various extensions to model high-bias conditions [Getreu76]. Under default conditions for certain parameters, this model automatically reduces to the simple E-M model.

The complete model for an *npn* transistor, as implemented in SPICE, is shown in Figure A.17 (The model of a *pnp* device is obtained by reversing the polarities of V_{BE} , V_{BC} , V_{CE} , I_C and I_B). Under normal operating conditions, the values of the current sources are given by the modified Ebers-Moll equations

$$I_C = I_S (e^{V_{BE}/(n_F \phi_T)} - e^{V_{BC}/(n_R \phi_T)}) \left(1 - \frac{V_{BC}}{V_A}\right) - \frac{I_S}{\beta_R} (e^{V_{BC}/(n_R \phi_T)} - 1) \quad (\text{A.34})$$

and

Figure A.17 SPICE model for an *n*pn transistor.

$$I_B = \frac{I_S}{\beta_F} (e^{V_{BE}/(n_F \phi_T)} - 1) + \frac{I_S}{\beta_R} (e^{V_{BC}/(n_R \phi_T)} - 1) \quad (\text{A.35})$$

where I_S is the saturation current, V_A the early voltage, and n_F and n_R the forward and reverse current-emission coefficients. As can be observed, the model superimposes the forward- and reverse-operation conditions. The *base-width modulation* effect is incorporated in the collector current through the *early voltage* V_A .

This model does not take into account the second-order effects occurring at low current levels (depletion-layer recombination) or high injection levels. These are properly accounted for in the Gummel-Poon model, which relates all terminal voltages and currents to the base charge. A complete description of the model is beyond the scope of this text.

The intrinsic capacitances and resistances are modeled by the expressions presented in Sections 2.2.4 and A.2.3. Q_{BE} and Q_{BC} represent the combined effects of the depletion and base charge for the *be* and the *bc* junctions, respectively.

Table A.2 lists the main SPICE model parameters. The parameters covering the parasitic resistive and capacitive effects have been transferred to a separate table for the sake of clarity (Table A.3).

In contrast to the MOS transistor case, not many parameters can be defined for the individual transistor. The only parameter that can be defined at the instantiation time of a device is the AREA factor, which determines how many of the bipolar junction transistors of type MODNAME are put in parallel to create that particular device. The default value is 1. This parameter is not used very often. The common practice is to provide a different model for every possible device dimension, which may be used. This is acceptable, because a typical bipolar digital design only uses a small range of device dimensions, contrary to the MOS practice, where transistor dimensions can vary widely, even in a single circuit.

Example A.3 SPICE Bipolar Junction Transistor Model

The model for a bipolar *n*pn transistor with an emitter area of $2 \mu\text{m} \times 3.75 \mu\text{m}$, implemented in a process technology with a minimum line-width of $1 \mu\text{m}$, is included. We use this as the generic bipolar transistor in the rest of the text, unless specified otherwise.

* SPICE Model for 2×3.75 npn transistor

Table A.2 Main SPICE bipolar junction transistor model parameters.

Parameter Name	Symbol	SPICE Name	Units	Default Value
Transport Saturation Current	I_S	IS	A	1.0E-16
Maximum Forward Current Gain	β_F	BF	-	100
Forward Current-Emission Coefficient	n_F	NF	-	1
Forward Early Voltage	V_{AF}	VAF	V	¥
Maximum Reverse Current Gain	β_R	BR	-	1
Reverse Current-Emission Coefficient	n_R	NR	-	1
Reverse Early Voltage	V_{AR}	VAR	V	¥
Corner for Forward Beta High-Current Roll-off ^a	I_{KF}	IKF	A	¥
<i>be</i> Junction Leakage Saturation Current ^a		ISE	A	1.0E-13
<i>be</i> Junction Leakage Emission Coeff. (low-current condition) ^a		NE	-	1.5
Corner for Reverse Beta High Current Roll-off ^a	I_{KR}	IKR	A	¥
<i>bc</i> junction leakage saturation current ^a		ISC	A	1.0E-13
<i>bc</i> junction leakage emission coeff. (low-current condition) ^a		NC	-	2
Ideal Forward Transit Time	t_F	TF	sec	0
Ideal Reverse Transit Time	t_R	TR	sec	0

a. Gummel-Poon Model Parameter

Table A.3 SPICE Parameters for parasitics (resistances, capacitances).

Parameter Name	Symbol	SPICE Name	Units	Default Value
Emitter Resistance	r_E	RE	Ω	0
Collector Resistance	r_C	RC	Ω	0
Zero-Bias Base Resistance	r_B	RB	Ω	0
Minimum Base Resistance		RBM	Ω	RB
Current where RB falls halfway to RBM		IRB	A	¥
Zero-Bias <i>be</i> -Junction Capacitance	C_{be0}	CJE	F	0
<i>be</i> -Junction Grading Coeff.	m_{be}	MJE	—	0.33
<i>be</i> -Junction Built-in Voltage	ϕ_{be}	VJE	V	0.75
Zero-Bias <i>bc</i> -Junction Capacitance	C_{bc0}	CJC	F	0
<i>bc</i> -Junction Grading Coeff.	m_{bc}	MJC	—	0.33
<i>bc</i> -Junction Built-in Voltage	ϕ_{bc}	VJC	V	0.75
Zero-Bias Collector-Substrate Cap.	C_{cs0}	CJS	F	0
<i>cs</i> -Junction Grading Coeff.	m_{cs}	MJS	—	0
<i>cs</i> -Junction Built-in Voltage	f_{cs}	VJS	V	0.75

```
.MODEL BF=100 BR=1 IS=1.E-17 VAF=50
+ TF=10E-12 TR=5E-9 IKF=2E-2 IKR=0.5
+ RE=20 RC=75 RB=120
+ CJE=20E-15 VJE=0.8 MJE=0.5 CJC=22E-15 VJC=0.7 MJC=0.33
+ CJS=47E-15 VJS=0.7 MJS=0.33
```

An example of an actual transistor instantiation is given below. Transistor *Q1* is an NPN device with its collector, base and emitter terminals connected to nodes 2, 1, 0, respectively. Transistor *Q2* consists of four transistors of type NPN, connected in parallel, and has its substrate connected to node 5.

```
Q1 2 1 0 NPN
Q2 2 1 0 5 NPN 4
```

A.3 The Bipolar ECL Inverter

While we learned in the previous section that the CMOS inverter has an almost ideal dc characteristic and low power consumption, it also became clear that the speed of the inverter is restricted by the fact that the maximum current is proportional to V_{DD}^2 (or V_{DD} under velocity-saturation conditions). This puts a firm upper bound on the achievable clock speeds. The bipolar device, on the other hand, is known to have an exponential relationship between collector current and base-emitter voltage, which means that a small change in voltage can provide a large current. It is therefore reasonable to assume that bipolar gates should be capable of achieving superior propagation delays.

Although the TTL gate (transistor-transistor-logic) has long been the flagship of the bipolar digital logic gates, its importance has been reduced dramatically in the last decade as CMOS has become more and more competitive. Most modern bipolar digital designs are implemented in a circuit style called emitter-coupled logic (or ECL). We will, therefore, concentrate on this logic family. Following the scheme of the CMOS discussion, we will first describe the ECL gate in a qualitative fashion, followed by a quantitative analysis of its static and dynamic behavior, as well as the power dissipation. A discussion of the effects of technology scaling concludes the section.

A.3.1 Issues in Bipolar Digital Design: A Case Study

Before engaging in a discussion of the ECL gate, some issues in the design of bipolar gates should be highlighted. The particular nature of the bipolar junction transistor and the lack of a good complementary device in most standard bipolar manufacturing processes translate into gate topologies that are dramatically different from the CMOS structures. The most important features can be summarized as follows:

- The high transconductance of the bipolar devices translates to a large variation in collector current for a small change in the input voltage.
- The input resistance of the transistor is finite. This means that a fan-out gate presents a current-sink for the driving gate. This is in contrast to CMOS designs, where fan-out presents only a capacitive load.

- The presence of an excessive amount of base-charge makes the transistor very slow when operated in the saturation mode. Saturating the device should, therefore, be avoided as far as possible.

To illustrate these observations, let us first examine a simple bipolar gate, called the *RTL (resistor-transistor logic)* gate. A schematic of such a structure is shown in Figure A.18. It can be readily seen from the circuit that with the input voltage V_{in} less than the turn-on voltage $V_{BE(on)}$ of the transistor, the transistor is in cut-off mode, the collector current I_C essentially equals zero, and the output voltage V_{out} is equal to the supply voltage V_{CC} . When the input voltage is increased above $V_{BE(on)}$, the transistor turns on and enters the forward-active mode. Further increasing the input voltage causes the output voltage to drop due to the increasing collector current I_C . This drop is swift, as the bipolar transistor is characterized by a *large gain in the forward-active mode*. With sufficient input voltage, when the output voltage has fallen sufficiently, the transistor will enter the saturation region. The output voltage V_{out} remains fixed at $V_{CE(sat)} \approx 0.1$ V, the saturation collector-emitter voltage. To enter this mode of operation, we require (as a first approximation) that $V_{in} \geq V_{BE(sat)} (\approx 0.8$ V). Observe that the above analysis uses the simple model for manual analysis of the bipolar transistor, presented in Figure 2.42.

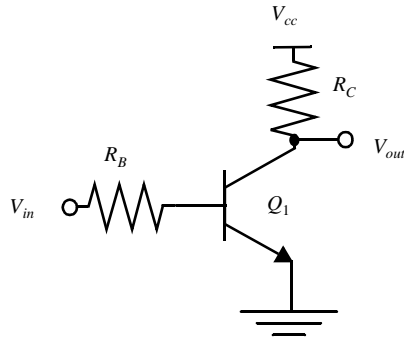


Figure A.18 Resistor-transistor logic (RTL) inverter.

From the above, it becomes obvious that the gate of Figure A.18 acts as an inverter. A rough sketch of the voltage-transfer characteristic is shown in Figure A.19, where linear interpolations are used to join the two major breakpoints of the graph.

Example A.4 VTC of an RTL Inverter

The VTC of an RTL inverter is derived for the following component and parameter values: $V_{CC} = 5$ V, $R_C = 1$ k Ω , and $R_B = 10$ k Ω . The bipolar transistor model presented in Chapter 2 is used for Q_1 . Initially, it is assumed that the fan-out of the inverter is set to 0.

- For $V_{in} < V_{BE(on)}$, Q_1 is in cut-off, $I_C = 0$, and $V_{out} = V_{OH} = V_{CC}$.
- In the low output mode, Q_1 is in saturation (or $V_{out} = V_{CE(sat)} \approx 0.1$ V), if the condition $V_{BE} \geq V_{BE(sat)}$ is fulfilled. The boundary condition on V_{in} for this to be valid can be derived by analyzing the circuit with Q_1 at the edge of saturation. This translates to the following condition on V_{in} ,

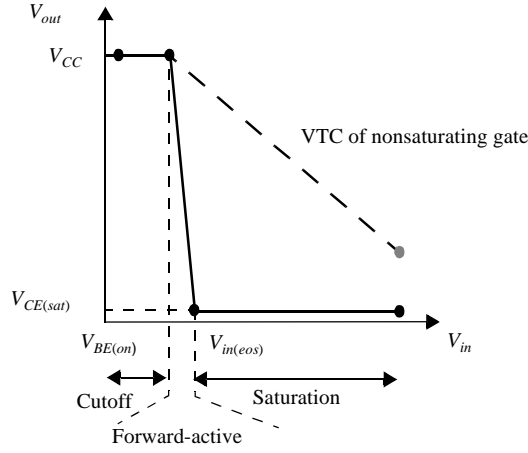


Figure A.19 Approximated voltage transfer characteristic of inverter.

$$I_B = \frac{I_C}{\beta_F} = \frac{V_{CC} - V_{CE(sat)}}{R_C \beta_F} \quad (\text{A.36})$$

$$V_{in(eos)} = V_{BE(eos)} + I_B R_B = V_{BE(eos)} + \frac{R_B}{R_C} \left(\frac{V_{CC} - V_{CE(sat)}}{\beta_F} \right)$$

Assuming $V_{CE(sat)} = 0.1$ V, $V_{BE(eos)} = 0.8$ V, and $\beta_F = 100$, Eq. (A.36) evaluates to $V_{in(eos)} = 1.29$ V.

The simulated VTC of the RTL inverter is plotted in Figure A.20 (consider only the case of zero fan-out at present). The observed breakpoints equal approximately 0.7 V ($\sim V_{BE(on)}$) and 1.4 V, which are close to the predicted values. The major discrepancy is the value of V_{OL} , which equals 0.5 V in contrast to the predicted value of 0.1 V. This difference is caused by the nonzero value of the series collector and emitter resistances ($R_E = 20$ Ω and $R_C = 75$ Ω). In the saturation mode, the collector current I_C equals approximately $V_{CC}/R_C = 5$ mA. This causes an extra voltage drop of $5 \times 10^{-3}(75 + 20) = 0.48$ V over the collector-emitter terminals of Q_1 .

It can be observed that the RTL inverter displays some major **dc problems**:

1. The VTC is *asymmetrical*. The NM_L is less than or equal to 0.2 V. Noise and disturbances on the *GND* line or input signal could easily result in faulty operation.
2. To obtain a reasonable low value of V_{OL} , it is necessary to drive the *transistor into saturation*. From the earlier remarks, this can be labeled as unfortunate as the excessive base charge has a significant impact on switching performance. One option is to choose the resistor values so that saturation never occurs for the voltage range of interest, which is equivalent to requiring that $V_{in(eos)} > V_{CC}$ in Eq. (A.36). This approach causes a deterioration of V_{OL} and reduces the gain in the transient region. The corresponding VTC is plotted in shaded lines in Figure A.19.
3. The *output impedance* of the gate is equal to R_C when the output is *high*. Reducing R_C is not a reasonable option, as this raises the value of the collector current I_C and

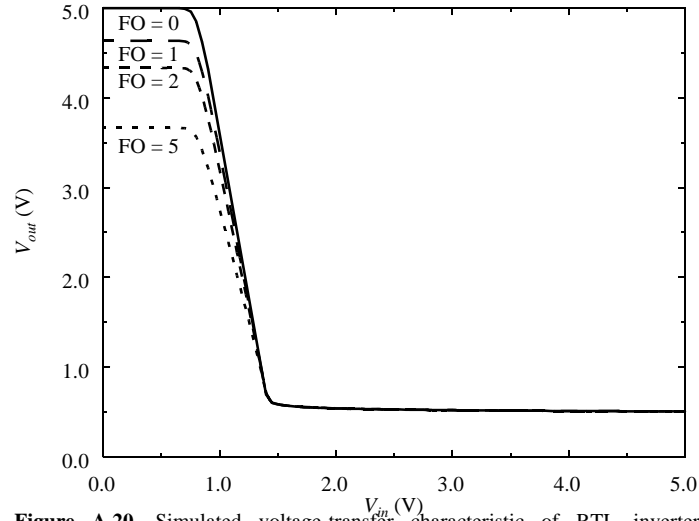


Figure A.20 Simulated voltage-transfer characteristic of RTL inverter for different fan-out configurations.

increases power consumption and switching noise.¹ Furthermore, in contrast to the CMOS inverter, the gate has a finite input impedance. When Q_1 is on, current flows into its base through R_B . The V_{OH} of the inverter (and the NM_H) is, hence, sensitive to loading as is illustrated in Figure A.21. This effectively restricts the fan-out of the inverter to approximately three gates.

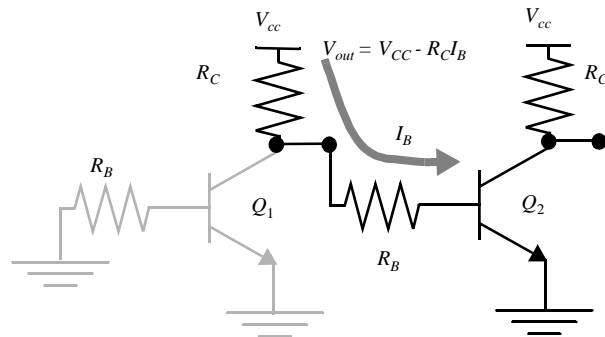


Figure A.21 Effect of fan-out on V_{OH} .

¹ *Switching noise* is the noise present on the supply rails resulting from the switching of large currents in the presence of parasitic resistances and inductors.

Example A.5 Effect of Fan-out on the VTC of the RTL Inverter

Suppose that the output of the RTL gate of the previous example is connected to N identical gates and that its output is high. In this mode, each of the connecting gates will sink an amount of current I_{in} into its base as illustrated in Figure A.21.

$$I_{in} = \frac{V_{out} - V_{BE(sat)}}{R_B} \quad (\text{A.37})$$

This results in the following value of V_{out} as a function of the fan-out N :

$$\begin{aligned} V_{out} &= V_{CC} - N \cdot R_C \cdot I_{in} = V_{CC} - N \cdot R_C \cdot \frac{V_{out} - V_{BE(sat)}}{R_B} \\ &= \frac{V_{CC} + N(R_C/R_B)V_{BE(sat)}}{1 + N(R_C/R_B)} \end{aligned} \quad (\text{A.38})$$

For large values of N , $V_{out} (= V_{OH})$ will eventually approach $V_{BE(sat)} \approx 0.8$ V, which means that the NM_H is reduced to zero (or is even negative)! For smaller values of N , the situation is not as bad, however. For instance, $N = 5$ results in an output voltage of $V_{OH} = 3.6$ V, as is confirmed by the simulations of Figure A.20.

The main problem of the RTL gate is, however, evident in its **transient behavior**. Consider, for instance, the t_{pLH} . The propagation delay can be decomposed into two elements:

1. Initially, Q_1 is in deep saturation. Turning off the device requires the removal of the base charge (through R_B). This is a slow operation.
2. As in the CMOS case, a major part of the switching time consists of charging up the load capacitance C_L from V_{OL} to V_{OH} . In the RTL case, this has to happen through the load resistance R_C , and the delay is determined by the time-constant $R_C C_L$. Making R_C small increases $I_{C(sat)} [(V_{CC} - V_{CE(sat)})/R_C]$. This raises both the power consumption as well as the first component of the propagation delay. Indeed, the charge stored in the base region is proportional to the collector current ($I_C = Q_F/\tau_F$, as discussed in Chapter 2).

Similar arguments can be made regarding t_{pHL} . The situation is not as bad in this case, as C_L is discharged by Q_1 instead of through a resistor. This is considerably faster. The dominant part of the propagation delay is the buildup of the base charge to bring Q_1 into saturation. Observe that avoiding saturation is not an option, since it results in a substantial degradation of the dc behavior.

Finally, note that the gate consumes static power when the output is low, because a direct path between V_{CC} and GND exists through the resistor and transistor.

Example A.6 Transient Response of the RTL Inverter

The transient response of the RTL inverter of the previous example (for a fan-out of 1) is simulated using SPICE and is shown in Figure A.22 (a manual analysis of the transient behavior of a bipolar circuit is complex and is, therefore, delayed until later in the chapter). The response is clearly asymmetrical: $t_{pLH} = 420$ psec, and $t_{pHL} = 165$ psec. The resulting propagation delay is equal to 292.5 psec, which is inferior to the performance of the CMOS inverter!

