IC Layout Basics
A Practical Guide

Christopher Saint

Judy Saint

McGraw-Hill
New York | Chicago | San Francisco | Lisbon | London
| Madrid | Mexico City | Milan | New Delhi |
San Juan | Seoul | Singapore | Sydney | Toronto

of Congress Cataloging-in-Publication Data

ristopher, date.
ayout basics : a practical guide / Christopher Saint, Judy Saint.
p. cm.
udes bibliographical references and index.
N 0-07-138625-4 (alk. paper)
ntegrated circuit layout. I. Saint, Judy. II. Title.
74.55.S35 2001
315—dc21 2001051309

raw-Hill
A Division of The McGraw-Hill Companies

nt © 2002 by Christopher Saint and Judy Saint. All rights reserved. Printed in
ed States of America. Except as permitted under the United States Copyright
976, no part of this publication may be reproduced or distributed in any form
y means, or stored in a data base or retrieval system, without the prior written
on of the publisher.

5 6 7 8 9 0 DOC/DOC 0 7 6 5 4 3 2 1

07-138625-4

nsoring editor for this book was Stephen S. Chapman, the editing supervisor
hen M. Smith, and the production supervisor was Sherri Souffrance. It was
mes New Roman by TopDesk Publishers' Group.

and bound by R. R. Donnelley & Sons Company.

v-Hill books are available at special quantity discounts to use as premiums
s promotions, or for use in corporate training programs. For more informa-
ase write to the Director of Special Sales, McGraw-Hill Professional, Two
aza, New York, NY 10121-2298. Or contact your local bookstore.

This book is printed on recycled, acid-free paper containing a minimum of
50% recycled, de-inked fiber.

mation contained in this work has been obtained by The McGraw-Hill
panies, Inc. ("McGraw-Hill") from sources believed to be reliable.
ever, neither McGraw-Hill nor its authors guarantee the accuracy or com-
ness of any information published herein and neither McGraw-Hill nor its
rs shall be responsible for any errors, omissions, or damages arising out
e of this information. This work is published with the understanding that
raw-Hill and its authors are supplying information but are not attempting
nder engineering or other professional services. If such services are
red, the assistance of an appropriate professional should be sought.

I dedicate this work to the insight of Dr. Alan Hughes.

Alan spotted the fact that I was not very capable in the clean room, and encouraged me to pursue activities that genuinely interested me. Since that time, I have been fortunate enough to not do a single day's work.

I play with some of the planet's biggest, most expensive toys, and people actually pay me for it. Thanks, Alan.

—Chris

Not only that, but Sue cooks up a mean barbecue.

—Judy

Contents

Introduction

IC layout is a very new field. Mask design has been with us for 30-odd years, but has only recently been considered a profession. People wanting to move into that profession—new college graduates and people wanting a career change—are required to know some extremely complicated principles. Likewise, experienced layout engineers find the complexity of modern IC processing requires an ever-increasing understanding of these fundamentals.

Finding a reference that covers everything a layout engineer needs to know has up until now been difficult. Isolated sections of various other works briefly touch on some of the fundamentals that a layout designer should know, but complete coverage, all in one place, has not been available.

IC Layout Basics: A Practical Guide seeks to change that. We want to give new layout designers all the theory and basic understanding they need to become productive and provide a resource they can refer to during their careers. We also want to give experienced layout professionals an increased depth of understanding of the components and techniques they use daily.

This book begins with basic semiconductor theory and continues through the development and construction of the common devices used in modern semiconductor processes. It gives the reader in-depth access to useful design equations, techniques and methods of performing IC layout that should prove invaluable throughout his or her career.

Information is presented in humorous, illustrated and well-paced segments.[1] Anecdotes and asides provide perspective and break the text at key points to increase motivation, without compromising the quality of the highly technical data.

Complete and readable. Welcome to *IC Layout Basics: A Practical Guide.*

Christopher Saint
Judy Saint

---

[1] This may be the first Bedtime Engineering Book.

IC Layout Basics

A Practical Guide

1

Basic Circuit Theory

Chapter Preview

Here's what you're going to see in this chapter:

- Review of basic circuit theory
- Materials that conduct, don't conduct, and partially conduct electricity
- How to make semiconducting material
- Two types of semiconducting materials—negative and positive
- The importance of the junction between these two materials
- Making switches using electric fields
- Putting two Complementary types of switches in series
- Using these Complementary switches as a decision-making circuit
- How to make logic circuits

And more . . .

Opening Thoughts for the Reader

You should already be familiar with most of the circuitry concepts in the first few pages of this chapter, as well as the idea of integrated circuits (IC). We will present a short review as a brief, common reference.

Most of an integrated circuit's functions are achieved by using electrical current in some way—steering current, switching current, or using current to develop a voltage. Much of this steering, switching and voltage creation use what are known as **semiconductor** materials.

Unlike a regular light switch that can only be on or off, a semiconductor switch can be on, off, or somewhere in between. This semiconductor switch is called a **transistor.**

In this chapter, we will build a transistor switch from semiconductor material, then use transistors to develop logic circuits.

Chip design begins with the process development team, continues through your circuit designers, and ends with you, the layout engineer.

You are integral to the successful manufacture of new chips. If you can design your layout with more knowledge, creativity and efficiency, you can save your company millions of dollars. Your chips will tend to work better than expected right off the wafer the first time. They will often be smaller than the design the next layout engineer might have drawn. You will catch and correct disastrous mistakes before production.

You can be immensely valuable to your company as a good layout engineer, particularly as the last person in the pipeline before actual production.

Conventions Used in This Book

- Diagrams will be drawn showing the width of a material as the vertical dimension and length as the horizontal.
- Current will be assumed to be flowing from the left edge to the right edge unless stated otherwise.
- The word "he," and all masculine references, shall include the word "she," or the appropriate feminine reference.[1]
- Illustrations are instructional only and do not portray all real elements or proportions actually involved in a process. We're keeping it simple.
- The reader is to retain a sense of humor, enjoy reading our book, and keep work as fun as possible. Always look for the undiscovered. There is a lot of it out there.

Basic Circuit Review

Following is some basic circuit theory for your quick reference. We only present an overview at this point. Readers are expected to already be familiar with basic circuit equations and concepts. If you need more help with these id see the bibliography for suggested further readings.

Opposites Attract—Likes Repel

Remember the phrase "Opposites Attract." Materials of opposite sign v (polarity) will attract each other, while like signs will repel. For exan atoms with positive charges will attract other atoms with negative chai even at a distance. These same positive atoms will repel atoms with like p tive charges, also at a distance.

Opposites attract.

Without this weird law of nature, the awesome circuits you will see in book would not work.

I want to know why electrons attract and repel at a distance. I mean reall why. How can one little electron have the faintest idea about the next doc neighbor electrons? In fact, positive and negative charges aren't even dif- ferent, there is only a different number of electrons. They shouldn't know anything about that. They can't count.

And why can't we see gravity? And magnets shouldn't work, either! And what really is past the infinity of space? And what is that creamy white stuff inside a Hostess cupcake? It's a frustrating world.—Judy

Units of a Basic Schematic

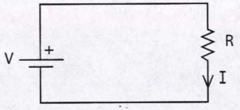


Figure 1–1. Voltage, resistance, and current all exist together in a circuit.

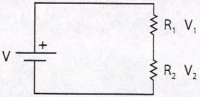
Voltage, *V,* is measured in **volts.**
Resistance, *R,* is measured in **ohms.**
Current, *I,* is measured in **amperes**, or **amps.**

---

[1] At my insistence. I get so distracted by all the his or her, he/she, s/he, he or she or she or he attempts. There must be a better way.—*Judy*

meter	$\frac{1000}{1}$	1000	$1e^{3}$
er	$\frac{1}{1}$	1	1
timeter	$\frac{1}{100}$	0.01	$1e^{-2}$
imeter	$\frac{1}{1000}$	0.001	$1e^{-3}$
rometer (micron)	$\frac{1}{1,000,000}$	0.000001	$1e^{-6}$
ometer	$\frac{1}{1,000,000,000}$	0.000000001	$1e^{-9}$
meter	$\frac{1}{1,000,000,000,000}$	0.000000000001	$1e^{-12}$
tometer	$\frac{1}{1,000,000,000,000,000}$	0.000000000000001	$1e^{-15}$

s Formulas



'igure 1–2. *Two resistors connected in series.*

voltage in a series circuit:

$$V_T = V_1 + V_2 + V_3 + \cdots \qquad \textit{volts}$$

resistance in a series circuit:

$$R_T = R_1 + R_2 + R_3 + \cdots \qquad \textit{ohms}$$

lel Formulas

current in a parallel circuit:

$$I_T = I_1 + I_2 + I_3 + \cdots \qquad \textit{amps}$$

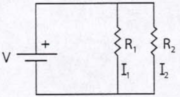


Figure 1–3. *Two resistors connected in parallel.*

Total resistance in a parallel circuit:

$$\frac{1}{R_T} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} + \cdots \qquad \textit{ohms}$$

Ohm's Law

Simply put, **Ohm's Law** states that voltage is equal to the current multiplied by the resistance.

$$V = IR \qquad \textit{volts}$$

Variations of this relationship are

$$I = V/R \text{ and } R = V/I \qquad \textit{amps, ohms}$$

Below is a convenient triangle to help you keep your VIR formulas the right way around.

- In the top of the triangle, you always see voltage.
- The bottom two corners are always current and resistance.
- You look at the triangle to remind yourself of the formulas.
- Use your finger to cover the item to be determined.
- The remaining two letters automatically form the appropriate calculation.

Don't you wish all formulas were this easy?

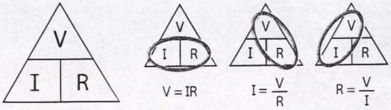


Figure 1–4. *Clever triangle method to remember the Ohm's Law formula variations.*

hoff's Laws

hoff's Voltage Law states that all voltage drops in a closed circuit d add up to the total voltage applied to the circuit. In other words, the nt you put in will equal all the voltage drops that occur in the circuit.

$$V_T = V_1 + V_2 + V_3 + \cdots \qquad \textit{volts}$$

hoff's Current Law states that all the various currents leaving a junc-hould add up to the total current entering the junction.

$$I_T = I_1 + I_2 + I_3 + \cdots \qquad \textit{amps}$$

neans that in any point having some current flowing in and some flow-it, these amounts must be the same. We cannot have more coming in than ng allowed to exit, for example.

eading about Laws is rather boring, but these relationships can be ated into algebraic equations. With equations, we can then solve for g parts.

the importance of having these rules. You get to solve for otherwise wn values. That, and being able to quote fancy names at dinner s.

an effectively consider capacitors and inductors as resistors, although esistance value is sensitive to the frequency of the voltage across them.

It

Ohm's Law to complete the chart below.

oltage *(volts)*	Current *(amps)*	Resistance *(ohms)*
5.2	0.25	
12		200
	0.003	3

nge all measures below into microns.

25 thousandths of an inch
2500 nanometers

(c) 5,000,000,000 femtometers
(d) 0.00045 meter

3. Given two resistors connected in parallel, what is the total resistance of the circuit if Resistor A is 100 ohms and Resistor B is 200 ohms? What if they were both 200 ohms? What if they were both 100 ohms? What if they were both x ohms?

4. In a closed circuit, we have a 12V source. Of our three devices in the circuit, one drops 6V and another drops 4V. How many volts does the third device drop? If someone challenged you on this point at a dinner party, whose work would you cite as your proof?

ANSWERS

1.

Voltage *(volts)*	Current *(amps)*	Resistance *(ohms)*
5.2	0.25	**20.8 ohms**
12	**0.06 A or 60 milliamps**	200
0.009V or 9 millivolts	0.003	3

2.

(a) 25 × 25.4 = 635 microns
(b) 2.5 microns
(c) 5 microns
(d) 450 microns

3. *100 and 200:*

$$\frac{1}{R} = \frac{1}{100} + \frac{1}{200} = 66.67\Omega$$

200 and 200:

$$\frac{1}{R} = \frac{1}{200} + \frac{1}{200} = 100\Omega$$

100 and 100:

$$\frac{1}{R} = \frac{1}{100} + \frac{1}{100} = 50\Omega$$

x and x:

$$\frac{1}{R} = \frac{1}{x} + \frac{1}{x} = 0.5x\Omega$$

4. The total voltage must equal the sum of all the dropped voltages.

$$V_T = V_1 + V_2 + V_3 \qquad \textit{volts}$$

$$12 = 6 + 4 + V_3 \qquad \textit{volts}$$

The third voltage drop must be 2V. We would quote nice Mr. Kirchhoff as our source.

Circuit Diagram Symbols

Below are the conventional symbols we will use to represent our circuit components.

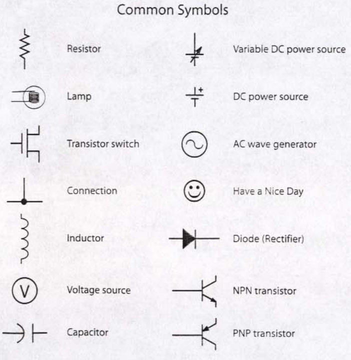


Figure 1–5. Common symbols.

Conductors, Insulators and Semiconductors

A **conductor** has plenty of free electrons that are able to move freely under the influence of a voltage. This is often referred to as a **sea of electrons.**

An **insulator** has no free electrons. The electrons are all held in place by stic bonds to other atoms. They aren't even allowed out on Saturday nights. Since tl must stay put, it is almost impossible for the material to **conduct electricity.**

A **semiconductor** is an insulator that is on the verge of being a conductor does not need much persuasion to conduct electricity. Hence the name, se conductor (*semi* meaning partial). For example, just raising a semiconduct temperature a few degrees can give it the ability to conduct a little electric We could also have called it a semiinsulator, but no one likes a word with t i's in the middle. Semiconductor it is, then.

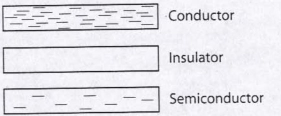


Figure 1–6. How well a material conducts depends on its number of free electrons.

If we can find a way to make a material start or stop conducting wheneve wish, then we can use that material to do useful things for us. We could u to turn on and off electrical devices, or to make logical decisions based o patterns. The possibilities seem endless when we are able to control con tion through a given circuit. This is the power given to us by using semi ductor materials.

Before we can start to understand the properties of a semiconductor, we to understand a few things about the nature of the atoms that form semi ductors. Since we are mainly concerned with trying to move electrons ar in a controlled manner, let's review some Atomic Theory.

Atomic Theory tells us that electrons can only exist in certain energy s surrounding their nucleus. These states are known as **shells.** These shell rather like the orbits of satellites around the Earth. I'm sure you've stu electron shells at some time.

In order to get an electron to move from one shell to the next, we ha add energy to the electron. Give it a shove, so to speak. As we add ener our electron, it will jump suddenly to the next available shell. Once an tron is at the correct energy level, we can use it to conduct electricity f mperatures near absolute zero, all the electrons from the silicon crystal ; are busy holding the crystal together. As the crystal's temperature rises, :oms in the crystal start to vibrate. By the time we get to room tempera- the atoms are vibrating enough to give some electrons sufficient thermal y to break free and jump into the conduction band.

measure the resistivity of pure silicon at room temperature, we see that l has a moderately high value, but it will conduct some electricity purely o the random thermally generated conduction band electrons that are ıble.

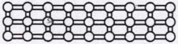
ɔure silicon is not actually very conductive in its raw form. In order to useful devices from silicon, we have to add some small, well-chosen rities that will allow more electrons to be freed at reasonable tempera- By controlling the addition of these impurities, we can conduct electric- a well-controlled manner.

ɪrocess of introducing these impurities is known as **doping.** We will dis- ɾarious methods of doping semiconductors in later chapters. The materi- ;ed to introduce these impurities are called the **dopants.** And, yes, these ; do lend themselves well to some good puns in the lab.

: next section, we will see how to use the doping process to make two rtant semiconducting materials. One that has extra electrons and another ıeeds electrons. We will then control the amount of current flowing :en these two materials. Eventually, we will see how a switch can be made this control.

pe Material

: have discussed, crystals are built of nice, fancy lattices. Rows and rows ɪms, all neatly lined up in a nicely ordered manner. All of the atoms are ɪ to each other evenly. In the heart of the crystal lattice, each atom's elec- are all shared with the surrounding atoms. There are no spares or deficits. ɾ electron has a bonding partner. That's pure silicon for you.



'igure 1–11. One layer of a silicon crystal.

rystal is a very good insulator in this state. There are very few randomly ated free electrons in our crystal that we can use for conduction.

However, once in awhile, we would prefer a few more electrons to be able to flow about the crystal. What can we do to free some more electrons from this nice structure?

Let's replace one atom of silicon with a different element altogether. We will use an element that will want to bond with the surrounding silicon, but will find that it has just one too many electrons for a comfortable fit.

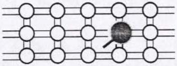


Figure 1–12. *One silicon atom has been replaced by an atom containing one extra electron. We see an extra electron that has no bonding partner.*

Just like a child frozen out during a round of musical chairs, we see a free, unattached electron shouting, "I'm Free! I'm Free! I can do whatever I want!" madly running about, right there in the middle of our silicon crystal.

We must choose this new atom correctly. If it's not the right size, we won't be able to make a good crystal with it. It will damage the lattice. The bonds will be out of place. In addition, the atom needs to have the right number of bonds, the right number of electrons.

By adding just the right impurity, we have ensured that we have a free electron under all conditions all of the time. Nothing is left to chance. We now have a free electron to do some useful stuff. We can put a voltage across this crystal and make this electron travel from one side of the crystal to the other.

Since the charge on our free, traveling electron is considered negative, this arrangement of atoms is known as **N Type** material. *N* for *Negative.* (Notice the word "considered." Have you ever realized there is no such thing, really, as negative? It's just a thought.)

P Type Material

Let's take that same silicon atom out of the chunk of crystal again, but this time replace it with an atom that has one *less* electron than silicon. In this case, we do not have enough electrons in the crystal. We will be one short of matching

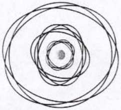


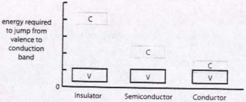
Figure 1–7. *Electrons stay in their shells unless given additional energy.*

ategorize electrons according to their shell, or band. Electrons that hold a ance together are said to be in the **valence band.** Electrons that have gh energy to move freely around are said to be in the **conduction band.**[2] :lectrons in the conduction band are the electrons that flow as electricity.

:onductor, the conduction and valence bands are either touching or overng. One minute you see the electron holding the atom together, and the minute it decides to jump into the conduction band. This means that the ·ons in the material can be easily encouraged to move around, to conduct. become conductive under the influence of a potential difference (voltage).

ı insulator, the conduction and valence bands are very far apart. The :y needed to push an electron from the valence band to the conduction is very high. So high, in fact, that the material will destroy itself before ectrons have enough energy to jump into conduction. That is why you do :e electrons free to move about in an insulator.

emiconductor, though, the conduction and valence bands are rather close ıer. We only need a small amount of energy to make our electron jump



'*igure 1–8.* *Electrons in the valence bands of semiconductors can be asily encouraged to jump into the conduction bands. Within insulators is much more difficult.*

ırs are in a band. The Brit plays a mean lead guitar. The American plays solid rock bass. . covered in authors' other book.) I can't believe McGraw-Hill is letting us write stuff like this in our book.

into the conduction band. Semiconductor material can easily be made into a conducting material this way.

Semiconductor Material

Silicon is an element with very little energy difference between its conduction band and its valence band. This small difference in energy makes it a very popular material for use in IC's. Not much effort is necessary to encourage the silicon's electrons into the higher band, freeing them up for conduction through the silicon.

Luckily, silicon is very abundant. It is found all over the planet on our sandy beaches in the form of silicon dioxide (an atom of silicon with two oxygen atoms, SiO_2). The silicon and the oxygen are attached to each other by their electron **bonds.** These sticky little connections refuse to allow electrons to leave the molecule. The straight lines in the diagram represent these shared electron bonds.

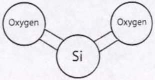


Figure 1–9. *Two atoms of oxygen bond with a silicon atom to form SiO_2.*

Now, if we were to leave this molecule unprotected in the wrong neighborhood at night, we would come back the next day to find the two oxygen atoms totally stripped from our molecule of sand. We would be left with just the one silicon atom. This can also be done in the lab.

Once stripped of the oxygen atoms, we can organize the silicon atoms into very large crystals, just like diamonds. This makes a pure material, whose electrons can be rather easily coaxed into loosening their grip, as we will see below.

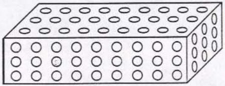


Figure 1–10. *Many silicon atoms form a nicely organized crystal.*

ever, there's that fence in the way. They can't get over to the other side. fence between the P and N Type materials is called a **Potential Barrier.** ing a piece of P Type material to a piece of N Type material is known as ing a **PN Junction.**

can't our electrons move across the PN junction, into the other material?

n we chose our two dopants, we were very careful. On the N side, we anted elements whose electrons had lower conduction energy levels. On side, we implanted elements whose electrons had slightly higher con- ion energy levels. Our extra electrons on the N side are in a lower energy than is desired by the choosy hole. The P material just has higher stan- s, that's all.

ve need to do now is add some energy into the system so we can push our electrons over the fence to fulfill their destiny. The more energy we add, nore can jump over to the P Type material. We can control the number of rons flowing by controlling the amount of energy we apply to the N Type rial.

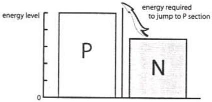


Figure 1–16. The free electrons in N material cannot fill the need in the P material without additional energy.

re can we get the additional energy that will kick them over the fence? , that's an easy one to answer. We add energy into the system by applying tage across our junction. Let's look at that next.

trolling Flow across the Barrier

ou increase the voltage across the junction, electrons in the N type mate- vill have enough energy to overcome the barrier. Current will start to flow. rons will go one way, the holes will go the other way, in a manner of cing. So, if you connect the battery to the correct side, and it's big enough, l get a current across the PN junction.

Remember that the direction of travel of electrons is directly opposite the direction of travel of current. **Conventional flow,** as seen in schematic diagrams, is literally backward from actual electron flow. People got it backward a long time ago and I wish someone would fix it.

Oh well. Current flows if you add energy. That's the main point. It just happens to be indicated in the opposite direction than the electrons actually flow.

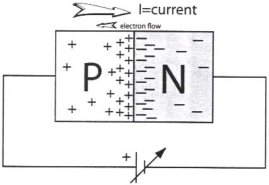


Figure 1–17. Applying voltage will allow N electrons to flow. Remember that conventional flow is opposite electron flow.

Rule of Thumb: If, after working too long, you suddenly think you have the current backward, you probably don't. Take a nap, then look at it again later.

As you begin to apply a positive voltage to the P Type material, the electrons in the N Type material will be attracted to a greater degree. The N Type electrons will be given more energy due to the greater attraction. As the positive voltage increases, the electrons will get more and more energy, and the difference between the two materials' energy bands will decrease.

As the energy bands get closer, thermal energy in the system will start to push electrons randomly over the brink. They will cross the junction. Thus, conduction has begun.

Increasing the applied voltage further will continue to increase the attraction of the electrons, until all the electrons have enough energy to cross the barrier.

ıis point, the junction is conducting fully. We now have a resistor. The cur- increases linearly with the voltage across it.

:n a junction is in conduction, it is said to be **Forward Biased.** Holes are ing forward across the junction, electrons are moving backward across the tion.

e increase our voltage even more, the current flow finally stops increasing linear fashion. Eventually, the amount of current levels off and becomes ;tant.

t level value is called the **saturation current.** You are getting as much cur- as you reasonably can.

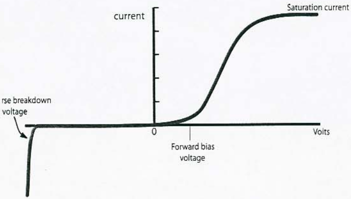


Figure 1–18. *After the initial forward bias voltage, the increase in current is linear until it reaches saturation.*

ʌ, if you were to connect the battery around the other way, you would be ucing the energy from the side that needs it. Instead of reducing the differ- e in energy bands, we would be increasing the difference even further. :hing would happen. Nothing moves. In this nonconducting condition, the ction is said to be **Reverse Biased.**

'ou increase the voltage enough in this reverse direction, however, a point l be reached where the junction gives up and breaks into sudden conduc- ı. This is the **Reverse Breakdown Voltage.** The reverse breakdown voltage ıld typically be quite high, though, on an IC. So, for most of our purposes, can consider a reverse biased junction to be nonconducting.