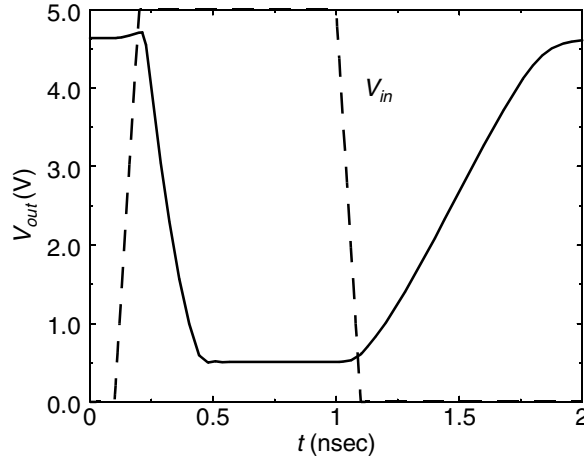


Figure A.22 Transient response of an RTL inverter (for a fan-out of 1).

An important part of this delay is due to the base-charge buildup and removal. A fast bipolar gate should avoid having its transistors going into saturation, since this is where the major base-charge buildup happens.

In summary, a more effective bipolar gate should address the dc and ac issues raised above. In the next section, we discuss how this is accomplished in the ECL gate.

A.3.2 The Emitter-Coupled Logic (ECL) Gate at a Glance

The ECL gate differs fundamentally from both CMOS and RTL inverters. It has been conceived with *utmost performance* in mind. One of the means to accomplish this is to keep the logic swing low. The ECL gate, hence, operates typically with a swing of only 0.5 V (compare this to the 3 to 5 V CMOS structures). Ensuring a reasonable noise margin under those conditions is nontrivial and requires careful circuit design.

The resulting gate is rather complex and is a composition of three components: *the differential pair, the output driver, and the bias network*. Each of these affects one particular aspect of the gate functionality and performance.

The Differential Pair (or Current Switch)

The function of the input stage of the ECL gate is to provide maximal noise margins, low noise sensitivity, and fast switching in the presence of a *small logic swing*. A schematic of this portion of an ECL gate is plotted in Figure A.23. This structure is well known in the analog world, where it is called an *emitter-coupled pair*. One of the two inputs is connected to a reference voltage V_{ref} that is generated by the bias network (discussed later). Assuming that both transistors Q_1 and Q_2 are identical, the circuit is biased in a completely symmetrical fashion if V_{in} is set equal to V_{ref} as typically would be the case in an analog design: $V_{be1} = V_{be2}$, $I_{C1} = I_{C2} = I_{EE}/2$ and $V_{out1} = V_{out2} = V_{CC} - R_C I_{EE}/2$. This bias condition

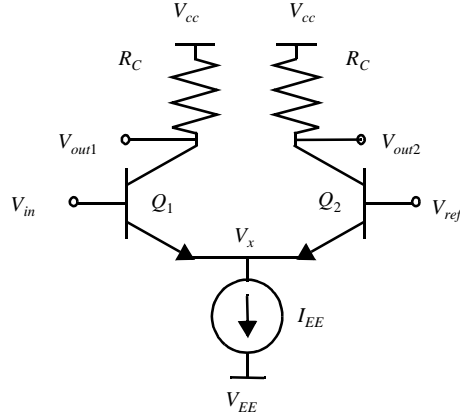


Figure A.23 The differential pair.

turns out to be rather useless for digital purposes, where a key requirement is that two easily distinguishable operation modes are produced.

Let us now increase (decrease) V_{in} by a small amount with respect to V_{ref} . This changes the current balance in the differential pair and increases (decreases) the fraction of I_{EE} that is routed through Q_1 . A critical observation is that this current fraction is an *exponential function* of the voltage difference between V_{in} and V_{ref} . This is expressed in Eq. (A.39), which is based on the familiar collector-current equation of the bipolar transistor (2.63) (with ϕ_T equal to kT/q , the thermal voltage).

$$\frac{I_{C1}}{I_{C2}} = \frac{e^{\frac{V_{in} - V_x}{\phi_T}} - 1}{e^{\frac{V_{ref} - V_x}{\phi_T}} - 1} \approx e^{\frac{V_{in} - V_{ref}}{\phi_T}} \quad (\text{A.39})$$

At room temperature, a **60 mV increase in V_{in} causes a tenfold increase of I_{C1}/I_{C2}** , while a 120 mV increase is sufficient to set I_{C1} to 100 I_{C2} . With I_{C2} only 1% of I_{C1} , it can be stated that, for all practical purposes, Q_2 is essentially off, and that Q_1 is carrying all of I_{EE} . As a result, approximate values for V_{out1} and V_{out2} can be derived.

$$\begin{aligned} V_{out1} &= V_{CC} - I_{EE}R_C \quad (\text{logic low}) \\ V_{out2} &= V_{CC} \quad (\text{logic high}) \end{aligned} \quad (\text{A.40})$$

A similar argument can be used to analyze the case where V_{in} drops below V_{ref} . Here, all the current is diverted to Q_2 , resulting in a high V_{out1} and a low V_{out2} . A number of conclusions can be drawn from this simple analysis:

1. The ECL gate provides *differential (or complementary) outputs*. Both the output signal (V_{out1}) and its inverted value (V_{out2}) are simultaneously available. This is a distinct advantage, as it eliminates the need for an extra inverter to produce the complementary signal. This also prevents some of the time-differential problems introduced by additional inverters. For example, in logic design it often happens that both a signal and its complement are needed simultaneously. When the complementary signal is generated using an inverter, the inverted signal is delayed with respect

to the original (Figure A.24a). This causes timing problems, especially in very high-speed designs. The differential output capability of ECL avoids this problem (Figure A.24b).

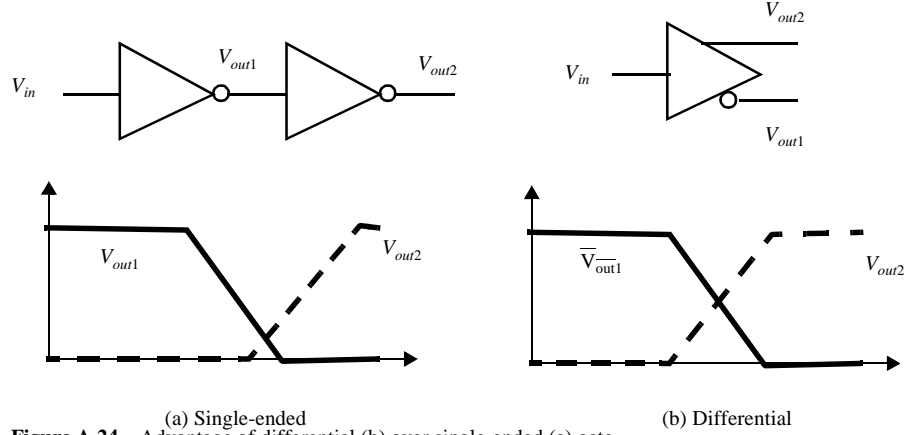


Figure A.24 Advantage of differential (b) over single-ended (a) gate.

2. Another advantage of the differential operation is an *increased noise immunity*, as common mode noise signals—i.e., signal disturbances common to both V_{in} and V_{ref} —are rejected to a large degree.
3. The *transition region* from one logic state to the other is *centered around V_{ref}* and is, to the first order, independent of the transistor parameters. To determine the boundaries of the transition region, an often-used simplification is to place them at the points where the current ratio drops to 1%. This is only a qualitative approximation, which does not strictly follow the definitions of V_{IH} and V_{IL} (gain = -1). A more detailed analysis is needed to produce precise values and will follow later. Using the simplified definition, the transition width evaluates to approximately 240 mV.

$$\begin{aligned} V_{IL} &= V_{ref} - 120\text{mV} \\ V_{IH} &= V_{ref} + 120\text{mV} \end{aligned} \quad (\text{A.41})$$

In the case of the static CMOS and the RTL gates, the logic swing approximates the voltage difference between the supply rails. This is not so for the ECL gate, where the difference between V_{OH} and V_{OL} is determined by the bias current I_{EE} and the load resistance R_C , as was illustrated in (A.40). The reduced voltage swing results in a faster operation, as the time to (dis)charge the load capacitor is reduced.

4. The *current drawn from the supply is always constant* and equal to I_{EE} . Therefore, the differential pair consumes a constant amount of static power equal to $I_{EE}(V_{CC} - V_{EE})$. This static power consumption precludes the use of ECL for the design of very complex VLSI circuits (> 100,000 gates), as the removal of the excessive heat becomes either impossible or excessively expensive. Therefore, the use of ECL-

style circuits is mostly confined to the design of very high-speed building blocks such as CPUs for mainframe computers.

On the other hand, the fact that the current drawn from the supply is constant and independent of the switching operation reduces the amount of switching noise on the supply lines. This is further aided by the low logic swing, which reduces the current spikes caused by charging and discharging the load capacitance. As a result, ECL circuits can operate reliably, even with small noise margins.

Furthermore, it can be observed that the output voltages (V_{out1} and V_{out2}) are defined with respect to V_{CC} and are, to a first degree, not influenced by the value of V_{EE} . Noise on the latter supply therefore has virtually no effect on the signal levels. On the other hand, V_{CC} has to be as clean as possible.

5. With a proper choice of the bias current, the input voltage levels, and the resistor values, it can be assured that *neither Q_1 nor Q_2 ever goes into saturation*. Actually, in normal operation, both transistors are always forward-active. This speeds up the device operation, as the time to alter the base charge is drastically reduced.

The Output Driver

The differential pair can be used in a stand-alone configuration. While it is very fast, this structure suffers from one of the problems that plagued the RTL inverter. Its VTC is very sensitive to *output loading*, since in the high state all current has to be provided by R_C . To alleviate the problem, an output driver is inserted between the outputs of the differential pair and the input of the next gate. The main task of this driver is to *reduce the output impedance* of the gate, so that ample current drive is available. At the same time, it acts as a *level shifter*. It aligns the output logic levels symmetrically around V_{ref} , making them compatible with the input logic levels of a fan-out gate. The emitter-follower circuit module performs all these functions and is, therefore, the ideal candidate for the output driver.

The revised circuit is shown in Figure A.25. It causes a level shift on the output signal equal to $V_{BE(on)}$. It also reduces the output impedance for the high logic level from R_C to approximately $R_C / (\beta_F + 1)$, thus providing ample current-drive capability (with β_F the forward-current gain of the bipolar device). To complete the circuit, a resistor R_B has been added at the input terminal. R_B serves as a current path to V_{EE} from the base of Q_1 or Q_2 . This helps to speed up the removal of the base charge Q_F during the switching of the differential pair. For a very high speed design, R_B also acts as termination for the transmission line formed by the interconnect. We will discuss this in depth in Chapter 8.

Design Consideration

The differential pair draws a constant current from the supply. This is obviously not true for the emitter-followers, whose current levels are substantially different between the high and the low output states. Large current peaks occur during the switching events, while (dis)charging the load capacitances. This induces large fluctuations in the supply voltages and can cause mal-

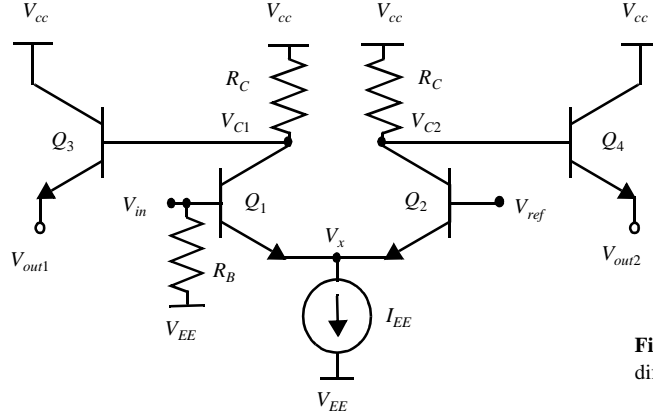


Figure A.25 ECL gate consisting of differential pair and emitter-followers.

functioning of the circuit, given the small noise margins. To avoid this cross-coupling, a common practice is to provide a *separate* V_{CC} supply for the output drivers.



The Bias Network

The purpose of the bias is to generate the reference voltage V_{ref} . The prime requirement for this reference is that it is to be centered as much as possible in the middle of the logic swing, independent of the operating temperature. As was mentioned in Chapter 2, the voltage drop over a forward-biased base-emitter junction is a strong function of temperature. Typically, V_{BE} changes linearly with the temperature, or $\Delta V_{BE} = -kDT$, with k approximately equal to 1.5 mV/°C. This drift causes the voltage-transfer characteristic of the ECL gate and its noise margins to vary as a function of temperature as well. The ambient temperature in ECL devices can be quite high, due to the large power dissipation of the gate. Therefore, the bias network has to be designed so that V_{ref} tracks those changes and the noise margins are kept symmetrical over a wide range of temperatures. An example of such a network is shown in Figure A.26. The value of the reference voltage is easily derived.

$$V_{ref} = V_{B5} - V_{BE(on)}$$

$$V_{B5} = V_{CC} - \frac{R_1}{R_1 + R_2}(V_{CC} - 2V_D - V_{EE}) \quad (\text{A.42})$$

Transistor Q_5 is connected in the emitter-follower configuration, reducing the output impedance of the reference network. Changes in temperature affect the $V_{BE(on)}$ of Q_5 and Q_2 . These changes are tracked to a first order by the change in the voltage drops across diodes D_1 and D_2 . Actual implementations of this circuit have shown that it ensures the centering of V_{ref} between V_{OH} and V_{OL} for a temperature range from -30°C to $+85^\circ\text{C}$.

In advanced ECL configurations, the bias network is also used to generate other reference voltages, such as the bias voltage for the current source. Fortunately, a single bias

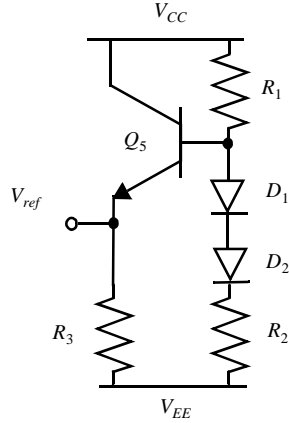


Figure A.26 Voltage reference circuit (as used in the ECL 10K series from Motorola)

network can serve multiple gates, reducing the overall power consumption and the area overhead.

After this qualitative analysis of the structure and properties of the basic ECL gate, we will now derive the quantitative dc and ac characteristics of the gate. Before proceeding, we would like to mention that there exist many different ECL families. The most famous are the ECL 10K and 100K series and the MECL I, II, and III circuits [Hodges88]. The major differences between these are the values of the resistors, the type of current source used and the construction of the bias network. Other variants of the gate are used in modern, integrated ECL circuits used in instrumentation circuits, CPUs, and even microprocessors (see, for example, [Jouppi93]). Instead of studying all those versions separately, we discuss only one single gate structure in detail. Similar approaches can then be employed to analyze the characteristics of related structures or even to design custom ECL gates, whose current levels are optimized for the required performance or power consumption.

A.3.3 Robustness and Noise Immunity: The Steady-State Characteristics

This section will analyze, in detail, how a typical ECL gate structure, as shown in Figure A.27, can operate with a low voltage swing, and yet maintain sufficient noise immunity. The gate topology is similar to the one presented before, with the exception of the current source. The latter is implemented as a resistance R_S in series with a transistor Q_6 , whose base is dc-biased at V_{CS} . To a first order, this produces the following bias current,

$$I_{EE} = \frac{(V_{CS} - V_{BE(on)} - V_{EE})}{R_S} \quad (\text{A.43})$$

For a manual analysis, we may assume that I_{EE} is constant over the range of input voltages. We will derive the VTC assuming a fan-out of 1, or, in other words, with the output stage loaded with the resistance R_B .

Suppose first of all that all transistors are operating in the forward-active region. Under those conditions, the following circuit equations are valid:

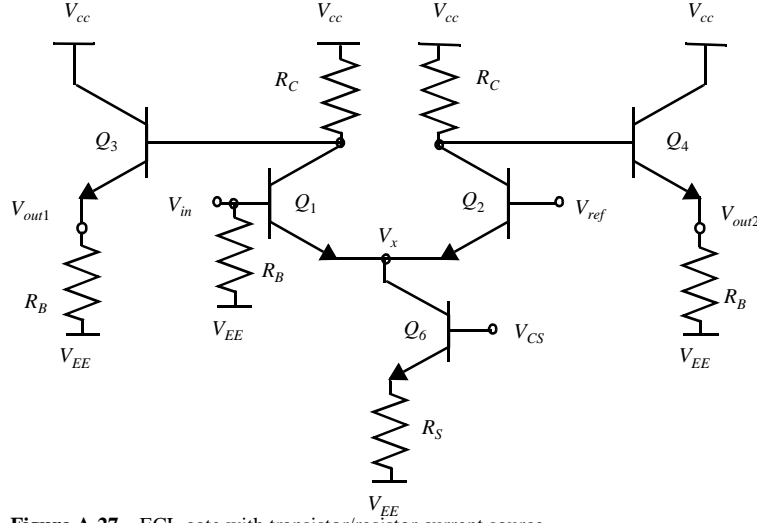


Figure A.27 ECL gate with transistor/resistor current source.

$$\frac{I_{C1}}{I_{C2}} = e^{\frac{V_{in} - V_{ref}}{\phi_T}}$$

$$I_{C1} + I_{C2} = I_{EE} \frac{\beta_F}{\beta_F + 1} \approx I_{EE}$$

$$V_{C1} = V_{CC} - R_C(I_{C1} + I_{B3})$$

$$V_{out1} = V_{C1} - V_{BE(on)}$$

$$I_{B3} = \frac{V_{C1} - V_{EE}}{(\beta_F + 1)R_B}$$
(A.44)

These equations can be solved to yield the following equations for I_{C1} and V_{out1} (similar equations can be derived for I_{C2} and V_{out2}).

$$I_{C1} = \frac{I_{EE} e^x}{1 + e^x}$$
(A.45)

with $x = (V_{in} - V_{ref})/\phi_T$, and

$$V_{out1} = \frac{(\beta_F + 1)R_B}{(\beta_F + 1)R_B + R_C} \left[V_{CC} - V_{BE(on)} - R_C I_{C1} + \frac{V_{EE} R_C}{(\beta_F + 1)R_B} \right]$$

$$\approx V_{CC} - V_{BE(on)} - R_C I_{C1} \quad (\text{for } R_B(\beta_F + 1) \gg R_C)$$
(A.46)

The important parameters of the voltage-transfer characteristic can now be derived.

Nominal Voltage Levels

For $V_{in} \gg V_{ref}$, $I_{C1} = I_{EE}$, and $V_{OL} \approx V_{CC} - V_{BE(on)} - R_C I_{EE}$. For $V_{in} \ll V_{ref}$, $I_{C1} \approx 0$, and $V_{OH} \approx V_{CC} - V_{BE(on)} - V_b$. V_b is the voltage drop over R_C , caused by the base current of Q_3 ($= R_C I_{B3}$). This factor can normally be ignored for manual analysis. The logic swing of the gate hence equals

$$V_{swing} = R_C I_{EE} \quad (\text{A.47})$$

Noise Margins

Deriving exact expressions for V_{IH} and V_{IL} is somewhat more involved, due to the heavily nonlinear current relations. As mentioned earlier, one common approach to circumvent this analysis is to use revised definitions of V_{IH} and V_{IL} . Specifically, V_{IL} is defined as the point where transistor Q_1 carries 1% of I_{EE} , while at V_{IH} Q_1 carries 99% of the total current. The corresponding values of V_{in} can be derived from Eq. (A.45):

$$\begin{aligned} \frac{I_{C1}}{I_{EE}} &= \frac{e^x}{1 + e^x} = \alpha = 0.01 \\ V_{IL, IH} &= V_{ref} \pm \phi_T \ln\left(\frac{\alpha}{1 - \alpha}\right) \end{aligned} \quad (\text{A.48})$$

At room temperature ($\phi_T = 26$ mV), we find that $V_{IL} = V_{ref} - 120$ mV and $V_{IH} = V_{ref} + 120$ mV. This results in a very narrow transition region of 240 mV.

The results ensuing from this approach tend to differ considerably from those obtained using the traditional unit gain definition. Consequently, it is worthwhile to derive more accurate expressions. In correspondence to earlier derivations, we use the small-signal approach. A number of simplifications have to be considered to make the analysis tractable.

- Combining the small-signal models of the differential pair and the emitter-follower yields rather intractable equations. The model can be simplified in an important way by observing that the input impedance of the emitter-follower is approximately equal to its load impedance, multiplied by $(\beta_F + 1)$. This technique, illustrated in Figure A.28, is called the *resistance reflection rule*² and will be quite useful for the computation of the propagation delay of the ECL gate later in the chapter. Furthermore, the gain of the emitter-follower is approximately equal to 1 and will have virtually no effect on the overall gain of the gate. Using this approach, we can eliminate the emitter-follower from the small-signal model and replace it by a resistor $R'_B = R_B(\beta_F + 1)$ in parallel with R_C .
- We assume that, at V_{IH} and V_{IL} , the differential amplifier is sufficiently balanced so that the small-signal gain of the amplifier approximates the gain of the amplifier under symmetrical biasing conditions. Although the actual situation deviates substantially from this assumption (at V_{IH} and V_{IL} , an important imbalance exists

² For an in-depth derivation of this rule, please refer to textbooks on analog circuit design, such as *Micro-electronic Circuits* by Sedra and Smith [Sedra87, p. 457].

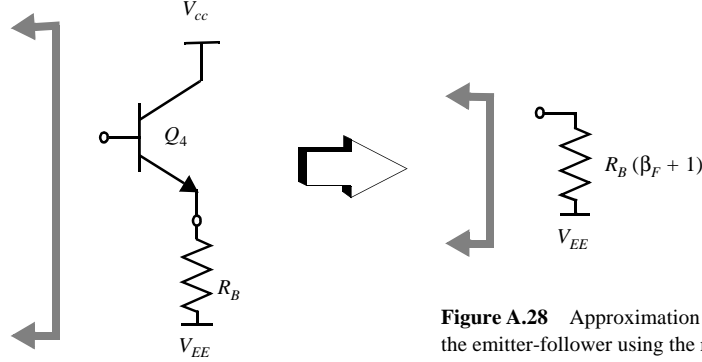


Figure A.28 Approximation of the loading effect of the emitter-follower using the reflection rule.

between I_{C1} and I_{C2}), it produces a reasonable approximation useful for manual analysis.

Under those assumptions the gain of the amplifier equals $g = -g_m R_L/2$, with R_L the load resistance of the amplifier ($= R_C \parallel R_B(\beta_F + 1)$) and g_m the transconductance of Q_1 ($= I_{C1}/V_T$).³ The values of V_{IH} and V_{IL} can now be found by setting g to -1 and plugging in the appropriate value of I_{C1} (Eq. (A.45)).

$$V_{IH,IL} = V_{ref} \pm V_T \ln \left(\frac{I_{EE}(R_C \parallel R_B(\beta_F + 1))}{2\phi_T} - 1 \right) \approx V_{ref} \pm \phi_T \ln \left(\frac{V_{swing}}{2\phi_T} - 1 \right) \quad (\text{A.49})$$

As expected, V_{IH} and V_{IL} differ only a couple of ϕ_T ($= 26$ mV at room temperature) from V_{ref} (for V_{swing} varying between 250 mV and 1V). Their values depend on the value of the logic swing $I_{EE}R_C$ in a logarithmic fashion, or, in other words, they are rather insensitive to voltage variations. It is worth observing that a more precise analysis, which does not employ the second assumption, yields only a slightly different equation:

$$V_{IH,IL} = V_{ref} \pm \phi_T \operatorname{acosh} \left(\frac{V_{swing}}{2\phi_T} - 1 \right) \quad (\text{A.50})$$

Switching Threshold

Finally, V_M can be determined by setting V_{in} equal to V_{out} in Eq. (A.46). This leads, once again, to a hopelessly complex expression that is difficult to solve analytically. Some interesting information can be obtained by realizing that the preferred value of V_M equals V_{ref} . Under this circumstance, I_{EE} is split equally over both branches of the current switch, and a dc bias condition for the gate can be derived:

$$V_{ref} = V_M \approx V_{CC} - V_{BE(on)} - \frac{R_C I_{EE}}{2} \quad (\text{A.51})$$

³ For more information on the small-signal model of the differential pair, please refer to standard analog design handbooks, such as [Sedra87, pp. 494–496].

Other dc bias conditions can be derived that help to define the values of the resistors and the current source. For example, R_C and I_{EE} have to be chosen so that transistors Q_1 and Q_2 are biased in the forward-active region over the complete range of input voltages.

The results of the above analysis are summarized in Figure A.29, where an asymptotic picture of the ECL voltage transfer characteristic is drawn. Notice that for large values of the input signal V_{in} transistor Q_1 eventually saturates, causing output V_{out1} to track V_{in} in a linear fashion (with $V_{out1} = V_{in} - V_{BE(sat)} + V_{CE(sat)}$). This operation mode is of no concern as it is out of the normal input voltage range.

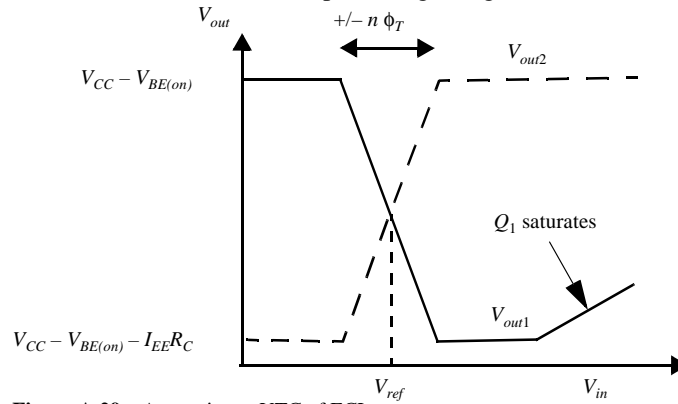


Figure A.29 Approximate VTC of ECL gate.

Example A.7 VTC of ECL Gate

An ECL gate was designed with the following parameters: $V_{CC} = 0$ V, $V_{EE} = -5$ V, $V_{ref} = -0.95$ V, $I_{EE} = 0.5$ mA, $R_C = 1$ k Ω , and $R_B = 50$ k Ω . The β_F of the bipolar devices equals 100, while I_S is set to 10^{-17} A. The bipolar transistors are fabricated using the bipolar process described in Chapter 2. All transistors are minimum-size devices with $A_{emitter} = 2.0$ $\mu\text{m} \times 3.75$ μm . The VTC of the device is computed using both manual analysis and SPICE.

First, the exact value of the $V_{BE(on)}$ of the emitter-follower transistors has to be derived. This voltage varies with the collector current. A reasonable approximation is obtained by setting I_C to $(5 \text{ V} - 0.7 \text{ V}) / 50 \text{ k}\Omega \approx 0.1$ mA, which equals the collector current of the emitter-follower in the high output state. This results in the following value of V_{BE} (for $\phi_T = 26$ mV):

$$V_{BE(on)} = \phi_T \ln\left(\frac{I_C}{I_S}\right) = 0.778 \text{ V}.$$

The important parameters of the voltage-transfer characteristic can now be computed.

$$\begin{aligned}
V_{OH} &= V_{CC} - V_{BE(on)} = -0.78 \text{ V} \\
V_{OL} &= V_{CC} - V_{BE(on)} - I_{EE} R_C = -1.28 \text{ V} \\
V_{IH} &= -0.95 + 0.026 \ln \left(\frac{0.5}{2 \times 0.026} - 1 \right) = -0.89 \text{ V} \\
V_{IL} &= -0.95 - 0.026 \ln \left(\frac{0.5}{2 \times 0.026} - 1 \right) = -1.01 \text{ V} \\
V_M &= V_{ref} = -0.95 \text{ V}
\end{aligned} \tag{A.52}$$

The VTC has also been computed using SPICE and is plotted in Figure A.30. The important parameters have been extracted and are compared with the manual results in Table A.4. An excellent correspondence is observed[.

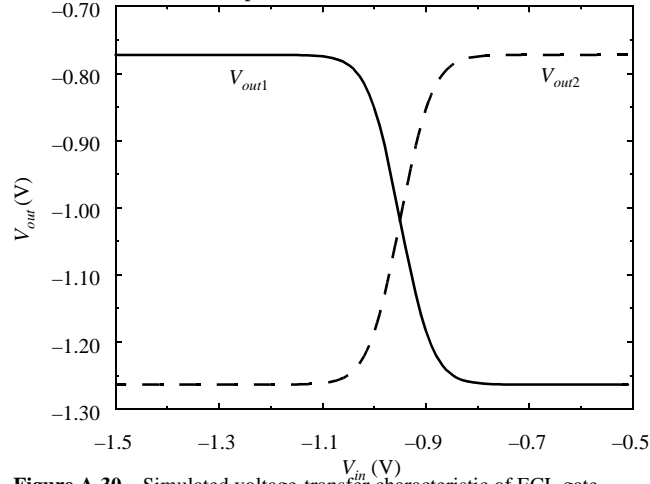


Figure A.30 Simulated voltage-transfer characteristic of ECL gate.

Table A.4 DC-parameters of ECL inverter (manual versus computed).

	Manual	SPICE
V_{OH}	-0.78	-0.77
V_{OL}	-1.28	-1.26
V_{IH}	-0.89	-0.87
V_{IL}	-1.01	-1.03
V_M	-0.95	-0.96
NM_L	0.27	0.23
NM_H	0.11	0.1

Design Considerations

Compared to the CMOS inverter, the small logic swing (0.49 V) and noise margins are apparent. The smaller swing has the advantage of increased switching speed. The smaller noise margins can be tolerated because the logic levels are robust and not particularly dependent on *process parameters*. For instance, the V_{BE} of a bipolar process is far more predictable than the threshold of an MOS transistor. The reduced voltage swing also helps to keep the noise levels down. In fast ECL circuits, however, with many simultaneous switching actions, it is essential to keep the noise within bounds.

At this point, it is worth observing that all voltage levels (such as V_{OH} and V_{OL}) in an ECL gate are defined with respect to V_{CC} and are only affected in a minor way by V_{EE} . It makes sense to assign V_{CC} the best available fixed potential, which is generally the ground potential of the power supply, which is, in general, the “most clean” supply rail. Therefore, ECL circuits will normally set V_{CC} to 0 V. This requires that V_{EE} be at a negative potential, which explains the choice of the supply voltages in the example ($V_{CC} = 0$ V, $V_{EE} = -5$ V). The standard value used for the 10k and 100k series was -5.2 V. Current ECL circuits operate at lower supply voltages ranging from -3 V to -4.5 V.

Finally, it is worth mentioning that the dc characteristics of the above gate are virtually unaffected by fan-out. This is due to the very low output impedance of the emitter-follower.



A.3.4 ECL Switching Speed: The Transient Behavior

Studying the propagation delay of the ECL gate is a considerably more involved task than the analysis of the CMOS inverter. The structure is inherently more complex, as it contains internal nodes, each of which could dominate the performance. This is illustrated in Figure A.31, which shows the equivalent circuit to be considered when computing the propagation delay between the input and the inverting output V_{out1} for a single fan-out. Getting an accurate prediction of the response is only possible through SPICE simulation. Unfortunately, this does not give an insight into the mechanisms governing the transitions. Such an insight can only be obtained by studying a tractable circuit model, obtained through significant simplifications:

- All capacitances are *linearized* over the appropriate voltage range and *lumped* together into a limited number of capacitances.
- Internal nodes are *eliminated* (as much as possible) using the *reflection rule* (introduced in the dc analysis).
- The transient response is *decomposed into several steps*, based on the dominant effects governing the response at each step. Although in reality **these steps do partially overlap**, this approach yields reasonable accuracy, while simplifying the analysis substantially. In particular, we isolate the switching of the differential pair from the (dis)charging of the load capacitors. This important assumption dramatically

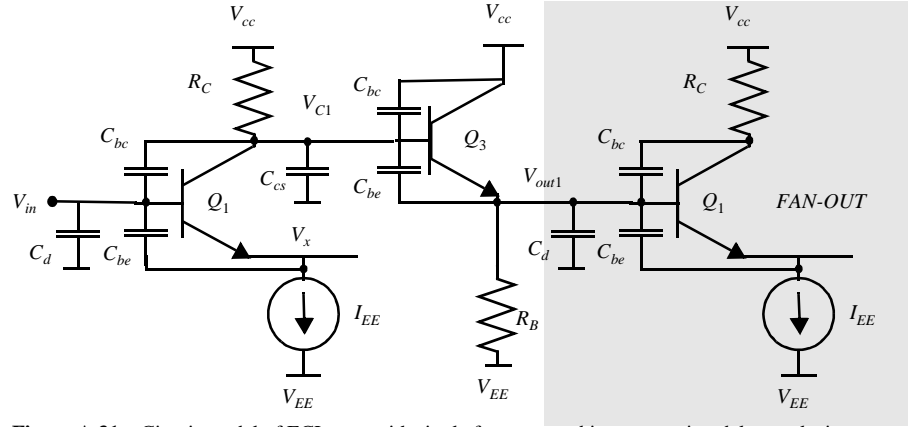


Figure A.31 Circuit model of ECL gate with single fan-out, used in propagation delay analysis.

simplifies the analysis, while enabling the identification of the dominant performance parameters.

Switching the Differential Pair

The analysis of the large-signal transient response of the current switch is intricate and yields complex expressions.⁴ Complicating the derivation is the impact of the base and emitter resistances, the feedback effect offered by the coupled emitters, and the coupling between the transistor terminals. To avoid this multiple-page numerical art-work, we use a simple heuristic model. In Chapter 2, it was derived that the transient response of a bipolar transistor is dominated by the base and depletion-region charge. The same is true for the differential pair. To illustrate this statement, let us study the simplified network of Figure A.32. Some of the most important transient parameters are annotated on the diagram. Notice that in this model, the collectors are kept at V_{CC} , which effectively eliminates the impact of the discharging of the load capacitances.

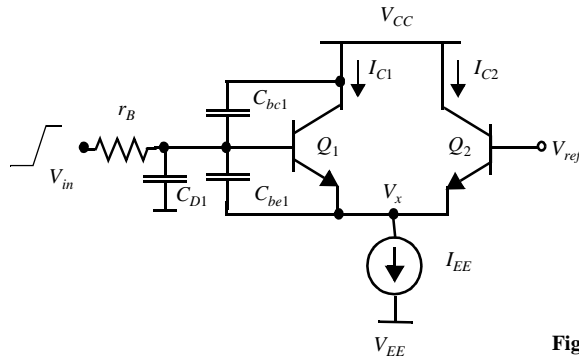


Figure A.32 Switching the differential pair.

⁴ For a detailed analysis, please refer to [Embabi93, pp.215–224]

The transient response of the current switch is simulated for varying values of the current source I_{EE} . The resulting collector currents for both branches are plotted in Figure A.33. The switching time is a strong function of I_{EE} for large values of the current source, which suggests that the base charge is the dominant factor under those circumstances. On the other hand, the delay is independent of I_{EE} for smaller current values. This means that the depletion charge of the be and bc junctions has become the most important factor. Varying the value of the junction capacitances in the simulation confirms this assumption. Observe the large spikes in the collector currents during the switching of the input signal. This is due to the capacitive coupling between base and emitter of Q_1 . A simple model can now be constructed for both mechanisms.

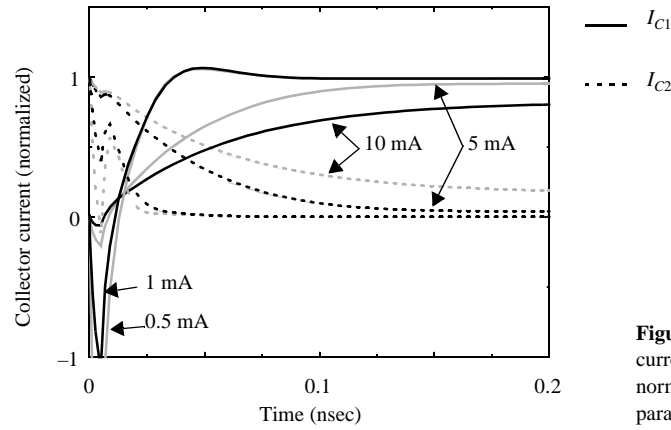


Figure A.33 Simulated collector currents of differential pair (final value normalized to 1), with I_{EE} as a parameter.

(Dis)charging the Depletion Capacitances (space charge). Based on our common practice, we first replace all nonlinear junctions capacitors by linearized equivalents. This is accomplished by multiplying the zero-bias capacitance with the K_{eq} factor (Eq. (2.18)), which is a function of the applied voltage range. The voltage at the base of Q_1 swings between V_{OL} to V_{OH} , while the emitter voltage moves between $V_{ref} - V_{BE(on)}$ to $V_{OH} - V_{BE(on)}$. The collector voltage stays constant at V_{CC} . Knowledge of these voltage ranges allows for the computation of the K_{eq} values:

$$\begin{aligned} K_{eq, be} &= K_{eq}(V_1, V_2) = K_{eq}(V_{OL} - V_{ref} + V_{BE(on)}, V_{BE(on)}) \\ K_{eq, bc} &= K_{eq}(V_1, V_2) = K_{eq}(V_{OL} - V_{CC}, V_{OH} - V_{CC}) \end{aligned} \quad (A.53)$$

The lumping of the junction capacitances into a single component requires some extra caution. The C_{bc1} undergoes a full V_{swing} during the transient period. The total voltage swing over C_{be1} , on the other hand, is limited to $V_{swing}/2$ (assuming that V_{ref} is placed in the middle of the voltage swing), as the emitter tracks the base once the transistor is on. Its impact is thus half of the other capacitor (an inverse Miller effect). Taking this into account, the following expression for the equivalent input capacitance can be derived over the V_{swing} voltage range.

$$C_{in,j} = K_{eq,bc} C_{bc1} + K_{eq,be} C_{be1} / 2 \quad (A.54)$$

The circuit diagram of Figure A.32 suggests that the (dis)charging of the depletion capacitances is dominated by the $r_B C_{in,j}$ time-constant. Achieving 90% of the final value requires approximately 2.2 time-constants, if the voltage waveform at the base is exponential (which is approximately the case). Under these conditions, the time to change the space charge is adequately modeled by the following expression:

$$t_{space} = 2.2 r_B C_{in,j} = 2.2 r_B (K_{eq,bc} C_{bc1} + K_{eq,be} C_{be1} / 2) \quad (A.55)$$

(Dis)charging the Diffusion Capacitance (base charge). The diffusion capacitance C_{D1} represents the excess-charge storage in the transistor base (introduced in Chapter 2). C_{D1} is placed at the base of the transistor and stores an amount of charge equal to ΔQ_F over the voltage range of interest. From Chapter 2, recall the following expression,

$$C_D = \frac{\Delta Q_F}{\Delta V} \quad (A.56)$$

During a single transition, the current through Q_1 evolves from virtually nonexistent to I_{EE} or $\Delta Q_F = Q_F = t_F I_{EE}$. The voltage swing at the base during that period is, obviously, V_{swing} . Hence,

$$C_{D1} = \frac{\Delta Q_F}{\Delta V} = \frac{\tau_F I_{EE}}{V_{swing}} \quad (A.57)$$

It is worth observing that the diffusion capacitance is proportional to the value of the current source. This proportionality is clearly manifested in the simulation results of Figure A.33. Remembering that the voltage swing can be expressed as a function of I_{EE} as well (Eq. (A.47)) yields another interesting relation. Eq. (A.58) states that the diffusion capacitance of Q_1 is only a function of the forward transit time τ_F and the collector resistance R_C (a design parameter!).

$$C_{D1} = \frac{\tau_F I_{EE}}{R_C I_{EE}} = \frac{\tau_F}{R_C} \quad (A.58)$$

Deriving an exact expression of the time to change the base charge is difficult, especially due to the impact of the emitter-coupling (and other factors, such as emitter resistance). The time is proportional, though, to the time-constant of the input circuit, formed by r_B and C_{D1} . This suggests the following model for t_{base} :

$$t_{base} = \alpha r_B C_{D1} \quad (A.59)$$

where α is an empirical factor, and depends upon circuit and device parameters. From simulations, we derive that α approximately equals 2 and 5 for the 50% and 90% points, respectively.