A reverse biased junction is very useful in an IC. We will see why later. But, it just might have something to do with the fact that it doesn't conduct except under extreme conditions. And that's the only hint you get.

Diode

This junction between P and N materials is what we call a **diode.** Current will only flow in one direction, as we saw in the last section. Notice that a direction arrow is used as the symbol for diode. That makes the symbol easier to remember.

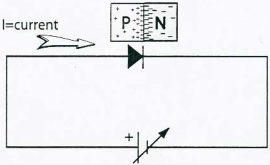


Figure 1–19. *We use the diode symbol to represent a PN junction.*

That's the basic semiconductor junction, its name, its symbol, and how it works. In the next section, we will use this one-way control to create some useful circuits.

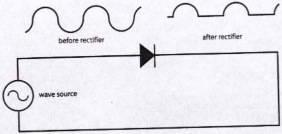
Diode Applications

A diode by itself has only a few useful applications. We can't do much at all with just a diode.

You could put a sine wave into the diode to pass just the positive portions on through to the rest of the circuit. Eliminating the negative portions is called **rectifying** the signal.

If you put a voltage across a diode, you could see what rectifying looks like. Look at the current as it flows through. You'll see current, then no current, then current, then no current, and so on. It's a good way to rectify a signal, but you lose half the power. That seems wasteful.

ld arrange diodes in such a way that the positive and negative cycles h rectified. We get the other half of the cycle back this way, and do not power. This is known as **Full Wave Rectification.** It is a very common que used in power supplies.



igure 1–20. Current can only flow in one direction through a diode. A iode can be used to rectify a voltage.

es, that is PN junctions, have one very useful effect in a circuit.

A PN junction gives us isolation.

u recall, by applying a voltage across a junction, we can cause current to in one direction. It just will not work in the other direction. And, as we discuss in the next section, this ability to isolate electrical flow is a mighty erful tool. That's the value of a PN junction. It isolates and controls our ent flow, so that we do not have current just flowing freely all over the e on our IC.

will need this isolation capability a bit later in this book. Wait for it . . . t for it . . .

can use a diode in a crystal radio. A Cat's Whisker (crystal radio) s a PN junction to demodulate amplitude-modulated radio waves. It's ectifier. It allows current to go one way but not the other way. In the radios, you had a point junction. You had to push this crystal down, so at's whisker is really a diode.

Semiconductor Switches

Enough talk. We want to make something that can actually do work for us. Something that will give us control.

We want to make a switch. We'll switch something on. We want to turn on a light bulb!

That's it: We want to turn on a light bulb!

Because we're power freaks.

Lighting a Bulb

We have our light bulb. We have our battery. How do we turn on our light bulb? We will put a switch in the way. Now we have control.

We can push the switch up, making a connection in the circuit, and the light bulb turns on. If we let the switch down, the connection opens, the bulb turns off. Push up again, it's on again. Off. On. Off. On.

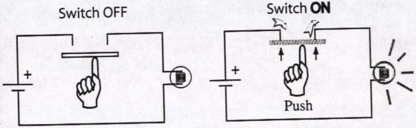


Figure 1–21. Semiconductor material can be used as an electronic switch that works just like pushing a contact bar into place.

So, by joining these two wires together across a contact bar, acting as a switch, the light bulb turns on.

Now, we can physically push all our switches against two contacts with our fingers to complete our circuits, but things might be easier if we could do this electronically. We just don't have enough fingers. Even if you invite some friends over to help.

t's take exactly the same light bulb circuit, but instead of our finger on the itch, we will place an independent voltage nearby. Make it a variable volt- e. In our circuit diagram, we will symbolize our switch with the little push , so that you can remember it easily.

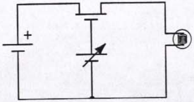


Figure 1–22. The symbol for a semiconductor switch looks a lot like the contact bar which we earlier pushed into place.

we can make a switch that would turn that light bulb on and off, depending n the voltage we have placed in the vicinity, we could control our light using ectricity instead of fingers. We just need a way to use voltage to allow or dis- low conduction, to open or close a switch.

ell, here's the answer. Put a semiconductor there. We have already seen that e conductivity of semiconductors can be controlled. Sometimes they will nduct, sometimes they will not. That sounds like a workable switch to me.

hat's what we will do. We will use semiconducting material as a switch.

n the next section, we will discuss why a semiconductor works like a finger ushing a switch, depending on a nearby voltage. We will not even require cur- ent to operate the switch. Total control. No energy spent. An electronics engi- eer's dream come true.

The Field Effect

emiconductors have very interesting properties. One particularly useful prop- rty is known as the **Field Effect.**

f we take a piece of N Type semiconductor material and apply a voltage across t, we will get some current to flow.

et's place a voltage *near* the semiconductor, not even touching it. Let's make t positive. The voltage attracts all the electrons in the semiconductor, even at distance. Opposites attract, remember.

The effect of the nearby voltage creates a gathered field of electrons in the N Type material. Such a positive or negative field generated from a distance is called Field Effect.

The electric field that the voltage exhibits has effectively increased the number of electrons near the surface of the semiconductor and consequently the resistance drops (more free electrons). If the resistance drops, we get more current. This still is not a switch though. We want to stop current flowing, not increase it!

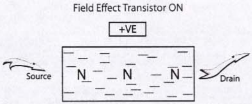


Figure 1–23. If the semiconductor material enjoys free electrons all about, electricity can flow in the source and out the drain.

Now let's place a negative voltage *near* the semiconductor material. All the electrons scream, "Eewwww, negative!" and they all push over to the other side. (Like Charges Repel.) Our negative voltage, just by being close to the semiconductor material, pushes all the electrons away. By pushing the electrons away, we are effectively creating holes, or P Type material, near our negative voltage area.

As you increase the negative voltage, eventually more and more holes come gathering. The generated positive holes are attracted to the negative voltage. (Opposites attract.)

When we finally have generated enough P Type material, we have in fact made two PN junctions across the middle of our semiconductor. One junction is forward biased, but the other is reverse biased. We have learned that electrons cannot cross a reverse biased PN junction easily. No current flows.

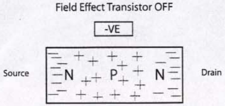


Figure 1–24. If a field of holes has been attracted to the center of the semiconductor material, the PN junctions prevent current flow.

At last, the switch we have been searching for! All we did was wave a negative voltage next to a piece of semiconductor and PN junctions pop up out of virtually nowhere. Now we are getting somewhere.

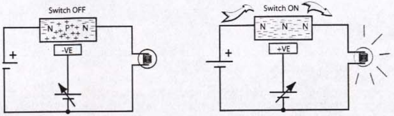


Figure 1–25. *Now we do not have to use our fingers. We can control switches electronically by varying the nearby voltage.*

Making an FET

We now have the capability of changing N Type semiconductor material into P Type and back again just by changing a nearby voltage. The effect of this nearby voltage turns on or off electron flow through the semiconductor. This device is called a **Field Effect Transistor (FET).**

Let's learn the various parts of the transistor we have just made.

When our transistor is ON, electrons flow in from one end, called the **source.** They fall out the other end, called the **drain.** The source and drain are interchangeable. The polarity of our voltage across the transistor determines which end is a source, and which end is a drain.

The terminal we use to apply our nearby controlling voltage is placed near the middle of the semiconductor, called a **Gate.** The name, *Gate,* is appropriate. It lets things pass through. The voltage on this Gate generates the field effect.

With small Gate voltages, we will only generate a small field. As a result, we will only generate a small P Type region. Some electrons can still get through, but not as many. With additional voltage, we can **invert** even more of the center region (change N to P) until there is so much P Type that no electrons can get through at all. Once the P field spans the entire depth of the semiconductor, it's called **pinched off,** like putting your foot on a hosepipe. As you pinch the hosepipe down, the water will stop flowing.

This is very different from a regular mechanical light switch. In a regular switch, the switch is either on or off. There is no middle ground. In a transistor, the transistor can be on, off, or somewhere in between.

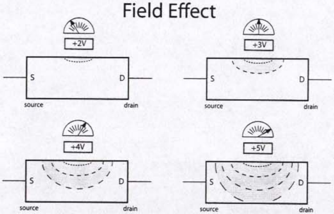


Figure 1–26. *The field in the semiconductor material gets larger as we increase the Gate voltage. At high enough voltage the field pinches off current flow altogether.*

Isolating Our Switches

Just one transistor is fun, but we need to find a way to make a bunch of trar sistors that we can turn on and off whenever we want.

Let's say we have three light bulbs we want to turn on individually. We nee three light bulbs. We need a battery. How can we most easily make three trar sistors? Putting each into its own plastic package is expensive. Let's see hov to easily make three transistors at the same time from a single piece of silicor

We also need to be sure that our three transistors are independent from eac other. After all, we do not necessarily want all of our light bulbs to turn on a the same time.

This sounds like the perfect job for a PN junction. You were just about to sa that very thing, weren't you? Reverse biased PN junctions can be used to isc late areas where you want to control the flow of current. In this case, to kee current from flowing at all. That will keep our three transistors separate fror each other, even though they share the same silicon bed.

Make a long strip of N in your silicon. Put P wherever you want to cut o one piece of N from the next piece of N. This gives you lots of transistors i a row from the same strip of N. Convenient, and part of your existing proces steps, as well. This saves you a lot of money. Well, not you personally. Yo company.

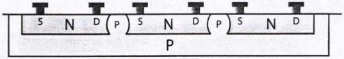


Figure 1–27. We can use PN junctions to isolate transistors from each other in a long strip of N.

k under the transistors. We have P Type material everywhere. Effectively, have a diode completely surrounding our transistor; left, right, ahead, ind, and underneath. A PN junction can give you isolation in all directions.

could use P to separate segments of N material, as mentioned above. Or, can start with something that's completely P Type throughout. Solid P Type, rywhere. Then we could embed smaller regions of N directly into the P rever we want a transistor. The N Type regions are automatically insulated n each other by the PN junctions created.

we need to do now is place our Gate material directly onto the surface of silicon to turn the N material on and off. The only problem with this is that our Gate would short to the middle of our transistor. Not a good idea. We ld just float our little voltage Gates over the top of each N region, but hav- something floating in midair hasn't yet been perfected. The Gates just tend all to the surface of the silicon.

need to place a very thin insulator between the silicon and our Gate. The ner the better, in fact. If we were to float our Gate too far away then the d effect will be too weak. Let's get the Gate in right close.

st happens that silicon dioxide is a very good insulator, even when it is y, very thin. We are building our transistors out of silicon, anyway, remem- Surprise, surprise, if you heat the silicon up, the oxygen in the air will t with the silicon, giving you silicon dioxide.

need to do some fancy stuff in the processing to control where you want ut it's easy to make. Pretty handy little insulator.

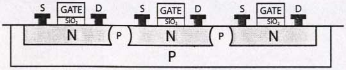


Figure 1–28. Each transistor will have its own Gate.

Up until now, we have been concentrating on N type transistors. There is nothing to stop us from making a transistor of P type material. A P Type transistor needs to be built in an N Type region, however. We need to isolate our P Type transistors from each other, of course. That's why they are built in an N region. The isolation from the diodes, again. Both N Type and P Type work well as transistor switches.

Well, there you have it. That's a basic field effect transistor, FET, taken from semiconductor theory. Now we have a way to turn current flow on and off electronically. From here, you can design decision-making circuits. We'll save that for dessert at the end of the chapter.

In the next section, we will see how the basic FET can be built more accurately, and in a way that makes our switching faster. Much faster. Much, much faster.

Enhanced versus Depletion

In the diagram below, we see an FET.[3] The P Type material isolates the N Type materials. The Gate floating on top controls the field in the semiconductor. If the voltage on the Gate is made negative, we see that the transistor is off, due to the field of P Type material created. The electrons that carry the current in our N Type material have been reduced or depleted in the center region of the transistor. A transistor that uses this method of operation is called a **Normally**

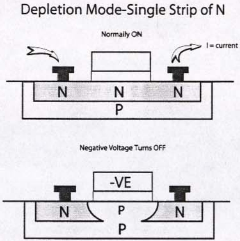


Figure 1–29. Just as we have built so far in this book, we see a Depletion Mode FET in the on and off conditions.

---

[3] You are supposed to say "Field Effect Transistor" every time you see FET, so I'm supposed to write "a FET," but I don't want to. Just letting you know that I do know my rules. But write to me anyway.

On, or a **Depletion Mode** transistor. They are simple, basic, straightforward strips of N Type material, just as we have studied up until now in this book.

One of the most important factors in modern transistor circuits is the speed at which the transistor can turn on and off. A transistor with a large field region in the center will take more time to turn on and off as there is more energy required moving the electrons and holes around a larger area. It is therefore desirable to try to make the center field region, the Gate region, as small as possible, in order to make the transistor switch as quickly as possible.

A Depletion Mode transistor such as we have made so far, suffers from a problem. Look at the diagram that illustrates the field effect. As the field increases, it spreads out from the Gate. The field gets bigger!! The exact opposite of what we want. A smaller field would switch faster. How can we overcome this problem? Now, here's where somebody was very clever.

Instead of making a long strip of N Type material, then placing the Gate over the top of it, let's set down the Gate *first* then use it to prevent the silicon underneath it from being implanted with N. Let me show you why this is so clever.

With the Gate already placed in position before we implant our N material, we have a natural protection in the middle region of our new transistor. We can now bombard the surface of our P Type silicon with N Type atoms and get a perfectly aligned pair of N Type regions on either side of our Gate, but not underneath. That's the ticket.

Sure, the Gate is hit by the N atoms as well, but the Gate material is thick and only the surface gets affected. Not a problem. For a more detailed description of how this atom bombardment or **Implantation** works, see the next chapter.

So, now we still have the P Type region directly underneath our Gate instead of an N Type region. This means that we have to use a positive voltage to invert the Gate region instead of a negative voltage.

Look at Figure 1-30. Just as before, the field effect spreads into the region under the Gate. With a positive voltage on the Gate, the underlying P Type material changes into N Type. The N region spreads out until it touches the N regions either side and we start to get a continuous N region formed. The transistor is now turned on.

This transistor starts naturally in an Off state and changes into an On state. With a positive Gate voltage, we are attracting electrons. This enhances the region between the source and drain.

This type of transistor is **Normally Off.** It is called an **Enhancement Mode** transistor, named for the enhanced region in the center. The other useful property that this type of transistor has is that its Gate region does not ger as we change the Gate voltage. In fact, the region under the Gate ends up smaller than the actual Gate itself, since the N Type source and drain regions actually creep underneath the Gate during processing.

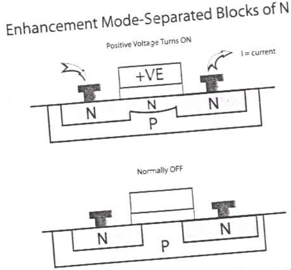


Figure 1–30. *If we make the Gate first, the N implant cannot fall underneath the Gate area, forming perfect alignment..*

This structure gives us exactly what we want. A transistor that has a small Gate region that does not vary as we use the transistor.

We will be using enhancement mode transistors as we continue in this book.

Even if our transistors are built using Enhanced Mode, with the Gate applied before implanting our silicon, we still can only turn current on and off. We want more. We want to make decisions. We want to combine these things in large arrays.

You say, "OK, I need to turn this bulb on AND that bulb on." Or, you might say, "I don't care which, either turn on this bulb or that other bulb, as long as one is on." You need AND and OR decisions.

Here come logic circuits. Now we're talking. It's all in how you hook these things up to each other. What feeds your Gate voltage can come from fancy places. Now, what sort of fancy place do you think I mean? I'll leave you wondering for awhile.

Let's see how to play with a handful of transistors to do some fancy stuff.

Complementary Switches

hat's even more useful than Enhancement Mode is what's called mplementary Metal Oxide Semiconductor, known as CMOS.

hat do we mean when we say the word Complementary? As you recall, an N pe transistor needs a positive voltage to turn it on, but a P Type transistor eds a negative voltage to turn it on. The two types of transistors are both on d off switches, but turn on and off in exactly the opposite fashion. In that y, the transistors complement each other. They work together as a perfectly ordinated team, just like Laurel and Hardy.[4] Most of the modern IC's today e CMOS as their basic technology.

Type and N Type devices are known as complementary transistors. If we ace these transistors next door to each other, we can start to build useful circuits. We have to wire them up correctly of course—just placing them next to ch other won't do us much good.

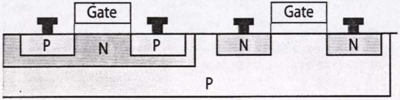


Figure 1–31. Implanting P around the left Gate and implanting N around the right Gate form complementary switches.

N Well and Substrate Contacts

here is one final thing left to consider before we can start using these devices.

ook again at the P type device. It has a region of N type material that is just tting there, not doing much. Likewise, the P type substrate material has been ft alone, not doing much. If we are not careful, these two regions could evelop some voltage. This voltage could come from stray currents that have aked out of our real devices. The PN junction formed from the P type evice's N region could become Forward Biased. All sorts of catastrophic ings could happen.

---

[4] [illegible] Laurel and Hardy into a technical book.

We need to ensure that these two regions can never become Forward Biased. The best way to do this is to intentionally Reverse Bias them. We need to connect the substrate to the most negative potential in our circuit, usually the negative supply; and connect the N region of our P type device to the most positive potential in our circuit, usually the positive supply.

The **N** type diffusion is fairly deep, like a water well. Therefore, the N region is usually referred to as an **N Well,** or **tub.** The device is built in a tub of N type diffusion. Every device in our circuit must have its well or substrate tied to the correct potential. Some technologies also have a second **P Well** for the N type devices to be built in. A two-well process is uncommon, though.

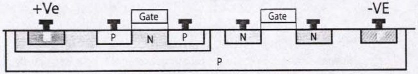


Figure 1–32. N Well connects to positive power supply. P Well connects to negative power supply.

The contacts we need to add are known as **N Well Contacts** and **Substrate Contacts.** We will discuss more about substrate in later chapters. Even with our well and substrate tied to the correct supplies, there is still a chance that circuit operation will cause the well/substrate junction to Forward Bias. This phenomenon is called **Latch-Up,** which can cause a chip to die horribly.

Now we're ready to do all sorts of things using Complementary switches. You'll be amazed at the work these little things can do. Let's see if we can get them to move a piano up a long outdoor flight of stairs.

Building Logic Circuits

We have messed around with light bulbs long enough. We will now see how to use transistors to create circuits that can perform binary logic functions.

Using Voltage as a Logic State

Up to this point, we have been using our transistors to turn on and off current. However, using current to represent a binary value is very wasteful of power. We would quickly wear down our battery if we used current to represent a binary value.

Using voltage to represent a binary value works much better. Remember that CMOS transistors are operated by voltage, so if we can design a circuit that uses transistors to switch voltage instead of current, then we can use these circuits to switch each other. If we can get our circuits to switch each other, then we can string all of our logic circuits together into much larger useful systems.

Switching voltage also has another advantage. Current will only flow in our circuit during the time that the transistors are switching. Once the transistors have changed state, no current will flow. That saves the battery. Now, let's examine some logic devices made using CMOS transistors.

CMOS Logic Circuits

If we want to make a circuit that uses binary logic, such as building a calculator, we can represent our binary values using voltage states.

Let's represent our binary one by the positive battery voltage, and our binary zero by the battery's negative voltage. The high and low voltage values become logic one and logic zero.

If we connect our CMOS devices in the manner shown in the diagram below, what happens? Trace through the diagram imagining a high voltage (positive) applied to the joined Gates. Use what you know about complementary switches, and see what you get. Then try it again imagining a low voltage (negative) applied to the joined Gates. Now, what do you get?

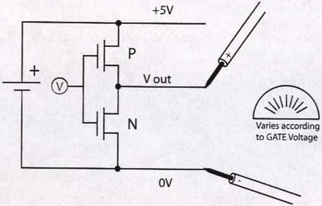


Figure 1–33. Using a Complementary set of switches with a common Gate voltage source.

The Gates of the N and P Type transistors are joined together so both Gate will have the same voltage applied at the same time. Remember the operatio of Enhanced Mode transistors: N Type devices turn on when a *positive* voltag is applied to their Gates, and P Type devices turn on when a *negative* voltag is applied to their Gates.

If we connect the two Gates to the positive supply—a logic one—wha happens?

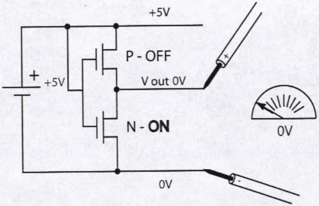


Figure 1–34. Inverter with high voltage as the common input to the Gates.

The N device will turn on because we have applied a positive voltage to i Gate. What happens to the P device? The P device remains in its normally o state. (The P device would need a negative voltage applied at the Gate in ord to turn on.) N is on. P is off.

The two devices have their source-drains connected. This means the drain one transistor is connected to the source of the other transistor. This gives u common output. If we measure the voltage at this common output point, will see the negative voltage of the battery—a logic zero—coming through t N device that is switched on.

If we connect the two Gates to the negative side of the battery—a logic zero what happens?

The P device turns on because we have applied a negative voltage to its Ga What happens to the N device? The N device remains in its normally off sta (The N device would need a positive voltage applied at the Gate in order turn on.) P is on. N is off.

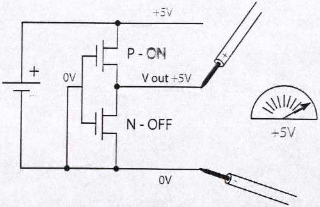


Figure 1–35. *Inverter with low voltage as the common input to the Gates.*

in, the two devices have their source-drains connected. So, if we measure voltage at the common output point, we will see the positive voltage of the ery—a logic one—coming through the P device that is switched on.

you notice that zero in gives one out, and one in gives zero out? It reverses logic state. So, using the above circuit, we have been able to create what is wn as an **Inverter.** The circuit inverts the logic state that is connected to its es. High input gives us low output. Low input gives us high output.

now have the ability to invert high and low states. Our first logic circuit. should be proud.

It

n a sheet of paper, without looking, try to reproduce the inverter chematic as it appears in the book. Follow through the logic of your rawing to convince yourself you have drawn it correctly. Then draw it gain, same thing.

ook back as often as you need to, but keep trying until you can jot it out asily time after time without looking.

ry the same thing at the end of the chapter, once you have seen systems at are more complex. If they make sense to you, you should be able to ason through the drawings of each circuit as you draw them.

Do that many, many times, and soon you will have to buy a new pencil. Plus, this is excellent mental preparation for becoming intimately fluent with the logic of these devices.

The Gate connections are usually thought of as the device input, and the common source-drain connection is usually thought of as the device output.

If we now connect two inverters together what happens? You can either play with the thought for a moment, or just read on. (Oh, just take the time. It makes life more fun to participate. Here, I'll help you. What if the output of the first became the input of the next? Stop. Think in steps. Draw a picture.)

The second inverter has its input taken from the output of the previous inverter. As far as the second inverter is concerned, it is still getting a zero or one applied, even though it is being supplied through transistors rather than directly from the battery.

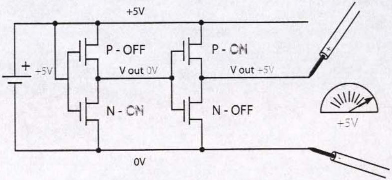


Figure 1–36. *You don't have to connect every input to a battery. It could come from the output of another device.*

Our final output has been inverted twice. High in gives us high out. Low in gives us low out.

An inverter is the simplest form of logic Gate. Every microprocessor has thousands of these inside. An inverter by itself, though, does not allow us to do very much. Changing between high and low is only interesting for so long, even if you connect two in a row. We need to get a bit more complicated before we can build a real microprocessor.

Below you will see diagrams of two of the more interesting logic circuits. Take a few minutes to mentally trace a high or low voltage through these systems.

ook at the table for each circuit. Both circuits have two Gate input voltages, and B. Either voltage could be high or low, independent from the other. The tput is called Z.

rify for yourself that the schematics truly result in NAND and NOR funcns as shown in the charts. High voltage is ON, shown in the chart as 1. Low ltage is OFF, shown in the chart as 0.

AND Gate

e **AND** function is defined as one that requires both A **AND** B to be at a gic one in order for the output to be at a logic one.

AND, however, means "not AND," meaning "opposite of AND." A NAND nction is the AND function that has results inverted. All the output ones from ND are zeros and all the output zeros from AND are ones. Then you have a AND. The only way to retrieve a logic zero is for both A and B to be at logic e. See the chart.