Propagation Delay of the Differential Pair. The overall switching time of the differential pair can now be expressed as the sum of the space- and base-charge components (assuming that both mechanisms happen consecutively). Remember from the simulations that the response is typically dominated by a single mechanism.

$$t_{dp} = t_{space} + t_{base} = r_B (2.2 C_{in,j} + \alpha C_{D1}) \quad (A.60)$$

Example A.8 Switching the Differential Pair

Consider the ECL gate of Example A.7. The bipolar transistor model defined in Chapter 2 provides the extra parameter values needed for the transient analysis:

$$r_B = 120 \, \Omega, r_E = 20 \, \Omega, \text{ and } r_C = 75 \, \Omega; C_{be} = 20 \, \text{fF}, C_{bc} = 22 \, \text{fF}, \text{ and } C_{cs} = 47 \, \text{fF}.$$

$$m_{be} = 0.5, m_{bc} = m_{cs} = 0.33, \tau_F = 10 \, \text{psec}.$$

With $I_{EE} = 0.5 \, \text{mA}$ (or $R_c = 1 \, \text{k}\Omega$) and $\tau_F = 10 \, \text{psec}$, C_{D1} equals $10 \, \text{fF}$. C_{be} and C_{bc} are defined to be 20 and $22 \, \text{fF}$, while their K_{eq} -factors evaluate to 3.35 and 0.75 , respectively (with $\phi_0 = 0.7 \, \text{V}$).⁵ This yields the following capacitance values:

$$C_{D1} = 10 \, \text{psec} / 1 \, \text{k}\Omega = 10 \, \text{fF}; C_{in,j} = 3.35 \times 20 / 2 + 0.75 \times 22 = 50 \, \text{fF}$$

The time it takes for the collector currents to reach 90% of their value is now approximated using Eq. (A.60) (and assuming that $\alpha = 5$).

$$t_{dp} = 120 \, \Omega \times (2.2 \times 50 \, \text{fF} + 5 \times 10 \, \text{fF}) = 20 \, \text{psec}$$

(Dis)Charging the Load Capacitances

To derive the second component of the transient response, we assume that the current switch has already reverted state. This means that the current-switch transistors are either off, or carry the complete current I_{EE} . Before addressing the analysis of the propagation delay, it is good practice to derive the capacitance values first, as we did for the MOS inverter.

Deriving the Load Capacitances. As illustrated in Figure A.31, the capacitance model of the ECL inverter is complex and consists of many contributions. To simplify the analysis, we gather all the individual contributions into two lumped capacitors: the collector capacitance C_C and the load capacitance C_L .

1. The collector capacitance $C_C = C_{cs1} + C_{bc1} + C_{bc3} + C_{D3}$.

The first three factors in C_C represent the depletion capacitances of the collector-substrate junction of Q_1 and the base-collector junctions of Q_1 and Q_3 . Since the emitter-follower transistor Q_3 is assumed to be always on and in forward-active mode, the voltage over the base-emitter junction is constant, and its depletion charge remains unchanged. This explains the absence of C_{be3} in the list. The equivalent junction capacitances are easily computed once the voltage excursions are known. The base of Q_1 is assumed to stay constant, while its collector voltage V_C swings between V_{CC} to $V_{CC} - V_{swing}$. The collector of Q_3 stays at V_{CC} . No Miller effect occurs, and the capacitances can simply be added, weighted with the appropriate K_{eq} factors.

C_{D3} models the change in base charge of Q_3 during a transition:

$$\Delta Q_F = \tau_F (V_{OH} - V_{OL}) / R_B = \tau_F V_{swing} / R_B$$

⁵ In the computation of K_{eq} for C_{be} , it is assumed that the maximum forward bias is equal to ϕ_0 .

where V_{OH}/R_B and V_{OL}/R_B approximate the collector currents of Q_3 in the high and low output state, respectively. The equivalent diffusion capacitance is derived by dividing ΔQ_F by the voltage swing of interest:

$$C_{D3} = \frac{\Delta Q_F}{\Delta V} = \frac{\tau_F V_{swing}}{R_B V_{swing}} = \frac{\tau_F}{R_B} \quad (\text{A.61})$$

Observe the similarity to Eq. (A.58).

2. The load capacitance $C_L = C_W + C_{in}$

The load of the gate equals the sum of the wiring capacitance C_W and the input capacitance of the fan-out gates. The latter is a complex combination of space- and base-charge capacitance, and varies with the direction of the transition. To simplify the analysis, we just assume a fixed, average value in this study.

Example A.9 Load Capacitances of an ECL Gate

Consider, again, the ECL gate of Example A.7. The values of the load capacitances are readily computed.

1. C_C —Due to the high value of R_B , C_{d3} evaluates to 0.2 fF and can be ignored. The K_{eq} factors for C_{bc1} , C_{bc3} , and C_{cs1} equal 0.84, 0.91, and 0.51, respectively (the low value for C_{CS} is due to the strong reverse bias on that junction, with the substrate at -5 V). This yields the following value for the collector capacitor.

$$C_C = 0.51 \times 47 + 0.84 \times 22 + 0.91 \times 22 + 0.2 \text{ fF} = 62.7 \text{ fF}$$

2. C_L —We assume a value of 60 fF here, which approximates the average input capacitance of a similar ECL gate.

Discharging the Load Capacitances. The time it takes to discharge the collector and output nodes is determined by two competing mechanisms: the discharging of the load capacitance C_L through the load resistor R_B , and the discharging of the collector capacitance by the current source I_{EE} . Both processes are illustrated in the (simplified) transistor network of Figure A.34. Depending upon the values of the resistances, capacitances, and current levels, either one of them can be dominant.

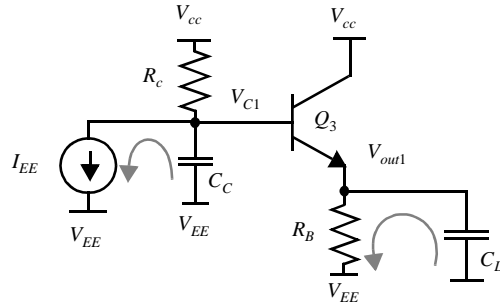


Figure A.34 Discharging the collector and load capacitances.

When $R_B C_L \ll R_C C_C$, the discharging of the collector node through I_{EE} dominates the performance, and the discharge time is proportional to the time-constant $R_C C_C$.

Observe that value of I_{EE} sets the dc voltage levels, but does not influence the discharge time. If this argument does not seem obvious, a useful exercise is to construct the Thevenin equivalent circuit of the network, consisting of I_{EE} , V_{CC} , and R_C .

A more accurate derivation takes the base current of Q_3 into account. Even for very small values of R_B , C_L , V_{out1} can never fall faster than V_{c1} , since this causes the emitter-follower to turn on. The nodes, therefore, discharge in unison. The collector and output nodes are tightly coupled as the output node follows the collector node with a (fixed) voltage difference of $V_{BE(on)}$. This operation is adequately modeled by the equivalent circuit of Figure

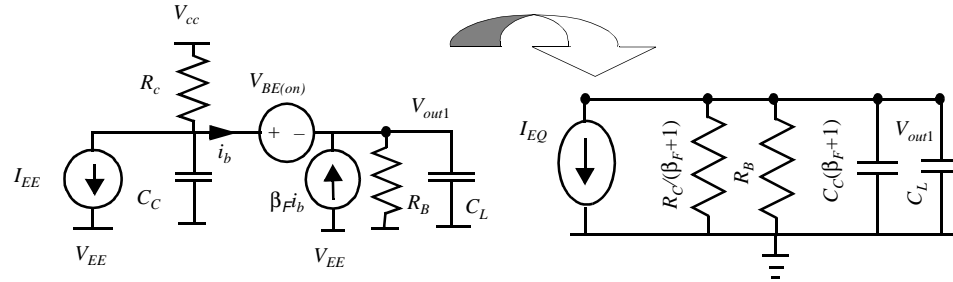


Figure A.35 (a) Simplified circuit Equivalent circuits for the computation of t_{pHL} , for $R_B C_L \ll R_C C_C$, with $I_{EQ} = I_{EE}(\beta_F + 1) - V_{EE}/R_B - (V_{CC} - V_{BE(on)})(\beta_F + 1)/R_C$.

A.35a (where the base-emitter diode is modeled as a voltage source), or by the even simpler Norton equivalent circuit model of Figure A.35b. The same circuit could have been derived in a single step by using the reflection rule to eliminate the internal node. Divide all impedances at the base of Q_3 by $(\beta_F + 1)$ and move them to the emitter node. Notice that this means that C_C is multiplied with $(\beta_F + 1)$. The discharge time is approximated by the time-constant of the resulting circuit:

$$t_{discharge1} = 0.69 \left(R_B \parallel \frac{R_C}{\beta_F + 1} \right) [C_L + C_C(\beta_F + 1)] \quad (A.62)$$

For large values of $R_C C_C$, this expression simplifies to $0.69 R_C C_C$, as was projected.

When $R_B C_L \gg R_C C_C$, the collector voltage drops faster than the output node, which means that the emitter-follower transistor Q_3 shuts off (until the output voltage gets clamped to its final value by the follower, as shown in Figure A.36). In this case, the propagation delay is dominated by the time it takes to discharge C_L through R_B . The effect of the clamping makes it necessary to use the average current approach to determine the delay, as the RC -technique used in Eq. (A.62) would predict a considerably different result.

$$I_{av} = \frac{1}{2} \left(\frac{V_{OH} - V_{EE}}{R_B} + \frac{V_{OH} - V_{swing}/2 - V_{EE}}{R_B} \right) = \frac{1}{R_B} \left(V_{OH} - V_{EE} - \frac{V_{swing}}{4} \right)$$

$$t_{discharge2} = \frac{C_L(V_{swing}/2)}{I_{av}} = \frac{0.5 C_L R_B V_{swing}}{V_{OH} - V_{EE} - \frac{V_{swing}}{4}} \approx 0.5 C_L R_B \left(\frac{V_{swing}}{V_{CC} - V_{EE}} \right) \quad (A.63)$$

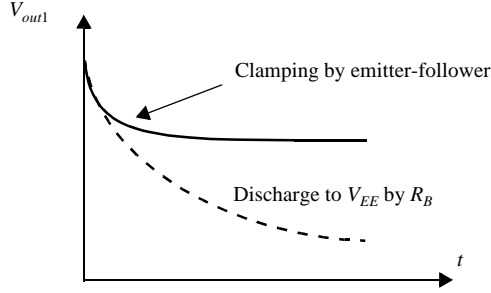


Figure A.36 Effect of voltage clamping on discharge waveform. When only the RC -network would be considered, the output would discharge to V_{EE} . This is prevented by the emitter-follower.

The actual discharge time is approximated by taking the worst case of both scenarios:

$$t_{discharge} \approx \max(t_{discharge1}, t_{discharge2}) \quad (A.64)$$

with $t_{discharge2}$ the time it takes to discharge the collector node and $t_{discharge1}$ the time to discharge the load capacitance through R_B .

Charging the Load Capacitances. A similar approach can be taken to estimate the time it takes to charge the load capacitances C_C (through R_C) and C_L (through Q_3). The process is illustrated in the circuit model of Figure A.37. To obtain accurate expressions for the

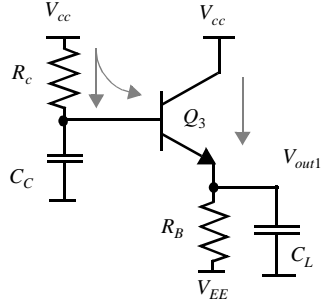


Figure A.37 Equivalent circuit for the computation of t_{pLH} .

delay, we resort again to the reflection rule and the use of equivalent circuits, as shown in Figure A.38a and b. The propagation delay of the resulting circuit is easily derived:

$$t_{charge} = 0.69 \left(\frac{R_C}{\beta_F + 1} \parallel R_B \right) [C_C(\beta_F + 1) + C_L] \quad (A.65)$$

which is similar to the expression obtained for $t_{discharge2}$ above.

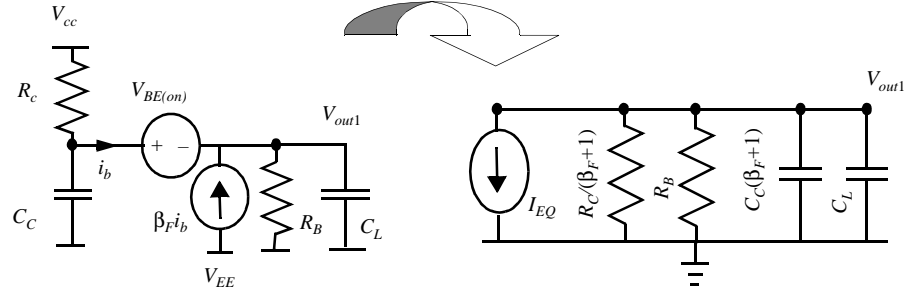


Figure A.38 (a) Simplified circuit. (b) Norton equivalent circuit of (a) for the computation of t_{pLH} , with $I_{EQ} = -V_{EE}/R_B - (V_{CC} - V_{BE(on)}) \times (\beta_F + 1)/R_C$.

Example A.10 ECL Transient Response

The total propagation delay of the ECL gate of Example A.7 can now be computed (recall that $R_C = 1\text{ k}\Omega$, and $R_B = 50\text{ k}\Omega$).

Consider first the t_{pHL} . The time to switch the differential pair was computed in Example A.8 and equals 20 psec. The second factor of the delay is the time to discharge the node capacitance. Since $R_B C_L \gg R_C C_C$, the output node time-constant dominates, and the discharge time is given by Eq. (A.63).

$$t_{\text{discharge}} = (0.5 \times 60\text{ fF} \times 50\text{ k}\Omega \times 0.5\text{ V}) / (-0.7 + 5 - 0.125\text{ V}) = 180\text{ psec}.$$

Adding the two components leads to the following approximation:

$$t_{pHL} = 20\text{ psec} + 180\text{ psec} = 200\text{ psec}.$$

The low-to-high transition consists of turning off the current switch and charging the load capacitances. The charge time is approximated by Eq. (A.65), with the collector node presenting the largest delay.

$$t_{\text{charge}} = 0.69 R_C C_C = 0.69 \times 1\text{ k}\Omega \times 62.7\text{ fF} = 43.3\text{ psec}.$$

Combining this with the delay of the differential pair yields

$$t_{pLH} = 20 + 43.4 = 63.3\text{ psec}$$

and

$$t_p = (200 + 63.3)/2 = 132\text{ psec}$$

The simulated transient response of the circuit is shown in Figure A.39. The extracted values of t_{pHL} and t_{pLH} respectively equal 163 psec and 92 psec, yielding $t_p = 127.5\text{ psec}$, which is consistent with the estimated results.

A Global View of the Transient Response

The results of this extensive analysis are summarized in Table A.5. Observe again that these expressions are first-order models and are by no means intended to be accurate.

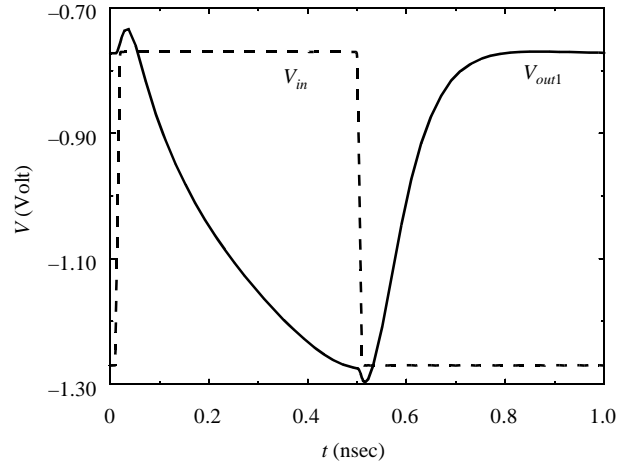


Figure A.39 Simulated transient response of ECL inverter.

SPICE simulations are indispensable if precise numbers are required. The most important function of the presented model is to provide a clear insight into the parameters dominating the performance of the ECL structure.

The analysis shows that the ECL gate exhibits asymmetrical high-to-low and low-to-high transitions. The difference between charge and discharge times is most obvious for large load values. In that case, the rise time is determined by the $C_L R_C \beta_F$ time constant,

Table A.5 Propagation delay of ECL gate: summary.

		Switching the differential pair	Capacitive (dis)charge
t_{pHL}	$R_B C_L \gg R_C C_c$	$r_B(2.2C_{in,j} + \alpha C_{D1})$	$0.5 C_L R_B \left(\frac{V_{swing}}{V_{CC} - V_{EE}} \right)$
	$R_B C_L \ll R_C C_c$		$0.69 \left(\frac{R_C}{\beta_F + 1} \parallel R_B \right) (C_C(\beta_F + 1) + C_L)$
t_{pLH}		$r_B(2.2C_{in,j} + \alpha C_{D1})$	$0.69 \left(\frac{R_C}{\beta_F + 1} \parallel R_B \right) (C_C(\beta_F + 1) + C_L)$

while the fall time at the output is set by $C_L R_B$. For C_L smaller or comparable to C_C , t_{pHL} and t_{pLH} are approximately equal, and the delays are dominated by the $C_C R_C$ time constant. The former case occurs most often in discrete ECL gates or in gates driving large capacitive loads. The latter situation occurs in gates internal to modules such as adders or CPUs. In these circumstances, the propagation delay is mainly dominated by the intrinsic capacitances, modeled by C_C .

Design Considerations

In the presented example, the propagation delay is dominated by the capacitive (dis)charge time. This causes the propagation delay to *increase linearly with the capacitive load*, as in the CMOS case. It is important to notice, however, that the ECL gate is less sensitive to capacitive loading than the CMOS counterpart. During a low-to-high transition, C_L is charged through the emitter-follower, which means that a substantial charging current is available (or the effective capacitive load is divided by the current gain of the emitter-follower). The discharge time for large load capacitances is dominated by the $R_B C_L$ time-constant, which can be made very small by reducing the value of R_B .