2 Input NAND GATE

A	0	0	1	1
B	0	1	0	1
Z	**1**	**1**	**1**	**0**

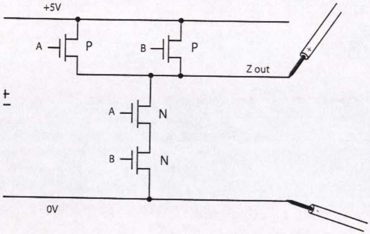


Figure 1–37. Two input NAND Gate.

2 Input NOR GATE

A	0	0	1	1
B	0	1	0	1
Z	**1**	**0**	**0**	**0**

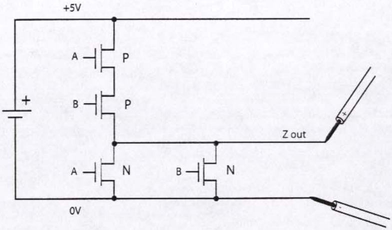


Figure 1–38. Two input NOR Gate.

NOR Gate

The **OR** function is defined as one that requires either A OR B, or both, to be at a logic one in order for the output to be at a logic one.

NOR means "not OR," meaning "opposite of OR." A **NOR** function is the OR function whose results have been inverted. All the output ones from the OR circuit are zeros and all the output zeros from the OR circuit are ones. Then you have the NOR Gate. See the chart.

Try these circuits by imagining all the combinations of A and B states. Trace mentally through the Gates as you run your fingers along the diagram. Decide your outputs. Verify the charts.

Closure on Basic Circuit Theory

The above theory is basic to all integrated circuits. We have taken some very simple concepts, strung them together one at a time, and gone from a simple atom to a complicated logic Gate.

The above material often requires three to six months of study in some college courses. Usually in great depth, and usually accompanied with lots of heavy formulas. Very heavy. Weigh the books, you'll see.

If you have been able to follow along, able to understand these concepts, then you will have a great background that will enable you to understand what you are trying to lay out. This understanding will stay with you throughout your professional life as a mask designer and enable you to be creative and productive. You are making a one-time investment in learning that will pay dividends the rest of your career.

> *I, personally, find it essential to understand the processing technology. When you've got 20 or 30 layers of really complicated process steps, understanding what each layer does lets you understand the technology and lets you understand how to come up with novel layout.*
>
> *You have to understand what you're drawing. You can't just blindly jump in and hope. Some of the most successful layout people I've known say to themselves something like, "Ok, if I combine this Diffusion, that Polysilicon, and that Metal, instead of having a resistor, diode and transistor, I can combine them into one fancy piece of layout that's half the size."*
>
> *If you can reduce size, you reduce cost. You get more in the same area. You get more bang for the buck. If you don't understand what the layers are, you get stuck using too many things in your layout.*
>
> *You don't have to understand every little bit of energy, like electrons jumping across a PN junction, but it's good background for making a transistor. It explains why the field effect does what it does.*
>
> *This is developmental background. You'll end up using it all. The better you understand this stuff the less aware you will be that you are using it. When it finally becomes intuitive, you'll think you've known this stuff all your life.*

That's enough theory, let's get into the real interesting stuff.

The next topic to discuss is *how* we get the extra atoms into the silicon. We don't just use a pair of tweezers. How we get them in determines what we draw as a layout engineer.

Here's What We've Learned

Here's what you saw in this chapter:

- $V = IR$ and other review formulas from series and parallel circuits
- Definitions of insulators, conductors, and semiconductors
- Why we dope silicon
- Definition of N Type, P Type material
- PN junction barrier as a rectifier, diode
- Using the PN junction to isolate transistors
- Making transistors efficiently by placing N regions into a large P region
- Semiconductor Switches
- Field Effect Transistors
- Complementary switches
- Using CMOS as a logic device
- Using voltage instead of current as the logic state determinant
- NAND and NOR logic Gates

And more . . .

Application to Try on Your Own

Trace through each schematic below to determine the Output State, given each set of input conditions in the chart. Write the result in the chart as a binary zero or one. Be careful. It might get tricky.

First Level: EASY

Easy 3 Input System

A	0	0	0	0	1	1	1	1
B	0	0	1	1	0	0	1	1
C	0	1	0	1	0	1	0	1
Z								

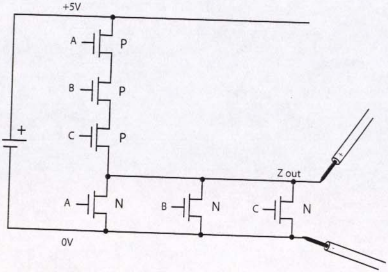


Figure 1–39. *Easy diagram.*

Second Level: MEDIUM

Medium 3 Input System

A	0	0	0	0	1	1	1	1
B	0	0	1	1	0	0	1	1
C	0	1	0	1	0	1	0	1
Z								

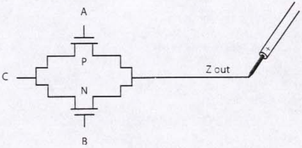


Figure 1–40. *Medium diagram.*

ird Level: HARD

Hard 3 Input System

A	0	0	0	0	1	1	1	1
B	0	0	1	1	0	0	1	1
C	0	1	0	1	0	1	0	1
Z								

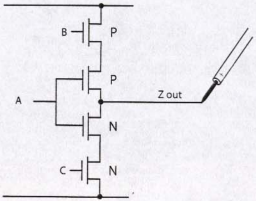


Figure 1–41. *Hard diagram.*

ANSWERS

1. Easy Circuit

A	0	0	0	0	1	1	1	1
B	0	0	1	1	0	0	1	1
C	0	1	0	1	0	1	0	1
Z	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**0**

This is an easy one. This circuit performs a NAND function. It is a 3-Input NAND Gate. All three P devices need to be ON (A, B, and C) in order for the output to go high. All other states turn on at least one N device.

2. Medium Difficulty Circuit

A	0	0	0	0	1	1	1	1
B	0	0	1	1	0	0	1	1
C	0	1	0	1	0	1	0	1
Z	**0**	**1**	**0**	**1**	**X**	**X**	**0**	**1**

In this configuration, we find that there is a situation where we don't know what the output can be. The situation is when both transistors are OFF (A = 1, B = 0). In this case, it doesn't matter what happens to C. We cannot see what C is doing through two OFF transistors. In all the other cases, one of the transistors is ON, and the output is just passed through unchanged.

This configuration is known as a Transmission Gate. It is most typically used with both transistors ON. Notice that A and B are the opposite states (inverted). Our inverter would drive this well.

3. Hard Circuit

A	0	0	0	0	1	1	1	1
B	0	0	1	1	0	0	1	1
C	0	1	0	1	0	1	0	1
Z	**X**	**1**	**X**	**X**	**X**	**0**	**X**	**X**

This configuration is most commonly used in output cells. The two middle devices can be thought of as an inverter, but the top and bottom transistor need to be ON in order for the inverter to "see" a supply voltage. When the B and C transistors are ON, the inverter acts normally. In all the other conditions, the output is effectively disconnected.

This disconnected state is sometimes referred to as Tri-State. It is a third state in a normally two-state system.

CHAPTER 2

Wafer Processing

Chapter Preview

Here's what you're going to see in this chapter:

- How we put impurities into a semiconductor
- What effects those impurities have
- How we can add material to a semiconductor
- How we can remove material from a semiconductor
- How we put these changes exactly where we want them
- How the wafer is processed and built upward
- How the processing steps determine what we draw
- How we can join processes for efficiency

And more . . .

Opening Thoughts on Wafer Processing

In the first chapter, we introduced an impurity into a silicon crystal to free some electrons. Then we sent those free electrons off to do work for us. These N Type and P Type materials were shown as regions that could be placed anywhere on a chip. Now it's time to learn how that placement happens.

In this chapter, we will see how chips are actually grown, slimed, blasted and gassed, depending on what we draw on our computer screens. We will build an IC from scratch to see how our drawings are used to create our product.

Starting with our substrate wafer, we can do three things: We can change it, add to it, or remove material from it. Once we understand how to make these changes, we will see how we can change, add or remove material from just the specific areas we want.

IC Layout

An IC is built sequentially in steps. The process steps work together to add layers to the IC. The chip gets thicker as the fabrication continues. Each step in the process has its own representative drawing. All these drawings relate to each other, forming three-dimensional components that operate together to form a working IC chip. IC **layout** is the process of creating the two-dimensional representations of the fabrication layers, which are then used to manufacture an IC chip.

When IC's were first being developed, the drawings were placed on clear plastic film using red see-through sticky tape. Each layer had its own film. They all had to line up with each other extremely accurately. Today we use **Computer Aided Drafting (CAD)** tools to draw each layer of our IC. The layers still have to line up extremely accurately, of course.

The question is, "How does what we draw in our CAD tool relate to what's really happening in the IC fabrication process?" In other words, how do our 2-D drawings result in 3-D products? Understanding the processes behind your rectangles will help you follow the proper design rules of manufacturing.

You will use one of several computer aided tools to check your layout against the process design rules. Sometimes, the output from those computer checkers is really weird—very hard to interpret. The results may mean a lot to the person who wrote the rules, but they may not mean too much to me as a layout engineer. That's why we need to know more about how our drawings are translated into what become real layers.

Here's a transistor, already built. Someone's done the layout for you. All you're worried about are three connections. You just need to wire to those three connections in metal.

Most layout work is done like this, just wiring pre-laid-out components using multiple metal layers. However, an understanding of what you are wiring is essential.

For example, you might think, "Ok, I don't care about this wire. It's just a metal, I can put that wherever I want. I can wire over the top of the transistor." Then when the chip comes back from the wafer fab you have a shorted out transistor. Oops.

"Well, it's blue—that's the same color. It's a different shade, but who cares?" Your layout will light up like a Christmas tree when you run your rules checks. You'll get lots of weird error messages like, "This is too close to this buried layer." To which you respond, "Huh?"

Finally when you ask your supervisors what this error message means, they will tell you that even though the layer you used is usually ok for wiring, it also is used to provide some other functions inside a transistor. If you run that layer over a transistor, you are in big trouble.

As you learn how IC's are manufactured, you should begin to develop a 3-D sense of what is happening with your rectangles. It is this understanding that separates the master layout engineers from mere mortals.

The Versatile Rectangle

The CAD tool layout is two-dimensional. Your chip is three-dimensional. As you draw these 2-D squares, rectangles, or other shapes, you should begin to see the shapes in their finished form, on top of each other, with thickness and connections. You have to think carefully about the results of what you draw. It can be tough at times, but you get used to it.

Sometimes, there are so many layers involved it's difficult to visualize those three dimensions. So, if you can, just understand what five or six out of 30-odd layers of the process really do, and just follow the rules for these five or six layers. The task becomes much easier.

Now, what does the CAD tool draw to achieve these layers upon layers?

Let's start with a side view of an FET.

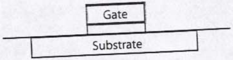


Figure 2–1. *Layers seen from the side.*

The Gate and oxide (center area) are not actually flat layers as our simplified drawing implies. They are materials that fall on top of each other. The edges of the silicon dioxide have depth. The Gate comes up and over the elevated sides of the oxide.

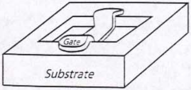


Figure 2–2. The Gate material in three dimensions shows undulations.

The Gate material undulates up and over the thick silicon dioxide walls, sort of blanketing wherever it falls. However, on your CAD tool you will not see the undulations. Also, notice that the Gate spreads out past the edge of the etched valley. That's why we see, in our top-down drawings, the Gate going past the edges.

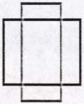


Figure 2–3. Top view, showing Gate material extending beyond the edges of the transistor.

So, you can see that two rectangles drawn on your computer screen, in reality form a sort of canyon, with blanketing Gate material flowing down, up and over the steep hills on the sides. One rectangle represents a hole in the oxide layer and the other rectangle represents a block of Gate material. In your CAD tool, though, they will look identical. Rectangles.

Ask yourself, "What am I really drawing? Does it represent a place where I want a hole to be dug down into the chip? Or, does it represent a place where I want a block of material to be layered on top? Or, does it mean I want part of the existing layer chemically altered but left in place?" These are quite different requests to ask of the manufacturing crew, yet in each case all we have to work with is a rectangle. How can they know what we want? How can we be sure of what we will get?

Let's make a chip from the bottom up to see how our rectangles can be interpreted. We will start by making a solid base for our chip to sit on.

Making a Wafer Base

To make our chips, we need a single crystal of silicon. It's only one crystal, but it's huge. The Hope Diamond is just a mustard seed compared to the size of one of these crystals.

The people in the lab have this rather clever way of making crystals. Have you ever done that experiment with sugar, where you put a whole bunch of sugar into water? (Or salt will do as well.) You dangle a piece of cotton[1] in this large vat of syrup and the crystal eventually grows? Well, that's a single crystal of sugar. They make silicon wafers the same way (but kids don't eat them when they're done).

You heat a big vat of silicon, until it is completely molten. Above it, you hang a rotating block with a small starter crystal attached. This is what they call a **seed crystal.** It's a fairly good-sized chunk, not microscopic. It's hefty enough to hold in your hand. They slowly lower this seed crystal into the vat, until it touches the surface of the molten silicon.

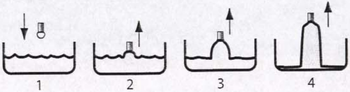


Figure 2–4. A seed crystal of silicon grows into a very large crystal of silicon.

Once the seed touches the surface of the silicon, the vat temperature is reduced. Gradually, the cooling silicon atoms start to attach themselves to the seed crystal. It's the same way with your crystal of sugar. They cling together as they cool.

Once the crystal starts to grow, the rotating block that is holding the seed crystal is slowly raised out of the silicon melt. Verrrrrrrry slowly. The crystal continues to grow as it is pulled out. Eventually, what began as our single seed crystal becomes a huge ingot. All the melted silicon in our crucible has finally attached itself. It's like a hanging cylinder now. We end up with this huge, enormous, single crystal of silicon, stretched out like a colossal salami.

---

[1] American: thread.

This method of crystal growth is known as the **Czochralski method** of crystal growth. The machine is commonly known as a **crystal puller.**

I once worked in a research lab where the guy in the lab next door to me had one of these crystal pulling machines. They take days to grow these things. Some of these crystals can be 8" wide, and way over your head by the time it's done. It's huge.

The lengthy crystal of silicon is then sliced into thin wafers, like slicing a loaf of bread. You end up with this thin, round crystal, sort of like a dinner plate. One slice is called a **wafer.** Each one of these wafers could be 250 microns thick.

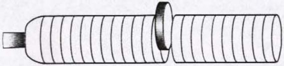


Figure 2–5. *The large crystal is sliced into wafers. Some of the circle is flattened to help workers keep lattice alignment accurate.*

The wafer is then cleaned, polished and checked for flatness and defects before we can use it. Our entire IC chip is then built on this thin wafer of what we call **substrate material.**

Just like diamonds, you cut silicon crystals along certain planes. The chips must be oriented in the same direction as the crystal lattice of the wafer to make the wafer easier to cut. The orientation of the lattice depends on the orientation of the seed crystal. You have to make sure the seed crystal orients properly in its setting. If you don't align the seed crystal properly, the wafer alignment will also be incorrect. Crystal pulling is a finely controlled process.

Some processing steps are dependent on crystal orientation. For example, there are some chemical etches that will etch preferentially faster on one edge of the crystal lattice than they will on another edge. Flat edges are ground along one or two sides of the wafer to let you know which way around it is. The flat side of the circle gives us a reference plane to align our chips so that we can control this preferential etching.

Wafers are also made of other materials besides silicon. One example is the semiconductor **gallium arsenide,** also known as **GaAs,** which is very, very brittle. You can gently push a scalpel blade into a wafer of GaAs, cleaving the wafer. It will break along the orientation of the crystal lattice giving us a good straight edge for alignment.

At last, we have our dinner plate—our wafer—a single, round slice cut from the crystal. Now, why just put one chip in the center if you can build more next to it at the same time? So, that's what we do. We place many chips next to each other, all across the wafer. One wafer could have hundreds of chips built on it. Of course, when we are done, we will cut all the chips apart.

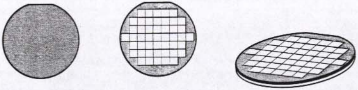


Figure 2–6. *One wafer is used to make many chips.*

We make our chips rectangular so that they align with the wafer's crystal lattice orientation. Since wafers are round, there is naturally some wasted surface area out at the edge.

Each one of these little rectangles will be an individual silicon chip. They could all be identical or variants of the same design. Or, if you are trying to do a test wafer, you may want to have four different chips, repeated, across the entire wafer, so you get a good sampling of test chips from different locations on the wafer.

The bigger the wafer, the more chips we can build at the same time. It takes the same effort to process a 1″ wafer as it does an 8″ wafer, so there's an economy of scale for larger wafers. That's our goal. We want to build as many silicon chips as we can on this big, thin wafer.

You might worry because silicon wafers are so thin, but silicon is comparatively mechanically strong.

You can drop, within reason, a box of silicon wafers and you'll get lucky and most will survive. If you drop a box of GaAs wafers, they'll shatter like glass. Total loss. I've seen it happen.

More than once.

Our original material in our big, molten vat is not actually pure silicon. Some of the impurities that we talked about have already been added. We started with our pure silicon, then mixed in our impurities while the silicon was still molten. We can add as many impurities as we want. It's a recipe. It's like knowing just how much baking soda you need to bake cookies.[2]

Stirrers help the impurities mix into the silicon. With that, and with the heat that has been added, the impurities eventually spread evenly throughout the silicon.

P Type material contains positive impurities. We also can have P+, which means we have added more P type impurities than normal. We can even have P++, which includes a whole bunch more. We can also have P−, which is not so many, and P−−. We can vary the level by controlling how much impurity we add.

N Type material contains negative impurities added to the silicon. Likewise, we can have varying levels of N, such as N−, N+, N−−. The more plusses in the name, the more conductive it is.[3]

We now have a starting base wafer that is big and easy to handle. We will next examine how to build the various layers that make our IC.

Changing Layer Composition

Processing steps fall into three main categories: We can change the surface material that we already have, we can add extra material, or we can remove material. Some process steps are a mixture of these three concepts. Let's begin by learning how to change the composition of our silicon base wafer.

Implantation

If you remember, we made PN junctions by introducing impurities into the substrate. Well, exactly how shall we convince those impurities to jump right in? Placing your atoms into the wafer surface one at a time with a fine pair of tweezers is kind of painful. So, let's use the modern engineer's totally sophisticated approach: Brute force and ignorance. Shotgun blast technique! Actually, this method of introducing impurities is called **implantation.** Same thing, as you'll see.

A source of the chosen impurity is placed above the wafer surface. Depending on the type of semiconductor we want, we could use boron, gallium, sulphur, or whatever atom we've chosen to implant. Our impurities are generated as

ions. An ion is an atom with some electrons removed, and is therefore positively charged, or with extra electrons and therefore negatively charged. The ions we will use have positive charge.

Many wafers, maybe 25 or so, are placed in a big chamber. The air is sucked out using large vacuum pumps.

Once generated, the ions are accelerated toward the wafer by putting an extremely high negative voltage across the wafer and the ion source. We're talking many thousands of volts! You can really get these atoms moving!

The positive ions are attracted to the wafer by the negative voltage. Magnets are used to focus and steer the ions. You end up with everything on the surface of the wafer totally blasted by all these ions.

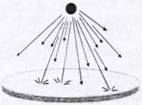


Figure 2–7. *As from a shotgun blast, our impurity is implanted into the surface of the wafer.*

They travel so fast they literally embed themselves into the surface of our silicon wafer, like bullets into cheese. The higher their speed the deeper they go. This process is called **ion implantation.**

Diffusion

Unfortunately, we have brutally forced atoms into our silicon crystal, damaging the crystal lattice. Since we rely on a good crystal lattice structure for our PN junctions to function correctly, we have to get our nice crystal lattice back somehow.

To repair the crystal lattice, the wafer is **annealed,** meaning they heat it. This helps all atoms loosen and settle with each other, forming a more consistent structure. This is rather like shaking a badly packed box of tennis balls to make them settle down evenly.

Heating has a secondary effect as well. If you drop some ink into a vat of water, it spreads out. The same thing happens to these implanted atoms. As we

---

[2] Well, once children learn the difference between a "tsp" and a "cp." I'll have to tell you that story sometime. I can't think about it right now. Back to the silicon.

[3] N++ is louder. It goes to eleven.

anneal, the atoms move in, down and outwards through the silicon just like the ink mixing in, down, and outward through the water. So, you might start with a very shallow implant, but when you anneal it, that makes the atoms diffuse outward. Your implanted atoms spread themselves through your material in all directions. That's called a **diffusion.**

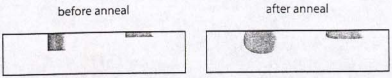


Figure 2–8. Implantation needs to be annealed, which causes diffusion.

Diffusion causes the impurities to spread. However, you do not just want them spreading out forever. The diffusion would be too weak. It would not be the right strength for you. Or perhaps it would not go down just a certain depth as you hope it will. This diffusion process has to be very controlled.

When you work with an actual process, someone says, "You can't run that layer there." When you ask why, they'll say, "Well, that's a diffusion and that's a diffusion. You'll short out."

Understanding your processes makes you the person explaining, rather than the person listening.

So, that's a basic diffusion. We bombarded atoms into the entire surface of our wafer. We annealed the whole thing to repair the crystal structure. We knew the mixture would diffuse slightly during annealing. We are left with implantation exactly as deep as we want, in exactly the right strength.

Adding a Layer

Besides changing the surface properties by implantation, perhaps we would like to add an entirely new layer of some other material. Let's look at various ways to add a new layer to our wafer.

Epitaxial Deposition

Some types of semiconductor devices need very good, thin layers of silicon, one on top of the other, in order to function correctly. (Our PN junctions need to be formed within a single crystal, as you recall.) Therefore, any new layers of silicon need to match the substrate crystal lattice. We cannot just use any method to do this, since some methods do not maintain the crystal lattice alignment. Growing another layer of silicon very slowly, though, keeps the necessary crystal lattice orientation.

Growing a new layer of silicon on top of another layer of silicon while maintaining the lattice structure is called **epitaxial deposition.** There are several ways to do this. Let's look at the most common method.

Chemical Vapor Deposition

Certain gases mixed at high temperature will react with each other to produce silicon. We could put our wafers nearby. Maybe they could catch a few of these silicon atoms falling out of the sky. In fact, we can help the atoms condense on our wafer by controlling temperatures. Growing a new layer using a mixture of gases this way is called **Chemical Vapor Deposition,** or **CVD.**

Here's the recipe for CVD:

Put your wafer in an oven. Well, a special kind of oven—a quartz tube that can withstand very high temperatures. At one end of the tube, arrange for highly reactive gases to be pumped in. The mixture of gases react with each other, encouraged by the high temperature they encounter. The reacted gases travel through the tube until they encounter our wafer. The temperature of the wafer is cooler than the gases, so the silicon in the mixture condenses on the surface. Our condensation gives us a nice epitaxial layer across the surface that aligns perfectly with the crystal lattice of the wafer.

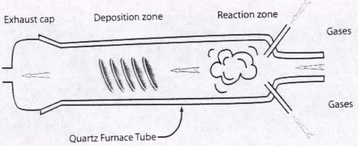


Figure 2–9. Vaporized silicon condenses on many wafers at one time.

If we vary the mixture of gases, we can vary the type of silicon grown. Sometimes the silicon layer might be P type, sometimes it might be N type. We can even vary the gas mixtures within a deposition run to get alternating layers of N and P type material.

Many wafers are coated at the same time. The gas mixes around and through the wafers evenly enough to coat them all.

That's it. You have built upward. You have deposited more silicon on top of whatever was previously your surface. A variety of materials can be used. We could deposit silicon on top of an oxide layer, for example.

From our section on diffusion, we know that when a wafer is annealed, atoms diffuse in all directions. If there is an epitaxial layer on top of implanted silicon, annealing will cause the buried impurities to diffuse upward, into the epitaxy layer. If we arrange things correctly, subsequent implants and diffusions can be made that will eventually join with a diffusion that is buried underneath an epitaxy layer. This essentially forms a contact connection to the buried region.

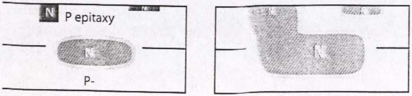


Figure 2–10. *Diffusion works for us, joining two N regions.*

Remember the FET we discussed in the previous chapter. The Gate is placed on top of thin oxide insulation. We can use CVD to quickly grow a relatively thick layer of silicon on top of this thin oxide layer to use as our Gate material. We want it grown quickly so the extra heat from the CVD process will not make our diffusions move around too much.

Whereas epitaxial layers, which are silicon on silicon, are grown slowly to maintain our crystal lattice, silicon that is grown quickly using CVD does not have a very good crystal structure. Like frost on a windowpane, it is made up of many different crystals that eventually join, rather than as one large crystal. This type of silicon is known as **Poly-Crystalline Silicon** (meaning *many crystals*). Poly-crystalline silicon is usually referred to simply as **Poly.**

Poly is used extensively in IC's to make FET Gates and resistors. Like our silicon, it, too, can be doped. The doping is usually to change its resistivity, whereas doping regular silicon is to set energy levels and transistor characteristics.

Depending on the gases introduced into a CVD furnace, we can deposit layers of almost anything that can be created in a chemical reaction. Oxide can even be deposited in very thick layers, as well as silicon nitride.

A variation of CVD is **Plasma Enhanced Chemical Vapor Deposition,** or **PECVD.** PECVD is very similar to CVD, but instead of high temperatures to start the chemical reaction, a **plasma** is used instead. Plasma is a state of matter formed when gases at very low pressures are subjected to high-frequency high-voltage. A fluorescent lighting tube contains a plasma, as do the Northern and Southern Lights.[4]

The plasma provides the energy for the chemical reaction between the gases without raising the wafer to extreme temperatures. The low temperature helps to keep our previous diffusions from diffusing further.

Oxide Growth

At times, we want to build wiring or other conductive materials over the top of our chip. We need to isolate our layers from each other, so that one metal layer does not short with another metal layer. An easy way to create a layer of isolation is to stick our wafer in a furnace with oxygen, and heat it up. The silicon on the surface turns to silicon dioxide. Silicon dioxide is a very good insulator. There you go, an instant layer of insulation. It's sort of like creating a quick layer of thick rust on iron.

Sputtering

High-energy plasma can also be used to help us deposit materials that cannot be deposited using CVD. Plasma made from inert gas, such as argon, can be used to knock atoms into submission, using a machine called a **Sputterer.**

Let's use a Sputterer to deposit a layer of metal.

We start with a chamber containing a large block of the desired metal, suspended above the wafers. The metal will be blasted onto the wafer to form the new surface. Here's how.

The air is sucked out, and a small amount of Argon let in. The Argon turns into high-energy plasma as described above. As the high-energy Argon atoms smash into our metal, they force metal atoms to separate from the block, flying out into the plasma. This is kind of like sandblasting the metal with Argon atoms.

---

[4] Now you can visit Alaska to see the Aurora Borealis, and write it off as business expense. Don't forget to invite us along.

The metal atoms become ionized by their ordeal and get attracted to the safety of the wafers waiting below. As the process continues, more and more metal atoms get blasted away, sticking to the wafer surface. Eventually the wafers are coated with the correct thickness of metal. There's your new layer. Clean, even, well-controlled thickness.

I like to think of sputtering as similar to snow falling. The metal atoms are able to fall into all sorts of nooks and crannies that you would not expect. That could be good. That could be bad. Your job is to know the difference, and control it. That's why they pay you the big bucks.

Evaporation

Another way of depositing metal is called **Evaporation.** Evaporation is exactly what you think it is.

Our wafers are loaded into yet another large chamber that has all the air sucked out of it. This chamber contains a large tungsten filament, rather like the coils of a light bulb. The coils have small chunks of the desired metal placed inside them.

As we pass current through the filament, it starts to glow. The coil gets so hot the metal inside melts. As the filament gets even hotter, the metal finally starts to evaporate. The evaporated metal atoms fly around in the hot gas. They eventually hit cooler surfaces and form a layer of condensation. They condense onto everything inside the chamber, including our chip.

Removing a Layer

There are chemicals that eat stuff. By pouring a certain wet, nasty chemical onto our wafer, we can cause a reaction, which dissolves the surface. Once dissolved, we can wash it all away in the kitchen sink, leaving underlying layers exposed. This eating-away is called **etching.**

We can remove metal by etching. We can remove oxide by etching. You name it, we'll remove it. We have powerful chemicals. That's how we remove a layer. We dissolve it.

Another way we can etch our newly formed materials is called **Reactive Ion Etching** or **RIE.** RIE is almost the opposite of sputtering. Instead of a block of metal being bombarded by Argon atoms we can reverse the polarity of the chamber voltage and sandblast our wafer instead. If we use a mixture of gases hat happen to react chemically with our new surface material, then the surface will be removed even faster.

Working in a wafer fab is interesting but can be dangerous. Some of the acids used in IC manufacture are not conducive to continued good health. One acid in particular, Hydrofluoric Acid or HF, is particularly nasty and will eat the calcium in your bones unless it is treated very, very quickly. I kept a tube of HF neutralizing jelly at home because I was so worried about HF.

Luckily, my boss realized that I was better at layout than making IC's and let me pursue a much safer career. Thanks, Alan!!!!

We now know how to change, add and remove layers. Next, we need to learn how to control exactly where we want each of these to occur on our chip.[5]

We don't necessarily want an entire layer added over the whole wafer. We don't necessarily want an entire layer etched away or an entire silicon layer blasted with implantation. Usually we prefer only small areas of the surface to be changed. That's the subject of our next section.

Photolithography

We can select specific stripes, rectangles or other areas of our wafer that are to receive the above-mentioned processing, by coating our wafer with a light-sensitive protectant known as **resist,** or **photoresist.** By using the properties of resist, we can sort of paint the areas that we want processed. The blasting, gassing and condensation that we do to the entire surface will only truly stay in the places which are not protected. Let's look at a more detailed explanation of how resist works.

We wet our wafer surface with this light-sensitive chemical. The resist is then baked hard. Resist that is not exposed to light remains intact, hard, firm and protective. However, where light does hit the resist, a change is made to its chemical structure. Resist that is exposed to light becomes dissolvable. Those exposed areas of the surface can then be washed away using the appropriate solvent. It's just like in the movies. When light hits the vampires, they dissolve.

Light chemically changes resist.

This process is known as **photolithography.** As the name implies, photolithography means *printing with light.*

---

[5] And you were just thinking,

We can shine light through a sheet of glass onto the surface of our wafer. The sheet contains an image of the regions that we want to protect from processing. The image falls onto the surface as shadows. Light that shines through the glass will expose the resist on all other portions of the wafer surface. Light that is blocked by the image will not get through, and will not change the hardened resist.

The glass sheet is known as a **mask.** Each layer we draw will have its own mask. The image on the mask is usually drawn in chrome, so that no light will be able to shine through.

Our wafer is nice and round, so we put the wafer on something that can rotate. From above, we drip resist onto the wafer. Because the wafer is spinning, the resist spreads out. We get this nice, even, thin film of resist.

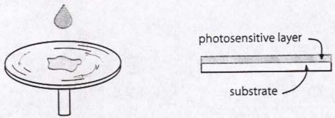


Figure 2–11. *Substrate is coated evenly with photosensitive resist by dropping the chemical onto the spinning wafer.*

First, we harden the layer of resist. Then, we expose the resist to light, shining the light through the mask. We then **develop** the layer.

Developing resist means washing it in a special dissolving solution. Developing dissolves areas that were chemically changed by exposure to light. Then we wash the dissolved areas away. They're gone. Eaten. Etched. Dug out. Vanished. All parts of the resist layer that were exposed to light are now gone.

Only the shaded areas of resist remain—the areas under the chrome. These areas of hardened resist now protect specific sections of our chip from whatever we do to the surface next. This is why the protective material is named resist. It resists the process that is about to be done to the chip.

For example, we could dip our chip in a strong acid. The regions protected by resist will remain untouched by the acid. Or, we could bombard the chip with ions. The regions protected by resist will remain untouched by the ions. Whatever we do next will be protected wherever we still have hardened resist on the surface.

However, in an acid bath, for instance, the areas no longer covered with resist will be attacked and eaten away. By removing certain areas of resist, we expose those underlying areas to the subsequent acid. After we are finished with our acid etching, we can remove the remaining resist by other means. We do not need to keep this layer of protection any longer than one step in the process. The pattern would be wrong for the next layer. We will build new protective layers of resist for each step, patterned as needed. That is why each layer of our chip requires its own chrome mask. Each layer has a different design.

Photoresist comes in two flavors. One flavor is naturally unaffected by our developer but becomes removable when exposed to light. This type of resist is called **positive resist** and is the most common type of resist used. We will assume we are using positive resist in this book.

The other flavor acts exactly opposite. The exposed regions do not dissolve. Instead, the exposed regions harden. They cannot be removed by the developer. So, the shaded regions are what naturally dissolve in our developer. This type of resist is known as **negative resist.** Negative resist is not as good at producing a high quality image and is harder to work with.

Photolithography is a fundamental part of every step in the IC fabrication process, whether you are implanting, etching, or making any other surface changes. Every step must only occur where we direct.

Every step requires coating with resist. Every step requires a mask. Every step requires exposure to light. Every step requires removal of exposed resist. Every step requires processing. Every step requires removal of the remaining resist in preparation for the next step. And all of that was only one step in processing. One layer. You might have 20 or 30 layers in a chip.

Building a Chip

We have now discussed most of the major processes used in IC processing. Let's examine how all of these processing steps fit together to build a chip.

Creating a Hole

It's time for us to practice. We will make a hole in an oxide layer. Yes, that sounds fun. We will apply an oxide-eating etch acid to make a hole.

In order to make a hole in a full layer of oxide we need to protect the majority of our surface with a coating of resist. We then create holes in the resist to let the acid-etch through.

Remember, we are using positive resist, so we need to shine light where we want to dissolve the resist. Therefore, we only want light to shine in the one spot that will become our hole in the oxide. The mask is covered entirely with chrome, except for the one place where light will shine through. Usually we want light to shine through many well-defined and intricate places. We will use a simple pattern, though, for our example. (I wanted to use a duck pattern, but I was outvoted.[6] We will use ABC.—*Judy*)

A mask that is mainly covered with chrome is known as a **Dark Field** mask. Light will not pass through the majority of the mask.

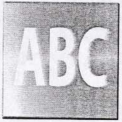


Figure 2–12. *Mask with holes through the chrome coating.*

1. We lay the mask down over the chip, then shine light at the mask.
2. Light only shines through our holes in the chrome.
3. The light causes a chemical reaction in the exposed resist.
4. We develop the resist. The exposed areas wash away.

We now have a hole in the resist, exactly where we allowed light to pass through the mask.

We can now put our chip into an acid bath. This particular acid eats oxide, but cannot damage the hardened resist that remains on the surface. The acid eats away the oxide only through the hole we have just created. Figure 2-14 shows the process in side view.

After we remove the remaining resist with a special resist-removing solvent, we start all over again for our next process step. Notice how the oxide has over-etched, making a hole that is bigger than our mask. Mask dimensions must be carefully determined taking over-etching into consideration.

---

te to the ladies: *outvoted* means you are giving in on something you don't care about so that owe you one. Don't tell the guys.

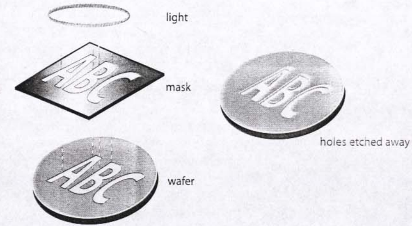


Figure 2–13. *Exposure to light causes the resist to dissolve when developed, leaving a hole.*

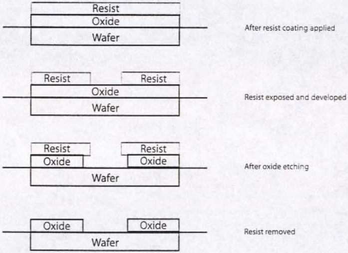


Figure 2–14. *The resist layer protects the oxide layer.*

Creating a Block

Now let's use resist and acid etch to build a lump on the surface.[7] As an example, let's build a Gate section on our chip.

---

[7] You can build upward by dissolving downward. Trust us. We're authors.

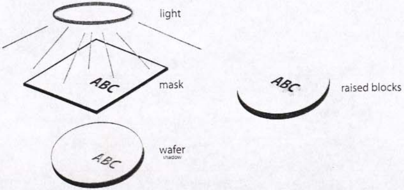


Figure 2–15. To build upward, deposit a full layer, then eat everything away except your shaded region.

Just as before, we use our positive resist. However, this time we have already covered the surface entirely with a new layer, in this case polysilicon Gate material. We need to etch most of it away. That's the difference. Instead of etching a small section out of a large field, we now will etch a large field, leaving standing just a small section.

We need to protect only small areas of the surface from the etching acid. So, our mask remains primarily clear with only small areas of chrome. This kind of mask is known as a **Light Field** mask. The majority of the mask is see-through.

Again, we expose, develop and etch. Voila. We have polysilicon Gate material left intact just where we want it. These small, protected areas are all that remain standing on the surface. The rest was eaten and washed away.

Finally, we use our special resist-removing solvent to remove the remaining resist so we can start all over again with the next layer. It seems like such a waste to coat the entire surface every time you want a little button to sit on the surface. That's a good reason to utilize existing layers whenever you can. Reduce processing steps. Reduce time and cost.

Notice, again, that the final Gate dimension is different from the original dimension of the chrome on the mask. The Gate has over-etched. The final dimension is smaller than drawn. These oversizings and undersizings need to be accounted for in mask making. If we draw a Gate at 1 micron, knowing it will undersize in processing by 0.1 micron, we can oversize our mask by 0.1 micron to compensate.

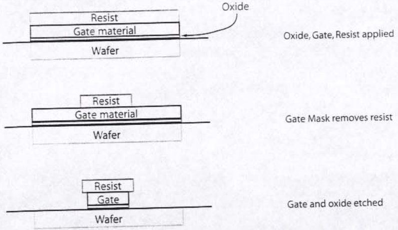


Figure 2–16. Side view of building a Gate by etching surrounding material.

Planarization

After so much digging and building to the surface of your chip, it can become very uneven, especially after a few layers. You can set a new, level plane on your wafer surface if you want. You can etch, grind or polish your surface to make it flat again. Making a flat plane for the next layer to sit on is called **planarization.**

Without this planarization process, a new layer might undulate down, up and over some pretty stiff turns and valleys. The turning causes stress at each bend. Your material stretches thin around tight angles. You would have to make your layer extra thick to keep it from breaking.

However, if the previous level has been planarized, you can lay down thinner materials. You can get smaller patterns, smaller dimensions. Smaller is better, faster, cheaper. Planarization improves chip performance.

Silicon Dioxide as a Mask

Unfortunately, resist does not hold up too well against being blasted with atoms. However, silicon dioxide does, especially if we make the oxide thick enough. Oxide might just make a better protection sometimes.

Plus, oxide is easy to make. We just heat our silicon in oxygen and it grows just like rust on iron (iron oxide). Instant layer of what will be a great implantation protection.

We can then use photolithography to cut holes in this new oxide, and use the oxide to protect the silicon underneath it. We then blast the entire surface with our implantation. These crazy little atoms would run right through resist, but our thick oxide layers form a nice mask for us. We can implant though the oxide holes.

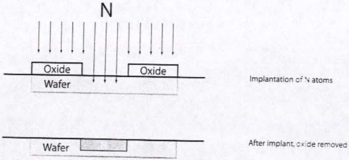


Figure 2–17. Using oxide layer to shield wafer from implantation.

Self-Aligned Gates

We want our transistors to switch as fast as they can. A transistor that is always ON takes a long time to turn OFF, since we need to create quite a large depletion region. Enhancement mode FET's, on the other hand, are much easier to turn on and off because the region under the Gate is much smaller. However, we need to align the Gate very accurately to the source and drain regions to make Enhancement mode transistors.

To create automatic alignment, we can use the Gate material itself as a mask to perfectly align our source-drain regions. Let's see how it works.

Let's take our bare wafer and stick it in an oven with oxygen for a rather long time. We'll end up with a fairly thick layer of silicon dioxide. We now expose and develop a layer of resist, then etch a hole in the oxide. The hole will go all the way through the oxide, exposing the silicon wafer below.

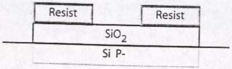


Figure 2–18. Patterning our resist above a layer of silicon dioxide.

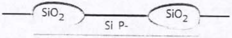


Figure 2–19. We have made two very thick regions of insulation.

We now have very thick silicon dioxide as an insulator exactly where we want it.

Let's put the wafer in the oven again for an extra couple of minutes. We will grow another layer of silicon dioxide in the hole we just etched. Remember that we can use a thin layer to get a chunk of Gate material very close to our silicon without actually touching it. That's what we want here, a thin insulation above the bare silicon.

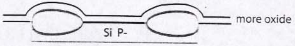


Figure 2–20. Getting ready to use some thin silicon dioxide (oxide) as an insulator under our Gate.

Next we deposit polysilicon for our Gate material. We cover our whole wafer with polysilicon. We then expose and develop our Gate mask. The Gate mask is a light field mask so the majority of the polysilicon is exposed to the Gate etch.

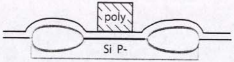


Figure 2–21. Polysilicon has been placed above our thin insulation.

The Gate etch leaves a very small sliver of polysilicon in the hole we made in the thick oxide. We now etch away the thin oxide that we grew. The thin layer only remains where it was protected by the Gate. We are left with our Gate strip on top of a thin layer of oxide. The thin layer of oxide has also been removed from the majority of the hole in the thick oxide so we now have the silicon in our wafer exposed again.

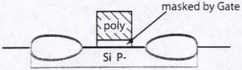


Figure 2–22. The Gate protected the underlying oxide layer from being etched away, just as the shadow from a mask would have done.

Ve are now ready for implantation. The Gate strip, the bare silicon and the ick oxide are all bombarded by the implant atoms.

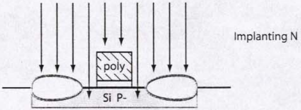


Figure 2–23. Atoms bombard the entire surface of our wafer.

he Gate and thick oxide protect the silicon underneath it. Only the bare sili-on on either side of the gate is implanted. We have ensured that our source-rain implant is perfectly aligned to our Gate strip. We have not had to perform ny tricky alignments.

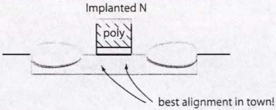


Figure 2–24. Self-aligned Gate.

he Gate has acted as a mask for us for the second time. These newly bom-arded N regions form along both edges of the Gate. They are perfectly ligned, too, since the Gate acted as an umbrella, protecting N bombardment from traveling to the region directly below. This is called a **self-aligned Gate.**[8]

In addition, we get a bonus when this implant is annealed. The source-drain regions will naturally diffuse underneath the Gate just a little, so we get a field region even smaller than we can define with light. It doesn't take much to turn this transistor off, since this field area is so small.

If we are clever, when we anneal, we will grow another silicon dioxide surface for ourselves at the same time. One heating process can anneal for us while it grows more oxide.

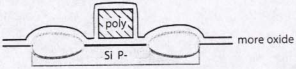


Figure 2–25. Why not form an oxide layer when we anneal? Two birds with one stone.

We do not care too much about the initial placement of the Gate itself, but we do care if N+ regions align well with the Gate in the end. You just plop the Gate down somewhere in the middle of a hole in thick oxide, blast some N+ material and everything aligns itself perfectly.

All CMOS processes are different. Some companies leave the thin oxide in place when the source-drains are being implanted, some don't. Some companies do not etch holes in thick oxide but use a Silicon Nitride layer to protect the wafer during Oxide growth. The examples given here are only provided to give some basic understanding of the process steps involved. If you are really interested in understanding every little step in the process that your company uses, go find the wafer fab operators and have them explain every step to you. Be careful carrying GaAs wafer boxes.

Closure on Wafer Processing

This chapter explains how the polygons we draw translate into the physical world. We have concentrated on CMOS devices just to see roughly how all chips are built.

---

[8] It's called the SAINT process (Self-Aligned Im . . . something or other). No, really, it is. Why would I make up something like that?

h process in the world is different. So, in Company A's process you may
y well be able to wire things over the top of a diffusion, in Company B you
not. Some companies allow wiring underneath capacitors, some companies
n't. It all depends on the process.

en you start work for a company, they'll sit you down, and walk you
ough their particular process, step by step. "This is a wire. You can do this
h it, not that. This is a diffusion. You can do this, can't do that." Whatever
r company's specific process rules involve, you now should be able to ask
right questions, and use the answers to build a mental picture as you do
r layout.

You can't get away without knowing your process.

on't mean *to be good* you can't get away without it. I mean you *can't* get
ay without it. Everyone who does layout needs to understand processing.

en I look at a layout, I am not thinking about the layers. I am thinking
out how the circuit works, where the current is flowing, or what volt-
es are at what points in the circuit.

to see a flow. Attain an underlying one-ness with your circuit. It's a
n-like state. "Be" the circuit. Say, "Ohm" real slow. The polygons will
ne to you.

you have worked in a wafer fab, you can actually see these layers build up
front of your eyes. It's good experience. If you know how the rectangles
come 3-D inter-reacting worlds, you can be the person who does the cutting-
ge, researchy type of stuff. With a good processing background:

- You can handle writing design rules. What does each layer do? How do they interrelate? What's causing what?
- You can learn to make shortcuts. A diode over here is made using the same process as something else over there—why not combine the two?
- You can generate full custom elements not in your design kit. You are not limited to copying and stretching from your existing library.

u do not need to master every subtle nuance, but it helps if you under-
nd some of the basics. It prepares you to hear what you are going to hear
the job.

With time and experience, you will be able to look at a circuit schematic and instantly visualize the layout in your mind. At that point, your job is not work; it's pure creativity.

Here's What We've Learned

Here's what you saw in this chapter:

- Silicon crystal wafer growth
- Using ion implantation to dope silicon
- Photolithography
- Light and dark field masks
- Epitaxial deposition to build upward
- Diffusions caused by annealing
- How to create self-aligned Gates

And more . . .

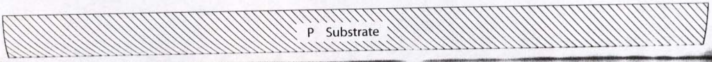
Chris' Disclaimer

There now follows a step-by-step flowchart of a typical CMOS process. The drawings are not accurate representations of the actual cross-sections of a real semiconductor process and are simplified to make the details easier to see. The intention is to give the reader a feel for the basics of semiconductor processing. Additionally, the process represented is generic and not 100% accurate. Again, this is deliberate and you should study the IC process that you will be using in depth in order to understand the subtleties required. In other words, please don't write to us telling us how blatantly inaccurate this process flow is. We know all about it!!

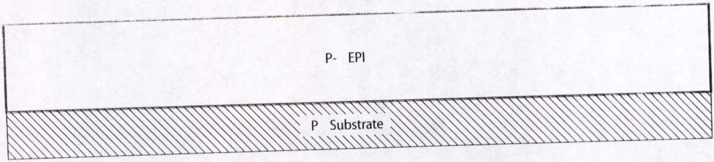
Judy's Disclaimer

Hey, I only draw what Chris tells me. And I like to get mail.