Compared to the propagation delay of the CMOS inverter (225 psec), a speedup of 1.8 is obtained for a single fan-out. Considerable improvements over this result can be attained. The gate design presented in the example is far from optimal. Its delay can be lowered by increasing the current levels in both the differential pair and the emitter-follower. This effectively *trades off an improvement in performance against an increased power consumption*. For a given logic swing, increased current levels mean reduced values of resistors R_C and R_B . From Table A.5, it is apparent that this translates to a reduction in capacitive (dis)charge times.

At this point, you probably wonder how low a delay can be obtained and if an optimum current level can be defined. In effect, increasing the collector current also increases the base charge in the bipolar transistors, and, consequently, the turn-on and turn-off times of the devices. This is clearly demonstrated in Eqs. (A.58) and (A.61), which state that the diffusion capacitances are inversely proportional to resistor values. At some current level, the diffusion capacitance becomes dominant, and the delay levels off and even starts to grow. Other second-order effects, such as the Kirk effect and the collector resistance come into play at higher current levels and cause a further degradation. To illustrate this behavior, we have plotted the simulated t_{pLH} of our ECL gate as a function of I_{EE} (we have kept R_B constant for this analysis). A

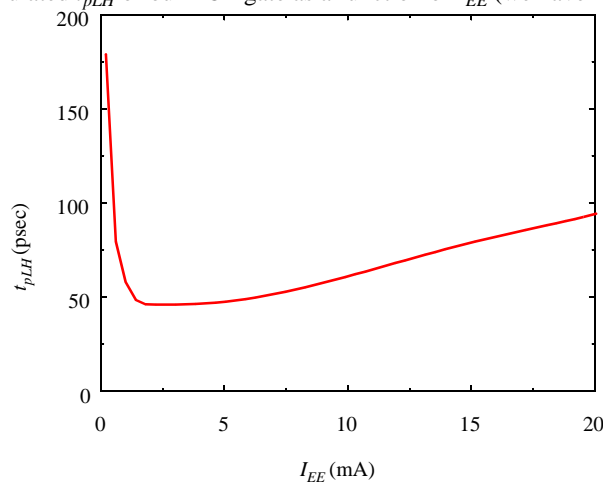


Figure A.40 t_{pLH} of ECL inverter as a function of the I_{EE} bias current.

sharp decrease in the delay can initially be observed due to a reduction in the dominant time-

constant $R_C C_C$. For higher current levels, the time to turn-off Q_1 starts to dominate, and the delay rises again (in accordance with $r_B C_{D1}$).

In summary, careful optimization of the gate structure can substantially reduce the delay. Combining this with a state-of-the-art bipolar technology (the parasitic capacitances can be as low as 10 fF in a 0.8 μm bipolar process) leads to ECL propagation delays ranging from 40 to 70 psec. This is between 3 to 4 times faster than its CMOS equivalent.



A.3.5 Power Consumption

The ECL gate clearly consumes static power. Sources of this static consumption are the emitter-coupled pair, the bias network, and the output stage. The dissipation of the latter depends upon the value of the pull-down resistor and the termination mechanism used. The power consumption of the bias network, on the other hand, can be distributed over multiple gates. These observations are summarized in the following equation,

$$P_{stat} = (V_{CC} - V_{EE}) \left(I_{EE} + \frac{I_{bias}}{N} + 2 \frac{\frac{V_{OH} + V_{OL}}{2} - V_{EE}}{R_B} \right) \quad (\text{A.66})$$

where N equals the number of gates serviced by a single bias network. The factor 2 in the power-consumption expression for the emitter-follower is due to the complementary outputs.

Furthermore, dynamic power is consumed during switching. During each switching event, both internal and external capacitances are (dis)charged, requiring a charge transfer equal to $V_{swing}(C_C + C_L)$. In contrast to the CMOS inverter, the voltage swing is considerably smaller than the supply voltage. This results in a slight modification in the expression for the dynamic power consumption (compare this to the expression in Eq. (3.42), derived for the CMOS inverter).

$$P_{dyn} = C_T (V_{CC} - V_{EE}) V_{swing} f \quad (\text{A.67})$$

with C_T the sum of all capacitances switched.

The static power factor generally dominates the overall dissipation. In fact, it can be shown that for the dynamic power to become the dominating factor, it is necessary to switch the gate faster than the propagation delay would allow.

Example A.11 Power Dissipation of the ECL Inverter

The power dissipation of the inverter of Example A.7 is analyzed. We ignore the power consumed in the bias network. The static power consumption is computed with the aid of Eq. (A.66).

$$P_{stat} = 5 (0.5 \text{ mA} + 0.162 \text{ mA}) = 3.3 \text{ mW}$$

Even when switching the gate at the maximum allowable speed ($f = 1/t_p$), the dynamic consumption is still smaller,

$$P_{dyn} = (60 \text{ fF} + 72.8 \text{ fF}) \times 5\text{V} \times 0.5\text{V} / 127.5 \text{ psec} = 2.6 \text{ mW}$$

Combining the two factors yields a considerable consumption of approximately 6 mW. The power-delay product of the gate equals 760 fJ, which is almost identical to the number obtained for the CMOS inverter. The reduced switching delay is offset by the increased power consumption. Notice, however, that PDPs as low as 120 fJ have been reported in state-of-the-art 0.6 μm bipolar processes. The high power dissipation obviously constrains the number of gates that can be integrated on a single die.

Design Consideration

Finally, one more observation is worth mentioning. During the qualitative analysis of the gate, we have already noted that the differential pair draws a constant current from the supply, hence introducing virtually no switching noise on the supply lines. This is also true for the bias network, but is not the case for the emitter-follower. Depending upon the values of the terminating resistor and the load capacitance, the switching causes large stepwise current variations in the collectors of the emitter-follower transistors. As will be discussed in later chapters, parasitic inductances can translate those current fluctuations into supply noise (ringing, voltage spikes), which might cause the circuit to fail. Therefore, ECL circuits often employ two V_{CC} 's: a “clean” V_{CC} which connects to the differential pairs and the bias networks, and a “dirty” V_{CC} , which feeds the emitter-followers. This separation avoids the feeding back of switching transients into differential pairs that could prove disastrous given the low noise margins.



A.3.6 Looking Ahead: Scaling the Technology

When discussing the influence of technology scaling on the performance of a CMOS process, we introduced the ideal scaling model. In this model, both dimensions and voltages are scaled in a similar way, keeping the electrical fields in the devices approximately constant.

Bipolar scaling is considerably more complex than MOS scaling, as more device parameters are involved in the design. Also, power-supply voltage and logic swing cannot be reduced much, since they already approach their lower limit at room temperature. The on-voltage of the base-emitter junction is a built-in parameter, which is virtually unaffected by scaling. The full-scaling model is therefore not appropriate for the ECL inverter, and the fixed-voltage scaling model must be used. Consequently, both electrical field strengths and current densities increase when scaling.

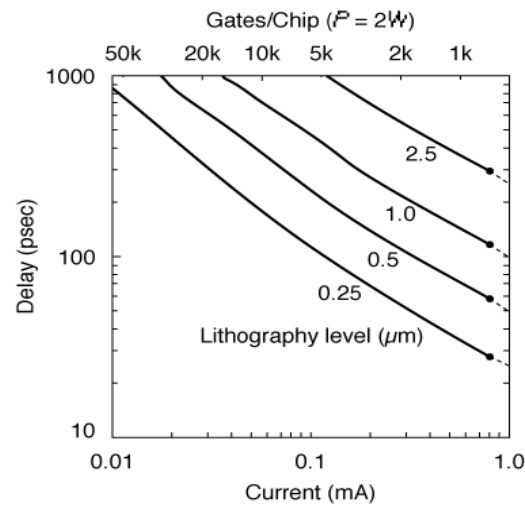
A typical bipolar scaling model is presented in Table A.6. Voltages and currents are kept to a constant level. This means that the current density increases with a squared factor. Capacitances, consisting of junction and diffusion capacitances, are scaled linearly, which requires a scaling of the base width as well as the doping levels in base and collector regions. Thus, bipolar scaling means not only shrinking the lateral dimensions, but also the vertical profile of the transistor.

This model is illustrated in Figure A.41, where the projected gate delay of an ECL circuit is plotted as a function of the current level and feature size. The plot also provides

Table A.6 Scaling model for bipolar inverter (from [Tang89]).

Parameter	Scaling Factor
A_E	$1/S^2$
w_B	$1/S^{0.8}$
V_{supply}, V_{swing}	1
J	S^2
I	1
C_d, C_j	$1/S$
t_p	$1/S$
P	1

insight in the number of gates that can be integrated on a single die (for a maximum consumption of 2 Watts).

**Figure A.41** Projected gate delay of scaled ECL circuit in function of current level and feature size (from [Tang89]).

A.4 Bipolar Gate Design

In the subsequent sections, we first discuss how the generic ECL gate, introduced in Chapter 3, can be extended to implement complex logic functions. A number of modifications to the generic gate are introduced to increase its performance or to improve its scalability. The section concludes with a short overview of some alternative bipolar logic styles.

A.4.1 Logic Design in ECL

To recapitulate, the basic structure of the ECL gate is repeated in Figure A.42. It consists

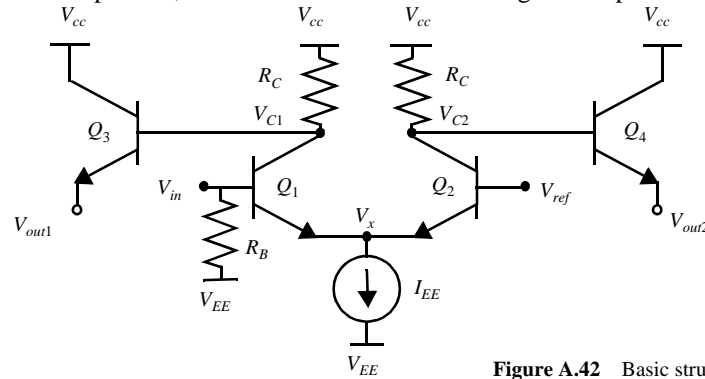


Figure A.42 Basic structure of the ECL gate.

of a bias network generating a reference voltage V_{ref} , a differential pair Q_1 - Q_2 , and a pair of emitter-follower output drivers. From a logic point of view, one of the useful properties of the basic ECL gate is that it uses a *complementary logic style*, which means that the complement of each logic signal is available. This is similar to the CMOS CPL design technique discussed earlier. This approach avoids the need for extra inverters, as is the case in inverting logic families such as complementary CMOS. It has been observed that for a given logic function, static CMOS takes approximately twice as many gates as ECL.

Example A.12 Advantages of Complementary Logic

A 4-bit ALU in CMOS requires 98 gates using NAND logic, while an ECL implementation can be conceived with only 57 gates. Even more important, the average number of cascaded logic stages in a typical signal path equals 6.1 for CMOS and 2.9 for ECL. This reduction of the critical path by a factor of 2.1 has a significant impact on performance [Masaki92].

One of the simplest gates to implement in ECL is the OR/NOR function, which can be realized by adding extra input transistors in parallel, as shown in Figure A.43. Turning one of the two input transistors on is sufficient to divert the bias current to the left branch of the current switch and to pull V_{C1} down. This is equivalent to realizing a NOR gate. The complementary node V_{C2} goes high at the same time, implementing the OR function. The static and dynamic characteristics of the standard ECL gate are only marginally affected by this modification.

Realizing an (N)AND operation is somewhat more complex. One approach to implement complex functions is to use the *wired-OR* configuration, as shown in Figure A.44. Simply wiring the gate outputs together (also called *dotting*) is equivalent to an “OR-ing.” It is sufficient for one of the connected outputs to be high for the combined result to be high. The resulting gate implements $(A + B) + (C + D) = (A + B) \cdot (C + D)$. The advantage of this approach is that complex gates can be constructed on the fly by simply connecting or wiring the outputs of basic logic OR structures. Observe that the combined gate is not complementary anymore, since only $(A + B) \cdot (C + D)$ is available, not its

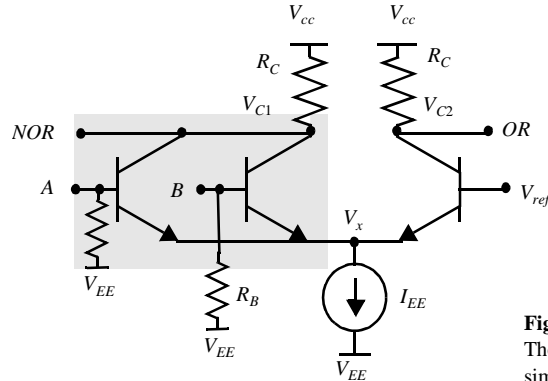


Figure A.43 Two-input OR/NOR gate in ECL. The emitter-followers have been omitted for simplicity.

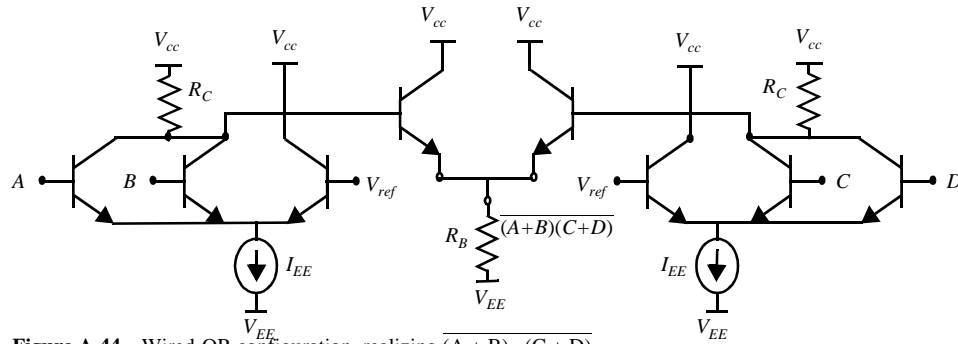


Figure A.44 Wired-OR configuration, realizing $(A + B) \cdot (C + D)$.

complement. Figure A.44 also shows that in case only one output of an ECL gate is used, it is perfectly reasonable to omit the collector resistance at the other side.

The ECL circuits discussed so far are of the so called *single-ended* nature. The central component of the gate is a current switch, with one side connected to the input(s) and the other side to a reference voltage centered in the middle of the logic swing. To ensure sufficient noise margins, it is essential that the low and high input levels differ a number of ϕ_T from the reference voltage. Remember from Chapter 3 that the transition region of the ECL inverter approximately equals 120 mV to 240 mV. For that reason, the logic swing of a single-ended family is normally chosen to range between 500 and 800 mV. The extra safety margin is necessary to accommodate variations in V_{ref} , supply voltage, process parameters, or temperature.

A.4.2 Differential ECL

Since the propagation delay of a digital gate is directly proportional to the logic swing, it is attractive to reduce the swing even more. This is made possible by a simple modification of the generic ECL structure. Instead of connecting the second input of the current switch to a reference voltage, it can be driven by the inverted value of the first input, as shown in Fig-

ure A.45. Such a gate has a reduced transition region. While one input goes up, the second

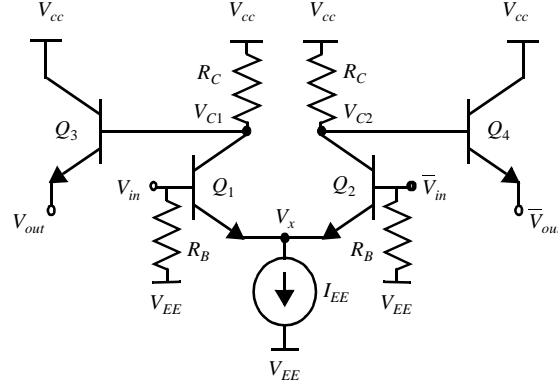


Figure A.45 Differential ECL buffer.

one goes down. This effectively doubles the voltage swing observed by the differential pair.

This observation is confirmed by a first-order derivation of the V_{IH} and V_{IL} of the inverter/buffer. The ratio of the currents between the left and right branches of the current switch can be expressed by Eq. (A.68).

$$\frac{I_{C1}}{I_{C2}} = \frac{I_S e^{V_{in}/\phi_T}}{I_S e^{\bar{V}_{in}/\phi_T}} = e^{\frac{V_{in} - \bar{V}_{in}}{\phi_T}} \quad (\text{A.68})$$

The current ratio is an exponential function of the difference between the signal and its inverse, which is twice as large as the difference between the signal and a fixed reference voltage, assuming that the inverse signal changes at the same rate. Using the alternative definition of V_{IL} and V_{IH} ($I_C = 1\%$ or 99% of I_{EE} , respectively), the width of the transient region can be computed as follows

$$\frac{I_{C1}}{I_{EE}} = \frac{e^{\frac{V_{in} - \bar{V}_{in}}{\phi_T}}}{1 + e^{\frac{V_{in} - \bar{V}_{in}}{\phi_T}}} = \alpha = 0.99 \quad (\text{A.69})$$

$$V_{IH} - V_{IL} = \phi_T \ln\left(\frac{\alpha}{1 - \alpha}\right) = 120 \text{ mV (at room temperature)}$$

As expected, a reduction by a factor of 2 with respect to the single-ended structure is obtained. This, in turn, allows for a reduction in voltage swing. Values as small as 200 mV are not uncommon!

Design Considerations

The differential approach offers several important advantages:

- The *sensitivity to supply noise* is reduced. Local drops in the supply voltages affect both V_{in} and \bar{V}_{in} in a similar way. This has no impact on the operation of the gate, since only the difference between input signals is important. This is not true in the single-ended case, where the input signal is compared to a fixed reference voltage that varies as a function of temperature, process parameters, and noise levels.
- In addition, the *switching noise* introduced on the supply lines is substantially reduced in the differential case as a result of the reduced voltage swing and the balanced load. For the single-ended case, it is assumed that only one side of the gate is connected to an emitter-follower, which causes large variations in the supply current during switching. The supply current in the differential case is approximately constant. This is confirmed in Figure A.46, which compares the supply current (or power) of the single-ended and differential logic structures for both a positive and negative signal transition (from [Greub91]).

Figure A.46 Instantaneous power consumption (or supply current) for differential and single-ended buffers (from [Greub91]). Parameters: $V_{CC} - V_{EE} = 5\text{ V}$, $I_{EE} = 400\text{ }\mu\text{A}$, $I_{emitter-follower} = 800\text{ }\mu\text{A}$, $V_{swing}(\text{differential}) = V_{swing}(\text{single-ended})/2 = 250\text{ mV}$.

- Because differential logic circuits tend to require a lower number of cascaded gates for a given function, they are generally faster than their single-ended counterparts.



The differential approach also has some potential deficiencies. The implementation of a fully differential style requires complementary logic networks similar to the DCVSL CMOS logic discussed in Chapter 4. This translates to a larger number of transistors, as demonstrated in Figure A.47, which shows the schematic diagrams of a three-input OR and AND gate. The OR/NOR structure requires six transistors in contrast to the four needed in the single-ended case.