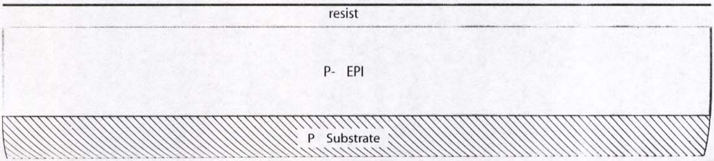
P Type Start Wafer



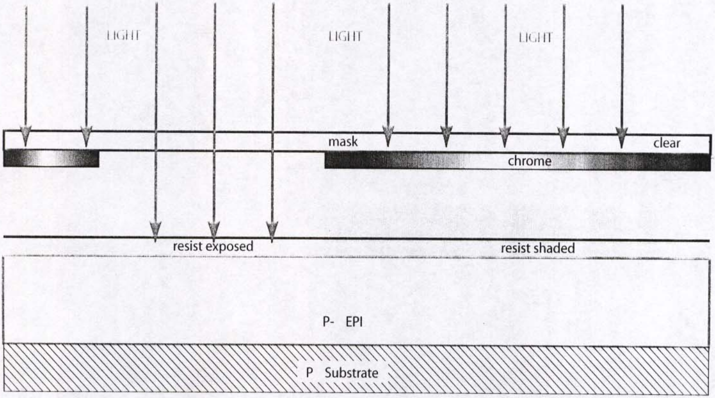
Grow P- Epitaxial Layer



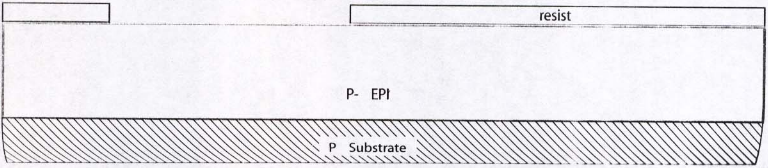
Spin Resist Coating



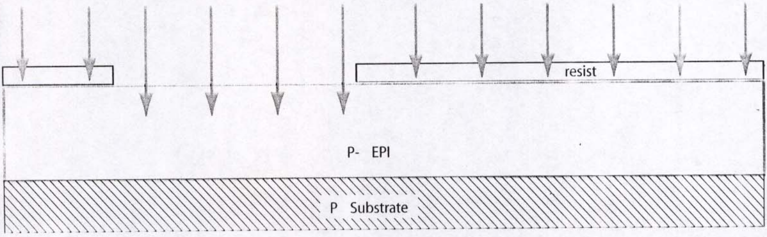
Expose N Well mask



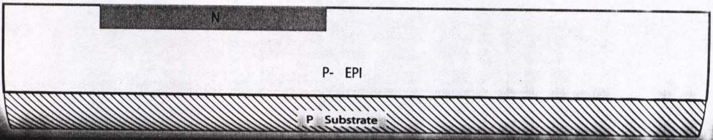
Develop resist



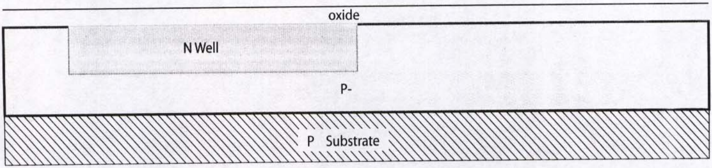
Implant N Well



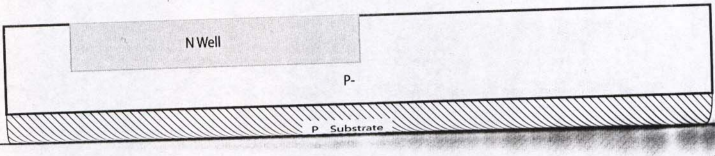
Remove resist



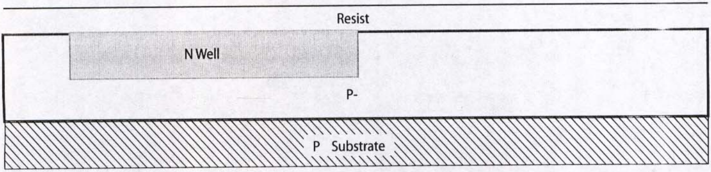
Anneal wafer -- gives us new oxide layer and diffuses N Well.



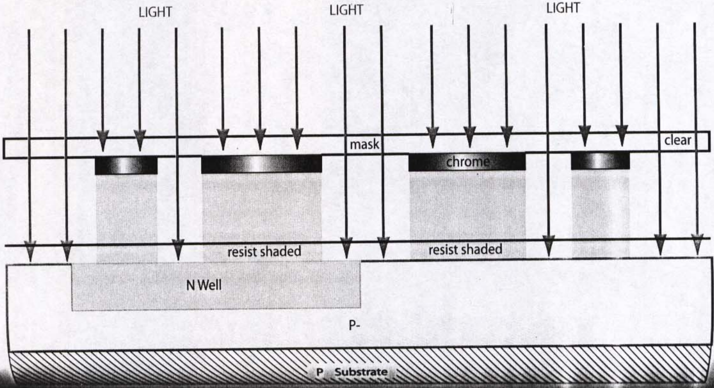
Remove oxide from Anneal



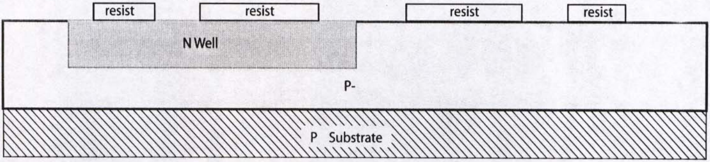
Spin resist



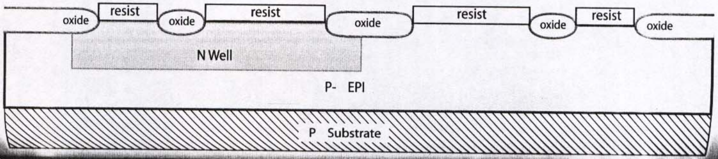
Expose resist with Active Diffusion Mask



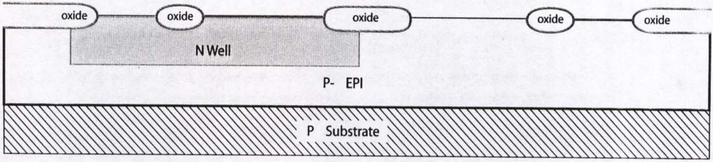
Develop resist



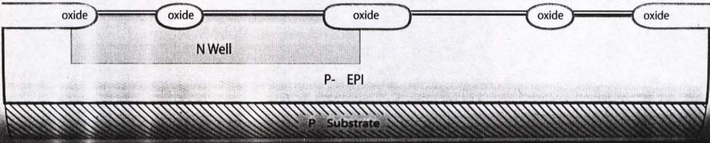
Grow oxide on exposed surfaces



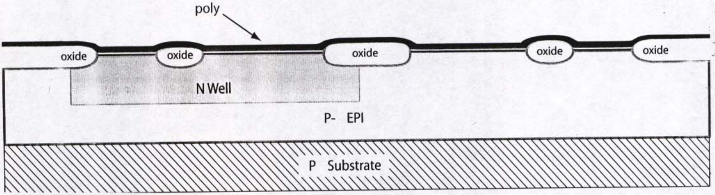
Remove resist



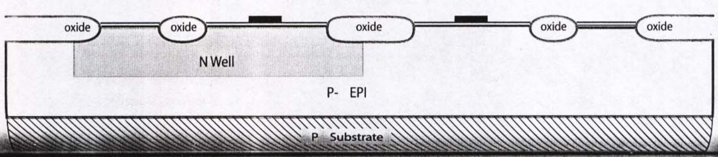
Grow thin oxide over silicon surfaces



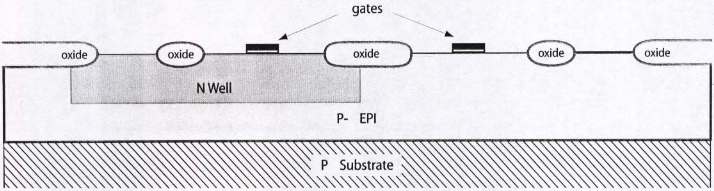
Deposit poly using CVD



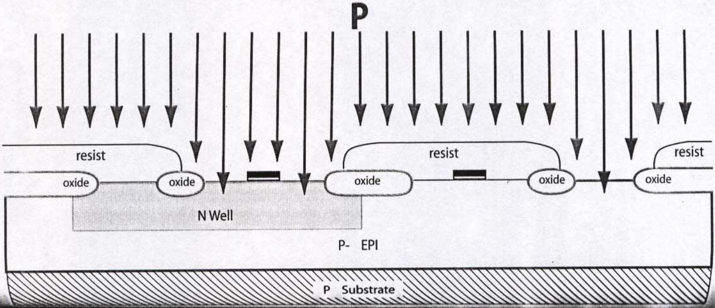
Spin resist
Expose resist using GATE mask
Develop resist
Etch poly



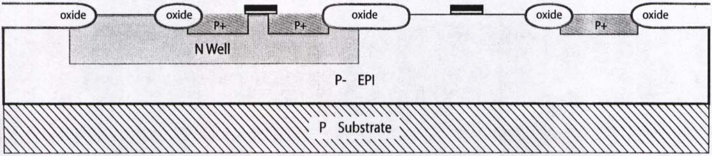
Remove thin oxide layer where exposed



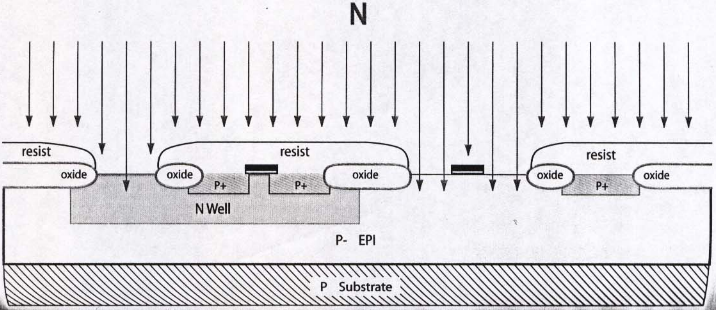
Spin resist
Expose with P implant mask
Develop resist
Implant with P



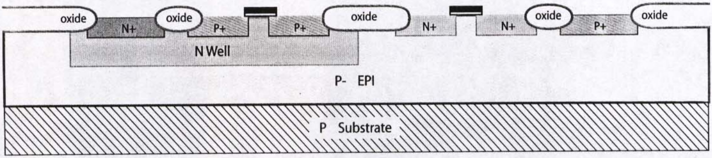
Remove resist
Anneal wafer
Oxide etch



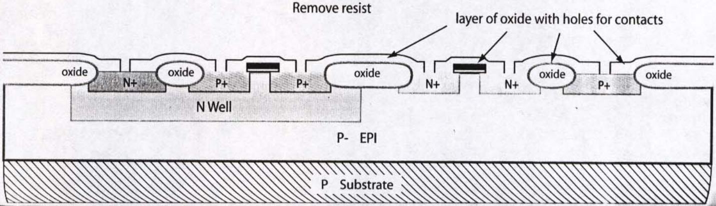
Spin resist
Expose with N implant mask
Develop resist
Implant N



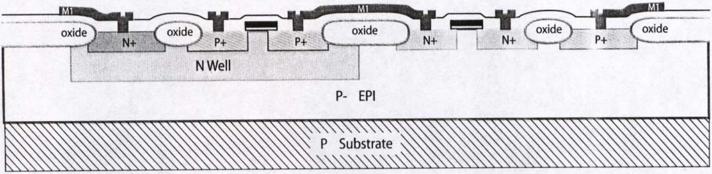
Remove resist
Anneal wafer
Oxide etch



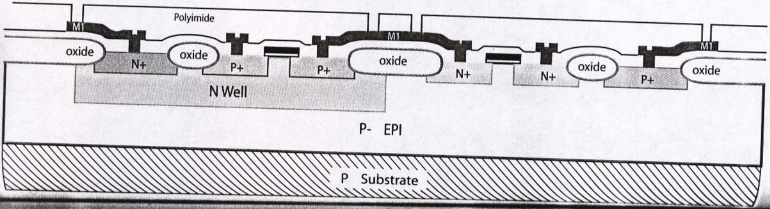
Deposit oxide using CVD
Spin resist
Expose Contact mask
Develop resist
Etch contact holes
Remove resist



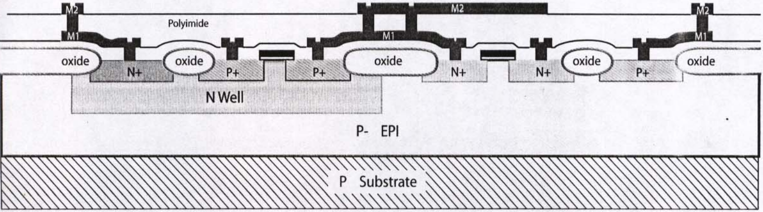
Deposit Metal 1 using sputtering
Spin resist
Expose Metal 1 mask
Etch Metal using RIE
Remove resist



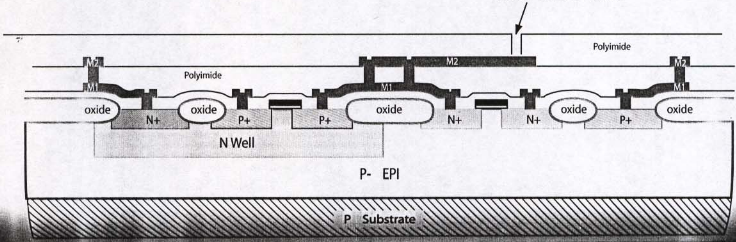
Spin Polyimide
Spin resist
Expose Via 1 mask (vias connect two metals)
Etch Via
Remove resist



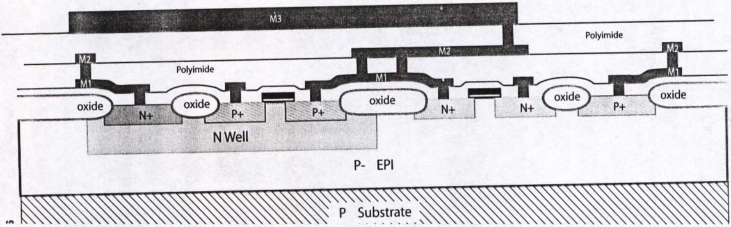
Deposit Metal 2 using sputtering
Spin resist
Expose Metal 2 mask
Etch Metal using RIE
Remove resist



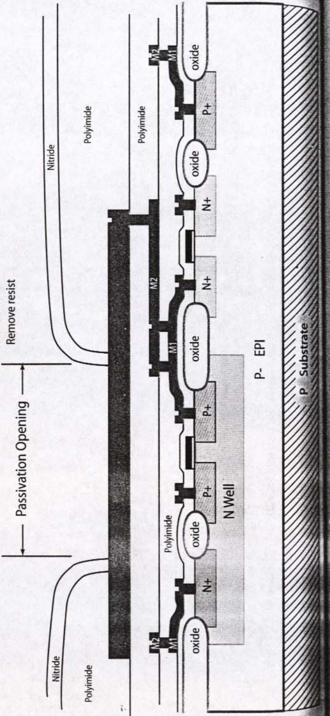
Spin Polyimide
Spin resist
Expose Via 2 mask
Etch Via
Remove resist


hole for Metal 3 contact

Deposit Metal 3 using sputtering
Spin resist
Expose Metal 3 mask
Etch Metal using RIE
Remove resist



Spin Polyimide
Spin resist
Expose Passivation mask
Etch Poly
Remove resist
Deposit Nitride using CVD
Spin resist
Expose Passivation mask again
Etch Nitride
Remove resist



CHAPTER 3

CMOS Layout

Chapter Preview

Here's what you're going to see in this chapter:

- Elements necessary for circuit computer simulations
- How specifications affect device sizes
- Why we split large devices into smaller ones
- How to reduce unwanted resistance
- How to draw stick diagrams as a design aid
- How to bleed unwanted voltages during processing
- How to prevent fatal backward electrical flow in your chip
- The visual relationship between schematics, diagrams and layouts
- How to save space when building multiple devices
- Interesting sample layouts for analysis

And more . . .

Opening Thoughts on CMOS Layout

In our first two chapters, we constructed individual devices. We built the various layers of each device from the ground up, creating a transistor, for example. Integrated circuits, as you know, are made from many transistors, though, not just one.

Now we will examine various techniques used to join these many transistors and other devices together to make real circuits. We will discuss the practical aspects of drawing our layout masks for these connections.

We will continue to use only CMOS transistors as we have done since Chapter 1.

Device Sizing

Here we see a figure of our basic transistor again. The transistor is built from a polysilicon Gate, placed over a region of thin silicon dioxide. You see we have drawn a rectangle to represent the edges of our Gate material. Likewise, we have drawn another rectangle representing the edge of our thin oxide area.

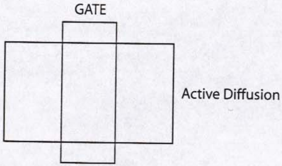


Figure 3–1. *Top layout view of an FET transistor showing Gate area and Active Diffusion area.*

The oxide layer is often referred to as the **active diffusion,** or the **active.** The active is where atoms will be implanted to create a transistor. The overlap of the Gate and Active layers determine the size of the device. Any surplus Gate material or active that lies outside the area of the other has no effect.

The question is, "How do we know how big to make our devices to achieve the circuit performance we require? How much overlap do we want?" Should we draw large rectangles or small rectangles?

SPICE

Before someone can design a circuit, he needs to know what the circuit must do. "What performance numbers do we need?"

For example, a circuit designer might be given a list of requirements that says, "I want an amplifier to have a voltage gain of 20, a frequency response of 20 Hz to 20 kHz and only take 2 mA from a 3.3-volt supply." This is the starting point. These numbers are what we call the **circuit specifications,** or **specs,** for short.

The circuit specifications enable designers to start the design process. With detailed information about your particular processes, the specs determine the necessary device sizes.

Many different circuit techniques can be used to design a component such as an amplifier. Our designer chooses a circuit topology that he thinks will give the required performance. Then he calculates the basic capabilities of the circuit, verifying that the circuit meets specs.

Designing circuits used to require only a first pass using basic calculations. Those were the simpler days. However, a transistor is a very subtle, complicated device with many different interacting elements. These interactions usually cause our circuit to perform differently than our basic calculations indicate. Adjustments to the design are expected, considering the complex interactions that cannot be fully appreciated from just the first design attempt. Our designer needs to build some confidence that his first-pass circuit design will function correctly.

In its infancy, the only way to verify IC design assumptions was to build the circuit and try it. As you may expect, there were many false starts. One chip design took a long time to finally get correct.

Today we can use computers to give us the confidence we need to commit our circuit design to silicon. We run our design through a computer program called a **simulator.** The simulation program indicates what the circuit does, how much current it requires, the frequency response, the gain, and so on. Like other simulation programs on the market, we see results without having to actually build the physical test situation. The computer pretends a test device has been built, then checks how it would operate for us if it were real.

These circuit simulation programs are commonly called **SPICE** (Simulation Program for Integrated Circuits Emphasis).[1] The original SPICE was developed by the University of California, Berkeley, in the 1970's. Running a simulation is cheaper and faster than building a chip that may not work!

There are now various flavors of SPICE on the market, written by different companies. You might see all sorts of brands available, as you see different brands of word processing programs or different brands of flight simulation games. All are just variations of the same type of program, developed by different companies, each with different advantages and disadvantages.

---

[1] or Several People In Cell Eleven.

In order for SPICE to accurately predict the complex workings of our circuit, we need more than just the specs and a first-pass schematic. We must create a mathematical representation of the circuit elements we will be using.

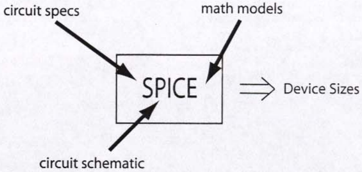


Figure 3–2. *Three elements used in SPICE.*

The mathematical representations of the circuit element behaviors are called **models.** However, you cannot just guess at the mathematical equations. You must accurately reflect the physics of the device and its electrical and physical characteristics.

A basic device could consume months to model. The basic device models can be tailored for each customer as well. Each customer builds many differently sized circuit elements, then records the actual characteristics of each. Once measured, curve-fitting software analyzes all the collected device data. From this data we can determine workable model parameters for an individual customer's device.

The Spice model is then tested to ensure that the simulator output reflects the real device output under all conditions. Once the models are verified against the real devices, we can confidently use the model to predict circuit performance. Spice model development is very complicated. Many years of effort have gone into creating accurate models of circuit elements for us.

The circuit design process now relies entirely on the output of these circuit simulators.

Now, there is another specialization you could pursue in your career. You might find that you have quite a flair for and interest in model development.

Keep learning new things. Dabbling beyond your job description keeps your job interesting, challenging and rewarding. If you direct your dabbling in areas that are interesting to you, every door that opens will be more fun than the last. That's how some people end up working in jobs they love. They open their own doors by dabbling and playing with new ideas that are fun for them.

Your SPICE simulation results tell you how the circuit performs with the given device sizes and values. Since we can adjust the input values so easily, the simulator allows designers to try some quick "what if?" scenarios. They can fine-tune the device sizes and values to optimize the design.

Once the simulations confirm our circuit function, we call the circuit complete. We have determined our optimal device sizes by observing the SPICE results. We now know the size of each device we will build on our chip.[2]

Use SPICE to establish device sizes.

This completes the first part of the IC process. We created the devices in a first-pass schematic. We then used SPICE to determine how big each device should be.

Modern CAD tools automatically create design rule correct devices directly from the circuit schematic. Gone are the days of manually calculating the length of each resistor from fundamentals. However, the *ability* to calculate device sizes from scratch is very valuable. There may be times when a CAD tool has not been set up to generate devices. In that case, it's all up to you!

Some layout engineers work 20 years in the field without ever calculating their own device sizes. What a waste of talent. I think a lot of managers don't think their layout people are capable, so they never give them the opportunity to learn more, to become more valuable. But it's part of our job. There is no reason why layout people can't work the numbers to determine or just to verify the schematic device sizes. If managers keep their layout engineers from the valuable and creative ends of the job, the layout designers soon become bored with their work. And understandably so.

---

[2] . . . if you're lucky. You'll think they've been finalized a number of times, but always three days later the circuit designers come back with another guess. Device sizes will change until two hours after you ship the final data. Yes, after!!!!

> *Without challenge and growth, pretty soon you have inexperienced circuit designers giving problematic designs to inexperienced layout designers. That's a sure recipe for disaster.*

Splitting Long Devices

Now that we know the device sizes, how well will they fit together?

Let's say our circuit designer has given us the circuit below.

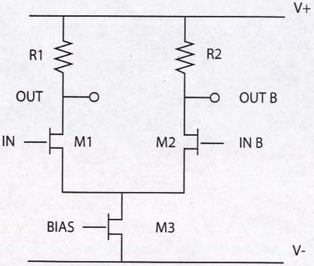


Figure 3–3. *Sample circuit. M1 and M2 are 200 × 1 microns.*

At this point in the layout process, we do not care what our circuit does, although this drawing happens to be a simple differential amplifier. All we care about next is making these components the size the circuit designer tells us. We will worry about the circuit functionality when we start to wire the components together.

The sample circuit has two inputs, two outputs, power supplies and a bias voltage to enable the circuit to function. Typically, in a schematic, the components are labeled to make it easier to identify. R is used for resistors, and M for MOS. There are 3 MOS transistors and two resistors in this diagram. It just happens that all of these transistors in this example are NMOS devices. Typically, though, some could be N, some could be P.

Let's suppose, for sizes, our M1 transistor is to have a width of 200 microns, and length of 1 micron. M2 will be the same. M3, let's say, is 60 microns. We will concentrate on the CMOS transistors in this example, so let's ignore the other components for now.

As mentioned above, the first task of a layout engineer is to build your devices. In the good old days, you would have a standard cell, which you would copy, then manually stretch to the correct size. This approach requires you to be very careful about following all the design rules required of each component. Copying, stretching and rule checking were quite repetitive exercises for making each component one at a time.

Today, components are built directly from information in the schematic. All you do is open the schematic diagram on your screen and say, "Make me that transistor." Voila, your transistor appears on the screen, correctly sized and correctly built to the design rules.

Whether by hand or automation, let's say you have made all three transistors for our sample amplifier schematic. After you make and place all your transistors, you realize, "Hang on a minute. 200 microns is enormous. And, it's 1 micron long!"

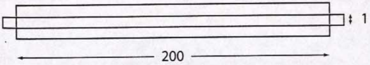


Figure 3–4. *Rather long, isn't it? Imagine if it was drawn to scale.*

Because of your experience, you will know to stop your layout right here, seeing the extreme dimensions of this item.

Long, thin transistors have problems.

Look at our FET cross section. Do you remember that we placed a tiny layer of insulating oxide between the Gate and the device's diffusion? That is the cause of one of our problems with long transistors.

However, we can do something about reducing the parasitic resistance, without changing the Gate area. We go back to our circuit designer and say, "Hey, when you made this transistor, it was long and skinny. You've got this big parasitic here. Why don't you break the transistor up into smaller ones and have less parasitic resistance?"

You sketch a few drawings quickly, adding, "Instead of having a single transistor like this:

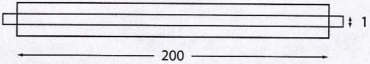


Figure 3–7. Long, thin transistor requested by circuit designers.

"Why don't we make four transistors of 50 microns each and hook them up in parallel? This arrangement will give us the same effective Gate width: 4 × 50 microns, that's still 200 microns."

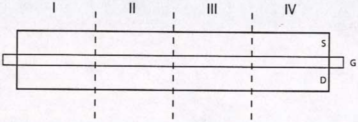


Figure 3–8. Splitting the long device. Source, drain and Gate will remain intact within each section.

You can think of four individual transistors as being equivalent to a single transistor if wired correctly. Each transistor terminal must be wired to the similar terminal of each of the other transistors. We have the same size transistor with the same effective Gate width, but with less parasitic resistance.

Each individual transistor has a Gate finger that is one fourth the width of the original device. This means that the Gate finger has a resistance that is one fourth of the resistance of the original device.