Differential logic structures are based on the same fundamental concept as the simple ECL inverter, namely, *current steering*. Based on the value of the inputs, the current

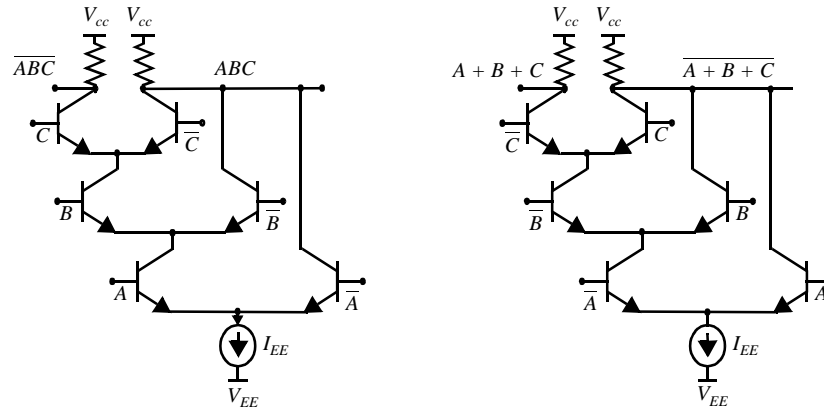


Figure A.47 (a) Three-input AND/NAND (b) Three-input OR/NOR
AND and OR gates in differential ECL (emitter-followers are omitted).

provided by the current source is guided along either the left or the right branch of the differential pair, causing one output to go low and the other to remain high. This is easily verified for the gates presented in Figure A.47. Be aware that a current path has to be available for all possible operation conditions!

While the idea seems simple enough, there is a catch, as illustrated in Figure A.48a. For performance reasons, Q_3 should not go into saturation (see Section 3.4.1). Consequently when Q_3 is on (or V_{in3} is high), it holds that $V_{C3} \geq V_{B3}$. To turn Q_2 on, V_{B2} must be at $V_{C3} + V_{BE(on)} \geq V_{B3} + V_{BE(on)}$ or, $V_{in2} \geq V_{in3} + V_{BE(on)}$. This means that all inputs connecting to transistors at level 2 (Figure A.48a) need a dc offset of at least $V_{BE(on)}$ with respect to inputs to level 3 devices.⁶ Similar offsets are needed when connecting to devices placed higher in the input network. Achieving these offsets is accomplished by adjusting the out-

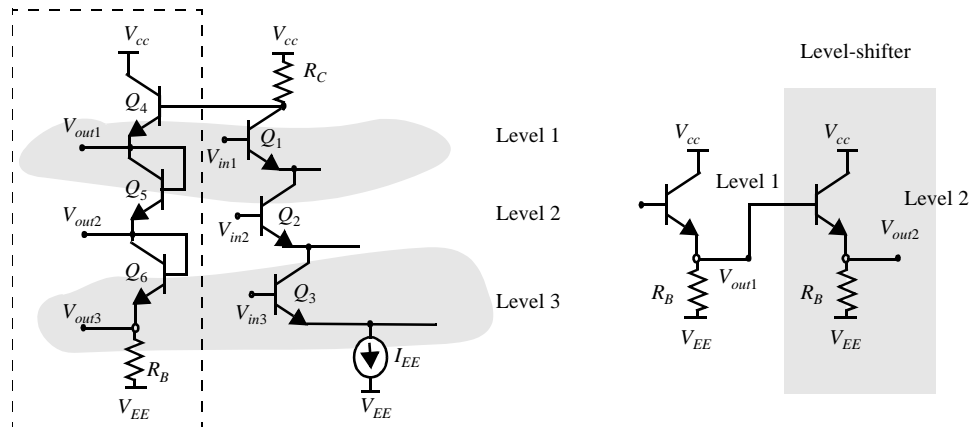


Figure A.48 (a) Multilevel emitter-follower output stage (b) Level 1-to-2 converter
Adjusting voltage levels in differential ECL using level shifters.

put stage of the (preceding) gate, as demonstrated in Figure A.48a. Connecting the base and collector of transistors Q_4 and Q_5 turns those devices into diodes with an on-voltage of $V_{BE(on)}$. By tapping the outputs at the various diode positions, suitable input signals can be generated for all required levels.

This level-shifting creates extra complexity, since multiple wires may need to be routed for the same signal, depending upon the fan-out of the gate. It also puts a restriction on the number of transistors that can be stacked. The propagation delay is a function of the output level, because signals lower on the stack have a higher delay. Adding too many layers results in an intolerable performance degradation. Furthermore, the number of layers is restricted by the available voltage range between the supply rails. For instance, the *level 3* signals in Figure A.48a range from $V_{CC} - 3V_{BE(on)}$ (high) to $V_{CC} - 3V_{BE(on)} - V_{swing}$ (low). A typical differential ECL library (e.g. [Tektronix93]) uses at most three layers.

Another way to address noncompatible signal levels, while avoiding the complex, multilevel output emitter-follower, is to insert level-shifting circuits whenever needed. A *level 1-to-level 2* converter is shown in Figure A.48b. In this way, all standard logic gates can be designed with a sole level 1 output, and level-shifting buffers are introduced only when connecting to multilevel gates.

One additional property of the differential logic style is worth mentioning. The single-ended ECL gate tends to exhibit radically different values for its output rise and fall times, because these are set by different circuit elements. This results in different values of t_{pLH} and t_{pHL} , because the switching point of the gate is determined by comparing the input signal with the fixed reference voltage. In the differential circuit, the switching is determined by the crossing of the input signal and its inverse. Although these signals may have very different slopes, this does not affect the delay of the gate, which is thus independent of the direction of the transition. Figure A.49 compares the delay of a single-ended versus a differential buffer as a function of the capacitance. While the delay of the single-ended version is a strong function of the transition direction, only one delay exists for the differential case. Notice also the smaller delay of the differential buffer caused by the reduced swing. Finally, observe that the value of the load capacitance has only a minor impact on the delay in contrast to CMOS.

Example A.13 Differential ECL Buffer

Figure A.50 shows the schematics and layout of a differential ECL buffer ([Tektronix93]). Its unloaded delay is 180 psec and increases by 330 psec for every pF of load capacitance for a total current of 0.6 mA. The input capacitance of the gate equals 50 fF. For a 5 V supply voltage, this translates into a PDP of 600 fJ (unloaded).

Example A.14 Differential ECL Gate

The circuit schematics of a four-input multiplexer is shown in Figure A.51. Verify its functionality by checking how the current flows for each of the combinations of the signals *A* and *B*. The logic swing at the output equals $400 \mu\text{A} \times 625 \Omega = 250 \text{ mV}$. To function correctly, the

⁶ Because ECL signals are normally referenced with respect to V_{CC} , it is common to refer to level 1 signals as those that are one V_{BE} below V_{CC} , level 2 as 2 V_{BE} s below V_{CC} , and so on.

input signals i_{0-3} must be at level 1 (or $V_{CC} - V_{BE(on)}$), while the B and A signals should be at levels 2 and 3, respectively. The static power consumption of the differential switch is easily computed and equals $400 \mu\text{A} \times 5 \text{ V} = 2 \text{ mW}$.

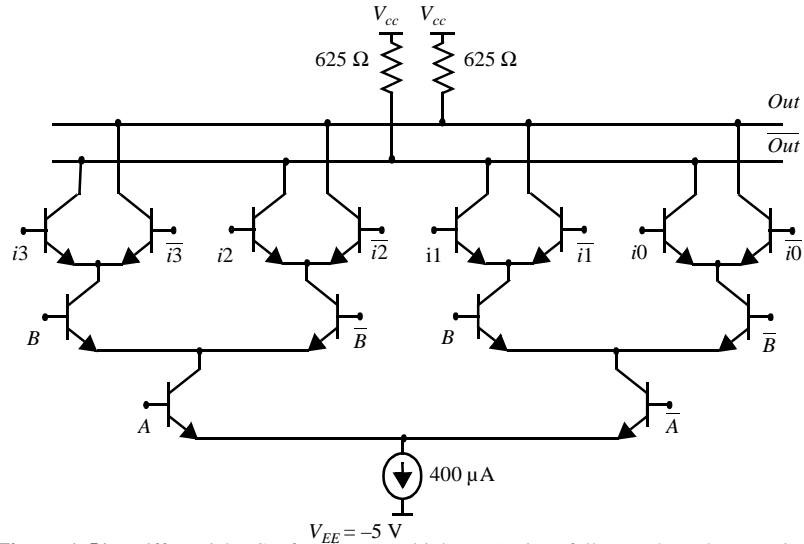


Figure A.51 Differential ECL four-input multiplexer (emitter-followers have been omitted); $V_{CC} = 0 \text{ V}$.

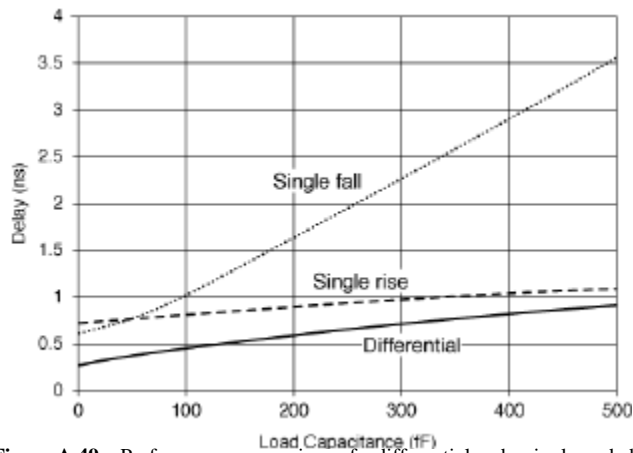


Figure A.49 Performance comparison of a differential and a single-ended buffer (from [Tektronix93]). Process parameters: $f_T = 12 \text{ GHz}$, $A_{E,min} = 0.6 \mu\text{m} \times 2.4 \mu\text{m}$, worst-case modeling. Both cells use $100 \mu\text{A}$ tree and $100 \mu\text{A}$ emitter-follower currents.

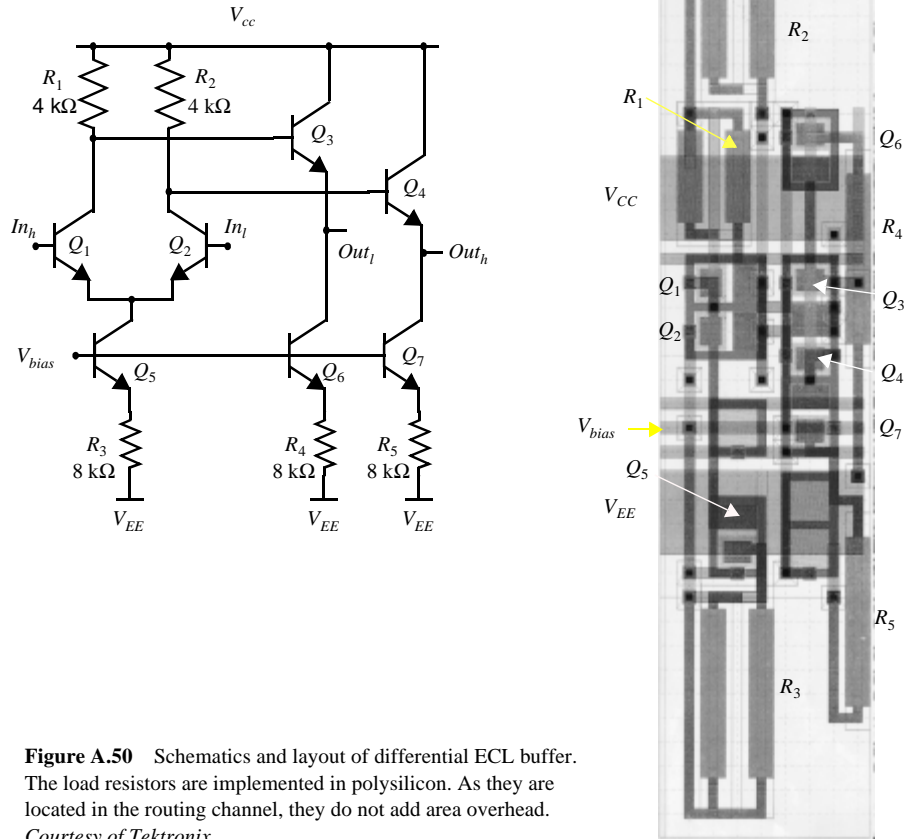


Figure A.50 Schematics and layout of differential ECL buffer. The load resistors are implemented in polysilicon. As they are located in the routing channel, they do not add area overhead. *Courtesy of Tektronix.*

Problem A.4 Differential ECL Gate

Derive the circuit topology for a two-input XOR gate in both differential and single-ended design styles.

A.4.3 Current Mode Logic

The emitter-follower output structure of the ECL gate makes it possible to drive large capacitive loads with small delay. Additionally, the output structure supports a large fan-out, which would otherwise be impossible due to the finite input impedance of the ECL gate, formed by the base of the nonsaturating differential pair input transistor in parallel with the pull-down resistor. On the negative side, the output stage adds a considerable area overhead, as is apparent in Figure A.50. It also consumes a substantial amount of power, which prevents the realization of complex integrated circuits. For instance, a 300 MHz 32-

bit microprocessor implemented in the single-ended ECL logic style [Jouppi93] has been reported. The power dissipation of the 486k transistor device equals 115 W! To disperse the excess heat generated at these power levels, special (and expensive) cooling approaches are an absolute necessity.

The capacitive loads on the internal nodes of a complex logic block such as an adder tend to be small, and the fan-out is generally restricted to 1 or 2. For these small loads, it is reasonable to eliminate the emitter-follower. When a bipolar gate only consists of a differential pair, the logic style is called *current mode logic* (CML). A differential version of a CML gate is obtained by eliminating transistors Q_3 and Q_4 , as well as the pull-down resistors in Figure A.45. Eliminating the follower results in a drastic reduction in power consumption, while keeping the propagation delay reasonable for small loads. Figure A.52 plots the propagation delay of a differential ECL and a CML buffer as a function of the logic swing for a fan-out of 1 (with and without the extra load of an interconnect wire of 0.5 mm long). CML is actually faster for small logic swings and small loads. Its delay increases rapidly with the logic swing, as the transistors start to saturate during the switching. On the other hand, the static power dissipation of the CML gate is only 2 mW versus the 10 mW of the ECL gate. CML is, therefore, only used within cells where the fan-out is small, the interconnect length is short, and power dissipation and cell area is an issue.

Figure A.52 Propagation delays of ECL and CML buffers versus logic swing for different loads (unloaded and loaded with a 500 μm interconnect wire). Design parameters: fan-out = 1, $I_{EE} = 400 \mu\text{A}$, $I_{\text{follower}} = 800 \mu\text{A}$ (from [Greub91]).

Example A.15 CML Gate Characteristics

A differential CML buffer (obtained by removing the emitter-followers in Figure A.45) has been designed using our generic bipolar technology. The value of the current source is set to 0.4 mA and $R_C = 625 \Omega$. The supply voltages V_{CC} and V_{EE} are set to 0 V and -1.7 V, respectively. This translates into a voltage swing of 0.25V with $V_{OH} = 0$ V and $V_{OL} = -0.25$ V. **Observe that the differential nature of the gate eliminates the need for output level-shifting, necessary in single-ended gates.**

The VTC of the gate is plotted in Figure A.53a. The simulated values of the logic levels are consistent with the manual analysis. The values of V_{IH} and V_{IL} (using the unity gain definition) evaluate to -0.085 V and -0.165 V, which results in both a low and a high noise margin of 0.085 V. Compared to the ECL gate of Example 3.13, all dc parameters (width of transition region, logic swing, and noise margins) have been halved. It should be mentioned that the dc analysis was performed with perfectly complementary inputs ($V_{in2} = -0.25$ V $- V_{in1}$).

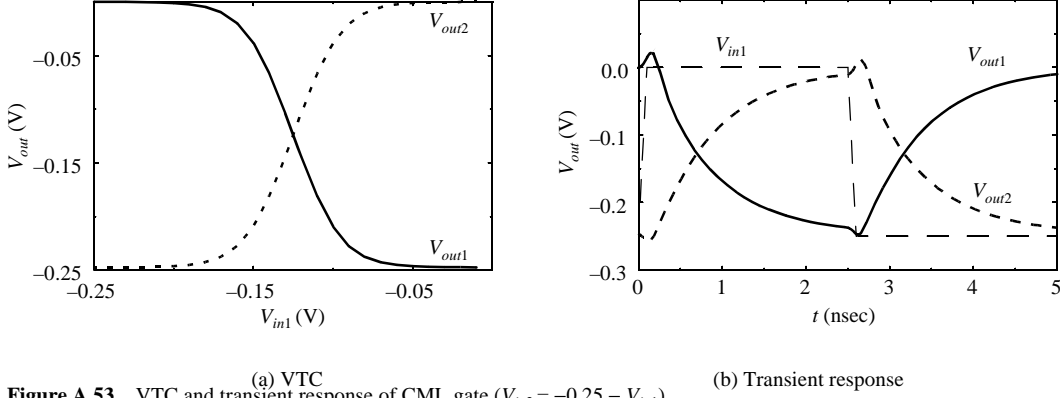


Figure A.53 VTC and transient response of CML gate ($V_{in2} = -0.25 - V_{in1}$).

For our standard bipolar technology, the load capacitance (for a single fan-out) can be computed by the same techniques used in Example 3.15 for the ECL gate. Assuming a fan-out capacitance of 60 fF, we can approximate the capacitance at the output node as follows:

$$C_L = C_{fan-out} + C_C = C_{fan-out} + (C_{cs1} + 2 C_{bc1}) = 60 \text{ fF} + 0.67 \times 47 \text{ fF} + 2 \times 1.01 \times 22 \text{ fF} = 136 \text{ fF}$$

Notice that C_{bc} is accounted for twice to incorporate the Miller effect. The weighting functions are the K_{eq} factors. Ignoring the turn-on (off) time of the transistors, the propagation delay can now be approximated:

$$t_p = 0.69 R_C C_L = 0.69 \times 625 \times 136 = 59 \text{ psec}$$

The simulated transient response of the circuit is shown in Figure A.53b. The values of t_{pLH} and t_{pHL} evaluate to 69 and 63 psec respectively, or $t_p = 66$ psec.⁷ This number rapidly degrades for larger fan-out values. For instance, for a fan-out of 3, the delay climbs to 102 psec!

The static power consumption of the gate is independent of the logic state and equals $0.4 \text{ mA} \times 1.7 \text{ V} = 0.68 \text{ mW}$. When switching the gate at the fastest possible speed, the dynamic power consumption evaluates to

$$P_{dyn} = C_L (V_{CC} - V_{EE}) V_{swing} / t_p = 0.98 \text{ mW}$$

⁷ These numbers are obtained using the traditional definitions of t_{pHL} and t_{pLH} . As mentioned earlier, these definitions do not make much sense for a differential logic family, where the propagation delay should be determined by measuring the time between the cross-over points of the differential input and output signals.

Observe that this expression takes into account that at least one collector gets charged for every transition. The differential nature of the structure effectively doubles the dynamic consumption.

Combining the obtained numbers results in a PDP of 102 fJ (for a fan-out of 1), which compares very favorably to the 750 fJ obtained for both the CMOS and ECL inverters discussed in Chapter 3.

A.4.4 ECL with Active Pull-Downs

The realization of a fast, low-power ECL-type gate is hampered by the following constraints or restrictions.

- In order to reduce t_{pHL} , the value of R_B has to be kept small; t_{pHL} normally dominates the transient performance.
- Small values of R_B result in high values of the static power consumption.

Scaling the technology to smaller lithographic dimensions has only a minor impact on this picture.

- The supply voltage of the traditional, single-ended ECL gate cannot be dropped much below $3 V_{BE}$ (≈ 2.5 V). V_{OH} equals $V_{CC} - V_{BE}$; the voltage drop between base and emitter of the input transistor adds another V_{BE} , and an additional V_{BE} is needed to ensure the proper operation of the current source of the differential pair. Since V_{BE} is a built-in voltage that is only marginally affected by technology parameters, this voltage is not reduced in scaled technologies. The power consumption/gate can hence only be reduced by reducing the current levels. This is in contrast to CMOS, where the supply voltage is expected to keep dropping as the technology feature sizes are reduced.
- R_B is not a function of technology parameters and is mainly set by the allowable power-consumption level. Fitting more gates on a die requires a reduction of the consumption/gate, which means a higher value for R_B . This adversely affects the performance (or keeps it constant at best), as $t_{pHL} \sim R_B C_L$. C_L might decrease proportional to the scaling factor, although this depends upon what portion of the load is composed of wiring capacitance.

Both of the above arguments, nonscalability of the supply voltage and the pull-down resistance, represent major obstacles to a continued performance improvement of the ECL logic family unless some substantial changes are made to the structure. Short of completely eliminating the emitter-follower (as is the case in CML), one approach is to replace the pull-down resistor by an active network [Chuang92]. Such an active pull-down network provides ample current when switching, but operates at reduced current levels in standby.

An example of such a circuit is shown in Figure A.54. This *ac-coupled active pull-down* circuit utilizes a capacitor to strongly turn on the pull-down *nnpn*-transistor during a negative-going output transition, and to turn it off during a positive one. The steady-state current of the device is set by the dc-bias network. This approach reduces the stand-by current while substantially improving the performance. The capacitor completely blocks dc

signals. Extra biasing devices are needed, however, to establish the steady-state current in the output stage. This approach has been used to realize the ECL buffers with a propagation delay of 23 psec, which is among the fastest achieved in room-temperature silicon [Toh89].

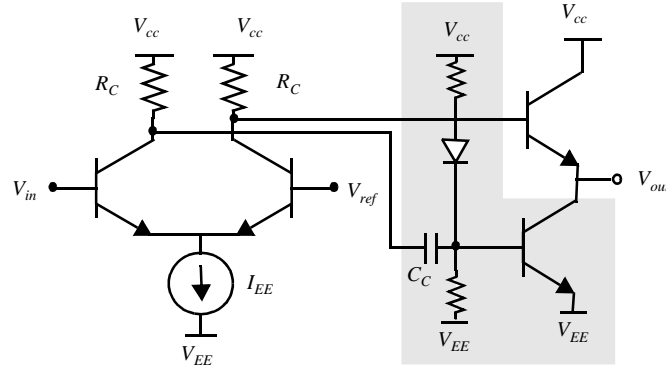


Figure A.54 Ac-coupled, active pull-down ECL buffer.

Another approach is to present a variable load to the emitter-follower. The circuit in Figure A.55 uses an NMOS device to achieve that goal. This approach can be seen as an introduction to the BiCMOS technology, which merges bipolar and MOS transistors in the same process and is treated in Section 5.3. The NMOS device is turned on more strongly during a high-to-low transition, providing more pull-down current and hence increased performance. In the high-output state, the gate-source voltage of the transistor and the static power consumption are reduced. It has been reported that for the same performance (of 200 psec), the structure presented consumes four times less power than the ECL circuit with the resistive pull-down [Chen92].

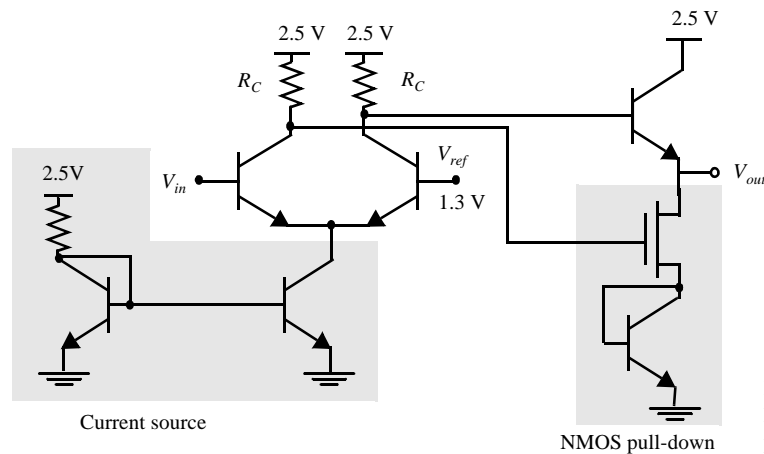


Figure A.55 2.5 V ECL gate with NFET pull-down.

Observe how the circuit of Figure A.55 is designed to operate at the *minimum possible supply voltage* (approximately three times V_{BE}). The signal swing is approximately equal to V_{BE} and is centered around 1.3 V ($1.5 V_{BE}$ below the supply voltage V_{CC}). The diode in the output stage reduces the dc standby current. The current source of the input

stage is implemented by a bipolar current mirror, which reduces the required voltage drop over the source. The reference voltage of 1.3 V is chosen to avoid the saturation of the current-source transistor.

Problem A.5 Active Emitter-Follower Load

Instead of using a resistive load for the emitter-follower circuit, a current source can also be employed, as shown in Figure A.56. The same circuit was shown earlier in the differential buffer of Figure A.50. Discuss the impact of this option on performance and power consumption.

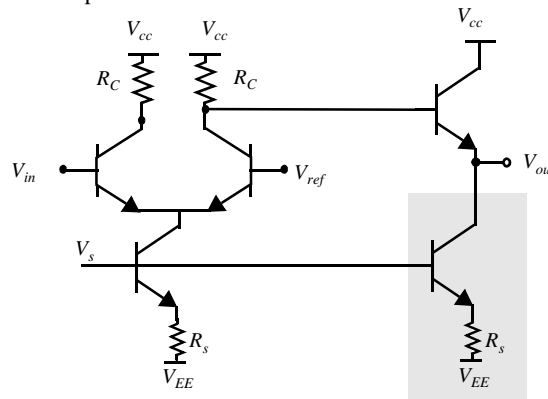


Figure A.56 Current-source pull-down.

Various other approaches have been proposed to improve the performance of the ECL structure or to reduce its power consumption [Chuang92]. While these techniques have resulted in spectacular performances at reasonable power levels, whether these approaches will help ECL to compete with deep submicron CMOS in times to come is an open question [Masaki92].

A.4.5 Alternative Bipolar Logic Styles

A large variety of bipolar digital gates have been proposed over time, covering a wide span of speed and power requirements. The most popular among them has been the *transistor-transistor-logic* (TTL) gate. TTL dominated the *discrete logic component* market, which supports individually packaged NAND and NOR gates, multiplexers or bus-drivers, for more than two decades. This picture has changed in the 1980s due to two major factors:

1. Discrete logic gates in static CMOS became competitive in speed at a lower power cost.
2. The advent of programmable logic components such as PLDs and FPGAs, discussed in Chapter 11, made it possible to *program* complex random logic functions (equivalent to hundreds of TTL gates) on a single component. This results in a large reduction in board real-estate cost, while adding flexibility.

The last factor in particular has influenced the demise of TTL. Due to its historical impact and for the sake of completeness, a brief examination of the basic TTL structure is worth the effort however. It consists of three units, as shown in Figure A.57:

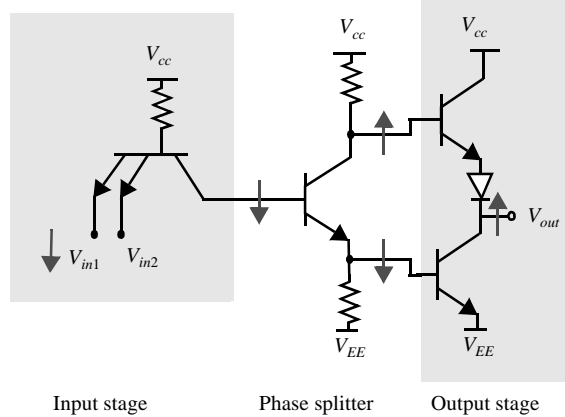


Figure A.57 Generic TTL NAND structure. The arrows denote the direction of the signal transitions for one of the input signals going low.

- The *input stage*, which consists of a multi-emitter bipolar transistor and performs an AND-ing of the inputs: both inputs of the transistor have to be high for its collector to be high.
- The *phase splitter*, which generates two signals with opposite phases. These signals are used to drive the output stage.
- The totem-pole (push-pull) *output structure*. Only one of the bipolar transistors of the structure is ON in steady-state mode, since the controlling signals have opposite phases. In contrast to the emitter-follower output stage, this structure has the advantage that no static power is consumed, while ample current drive is still available. The disadvantage is the need for a phase splitter. Saturation of the output transistors also degrades the performance. Elaborate techniques (using, for instance, Schottky Barrier diodes—see Appendix D) have been devised to avoid saturation.

Because the TTL gate is an intricate composition, its actual operation is complex, but not particularly fast. A detailed analysis of its behavior would lead us astray and would contribute little to the understanding of contemporary digital design approaches. The interested reader should refer to the numerous available textbooks and reference works that treat TTL design in extensive detail (e.g., [Hodges88]).

One legacy of the TTL era has endured. The TTL logic levels have become a de facto standard due to their widespread usage. Input and output signals of integrated circuits must often still comply with this standard. An overview of these requirements is given in Table A.7.

Table A.7 TTL signal requirements, valid for popular TTL families such as low-power Schottky (LS) and fast (F) TTL (from [Buchanan90]).

$V_{IL}(\text{max})$, V	$V_{IH}(\text{min})$, V	$V_{OL}(\text{max})$, V	$V_{OH}(\text{min})$, V
0.8	2.0	0.5	2.5

Besides ECL and TTL, a number of other bipolar logic families have emerged and vanished over the years. Three examples are integrated injection logic (I²L or MTL), Schottky transistor logic, and integrated Schottky logic ([Hodges88]). A detailed discussion is not warranted since none are actively used in current integrated circuit designs.

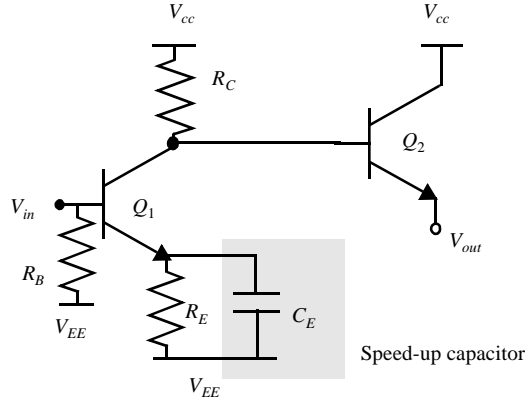


Figure A.58 Nonthreshold logic (NTL) gate structure.

One gate structure, called *nonthreshold logic* (NTL), is gradually gaining some acceptance in high-performance designs, due to its very low power-delay product ([Ichino87]), and is therefore worth analyzing. The basic structure of an NTL gate is shown in Figure A.58. It consists of an input (logic) stage and an emitter-follower output structure. The input structure resembles the RTL gate, briefly discussed in Chapter 3, where it was concluded that its performance suffers from the saturation of the pull-down device. This is avoided in the NTL gate by adding the emitter-degenerating resistor R_E , which ensures that Q_1 stays in the forward-active mode for the voltage input range of interest. The emitter-follower is added to provide additional fan-out drive capability, eliminating another deficiency of the RTL gate. The dc parameters can be derived by ignoring the base current of the emitter-follower.

$$\begin{aligned}
 V_{OH} &= V_{CC} - V_{BE(on)} \\
 V_{OL} &= V_{CC} - V_{BE(on)} - \frac{R_C}{R_E}(V_{OH} - V_{BE(on)}) \\
 &= \left(\frac{R_E - R_C}{R_E}\right)V_{CC} + \left(\frac{2R_C - R_E}{R_E}\right)V_{BE(on)}
 \end{aligned} \tag{A.70}$$

The gain of the gate in the transient region is approximately proportional to $(-R_C/R_E)$, as can be derived using the small-signal approach. A large value of the ratio enhances the voltage gain and the noise margin, but increases the gate delay. $R_C = 2R_E$ seems to be a good compromise. Under those conditions, V_{OL} evaluates to $3V_{BE(on)} - V_{CC}$, and the total signal swing equals $2V_{CC} - 4V_{BE(on)}$. As V_{OL} has to be larger than 0, this translates to the constraint that V_{CC} has to be kept smaller than $3V_{BE(on)}$. Ensuring that Q_1 does not saturate over the input range of interest puts a lower bound on V_{CC} of approximately $2V_{BE(on)}$.

A remarkable feature of this gate is that the voltage-transfer characteristic displays only a single strong nonlinearity, corresponding to the turning on of Q_1 . For the rest of the operation range Q_1 stays in the forward-active mode, because saturation is avoided during

normal operation. Hence the name *nonthreshold* logic. The lack of a well-defined low output level makes this approach particularly sensitive to noise. The attentive reader will have observed already that the gate retains the regenerative property as long as $R_C > R_E$.

Example A.16 VTC of an NTL Gate

The simulated VTC of an NTL gate is shown in Figure A.59 for the following parameters: $V_{CC} = 1.9$ V, $R_C = 2 R_E = 0.5$ k Ω . Assume that $V_{be(on)} = 0.75$ V. The manually derived values for V_{OL} and V_{OH} (0.35 V and 1.15 V, respectively) closely correspond to the simulated values marked on the simulation. This corresponds to a signal swing of 0.8 V. The value of V_{IL} equals 0.8 V, while V_{IH} cannot be defined.

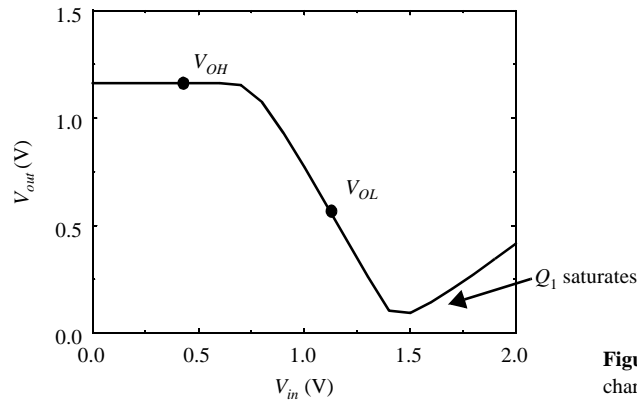


Figure A.59 Voltage-transfer characteristic of an NTL gate.

While the emitter degeneration acts as a negative feedback and helps create the non-threshold characteristic, its effects are not desirable from a transient point of view. Raising the voltage at the base of Q_1 also raises its emitter voltage, reducing the current available for discharging the collector capacitance. The effects of this negative feedback can be reduced by adding an extra decoupling capacitor C_E at the emitter node, as shown in Figure A.58. The larger the capacitance, the larger the decoupling effect and the speed-up. This capacitor can be implemented with a reverse-biased diode.

Example A.17 Transient Response of the NTL Gate

The simulated transient response of the NTL gate is shown in Figure A.60 (for a fan-out of 1 and a pull-down resistance of 2 k Ω). The response is shown without and with a decoupling capacitance of 0.3 pF. It can be observed that the decoupling capacitance only affects the high-to-low transition. The simulated value of the propagation delay with decoupler equals 41 psec.

The static power consumption of the gate can be derived. For a low input, only the emitter-follower consumes static power, while both the logic and the output structures are consuming power for a high input. Adding all factors results in a static consumption of 2.28 mW. The dynamic consumption, when clocked at the maximum possible rate, equals 2.3 mW. Reduction of the consumption is possible by reducing the voltage swing, the supply voltage,

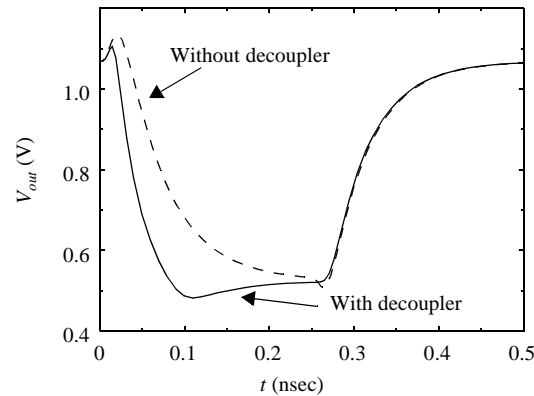


Figure A.60 Transient response of an NTL gate (without and with an emitter-decoupling capacitance of 0.3pF).

and the circuit performance. For this particular version of the gate, the PDP evaluates to 376 pJ.

The NOR operation is the logic function most easily implemented in NTL, and is constructed by a parallel connection of the input transistors. More complex functions can be realized by transistor stacking. This requires extra level shifters, while saturation should be carefully avoided.

REFERENCES

- [Antognetti88] P. Antognetti and G. Masobrio (eds.), *Semiconductor Device Modeling with SPICE*, McGraw-Hill, 1988.
- [Banzhaf92] W. Banzhaf, *Computer Aided Analysis Using PSPICE*, 2nd ed., Prentice Hall, 1992.
- [Chen90] J. Chen, *CMOS Devices and Technology for VLSI*, Prentice Hall, 1990.
- [Getreu76] I. Getreu, "Modeling the Bipolar Transistor," Tektronix Inc., 1976.
- [Gray69] P. Gray and C. Searle, *Electronic Principles*, John Wiley and Sons, 1969.
- [Gray93] P. Gray and R. Meyer, *Analysis and Design of Analog Integrated Circuits*, 3rd ed., John Wiley and Sons, 1993.
- [Haznedar91] H. Haznedar, *Digital Microelectronics*, Benjamin/Cummings, 1991.
- [Hodges88] D. Hodges and H. Jackson, *Analysis and Design of Digital Integrated Circuits*, 2nd ed., McGraw-Hill, 1988.
- [Howe95] R. Howe and S. Sodini, *Microelectronics: An Integrated Approach*, forthcoming, Prentice Hall, 1995.
- [Hu92] C. Hu, "IC Reliability Simulation," *IEEE Journal of Solid State Circuits*, vol. 27, no. 3, pp. 241–246, March 1992.

- [Hu93] C. Hu, "Future CMOS Scaling and Reliability," *IEEE Proceedings*, vol. 81, no. 5, May 1993.
- [Jensen91] G. Jensen et al., "Monte Carlo Simulation of Semiconductor Devices," *Computer Physics Communications*, 67, pp. 1–61, August 1991.
- [Ko89] P. Ko, "Approaches to Scaling," in *VLSI Electronics: Microstructure Science*, vol. 18, chapter 1, pp. 1–37, Academic Press, 1989.
- [Muller86] R. Muller and T. Kamins, *Device Electronics for Integrated Circuits*, 2nd ed., John Wiley and Sons, 1986.
- [Nagel75] L. Nagel, "SPICE2: a Computer Program to Simulate Semiconductor Circuits," Memo ERL-M520, Dept. Elect. and Computer Science, University of California at Berkeley, 1975.
- [Sedra87] A. Sedra and K. Smith, *Microelectronic Circuits*, 2nd ed., Holt, Rinehart and Winston, 1987.
- [Sheu87] B. Sheu, D. Scharfetter, P. Ko, and M. Jeng, "BSIM: Berkeley Short-Channel IGFET Model for MOS Transistors," *IEEE Journal of Solid-State Circuits*, vol. SC-22, no. 4, pp. 558–565, August 1987.
- [Sze81] S. Sze, *Physics of Semiconductor Devices*, 2nd ed., John Wiley and Sons, 1981.
- [Thorpe92] T. Thorpe, *Computerized Circuit Analysis with SPICE*, John Wiley and Sons, 1992.
- [Toh88] K. Toh, P. Koh, and R. Meyer, "An Engineering Model for Short-Channel MOS Devices," *IEEE Journal of Solid-State Circuits*, vol. 23, no. 4, pp. 950–957, August 1988.
- [Tsividis87] Y. Tsividis, *Operation and Modeling of the MOS Transistor*, McGraw-Hill, 1987.
- [Yamaguchi88] T. Yamaguchi et al., "Process and Device Performance of a High-Speed Double Poly-Si Bipolar Technology Using Boron-Poly Process with Coupling-Base Implant," *IEEE Trans. Electron. Devices*, vol. 35, no. 8, pp. 1247–1255, August 1988.
- [Weste93] N. Weste and K. Eshragian, *Principles of CMOS VLSI Design: A Systems Perspective*, Addison-Wesley, 1993.