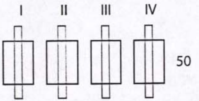


Figure 3–9. Splitting one long transistor into four reduces R.

In addition, the Gate fingers are wired in parallel. If you look ahead to our basic resistor equations, four equal resistors in parallel yield an overall resistance of one fourth of the individual resistance value. The overall effect of this splitting technique gives us a parasitic resistance that is one sixteenth of what we suffered originally.

Since our parasitic resistance has been reduced, our RC time constant has been reduced. The transistor will now operate much more effectively. There is no limit to this technique. Depending on the size of the device and other factors, we could split our transistor into even more segments.

In summary, we split a long transistor into four smaller transistors. We did not disturb the effective Gate width. However, we have dumped a substantial amount of parasitic resistance. The original parasitic capacitance still exists, because the Gate overlap cannot be changed. But, some improvement is better than none. When one of you figures a way to reduce capacitance without altering the specs, well, that will be another section for this chapter, won't it?

Source-Drain Sharing

Let's look at our long, thin device in more detail. We have labeled the source and drains *A* and *B* (interchangeable) and the Gate finger *C*.

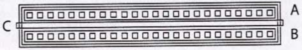


Figure 3–10. CAD drawing of a long, thin transistor. Either A *or* B *could be the source or the drain.* C *is the Gate.*

After splitting the device into four smaller devices to reduce its parasitic resistance, take a look at each individual transistor. We have *A*, *B* and *C* in each of the four transistors; that is, a source, a drain, and a Gate.

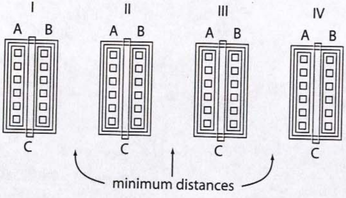


Figure 3–11. *Four transistors side by side. Differing nodes must be placed a minimum distance apart.*

We need to wire all the *A*'s together, all the *B*'s together and all the *C*'s together in order for our device to act as a whole.

We could wire the devices as in this figure:

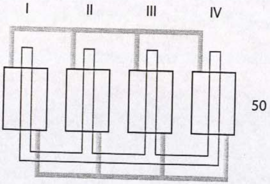


Figure 3–12. *Joining four devices. Sources are connected above (gray). Drains are connected below (gray). Gates are also connected below (blue).*

Everyone is keen on you saving lots and lots of space. In this field, chip area is directly proportional to cost. The smaller you can make your chip, the less your company pays to make it, the higher the profits. So, in general, you will make your layouts as compact as possible. This enables you to get as many transistors on a single chip as you possibly can.

The figured connection scheme above is perfectly legal for the four transistors we wish to connect. However, our proximity rules force the transistors apart. Differing nodes must be located a minimum distance from each other. We waste a lot of space using this connection scheme.

Time to get clever again. Here's how. The source and drain of a CMOS transistor are interchangeable. You can flip the device around, and it is still the same device. We are careful with the wiring, of course.

Let's flip two of our transistors. Redraw the same four transistors as *A-B-B-A-A-B-B-A* (no relation to the singing group). The second and fourth transistors have been flipped left for right, like pancakes.

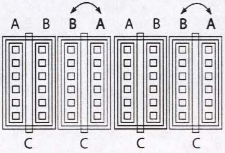


Figure 3–13. *Flipping transistors is legal. Here we have flipped two.*

Now, the two *B*'s are pointing toward each other, and the two *A*'s are pointing toward each other. The device wiring has now been made easier and we can move the transistors closer to each other. We are almost at the end of our journey.

The final part of our sample exercise is called **source-drain sharing.** We can pick up the first two transistors and smoosh them together such that the individual source-drain regions merge. The combined region is both a source for one transistor and a drain for the other, at the same time. Both were labeled *B*.

We can continue sharing all of the common source-drain regions until all of the nodes between the transistors have been partnered. Not only have we eliminated the spaces between the transistors, but we have combined parts of devices as well. Talk about your space savers.

Source-drain sharing can be done between any two devices located next to each other which bear a common terminal.

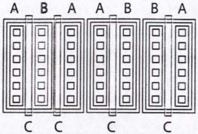


Figure 3–14. The first two adjacent B *nodes have been shared.*

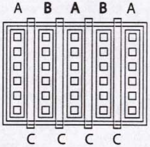


Figure 3–15. Common nodes next to each other have all been combined.

Try It

You can share like source-drain terminals if they are the same piece of wire. If one is an *A* and one a *B* then you cannot share.

You might have a circuit diagram that looks like:

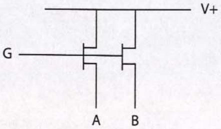


Figure 3–16. Exercise in sharing.

What items do you suppose can be shared? Do the layout.

ANSWER

The only terminal you can source-drain share is the V+. *You have to keep the* A *and* B *terminals separate because the diagram tells you to. (They are not connected.) If you draw a layout you get:*

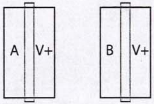


Figure 3–17. Non-sharing layout.

You cannot join these as they are drawn. The wrong nodes are in the middle. Let's flip the transistor on the right. That puts the V+ *in the middle, so you can join them. Did your layout look like this:*

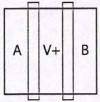


Figure 3–18. Good.

Device Connection Technique

Once we have smooshed all the *A*'s and *B*'s together, we must join all the like terminals with wire. The diagram below shows the wiring partially completed. All the *A* terminals have been connected in metal deposited in strips above the device. Notice the fingers of metal reaching down each terminal, all connected together at the top of the diagram. Notice the "M"-shaped piece of metal shown in gray.

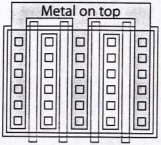


Figure 3–19. Joining the A*'s. The sources are now connected properly using a metal covering the tops of the desired regions.*

We will use the same technique to join the two *B* terminals. We deposit metal fingers, which reach up into each *B* terminal, joining together at the bottom of the diagram. Notice the "U"-shaped piece of metal, also shown in gray.

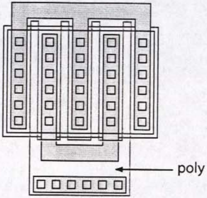


Figure 3–20. All sources connected, drains connected and Gates connected. This has the same effective Gate width as one long, thin transistor.

The Gate connections are slightly different, however. Polysilicon can be used in some special instances to perform circuit wiring. Poly conducts. We can extend the transistor Gate fingers out from the device, then use polysilicon to connect them.

Poly can be used as wire.

Using polysilicon as a wire is only recommended for very short distances. There are some dangers involved. Polysilicon is much more resistive than metal. After only a short distance, a significant wire resistance can develop. If the wire you are trying to build carries current and has a high resistance, there is a very likely danger that the circuit will cease to function. Use polysilicon wisely!!! That means watch your distance, your current and your resistance. In this example, we have also added some metal to the Gate contact so that any further wiring can be connected here, using more metal.

There you have our original 200 × 1 transistor. The two layouts in Figures 3-10 and 3-20 are identical in function, but our final drawing has a parasitic that is a lot smaller, operates more quickly, and uses chip real estate more effectively. Layout designers make choices. Remember that, it's what makes your job creative.

Layout designers make choices.

Following are some alternate ways of wiring the four transistors with some advantages worth noting.

If you work with the transistors auto-generated from your design tool, they will normally look like the CAD figures above. Each transistor will have many little square contacts stitched the entire length of its source-drain region.

However, if you want to save yourself some more room, you could throw away some of the contacts and wire your connections directly over your device.

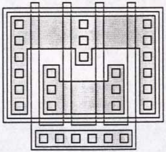


Figure 3–21. Bringing the metal inboard.

This effectively brings the jutting pieces of metal inboard. You do not need contacts all the way up the entire width, do you? We still have plenty that are usable. The original reason for all those contacts was to reduce the contact resistance of the device. A few contacts will likely suffice, but be careful. If you remove too many contacts, your contact resistance could be too high for your needs.

Try It

- Compare layouts. Can you see that these two drawings create equivalent devices? Can you see how we have reduced our device size by running metal over the top of the diffusions, instead of running it outside?

The strip of metal running over the top of the device is the same metal that was running outside before. The A*'s are still joined by metal that looks something like an "M." The* B*'s are joined with metal that still resembles a "U." The* C*'s are joined in polysilicon with a metal strip along the bottom.*

There is a third option. Examine this drawing a moment before reading on. What looks different?

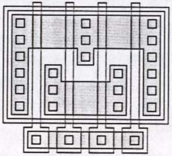


Figure 3–22. *Gate connections separated, connected in metal.*

We have the same two metal pieces for the *A* and *B* terminals, brought inboard again. No difference there. The difference is that we have connected the individual polysilicon Gate fingers with metal wiring. This method of Gate connection is the safest. It uses the minimum amount of polysilicon and ensures that we can safely connect to, from and through the Gate of the device without having to worry too much about the kind of signal present.

We have used circuit schematics to generate the basic components of our layout. Using source-drain sharing and device-splitting techniques. We reduced a parasitic at the same time. This is only the beginning of parasitic reduction. We will learn to strive at every turn to ensure that our parasitic elements have the least effect on our circuit's functionality.

Source-drain sharing, device splitting and parasitic reduction are fundamental techniques used throughout CMOS layout. You can use these techniques on many devices other than our small example. Keep your eyes open for

We have generated our devices with the optimum size and shape. But, all we have are a bunch of devices, floating around by themselves. We now need to examine ways to connect our devices and join the circuit blocks together.

Compact Layout

Most integrated circuits are designed using very small, easy to handle, easy to understand circuits. These little circuits are then wired together to create a larger, more complex circuit.

■ **Rule of Thumb: Build large projects from small, easy to understand blocks.**

This approach to circuit design makes layout much easier. Instead of trying to determine how to wire twelve million transistors in one go, you start with a circuit that has twelve and creep up on the larger problem.

The goal is to compact your layout. Part of that plan is to make your devices as rectangular as possible using device splitting. A rectangular circuit layout is much easier to use in conjunction with one million other rectangular circuit blocks than irregularly shaped lumpy circuit layouts.

Look at the following layout, for example. It has a big device poking out each side. If you were to use that set of transistor configurations, then you have wasted space. Imagine trying to fit dozens of these circuits together without wasting a drop of chip space. Nightmare. Probably impossible with most lumpy shapes.

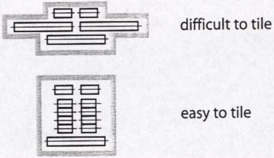


Figure 3–23. *Keep your device sets rectangular for efficient tiling.*

If you were to reshape the two middle devices by splitting them up into times-fours (four parallel devices), then you stand a much better chance of fitting all

Rule of Thumb: Try to make your device sets rectangular.

There will be times, however, when you just *have* to make your layout some weird, irregular shape. In this case, you need to examine the other circuits that interact with this piece of layout. If you are lucky, you can plan the device placement of the other circuits in such a way as to end up with a nice rectangular layout for the combined circuit.

Stick Diagrams

How do we easily get from a circuit diagram to the most efficient source-drain shared layout? A useful tool is a **stick diagram.** A stick diagram is a very simple representation of the devices and their interconnections. The diagram is a halfway sketch between the schematic and the final layout.

Once the stick diagram is complete, it gives you the device placement and connections. All you do after that is implement your stick diagram with real devices and wiring.

Let's use some of our old, faithful inverter circuits as an example.

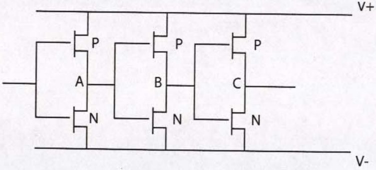


Figure 3–24. Sample schematic. P's on top, N's on the bottom, as is the convention.

In CMOS layout, we typically build all of the P Type devices in a common N well region. Each device must be built in an N well, which are all usually connected to the same point anyway. So, it makes sense to build the devices in a common N well. This common N well technique also reduces circuit area, as N well spacing rules are usually very large. Otherwise, the device spacing will be defined by N well to N well spacing rules, not transistor to transistor spacing rules.

Similarly, N Type devices are built in a common area, either a P well or P-Substrate.

Since we use common areas, all P Type devices are generally close to each other, and all the N Type devices are generally close to each other. Consequently, we can draw P diffusions as a horizontal stick across the top of the diagram, and N diffusions as a horizontal stick across the bottom of the diagram. In a stick diagram, you represent the polysilicon, diffusion and wiring as simple lines on a piece of paper. Where the lines for the polysilicon and the diffusion cross, we have a transistor. Since we represent everything with sticks, you can see where we get the name for this method.

Figure 3–25. Diffusions drawn as sticks. P devices are generally placed together along the top, N devices along the bottom.

Luckily, most schematics are drawn in this fashion already: P devices are drawn at the top, N devices at the bottom. This helps us create our stick diagram, since some of the device placement work has already been done for us by our circuit designer. If the connections on the schematic are short, it is likely that out layout connections will follow the same kind of placement.

Our device polysilicon Gates are then drawn as vertical lines that cross the diffusion. Each crossing makes a transistor. (Poly laid over diffusion, get it?)

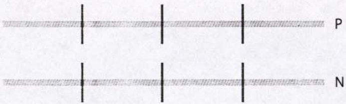


Figure 3–26. Gates drawn for six transistors.

Finally, our wiring is represented by lines that join the various device terminals. Device contact points can be represented any way you prefer. We will be using a small "x" to show where we want to make contact.

Now we start to work on the source-drain sharing. We know our first P Type transistor has a terminal marked *A* and a terminal marked *V+*. Let's jot these down on either side of the first Gate.

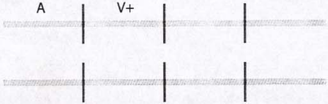


Figure 3–27. *Beginning to place some nodes by trial and error.*

Then proceed to jot down the rest of the transistors and their terminal names. Your first attempt might work very well, or your first attempt might not be able to source-drain share at all well.

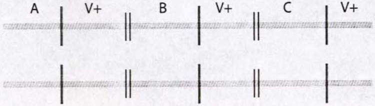


Figure 3–28. *Two breaks needed in the diffusion. Not a good design.*

In the above stick diagram, we have not done a very good job. We would have to **break** the diffusion into multiple pieces in order to layout the transistors in these orientations. Breaking the diffusion is denoted with two parallel lines across the break point. The split is forced into the diffusion to maintain adequate distances between unlike components. Sometimes breaking the device is necessary if the sources and drains cannot be shared. Each break in the diffusion pushes the transistors apart and so wastes some circuit area. The perfect solution would be no breaks at all.

We need to reduce the size of this layout. Let's examine the device placement to see if source-drain sharing can remove some breaks. We could combine the *V+* terminals.

Examining the terminal order with reference to source-drain sharing, we have the diagram below.

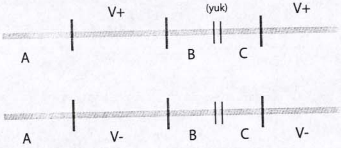


Figure 3–29. *Source-drain sharing does not always work perfectly.*

The above diagram is as far as we can go. There is no better source-drain sharing possible, regardless of how we flip the devices. On larger circuits, this source-drain sharing exercise is often repeated many times until you are convinced it can simplify no further. You may also need to change the order in which the devices are placed in order to get the most compact solution.

> *As a young lad, Chris used to work jigsaw puzzles all the time. If I'm working a jigsaw puzzle, I won't let him near it. He can't help it. They just fall into place for him. I think he knows how they'll all fit together even before I open the box, just by putting it next to his ear. —Judy*
>
> *This shows you what a lot of practice can do. Now when a circuit designer brings me a circuit I can often picture what the layout will look like even before I get anywhere near a stick diagram. Practice is the key. Lots of practice helps you see the solutions. They start to fall into place for you. —Chris*

In the first attempt above, I went for broke and tried to get a solution in one pass. The design was poor, with no sharing and two breaks in the diffusion. I quickly saw that I needed to change a few things. In this case, I shared in one place where terminals are the same, the *V+*'s. You just have to start somewhere, knowing it may not be the best pattern. But it gets you started.

Once you have worked out your source-drain sharing, you have a start for your device placement. You can join the rest of the terminals. In digital circuits, it is very typical for each P Type transistor to have a corresponding N Type transistor. Keep the pairs close to each other, and the Gates can be wired with a nice short piece of polysilicon.

Rule of Thumb: Diagram P and N pairs of transistors close to each other.

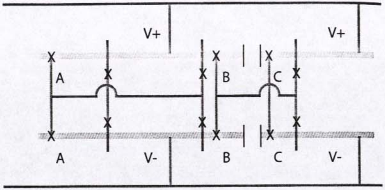


Figure 3–30. Finishing a stick diagram with connections.

Once you have all of your connections drawn, you have the template you will use for your layout.

In this example, we can improve the design further. Notice the last transistors (devices connected to terminal C) could be flipped to remove a wire crossover. The final stick diagram could look like this.

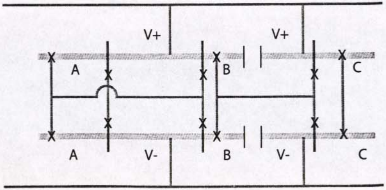


Figure 3–31. Stick diagrams are quick ways to keep toying with your design, until you find an optimal solution.

I flipped the last device over to make an easier connection between the output of the inverter made from device *B* to the inverter made with device *C*. Even after you finish your optimal source-drain sharing layout, you can still keep rearranging to make connections nicer.

Do you teach new hires stick diagramming? —Judy

No. Not for its own sake. Most people are already familiar with the method. Every now and again, however, I have worked with people who have made a career change into layout and are not familiar with stick diagrams. I teach them the technique when they need it.

Do you use stick diagrams?

Yes. I don't use it on a day to day basis, but sometimes you just have to sit down with a circuit and diagram it out. I personally use a hybrid stick diagram.

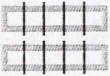


Figure 3–32. Stick diagram with regions.

By hybrid stick diagram, I mean that I use a rectangle for the diffusions instead of a stick. It gives me more of a feel of the devices I am working with. It's just a little closer to your final layout.

Stick diagramming is a useful technique, but on the simpler circuits, I just don't bother. Once you get used to doing layout, you do not need stick diagrams for every job. However, certainly for junior people, that's what they teach.

On the other hand, I have seen people with 20 years experience still use stick diagrams routinely.

We have drawn our basic stick diagram. We have sized our devices with the help of our SPICE simulator. That's it. Houston, we have a floorplan. We can start using our CAD tool to make a real layout.

Before you start your wiring, take the stick diagram, or the floorplan layout, to the circuit guys.[3, 4] Tell them, "Ok, here's what the devices look like for the circuit you designed. Do you like this? Do you like how they are placed relative to each other? Now that you can see what your devices look like from the circuit diagram, do you want to make some changes before I go off and wire this up?"

They might say, "Hmm, if we changed the size of a particular device we can rearrange things and make everything a lot smaller."

Or, "Oh, I didn't realize the devices were going to be that big. They don't need to be that big. Let's make them smaller."

Or, "If we move that device 20 microns to the right, we can move those resistors in there, and make the whole thing a lot more rectangular."

It is always worth going back to your circuit designer even if it is just to give yourself some feeling that, yes, you are doing things right and you are on the right track.

I strongly recommend that layout people do not work in isolation. Get help from the circuit people. Do not just go off blindly and do what you think is right. You will end up throwing away many hours of work if the circuit designer does not like the way you have placed the circuit's devices. If you get some buy-in from the circuit designer from day one then no one will be surprised by the final layout.

■ **Rule of Thumb: Keep the circuit designers in your communication loop . . . from the beginning.**

Well and Substrate Ties

We thought we could fire up our CAD tool and start wiring up devices with wild abandon. Not quite yet! Let's consider a few final issues before we commit polygon to database (modern equivalent of pen to paper).

Let's look at the device cross-section we drew earlier.

---

[3] Guys is California for people.—*Judy*

[4] No, it's not. It's British. Originated with Guy Fawkes, our famous gent with the kegs of gunpowder.—*Chris*

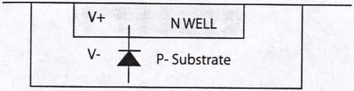


Figure 3–33. *PN junctions are diodes. Accidental forward biasing would be disastrous.*

In the above figure, we see an N well in a P− substrate. We have a PN junction formed by the N well and the Substrate. A PN junction is a diode, remember. If the voltage on the N well drops and the voltage on the substrate rises, there is a possibility you could forward bias that diode. You have to make sure you can never forward bias that diode. One sure-fire way to prevent the diode from becoming forward biased is to make sure it is always reverse biased.

The easiest way to do this is to connect the N well to the most positive supply and the Substrate to the most negative supply. These connections are known as **Well Ties** and **Substrate Ties.**

This is a PMOS device. We see that we have two well ties, one on either side of the FET. The well ties are simply N+ doped diffusions that are built in the N well. The N+ doping is low resistance which opens a window in the thick oxide to allow contact holes to be etched, and metal deposited.

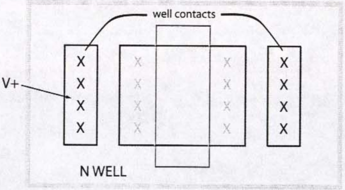


Figure 3–34. *Well ties on either side of an FET. Made of N+, connected to V+.*

The more well tie-downs you place, the less chance there is for the PN diode to become forward biased.

Rule of Thumb: There is no such thing as too many tie-downs.

The more the merrier.

The substrate ties are simply P+ doped diffusions that are built in the substrate. Just like an N+ diffusion gives us a low resistance contact to the N well, the P+ doping gives us a low resistance contact to the substrate.

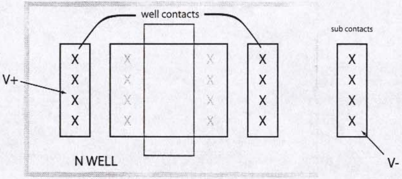


Figure 3–35. Substrate is tied to V−, seen outside the N well to the right.

In CMOS layout, you will typically see that the well is absolutely covered with N well ties.

Rule of Thumb: Wherever there is any spare space in an N well, put in a well tie. Wherever there is any spare space in the substrate, put in a substrate tie.

All this effort helps stop the parasitic diode between the well and the substrate from turning on. Remember that if the PN diode does turn on, it is a potential chip killer. This is a phenomenon known as latch-up. (We talk more about substrate latch-up, and more ways to avoid that, in our other book. This book deals with basics. Hence, the title. The companion book deals with essential techniques of layout, such as dealing with latch-up. And floorplanning, tools, verification, and lots of other advanced concepts you need to know.)

There are rules for your substrate and well ties. Rules that tell you how often you must have a well contact, and how close that well contact must be to your transistor. There are also rules as to how often you need substrate contacts. Not only how close the tie must be to N Type transistors, but also how close they can be placed to any transistor. For example, "You must have an N well contact at least every 50 microns." These design rules will be checked by the CAD tool you are using.

Well Tie Schemes

Let's examine some possible well tie connection schemes.

You might only be able to get ties at each end of a long thin well such as this one.

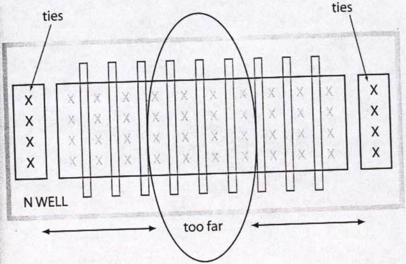


Figure 3–36. This design places some FETs too far from the nearest well tie. Living dangerously here. Your center PN junctions might blow up.

You may find that when you run your design rule checks that the transistors in the middle of the well might be too far from the nearest tie. If something like this happens, you might have to split the device and put some ties in the middle. Remember that the N well has some resistance associated with it. That resistance may have a voltage develop across it that can cause the PN diode to turn on.

An alternative is to increase the size of the well at the top of the device and put some well ties there.

Or, you might ring the whole well in well ties.

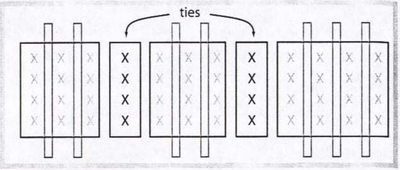


Figure 3–37. *Placing ties in the middle.*

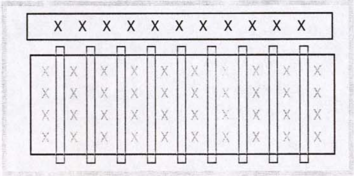


Figure 3–38. *Placing ties along the top.*

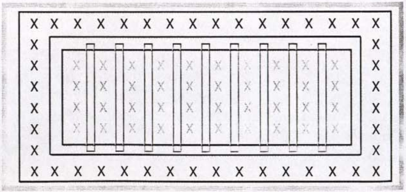


Figure 3–39. *Ties circling the device.*

If you do your wiring first, then later place your well ties, some of these rules may pop up and bite you. So tie down your well and tie down your substrate before you do any real wiring. You are then almost guaranteed that everything is ok.

- **Rule of Thumb: Place your well ties and substrate contacts before you do any wiring.**

Maybe even run the design rule checks before you begin wiring to make sure you have your wells and substrate tied down correctly. Wiring should be the last thing you do.

All of the above methods can be used for substrate tie-downs as well.

The Antenna Effect

CMOS Transistor Gates are very fragile and easy to break. We must pay very careful attention to how the device Gates are wired. We could introduce potential problems into the circuit that could kill the chip during wafer processing.

We worry about several processing issues during layout. One of these issues is known as the **antenna effect.** The antenna effect occurs during the polysilicon Gate etch process. Polysilicon is etched using reactive ion etching, RIE. (We talked about reactive ion etching earlier.)

Here is a diagram of our Gate, ready for etching in an RIE chamber.

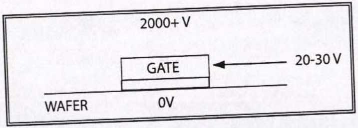


Figure 3–40. *In the RIE chamber we might see substantial voltage where we do not want it. Bbmmmmffffff! We're toast.*

There are more than 2000 volts placed across the RIE chamber. As the back o the wafer is touching the base of the RIE chamber, the wafer is effectively a zero volts. Our Gate, being caught in the middle, has a voltage somewher between zero and 2000 volts. As the polysilicon etches, a charge builds on th polysilicon that remains unetched. If we have a large single chunk of polysil con, then the charge build-up can be quite substantial. This charge build-up can translate into a reasonable voltage.

If you have too much voltage on polysilicon that is used as a transistor Gate, the Gate oxide could become damaged and kill the transistor.

If you use polysilicon to join all the Gates of a big transistor with, say, 500 Gate fingers, you will be creating a large chunk of polysilicon, won't you? If you can separate the Gate fingers into smaller chunks then you reduce the amount of voltage each chunk picks up. Much safer. Little bits probably won't fry under the voltage as a large chunk would. So, instead of wiring all your Gates together in poly, break them individually and wire them in metal. This is a much safer, more robust, more reliable way of making Gate connections.

Here we have four transistors. Their Gates each extend into a slightly wider chunk of poly with a contact. Metal connects all four Gates.

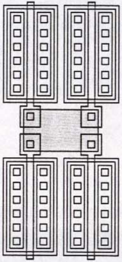


Figure 3–41. *Four Gates meet in the center. The small poly extensions are connected with a chunk of metal.*

Design rules usually specify the maximum size for a single Gate structure. The antenna effect is directly proportional to the Gate area. By Gate area, remember, we mean the amount of overlap of the Gate stripe and the active diffusion. If the Gate area is too big, you will have to break it up.

Splitting large poly areas into smaller poly areas protects the poly during the poly etch.

A secondary processing issue that changes how you wire your circuit is very similar to the antenna effect. We also perform our first metal etch in an RIE machine. Just as the voltage built up when we etched the polysilicon, we have a similar voltage build-up when the metal is etched. CMOS Gates are connected with first metal, so the voltage build-up on the metal conducts to the transistor Gates. We need to protect the transistors. The voltage needs to be kept to a minimum.

A reverse-biased PN junction can give us some protection. We discussed the PN junction in Chapter 1. Remember that if we reverse bias the junction enough, it will start to conduct at what is known as the breakdown voltage. This reverse breakdown voltage is much higher than any voltage that will normally occur in normal operation, but is low enough to protect the Gate of a transistor. A small diode in the substrate attached to the metal connection of a transistor Gate will limit how high the voltage can get. This diode is known as a **Gate Tie Down,** or **NAC Diode (Net Area Check),** and is very common in most modern CMOS processes.

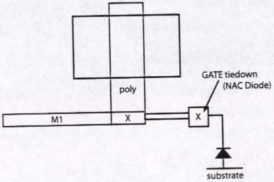


Figure 3–42. *Providing a tie-down to the substrate shorts any dangerous voltage buildup.*

Gate tie-downs protect poly during metal one etch.

Not all Gates require a Gate tie-down. If a Gate is connected directly to th source-drain region of another device in metal one, then it is tied down by th diode to the substrate of the other device. A good example of this is the outp of one inverter driving the input of another inverter. The first inverter's Gat need a tie-down, as they are floating, meaning they are not connected to an thing but each other. The second inverter input is directly connected in met one to the first inverter output, so it is automatically tied down.

Try It

Why not just make a tie-down for poly to protect it during poly etching? Think about it before reading the answer.

ANSWER

You can't. Ok, you can tie down poly to protect it against metal etching, but there's nothing you can do to protect poly against its own etching except split it up. You can't ground it because the field is needed for the etching process. Grounding it would defeat the process altogether.

Finally, we are now in a position to put all of our learning together and start layout. We have a schematic. We have created a stick diagram. We thought about where we would like to put our well and substrate ties. We have found which Gates need to be tied down to protect them from the antenna effect. Now for the wild abandon.

Wiring with Poly

As we mentioned earlier, you can wire with polysilicon. But, be careful when using poly as an interconnect. Wiring Gates with poly is the only safe option, but we have to be careful about obeying the antenna effect rules.

A Gate will only draw current when the parasitic Gate capacitor charges or discharges. Most of the time, a Gate only holds a voltage. If we wire Gates to each other with long runs of poly, we can quickly form a reasonable resistance. Just as a long, skinny transistor has a large parasitic resistance, the poly interconnect has large resistance also. The RC time constant of the Gate will be longer. Our circuit runs slower than we want it to.

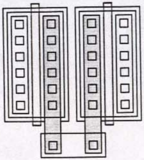


Figure 3–43. What if we connected metal source-drains with poly?

Sometimes our circuit is so complicated that we cannot wire it using metal alone. We can use poly as an underpass to get a signal out of a tight spot. We must understand the connection thoroughly if we want to use poly underpasses. Know what the connection does. Usually we make these poly underpasses as short as possible, so the extra resistance does not hurt our circuit.

What if we try to connect metal source-drains in poly?

In the figure above, we see a source-drain connection brought outside the devices. We have come off the transistor in metal, through poly, and back to the second device in metal again. We are using poly as an interconnect. There is a good likelihood some current will flow through this piece of interconnect. Therefore, we are much better off joining source-drains in metal. This is not a good place to wire with poly.

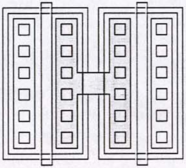


Figure 3–44. Connection is metal, due to expectation of substantial current.

Or, even better, source-drain sharing.

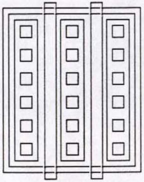


Figure 3–45. Source-drain sharing, no connections.

Rule of Thumb: If you want to distribute a voltage (just turning things on or off), you can use poly. If you want to distribute a current, use metal.

You can use poly as an interconnect, but typically only to connect the Gates, since there is not much current through the Gates.

Diagram Relationships

Could you—I mean you, personally, the reader—at this moment—I mean right now—take a circuit diagram, and make a stick diagram from it? If the device connections were easy enough, could you make a layout? If you are new to layout, you might be getting confused with the various drawing views we have been using.

Before trying some layout on our own, review how to move between the different views. Look at the various drawings for a few circuit examples to help you. Think in terms of how they relate to each other. Let's examine the various views for an FET side by side.

As mentioned before, active diffusion is a hole in the oxide. That defines where our transistors are built. The rectangle you draw is the active. You only draw one oxide hole for the entire Active area.

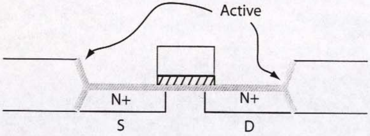


Figure 3–46. *The Active refers to the entire hole in the oxide.*

These drawings all represent the same transistor. Learn to translate quickly between them. Find the Active, drawn in blue, in each of the four views. Then find all the other components and connections in each view. How do the drawings relate to each other?

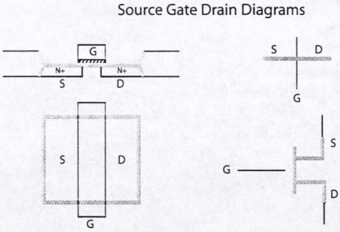


Figure 3–47. *Comparing four diagrams, each representing an FET. From top left, clockwise: cutaway side view, stick diagram, schematic, and top down layout.*

Try It

This is another good exercise to reproduce on your own without looking. Look back if you have to, but keep practicing until you can draw all four types of diagrams easily, knowing what each line represents. Once you master the FET, try drawings that are more complex.

Closure on CMOS Layout

I've gone as far as I want to go in CMOS layout for this book. All the basic techniques required for CMOS layout are in this chapter. You cannot really master CMOS layout from a book, though. You have to get in there and move those polygons.

Practice. Practice. Practice.

Here's What We've Learned

Here's what you saw in this chapter:

- Models, schematics and specifications for computer simulations
- Determining device sizes

- Avoiding resistance parasitic by splitting
- Better fitting techniques
- Stick diagrams as a design tool
- Bleeding voltages with tie-downs and contacts
- Prevention of forward biasing the inherent substrate diode
- The visual relationship between schematic, diagrams, and device
- Source-drain sharing

And more . . .

Applications to Try on Your Own

Using the circuit diagrams shown below, make a stick diagram from each. Then draw layout diagrams for #2 and #3.

1.

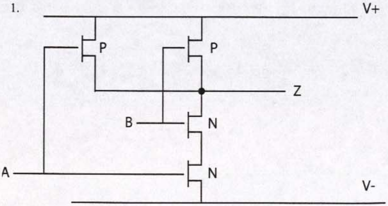


Figure 3–48

2.

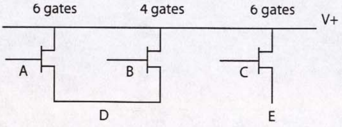


Figure 3–49

3.

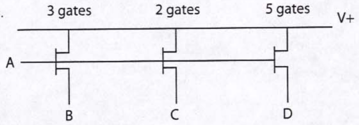


Figure 3–50

Answers

Many variations are possible. We provide only one possible answer for each diagram.

1. Possible stick diagram for #1.

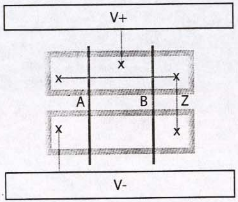


Figure 3–51

2. Possible stick diagram for #2.

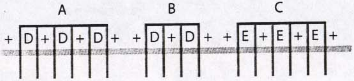


Figure 3–52

Possible layout diagram for #2.

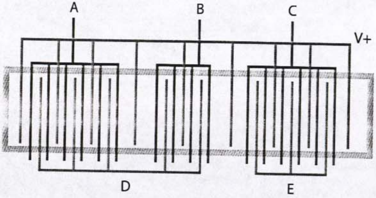


Figure 3–53

3. Possible stick diagram for #3. This stick diagram shows two breaks in the diffusion. This is not the best layout. It is probably someone's first try.

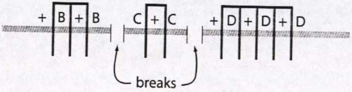


Figure 3–54

Possible layout diagram for the poorly designed stick diagram above. Notice the breaks in the diffusion.

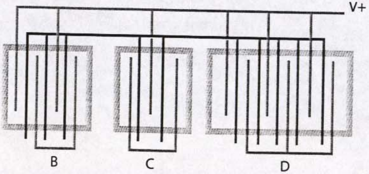


Figure 3–55

Better stick diagram for #3, probably after giving it some thought. Notice we are source-drain sharing all across. No breaks. Much better plan.

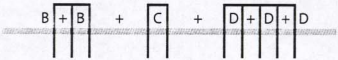


Figure 3–56

Possible layout diagram for the improved version above. Notice how much more compact is this diagram than the first attempt.

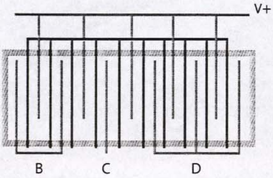


Figure 3–57

Extra to Consider

I've given you some very different types of practice cells to consider so that you can see what's happening. Often, in layout courses, they make you concentrate on the same type of simple cell until you can't stand to look at it any more.

Look at the differences in the examples I've provided. Learn from the differences. Seek more unique layouts to study.

Then, as I said before, practice, practice, practice.

CHAPTER 4

Resistance

Chapter Preview

Here's what you're going to see in this chapter:

- What each variable in the resistance equation means, and why it's there.
- The importance of and reason behind the units used in the calculations.
- How a resistor is fabricated, and why you need to know this.
- How to build resistors to compensate for errors in manufacturing.
- A few alternate resistor designs.
- Practical Rules of Thumb for resistor design.
- Resistors for extreme needs.
- How and why to check other calculations before doing your layout.
- Practice situations for you to try on your own.

And more . . .

Opening Thoughts on Resistance

Understanding the intricacies of resistor layout will allow you to catch mistakes, improve CAD tools, and read unfamiliar layouts. It allows you shortcuts, like building a diode with a resistor in series without having a separate diode and separate resistor. You can modify circuit elements in your layout to get new creations that do some unique tricks for you.

You will need to know how resistors are constructed if you want to get into such advanced areas as rules file writing or extractions files. For example, if you are trying to write a rules file to extract a specific resistor type from a layout, you have to know what uniquely defines the various chunks. For example, if I find the P, the P+, and the active touching my poly layer, then I have an FET.

I recognize quickly it is not my certain resistor and I can throw all that away. If I have my poly layer without an active touching it, then I know I have a resistor.

Many people just use the automatically generated components. However, you might need to develop resistor layout for a brand new process. For example, someone might say, "Ok, go lay out a bunch of diffusion resistors." Oh, crikey, there are no rules! No one has done it before! You have to understand what the layers are, how they're used, and where they fit in.

> *If you want to become the superstar of the startup company with 100,000 stock options and chance of being a millionaire in 5 years, you have to know this information. Just a regular layout person who places cells and joins dots won't cut it.*

Introduction to Resistance

There are two types of materials in this world—conductors and insulators. A **conductor** has the ability to allow electric current to flow through it. An **insulator** cannot allow electric current to flow through it.

Under extreme conditions, an insulator can break down, thereby allowing current to flow. This usually has catastrophic consequences, though.

There are good conductors and poor conductors. The extent to which a material can conduct electricity is characterized by giving the material a value, called a **resistance value.** Some conductors have a very high resistance, so high that for all practical purposes they can be considered insulators.

Everything has some level of resistance. Here are some examples:

- Metal has low resistance—great conductor, very poor insulator
- Air has high resistance—poor conductor, fair insulator
- Skin has moderate resistance—fair conductor, poor insulator

Likewise, in an Integrated Circuit (IC), every material used in the chip has its given resistance value. Some values are low. Some values are high, just as in the rest of the world.

Given a chip design project, then, the question becomes, "How do you make the resistors you want out of semiconductor materials used on the chip?"

To make the required resistive components from the materials available on a chip is a matter of learning to control values. Controlling values comes from the ability to calculate values. And that comes from measuring. The next section will show you how to measure resistance values.

Measuring Resistance

Once you learn to calculate resistance values, you will be better able to understand and control your circuit, thereby avoiding errors before the chip is built. This is a primary reason why some layout designers' chips tend to always work the first time. It's a valuable tool. Let's start with a few basic concepts about measurement.

Width and Length

Let me show you one of the easiest methods to calculate resistance value.

Integrated Circuit (IC) chips contain many types of materials such as polysilicon, oxide, various diffusions of basic CMOS transistors, and metal. A popular resistor material is **polysilicon,** also known as poly. We'll refer to poly often for examples in this book. Typically, all chip materials, including poly, are made in thin sheets.

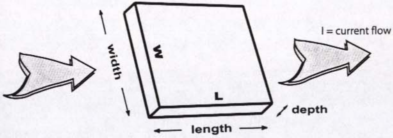


Figure 4–1. *Sheet of poly. Remember the conventions we will use throughout this book: Width will be the vertical dimension. Length will be horizontal. Current will flow left to right.*

Let's assume that we pass a current through a sheet of polysilicon, as shown in the diagram. If the sheet were thick, there would be more room for current to flow. Therefore, a thicker chunk would have a lower resistance value.

If it were very thin, it would have less ability to carry current, because there is less room for the current to flow through the material. Therefore, thinner sheets have higher resistance values. Other factors, such as type of material, length, and width also change resistive values.

These values of resistance are what we need to measure. We need to put exact numbers on these resistance levels so that we can make and exactly control resistance in our circuits.

Film thickness is considered constant within a given IC process. It's just one of those things we cannot expect to change. So, for a given material, the only parts we get to vary are width and length.

The only parts we get to vary are width and length.

Let's use only those two dimensions to calculate resistance. The next section will show us how to do that.

Concept of Squares

Let's make a resistor out of poly. To be easy on ourselves, let's begin by making the shape square. The width equals the length exactly.

If you run a current through our square of poly, and measure the voltage at both left and right edges, you can calculate a certain resistance value. Let's say we have measured this square of resistive stuff and just for argument's sake, let's say it comes out to be 200 ohms.

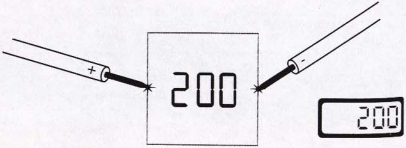


Figure 4–2. Measuring one square of resistive material.

Now let's wire two of these squares next door to each other. Each is 200 ohms. We would have a resistor with a total value of 400 ohms plus whatever resistance comes from the metal connections (*wires*).

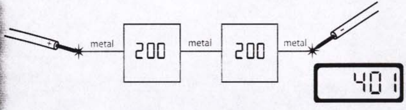


Figure 4–3. Resistance values in series add. Don't forget the wire adds some resistance as well.

But hold on a tick. What's the point of having all those metal connections between them? All materials have resistance, so these bits of connecting wires could actually change the total resistive value, from beginning to end.

If you were to make one resistor from both resistors pushed together, we rid ourselves of some wire. By basic laws of electronics, we add the two ohms together and get 400 ohms total for the pair. We no longer have any pesky wire to alter our accuracy.

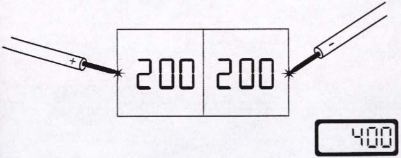


Figure 4–4. We have more control of our resistance values without the wires.

What if we were to wire four of these squares in a box pattern? Each square is exactly like the original, so each square is still 200 ohms.

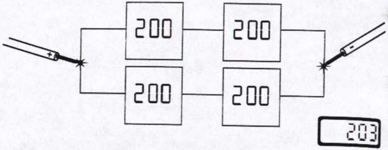


Figure 4–5. *Calculating four 200-ohm resistors with parallel and series considerations.*

Take it one step further. Just as we combined two resistors together earlier, why not join the top two and the bottom two together as well?

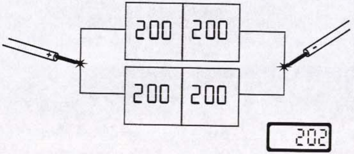


Figure 4–6. *Combining resistors to eliminate extra wires.*

In fact if we combine them in series like this, why not just combine the whole lot of them in parallel as well—put all four of them together? While we are at it, let's just get rid of all that extra wire and space.

Now let's remember some of our basics.

> 200 + 200 = 400 on top, and 200 + 200 = 400 on the bottom. So, we're equivalent in this shape to two 400's in parallel.

If we remember our parallel basics, we know that

$$\frac{1}{r} = \frac{1}{r_1} + \frac{1}{r_2} + \cdots \qquad \textit{ohms}$$

(See earlier chapter for help with this formula.)

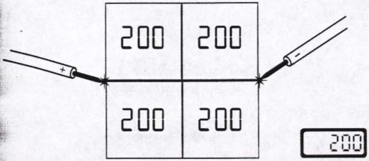


Figure 4–7. *Combining all four resistors in series and parallel still forms a square. Notice the total resistance value is the same as that of each smaller square.*

There are special cases for this equation. If r_1 and r_2 happen to be the same, 400 and 400 in this case, it just happens to pop out at 200 ohms for the total. When you have two identical resistors in parallel, it happens to halve the resistance value. Use the equation. Do the math.

So, we've taken four squares of 200 ohms material, arranged them into a bigger square, and we still get 200 ohms resistance over the whole arrangement. It's just one of those magic special cases. We could keep doing this, making larger and larger squares out of smaller squares, but the total resistance would always be 200 ohms regardless of the size of the square. The math always works. Try it.

So, the total value is still equivalent to the value of each original square resistor. Although it is four times the previous area, it is still 200 ohms and it is still square. So people tend to talk about resistance in what we call **ohms-per-square.** (The Greek letter omega and the shape of the square are used as symbols. 200 ohms-per-square looks like 200 Ω/□.)

We just count squares. All squares of the same material in the same process have the same value. It's just magic. (Oh, alright then, it's physics.)

Ohms-per-square is the basic unit used for IC resistance.

The ohms-per-square value is also known as the **sheet resistivity** of the material, as we will discuss in more detail shortly. This number will be used in your resistance formulas.

For dimensional analysis, the units are ohms in the numerator, squares in the denominator.

$$\frac{\textit{ohms}}{\textit{squares}}$$

If you use ohm-per-square values, you don't have to worry about how deep the material is. The only dimensions you vary are length and width. If the depth varies, all bets are off. However, in the same process, depth changes are just not supposed to occur. So, we don't worry about depth variation.

We can simply count the number of squares regardless of the size of the squares. Any size square will have the same resistive value as any other size square of this same material in this process. If you lay out a 1-micron by 1-micron resistor, it will have the same resistance as if you had made it 4 meters by 4 meters.

A square could be 5 miles by 5 miles, or 2 cm by 2 cm or .1 micron by .1 micron. It all comes out the same resistor value for the same material in the same manufacturing process.

Ohms-per-Square Values

To recap, ohms-per-square is the basic unit used for resistance. Ohms-per-square is also known as the sheet resistivity of that certain material. The number of ohms-per-square for a certain substance tells you how resistive that material is to electrical flow. Remember that you do not have to worry about thickness.

So let's say you have a long resistor made of eight of our previous squares. The material was found to have a resistivity of 200 ohms-per-square in our example. You then have 200 ohms-per-square multiplied by 8 squares.

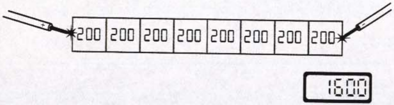


Figure 4–8. How many squares? Multiply the number of squares by the ohms-per-square value. There you are—the value of the entire resistor.

Before we multiply, let's look at the dimensional analysis just to double-check the reasonableness of our method:

$$\frac{ohms}{squares} \cdot squares = ohms$$

We see ohms-per-square is used in one factor and squares used in the second factor. We can cancel the unit squares since it appears in both numerator and denominator. Thus we are left with ohms for our resistance value, exactly the units we want. Since the units cancel correctly, we know our formula will work. Now we are free to use the formula with numbers:

$$\frac{200\ ohms}{1\ square} \cdot 8\ squares = 1600\ ohms$$

The method of counting squares then multiplying by the ohms-per-square value is a useful way to estimate or actually calculate the resistance of any IC material. The material could be poly, could be metal, could be anything you want.

Depending on the material, of course, the ohms-per-square value will change. The value of the material you are working with could be very low. For example, Gate poly is 2–3 ohms-per-square. (*Gate poly,* see previous chapter.)

You can use this ohms-per-square method to calculate resistor values regardless of the dimensions of the resistor. Any width and length can be recalculated into a number of squares. Let's say you have a strip of material that is, for example, 80 whatevers by 10 whatevers (could be any unit).

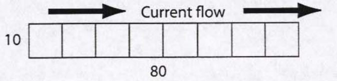


Figure 4–9. Number of squares can be calculated from any rectangle. Divide length by width.

We can divide the flow length by the width to find we have 80 / 10 = 8 squares. So, the number of squares you have can be calculated:

$$squares = \frac{L}{W}$$

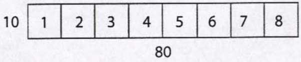


Figure 4–10. 10 × 80 resistor divided into 8 squares.

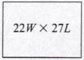
Unlike in junior school, not all your answers need to be whole numbers. You might have 4.28 squares, for example.

Try It

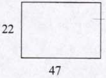
1. How many squares is each of these resistors? (Assume current flows left to right in each diagram.)