A.5 Exercises and Design Problems

For all problems, use the device parameters provided in Chapter 2 (Sections 2.2.5, 2.3.5, and A.2.5) and the inside back book cover, unless otherwise mentioned. Also assume $T = 300$ K by default.

1. [E, SPICE, 2.23]
 - a. Consider the circuit of Figure A.61. Using the simple model, with $V_{Don} = 0.7$ V, solve for I_D .
 - b. Find I_D and V_D using the ideal diode equation. Use $I_s = 10^{-14}$ A and $T = 300$ K.
 - c. Solve for V_{D1} , V_{D2} , and I_D using SPICE.
 - d. Repeat parts b and c using $I_s = 10^{-16}$ A, $T = 300$ K, and $I_s = 10^{-14}$ A, $T = 350$ K.
2. [M, None, 2.2.3] For the circuit in Figure A.62, $V_s = 3.3$ V. Assume $A_D = 12 \mu\text{m}^2$, $\phi_0 = 0.65$ V, and $m = 0.5$. $N_A = 2.5 \text{ E}16$ and $N_D = 5 \text{ E}15$.
 - a. Find I_D and V_D .
 - b. Is the diode forward- or reverse-biased?
 - c. Find the depletion region width, W_j , of the diode.
 - d. Use the parallel-plate model to find the junction capacitance, C_j .
 - e. Set $V_s = 1.5$ V. Again using the parallel-plate model, explain qualitatively why C_j increases.

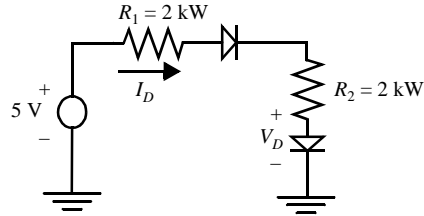


Figure A.61 Resistor diode circuit.

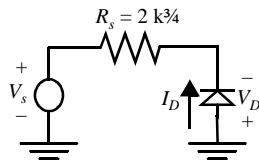


Figure A.62 Series diode circuit

3. [C, None, 2.2.3] For the circuit in Figure A.63, sketch $i_R(t)$, $v_d(t)$, and $v_o(t)$ showing quantitative values for the asymptotes and time constants. Assume $V_{D(on)} = 0.7$ V and the minority carrier transit time $\tau_T = 50$ nsec. Further assume a short-base model with a high hole-injection efficiency. Neglect the effects of junction capacitance.

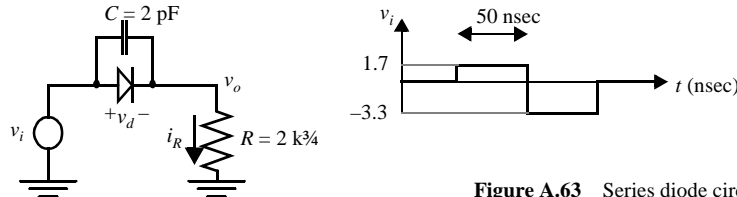


Figure A.63 Series diode circuit

4. [E, SPICE, 2.3.2] Figure A.64 shows NMOS and PMOS devices with drains, source, and gate ports annotated. Determine the mode of operation (saturation, triode, or cutoff) and drain current I_D for each of the biasing configurations given below. Verify with SPICE. Use the following transistor data: NMOS: $k'_n = 60 \mu\text{A}/\text{V}^2$, $V_{T0} = 0.7$ V, $\lambda = 0.1 \text{ V}^{-1}$, PMOS: $k'_p = 20 \mu\text{A}/\text{V}^2$, $V_{T0} = -0.8$ V, $\lambda = 0.1 \text{ V}^{-1}$. Assume $(W/L) = 1$.
- NMOS: $V_{GS} = 3.3$ V, $V_{DS} = 3.3$ V. PMOS: $V_{GS} = -0.5$ V, $V_{DS} = -1.5$ V.
 - NMOS: $V_{GS} = 3.3$ V, $V_{DS} = 2.2$ V. PMOS: $V_{GS} = -3.3$ V, $V_{DS} = -2.6$ V.
 - NMOS: $V_{GS} = 0.6$ V, $V_{DS} = 0.1$ V. PMOS: $V_{GS} = -3.3$ V, $V_{DS} = -0.5$ V.

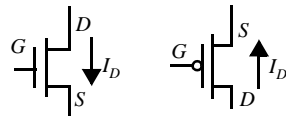


Figure A.64 NMOS and PMOS devices.

5. [E, SPICE, 2.3] Using SPICE plot the I - V characteristics for the following devices.
- NMOS $W = 2.4 \mu\text{m}$, $L = 0.6 \mu\text{m}$
 - NMOS $W = 12.8 \mu\text{m}$, $L = 3.6 \mu\text{m}$
 - PMOS $W = 2.4 \mu\text{m}$, $L = 0.6 \mu\text{m}$

- d. PMOS $W = 12.8 \mu\text{m}$, $L = 3.6 \mu\text{m}$
6. [E, SPICE, 2.3] Indicate on the plots from problem 5.
- the regions of operation.
 - the effects of channel length modulation.
 - Which of the devices are in velocity saturation? Explain how this can be observed on the I - V plots.
7. [M, None, 2.3] Repeat problem 5 using hand analysis using the LEVEL 1 model. Explain the differences with the results of problem 5 and suggest approaches to reduce the discrepancies.
8. [E, None, 2.3.2] You are given an NMOS device with the following process parameters: $N_A = 10^{16} \text{ cm}^{-3}$ (substrate), $N_D = 10^{20} \text{ cm}^{-3}$ (gate), $N_{SS} = 10^{11} \text{ cm}^{-3}$ (surface states), and $x_j = 0.5 \mu\text{m}$. Assume there has been no threshold channel implant.
- Due to process variations in t_{ox} you observe values of V_{T0} between -0.1 V and 0 V . What is the range of oxide thicknesses?
 - You desire to make a low- V_T process with $V_{T0,min} = 0.3 \text{ V}$. What channel implant dose is required?
9. [M, None, 2.3.2] Given the data in Table A.8 for an NMOS transistor with $k' = 20 \mu\text{A/V}^2$, calculate V_{T0} , γ , λ , $2|\phi_f|$, and W/L :

Table A.8 Measured NMOS transistor data

	V_{GS}	V_{DS}	V_{BS}	$I_D (\mu\text{A})$
1	3	5	0	1210
2	5	5	0	4410
3	5	10	0	5292
4	5	5	-2	3265
5	5	5	-5	2381

10. [M, None, 2.3] The data points in Table A.9 were measured for an MOS transistor:
- What type of device is this: PMOS or NMOS, $V_T > 0$ or < 0 ?
 - Is this device velocity saturated? Why or why not.
 - Derive the values of $k = k' (W/L)$, V_{T0} , and λ . Determine the operation region for each of the rows in the table.
 - Suppose that t_{ox} is reduced. How do k , V_{T0} and λ change?

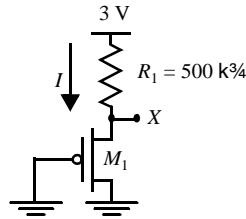
Table A.9 Measured MOS transistor data.

	$V_{GS}(\text{V})$	$V_{DS}(\text{V})$	$V_{SB}(\text{V})$	$I_D (\mu\text{A})$
1	-2	3	0	0
2	0	1	0	47.86
3	0	5	0	62.11
4	1	3.5	0	192.52
5	1	5	0	208.76

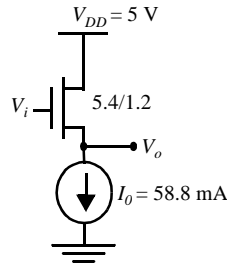
Table A.9 Measured MOS transistor data.

	$V_{GS}(\text{V})$	$V_{DS}(\text{V})$	$V_{SB}(\text{V})$	$I_D (\mu\text{A})$
6	5	5	0	1182

11. [M, None, 2.3.2] Consider the circuit configuration of Figure A.65.
- Write down the equations (and only those) which you need to determine the voltage at node X. Do NOT plug in any values yet. Assume that $\lambda_p = 0$.
 - Draw the (approximative) load lines for both MOS transistor and resistor. Mark some of the significant points.
 - Determine the required width of the transistor (for $L_{eff} = 1.2 \mu\text{m}$) such that X equals 1.5 V.
 - We have, so far, assumed that M_1 is a long-channel device. Redraw the load lines assuming that M_1 is velocity-saturated. Will the voltage at X rise or fall?

**Figure A.65** MOS circuit.

12. [M, None, 2.3.2] The circuit of Figure A.66 is known as a *source-follower* configuration. It achieves a DC level shift between the input and output. The value of this shift is determined by the current I_0 . Assume $L_D = 0.15 \mu\text{m}$, $\gamma = 0.543$, $2|\phi_t| = 0.6 \text{ V}$, $V_{T0} = 0.74 \text{ V}$, $k_C = 19.6 \mu\text{A/V}^2$, and $\lambda = 0$.
- Derive an expression giving V_i as a function of V_o and $V_T(V_o)$. If we neglect body effect, what is the nominal value of the level shift performed by this circuit.
 - The NMOS transistor experiences a shift in V_T due to the body effect. Find V_T as a function of V_o for V_o ranging from 0 to 3 V with 0.5 V intervals. Plot V_T vs. V_o .
 - Plot V_o vs. V_i as V_o varies from 0 to 3 V with 0.5 V intervals. Plot two curves: one neglecting the body effect and one accounting for it. How does the body effect influence the operation of the level converter? At V_o (body effect) = 3 V, find V_o (ideal) and, thus, determine the maximum error introduced by body effect.

**Figure A.66** Source-follower level converter.

13. [M, SPICE, 2.3.2] Problem 13 uses the MOS circuit of Figure A.67.

- a. Plot V_{out} vs. V_{in} with V_{in} varying from 0 to 5 volts (use steps of 1 V). $V_{DD} = 5$ V.
- b. Repeat *a* using SPICE.
- c. Repeat *a* and *b* using a MOS transistor with $(W/L) = 4.8/1.2$. Is the discrepancy between manual and computer analysis larger or smaller. Explain why.

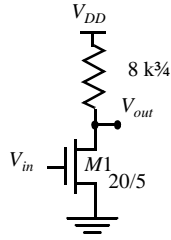


Figure A.67 MOS circuit.

14. [M, None, 2.3.3] Compute the gate and diffusion capacitances for transistor *M1* of Figure A.67. Assume that drain and source areas are rectangular, and are $20\ \mu\text{m}$ wide and $5\ \mu\text{m}$ long. Use the parameters of Example 3.5 to determine the capacitance values. Assume $m_j = 0.5$ and $m_{jsw} = 0.33$. Also compute the total charge stored at node *In*, for the following conditions:
 - a. $V_{in} = 5$ V, $V_{out} = 5$ V, 2.5 V, and 0 V.
 - b. $V_{in} = 0$ V, $V_{out} = 5$ V, 2.5 V, and 0 V.
15. [C, SPICE, 2.5] Though impossible to quantify exactly by hand, it is a good idea to understand process variations and be able to at least get a rough estimate for the limits of their effects.
 - a. For the circuit of Figure A.67, calculate nominal, minimum, and maximum currents in the NMOS device with $V_{in} = 0$ V, 2.5 V and 5 V. Assume 3σ variations in V_{T0} of 25 mV, in k' of 15%, and in lithographic etching of $0.15\ \mu\text{m}$.
 - b. Analyze the impact of these current variations on the output voltage. Assume that the load resistor also can vary by 10%. Verify the result with SPICE.
16. [M, None, 2.3.4] Short-channel effects:
 - a. Use the fact that current can be expressed as the product of the carrier charge per unit length and the velocity of carriers ($I_{DS} = Qv$) to derive I_{DS} as a function of W , C_{ox} , $V_{GS} - V_T$, and carrier velocity v .
 - b. For a long-channel device, the carrier velocity is the mobility times the applied electric field. The electrical field, which has dimensions of V/m, is simply $(V_{GS} - V_T) / 2L$. Derive I_{DS} for a long-channel device.
 - c. From the equation derived in *a*, find I_{DS} for a short-channel device in terms of the maximum carrier velocity, v_{max} .
 - d. Based on the results of *b* and *c* describe the most important differences between short-channel and long-channel devices.
17. [C, None, 2.3.4] Another equation, which models the velocity-saturated drain current of an MOS transistor is given by

$$I_{dsat} = \frac{1}{1 + (V_{GS} - V_T)/(E_{sat}L)} \left(\frac{\mu_0 C_{ox}}{2} \right) \frac{W}{L} (V_{GS} - V_T)^2$$

Using this equation it is possible to see that velocity saturation can be modeled by a MOS device with a source-degeneration resistor (see Figure A.68).

- a. Find the value of R_S such that $I_{DSAT}(V_{GS}, V_{DS})$ for the composite transistor in the figure matches the above velocity-saturated drain current equation. *Hint: the voltage drop across R_S is typically small.*
- b. Given $E_{sat} = 1.5 \text{ V}/\mu\text{m}$ and $k' = \mu_0 C_{ox} = 20 \mu\text{A}/\text{V}^2$, what value of R_S is required to model velocity saturation. How does this value depend on W and L ?

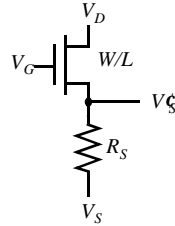


Figure A.68 Source-degeneration model of velocity saturation.

18. [C, SPICE, 2.2.3] The circuit in Figure A.69 shows an NMOS device with its parasitic diode. The NMOS transistor has $k_n = \mu_n C_{ox} W/L = 6000 \mu\text{A}/\text{V}^2$, $V_{T0} = 0.7 \text{ V}$, and $\gamma = 0$, and is off ($V_g = 0 \text{ V}$) until time t_1 when the gate voltage steps to 5 V . The parasitic diode has $A_D = 100 \mu\text{m}^2$, $\phi_0 = 0.65 \text{ V}$, and $m = 0.33$. $I_{BIAS} = 500 \mu\text{A}$.
 - a. Solve for the initial conditions on I_D , V_D , and I_{MI} (at time t_1).
 - b. Find the minority carrier charge in the diode, Q_D .
 - c. Find the time for the NMOS device to remove this excess minority carrier charge from the diode. During this interval, assume that V_D is constant, and model the NMOS device as a constant current source. Verify your results with SPICE.
 - d. Find the final values of I_D , V_D , and I_{MI} (after the NMOS device has been on for a long time).
 - e. Find the diode junction capacitance, C_j , linearized over the values of V_D found in parts a and d.
 - f. Find the time to reach 90% of the final value of V_D . During the diode space-charge removal, model the NMOS device as a current source of value equal to the average current at the interval endpoints. Compare your results with those of SPICE.

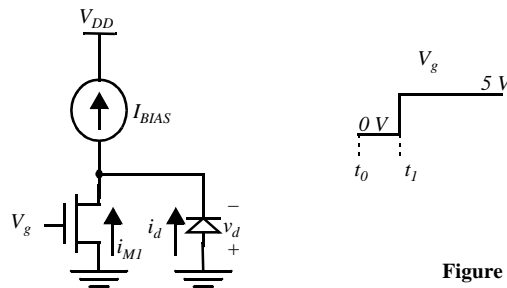


Figure A.69 NMOS with parasitic diode.

19. [E, SPICE, 2.3.2] Figure A.70 shows *nnp* and *pnp* devices biased with bias resistors R_B and R_C . Determine the mode of operation (forward-active, saturation, or cut-off) and collector current I_C for the resistor values given below. Verify with SPICE. Use the following transistor data. *nnp*: $\beta_F = 100$, $V_{BE(on)} = 0.7 \text{ V}$, $V_{BE(sat)} = 0.8 \text{ V}$, $V_{CE(sat)} = 0.2 \text{ V}$. *pnp*: $\beta_F = 30$, $V_{BE(on)} = -0.7 \text{ V}$, $V_{BE(sat)} = -0.8 \text{ V}$, $V_{CE(sat)} = -0.2 \text{ V}$.
 - a. $R_B = 5 \text{ k}\Omega$, $R_C = 1 \text{ k}\Omega$.

b. $R_B = 50 \text{ k}\Omega$, $R_C = 0.5 \text{ k}\Omega$.

c. $R_B = 10 \text{ k}\Omega$, $R_C = 250 \text{ }\Omega$.

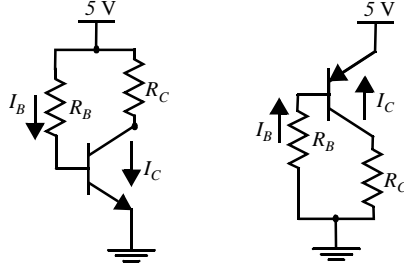


Figure A.70 *npn* and *pnp* devices with their biasing schemes.

20. [M, None, 2.5] Assuming 3σ process variations of 5% in R_C , 10% in R_B , and 20% in β_F , repeat Problem 19. For each case, find the region of operation for both devices, and the minimum and maximum values for I_B , I_C , and V_{CE} .
21. [M, None, 2.4.3] In this problem, we consider the delays involved in turning on the bipolar transistor. Sketch $v_{BE}(t)$, $Q_F(t)$, and $i_C(t)$ for the circuit of Figure A.71. Annotate the critical delay components on your graphs. Use the transistor data of Example A.2.

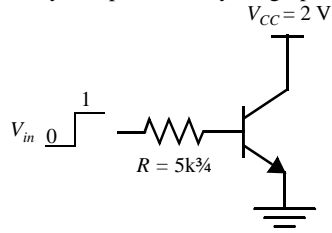


Figure A.71 Circuit for analyzing turn-on transient of bipolar transistor.

22. [C, None, 2.4.3] Consider adding a capacitor in parallel with the base resistor in the circuit of Problem 21 (Figure A.72). The capacitor is referred to as a speed-up capacitor, and it allows the output current to achieve its steady-state value almost instantaneously. How? What value of C is required to achieve this effect. Repeat problem 21 using this value for C .

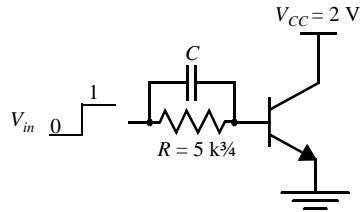


Figure A.72 Bipolar transistor circuit with speed-up capacitor.

23. [M, None, 2.4.4] The deviation in the actual collector current of a bipolar transistor, I_{CA} , from the ideal value, $I_{CI} = I_S e^{+V_{BE}/\phi_T}$, due to parasitic and high-level injection effects can be modeled as a base resistance, R_B .
- a. Derive an expression for R_B as a function of I_{CI} and I_{CA} .
- b.