(a) Width = 5 Length = 65

(b)



(c)



(d) Width = 2 Length = 200
(e) Width = 200 Length = 2

ANSWERS

Length (horizontal) divided by width (vertical) equals the number of squares.

1. (a) $\frac{65}{5} = 13$ (b) $\frac{27}{22}$ (c) $\frac{47}{22}$

(d) $\frac{200}{2} = 100$ (e) $\frac{2}{200} = 0.01$

Typically, for each manufacturing process you have a book of values, probably locked in that small floor safe behind you. The manufacturer might call this the **Design Manual,** the **Process Manual** or the **Rulebook.**

This is where you will look up the resistivities in terms of ohms-per-square for your material. In your book, this will be called either sheet resistivity or **sheet rho.** (The Greek letter rho is pronounced "row," as in "rho, rho, rho your boat"—which only works in the English translation of this book, by the way. I can't wait to see how the translators handle that one.) The symbol for rho is ρ.

You will find ohms-per-square listed as sheet ρ.

The process manual will have a sheet rho value for every material you can use in each process. Typically, the manufacturer provides this book of values for you.

The same material processed by different manufacturers could get different values, depending on thickness for example. Your book has all the magic numbers for all the materials for their processes, which they have acquired by extensive testing.

How to Determine Ohms-per-Square

Companies make test chips using big, huge, square resistors. They pass a current through each one, and measure the voltage on what they call a **4-point measurement.** You pass a given current through the square and measure the voltage difference. (No current goes through the terminals that are actually measuring the voltage.) This is also known as the **Force/Sense** method of measurement.

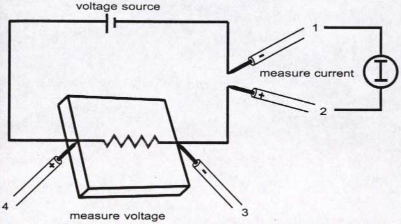


Figure 4–11. Testing a large square to find ohms-per-square value for the material using 4-point measurement.

So, you have measured the voltage and you know the current. Using $V = IR$, you can calculate the resistance. Since it is a square, hey presto, there's your ohms-per-square value for that material.

Sometimes as a layout designer, you will be designing the test masks that will be used to determine these sheet rho values, so you have to know a bit about how these things are made.

Changing Typical Resistive Values into Useful Values

When somebody is designing an IC process, they typically will go to the circuit designer and say, "If you had your wishes in the world. what kind of resistance values would you like to see?"

And they'll say, "You know, I could use some really low value ones, some medium value ones, and some really high value ones."

The process engineers then look at the various materials they already have in the process, and say, "Ok, now, which of these can we use to get all those different values?"

Some typical IC resistive values:

- Gate poly—2 to 3 Ω / □
- Metal—20 to 100 mΩ / □ (small resistivity; good conductor)
- Diffusion—2 to 200 Ω / □

You can use any material within an IC process. Anything can be made into a resistor, but some materials are better for resistance than others for various reasons. Poly is a very good material because it's free, and has a well-controlled length and width. By free, I mean the layer is already there for you to use, since you use poly within a MOS transistor anyway. There will be no additional expense to deposit extra layers of new material just to make a resistor. You already have diffusions, as well, because they are part of a transistor.

Typically, though, the raw values of IC materials are really low, like 2 to 200 ohms-per-square. If you just make a resistor out of your raw material and you measure it, you'll probably find it's too low to be useful.

Some useful ranges for resistor values:

10 to 50 ohms
100 to 2K ohms
2K to 100K ohms

So, if you are trying to make a 100K-ohm resistor out of something that's only 2 ohms-per-square it will be huge. Enormous. It'll be 50,000 squares. Or let's say you try to make a 50-ohm resistor out of something that's 200 ohms-per-square. It'll be tiny. Or really, really wide. In order to get a reasonable length out of it you would have to make it hugely wide.

What we can do in this case is add extra layers or extra process steps to get ohms-per-square values that will be useful. Your circuit designers may say, "Make this diffusion" and "Put more (or less) implant into it." (*Implant,* see previous chapter.) So, with extra processing steps we can modify resistor values up into useful ranges.

Now we are entering the real world. In the next section, we will learn how these real world values are calculated.

So, to sum up, if we know the sheet resistivity of a material, its length and its width, then we can calculate the basic resistance value for that resistor, using the following formula:

$$R = \frac{L}{W} \cdot \rho \qquad \textit{ohms}$$

R = resistance (ohms)
L = length (microns)
W = width (microns)
ρ = sheet resistivity (ohms-per-square)

Deriving the Resistor Formula Using a Poly Resistor

You have just seen why resistance material is measured in ohms-per-square, also known as sheet rho. Now we will use a poly resistor to understand the different parts of the complete resistance equation for a real world IC resistor. You will notice that many elements of the formula are given in the ohms-per-square units we just learned.

Basic Resistor Layout

Let's follow the manufacturing process involved in making a polysilicon resistor, in order to understand all the components.

The **substrate** material is the bottom layer of raw silicon. On the substrate, we deposit a poly layer, which we will use to resist electrical flow.

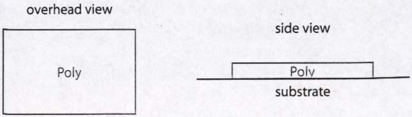


Figure 4–12. *Polysilicon layer.*

We need to connect our bits of poly material somehow into the path of electrical flow. So, on top of the poly layer we lay down an oxide layer. This good insulator will protect further layers from contacting the poly in unwanted places.

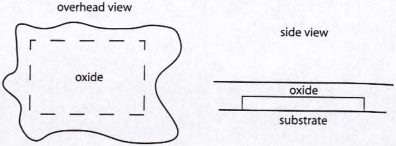


Figure 4–13. Layer of oxide covers the poly layer.

Then we can etch through the oxide, making **contact** holes. These holes dig through to the poly in exact locations just where we want to make contact with the poly. Hence the name *contact* holes.

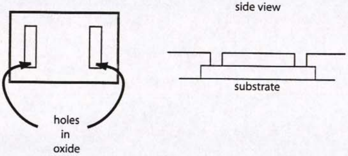


Figure 4–14. Holes cut through the oxide to expose the poly.

We deposit some metal into these etched holes. The metal fills the contact hole, contacting the poly. In the diagram, you will notice one etched contact hole at one end of the poly and another etched contact hole at the other end of the poly. Making one of these a positive contact and the other a negative contact allows the current to flow through the poly resistor on its way through our circuit. (The metal does not touch the poly anywhere except through the holes in the oxide layer, which have been carefully placed.)

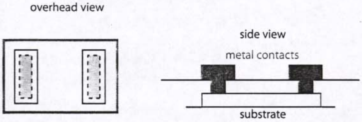


Figure 4–15. Metal is deposited into the holes to make contact with the poly.

So, the actual contact is merely a hole etched down through the oxide layer, filled with metal.

This is the simplest form of a resistor. Its basic resistance value is found from the width of the material and the length between the two contact holes.

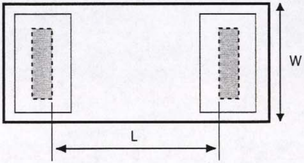


Figure 4–16. Simplest form of a resistor.

Using the equation developed earlier, let's put subscripts on the length, width and sheet rho indicating we mean the body length, the body width and the sheet resistivity of the body material. We will continue using subscripts in this manner as we continue to develop the formula. So, our formula looks like this:

$$r_b = \frac{L_b}{W_b} \cdot \rho_b \qquad \text{ohms}$$

r_b = *resistance (ohms)*
L_b = *length of body (microns)*
W_b = *width of body (microns)*
ρ_b = *sheet resistivity of body material (ohms-per-square)*

Try It

Use the formula $r_b = \frac{L_b}{W_b} \cdot \rho_b$ to determine the following resistor values.

1. $L = 10$ microns
   $W = 10$ microns
   $\rho = 200\ \Omega / \square$

2. $L = 129$ microns
   $W = 2.5$ microns
   $\rho = 500\ \Omega / \square$

3. $L = 10$ microns
   $W = 65$ microns
   $\rho = 350\ \Omega / \square$

4. $L = 5$ microns
   $W = 200$ microns
   $\rho = 10\ \Omega / \square$

5. $L = 96$ microns
   $W = 12.2$ microns
   $\rho = 956\ \Omega / \square$

ANSWERS

Divide L by W, then multiply by ρ.

1. $\frac{10}{10} \cdot 200 = 200\ \Omega$
2. $\frac{129}{2.5} \cdot 500 = 25{,}800\ \Omega$
3. $\frac{10}{65} \cdot 350 = 53.85\ \Omega$
4. $\frac{5}{200} \cdot 10 = 0.25\ \Omega$
5. $\frac{96}{12.2} \cdot 956 = 7522.62\ \Omega$

Contact Resistance

If we make and measure a square resistor, we can determine the ohms-per-square of the body material quite easily, especially if the square is very big.

Early engineers made lots of square resistors of different sizes—10, 20, 30, 40 microns wide, for instance. They plotted a graph of square size versus resistor value. The graph was expected to be constant, since they knew that all squares of the same material have the same resistance value.

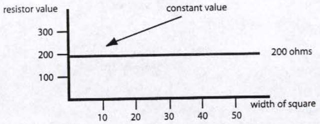


Figure 4–17. The resistance of a square of poly should stay constant even as your square increases size, as discussed previously.

What we see in reality, however, is that the smaller resistors, as measured through their metal contacts, give a measured value that is higher than expected. This discovery was quite surprising.

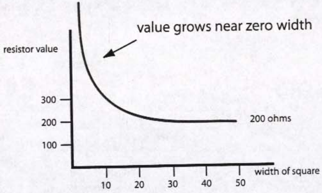


Figure 4–18. In the real world, as the square gets small, you can see the resistance value does not remain constant, contrary to our magic rule of squares that we learned earlier.

We have resistors with the same sheet resistivity, all squares, but we get different resistance values. In fact, this change in resistance value becomes quite dramatic at small widths. This shouldn't happen unless something else is contributing to the resistance besides the poly. In fact, that is the case.

Remember that every material in the world has some degree of resistance. As we look at our resistor cross section again, you can see that the layer of poly is the main resistance in this pathway. Notice also, though, that you can see small resistances up each left and right chunk of metal. Although it is very tiny, it is there.

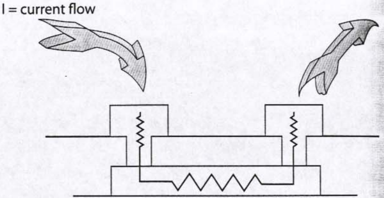


Figure 4–19. *Resistance is shown through the small metal contacts as well as through the poly.*

Since we have discovered two very small resistance values, we must account for them in our equation. We will denote the resistance of the contact material as r_c and the resistance of the body material as r_b.

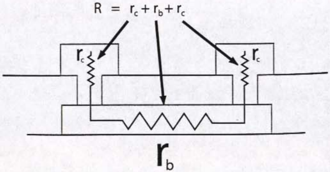


Figure 4–20. *Symbols* r_c *and* r_b *used to represent resistance values.*

So, our equation for the total resistance becomes:

$$R = r_b + 2r_c \qquad \textit{ohms}$$

R = total resistance (ohms)
r_b = resistance of the body (ohms)
r_c = resistance of the contact metal, one on each end (ohms)

We know how to calculate r_b. We learned that earlier when we discussed ohms-per-square. But what about the contacts?

Contacts do not lend themselves well to the idea of squares. For one thing, contacts are considered a fixed length. After all, why would you bother to lengthen a contact when all the electron flow goes out only from one edge anyway? Not only that, but contacts have an odd-shaped cross section. We find that ohms-per-square just does not work for calculating contact resistance.

Let's look at an example to help us learn how to measure contact resistance.

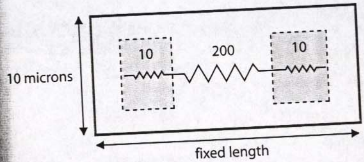


Figure 4–21. *Each contact resists electrical flow.*

Start with a resistor that has a width, say, of 10 microns. Let's say we find its contact resistance, which we will symbolize as r_c, to be 10 ohms for each contact, in addition to the 200-ohm resistance we measure in the poly.

Now let's double the resistor width, but keep its length fixed. We now have a 20-micron resistor, twice the width, of the same material and same length. We find it has a contact resistance of only 5 ohms when measured. Likewise, the overall body resistance is now only 100 ohms.

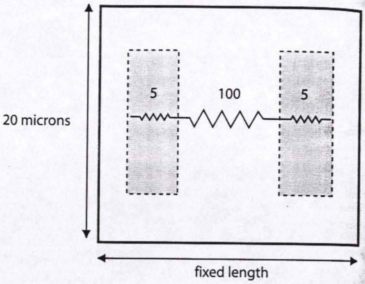


Figure 4–22. Twice as wide, the resistance from the contact measures only half what it was before.

When we doubled the width of the contact, we doubled the available path for electrons to flow. Doubling the width of the contacts in this way is exactly like having two 10-ohm resistors in parallel. This naturally gives us less resistance. If you combine two 10-ohm resistors in parallel, you have only 5 ohms resistance overall.

If the contact width were very small, on the other hand, the contact resistance would be comparatively high. This explains the sharp increase in the resistance value as the resistor width goes to zero. Remember that the length of a contact is not relevant, so lengthening the contact to keep it square cannot be used to compensate.

If the contact width gets larger, the contact resistance goes down. If the contact width gets smaller, the contact resistance goes up. This illustrates the rule that the contact resistance is inversely proportional to its width. We can use this rule, with what we know about dimensional analysis, to construct a method of calculating contact resistance value.

Here is what we know about contact resistance so far:

- We know we must end up with ohms as our final unit, since we are trying to find resistance, and resistance is measured in ohms.
- Therefore, we need some sort of ohms unit in the numerator.

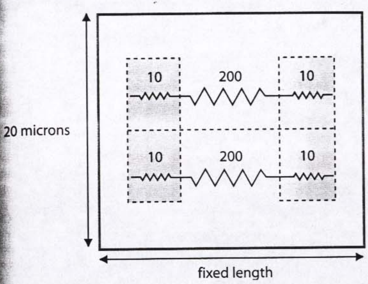


Figure 4–23. Widening the contact is equivalent to placing contacts in parallel. Resistance value is inversely proportional to width.

- We expect to divide by the width of the contact, having seen that the resistance is inversely proportional to width.
- Therefore, we know we will see microns in the denominator, since width is measured in microns.
- In that case, we must place some sort of micron unit in the numerator, since the micron units must completely cancel. Otherwise, we would be left with microns as part of the result.

Therefore, our numerator must be ohms times microns. Our denominator must be microns.

There you have it. We will define a resistance factor for the contact material, which we will express as ohms times microns. We use ohm-microns to distinguish that numerical value unique to the process, which yields a correct width-dependent resistance value. We will use the symbol R_c to denote this ohm-micron factor for the contacts.

$$R_{contact} = \frac{R_c}{W_c} = \frac{ohm \cdot microns}{microns\ (width)} \qquad ohms$$

$R_{contact} = r_c$ = *total resistance due to contact (ohms)*
R_c = *resistance factor attributable to contact (ohm-microns)*
W_c = *width of contact (microns)*

In the example above, since we know our contact resistance is 10 ohms, and we know our width is 10 microns, our contact resistance factor must be 100 ohm-microns.

$$10\ ohms = \frac{100\ ohm \bullet microns}{10\ microns}$$

Ohm-micron values are calculated in this way for you by the manufacturer of the process.

But, what exactly is an ohm-micron? We have seen why the units work in the equation, but the concept is man-made. You can physically see what a micron is. It's a small length. You can see what a liter is. It's a certain size jar full of liquid. Likewise, you already have a sense of a mile, meter, gram, hour and other units. However, some units do not lend themselves well to visualization. Ohm-micron is one such unit. It's unreal, like the unit of acceleration, meter per second per second. Have you ever stopped to ask yourself what a square second is? We work with these all the time, but on some occasions, they provide convenient, workable units more than they actually represent the real world.

Ohm-micron is merely a unit we use for resistance factors. One can try to visualize a unit made of length multiplied by resistance as a fun mental exercise, but it is not necessary. I can think of better reasons to lose sleep. Like watching movies full of explosions, jig-sawing wood for my latest custom guitar, teasing the kids, or chasing the wife.

Remember to first work your equations in units only, to verify that you are obtaining the correct units expected in your answer. If you work calculations only with numbers, people will die, cellular phones will not work, and space ships will crash into Mars.

To check our learning, let's practice a few calculations.

Try It

Use the above ohm-micron equation to calculate the contact resistance for the following resistors.

1. Given a contact resistance factor of 100 Ω-microns, calculate the ohms through each of these resistors:

   (a) Width = 5 microns
   (b) Width = 10 microns
   (c) Width = 22 microns

2. Given a contact resistance factor of 22 Ω-microns, calculate the ohms through each of these resistors:

   (a) Width = 2 microns
   (b) Width = 22 microns
   (c) Width = 50 microns

ANSWERS

Divide Ω-microns by width (given in microns).

1.

(a) $\frac{100}{5} = 20\ \Omega$

(b) $\frac{100}{10} = 10\ \Omega$

(c) $\frac{100}{22} = 4.54\ \Omega$

2.

(a) $\frac{22}{2} = 11\ \Omega$

(b) $\frac{22}{22} = 1\ \Omega$

(c) $\frac{22}{50} = 0.44\ \Omega$

Using the ohm-micron factor, R_c, to determine our contact resistance makes our total resistance equation look like this:

$$R_{total} = \frac{L_b}{W_b} \bullet \rho_b + 2\,\frac{R_c}{W_c} \qquad ohms$$

R_{total} = *total resistance (ohms)*
L_b = *length of body (microns)*
W_b = *width of body (microns)*
ρ_b = *sheet resistivity of the resistor body (ohms-per-square)*
R_c = *resistance factor of the contact (ohm-microns)*
W_c = *width of the contact (microns)*

NOTE: The contact width may not necessarily be the same width as the resistor. Depending on the design rules of a process the contact may need to be smaller than the resistor.

Try It

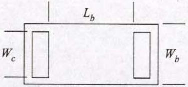
1. Use these given values to calculate how long to make each of the five resistors in the chart below.

Given: $R_c = 25\ \Omega$-microns
$\rho_b = 1200\ \Omega / \square$
$W_c = W_b - 1$ micron

#	W_b (microns)	R_{total} (ohms)	Length between contacts
1a	20	1200	
1b	2	10,000	
1c	10	850	
1d	50	200	
1e	15	425	

2. Use these given values to calculate the contact resistance, the body resistance, and the total resistance of each of the five resistors in the chart below.

Given: $R_c = 60\ \Omega$-microns
$\rho_b = 270\ \Omega / \square$



#	L_b	W_b	W_c	$R_{contact}$, R_{body}, R_{total}
2a	100	10	8	
2b	20	12	10	
2c	45	6	4	
2d	6	20	18	
2e	27	50	48	

ANSWERS

1. Use the formula given in the above text. Solve for *L*.

1a. $1200 = \frac{L_b}{20} \cdot 1200 + 2\,\frac{25}{20 - 1}$

Resistor body length = 19.96 microns

1b. $10{,}000 = \frac{L_b}{2} \cdot 1200 + 2\,\frac{25}{2 - 1}$

Resistor body length = 16.58 microns

1c. $850 = \frac{L_b}{10} \cdot 1200 + 2\,\frac{25}{10 - 1}$

Resistor body length = 7.04 microns

1d. $200 = \frac{L_b}{50} \cdot 1200 + 2\,\frac{25}{50 - 1}$

Resistor body length = 8.29 microns

1e. $425 = \frac{L_b}{15} \cdot 1200 + 2\,\frac{25}{15 - 1}$

Resistor body length = 5.27 microns

2. Use the formula given in the above text. Solve for R_{total}.

2a. $R_{total} = \frac{100}{10} \cdot 270 + 2\left(\frac{60}{8}\right)$

$R_{total} = 2700 + 2(7.5)$
$R_{total} = 2715$

$R_{body} = 2700\ \Omega$
$R_{contact} = 7.5\ \Omega$ each
$R_{total} = 2715\ \Omega$

2b. $R_{total} = \frac{20}{12} \cdot 270 + 2\left(\frac{60}{10}\right)$

$R_{total} = 450 + 2(6)$
$R_{total} = 462$

$R_{body} = 450\ \Omega$
$R_{contact} = 6\ \Omega$ each
$R_{total} = 462\ \Omega$

2c. $$R_{total} = \frac{45}{6} \cdot 270 + 2\left(\frac{60}{4}\right)$$

$$R_{total} = 2025 + 2(15)$$
$$R_{total} = 2055$$

$$R_{body} = 2025\ \Omega$$
$$R_{contact} = 15\ \Omega \text{ each}$$
$$R_{total} = 2055\ \Omega$$

2d. $$R_{total} = \frac{6}{20} \cdot 270 + 2\left(\frac{60}{18}\right)$$

$$R_{total} = 81 + 2(3.33)$$
$$R_{total} = 87.66$$

$$R_{body} = 81\ \Omega$$
$$R_{contact} = 3.33\ \Omega \text{ each}$$
$$R_{total} = 87.66\ \Omega$$

2e. $$R_{total} = \frac{27}{50} \cdot 270 + 2\left(\frac{60}{48}\right)$$

$$R_{total} = 145.8 + 2(1.25)$$
$$R_{total} = 148.3$$

$$R_{body} = 145.8\ \Omega$$
$$R_{contact} = 1.25\ \Omega \text{ each}$$
$$R_{total} = 148.3\ \Omega$$

"Why would a layout designer want to do these calculations?"—Judy

The folks in the lab come along and have a new process. "We need you to put together the software such that all little Johnny has to type is that he wants a 1K resistor, 5 microns wide. Go ahead and calculate the resistor for him."

Someone has to make these things to fit the values. In the old days, I had to calculate the values of each resistor we were given. I had a little spreadsheet, took the book numbers, figured out how long a resistor had to be in order to get the value we want, given constants. Basic algebra, if you know these things.

I would make myself one resistor based on the rules. I would be told that these numbers come from this kind of structure, so all I'd do is take it and stretch the darn thing or grow its width. I'd be told the circuit designer wants 1K in there and 5 microns wide. So, I'd have to calculate the length. I'd have to make it. We usually have to find L, *not* R.

By the old days, I'm only talking 5 or 10 years ago. Now today, you fire up the software, hit the "Generate me a resistor" button, enter the value and the width you want, and it creates the layout for you. But, someone had to program that in. Someone had to understand it. Typically, those people are layout engineers who know their stuff and have moved on.

So, this information is good for someone to know who wants to move on or do some more softwarey stuff. You may be working in a research lab, there may not be a design kit built for you. You may have to do all this from scratch.

Changing Body Material

So, that's the simple resistor. I'm going to make it more complicated now.

Typically, when poly is used on an FET, it is desirable to have as low a Gate resistance as possible to make the FET operate as fast as possible.

If you remember, Gate poly is only about 2 to 3 ohms-per-square. This low value works well for Gates, but useful ranges for resistors are usually much more than that. You may want something in the 100's of ohms. You may really want 200 or maybe 250 ohms-per-square.

So we say, "Ok, let's change the resistor." You start with the same layout we've already talked about—poly, contact and metal. Here's our simple resistor drawing again:

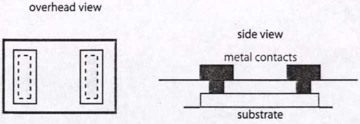


Figure 4–24. Our basic resistor.

Changing the body material effectively can raise the resistance. This will help give you a higher, more useful value of resistance.

There are various ways to change the body material. One way would be to deposit a different layer of poly that has different resistive characteristics.

However, why not change the body using materials already being deposited onto the chip? This saves having to use extra processes and adds no cost.

Let's open a window to work with in the middle of the poly to give us a higher ohms-per-square value.

One way to make a region of higher resistance is to implant extra stuff in the poly, discouraging the easy flow of electrons. Another way to increase resistance is to etch some of the poly away, making it thinner. So, either way, you change the chunk in the middle in order to increase the resistance value.

We will now call this changed chunk the **body** of the resistor. The dimensions of the resistor, of course, are still the same. We have simply altered the material in the center.

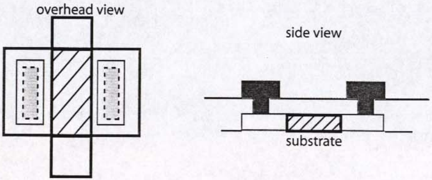


Figure 4–25. Implanting a higher resistance material into the poly to raise the resistance of the path creates a body portion in the center.

So, we have a new, specially treated area that we consider the body. There's usually a design rule that says how close the contact can be to the changed body bit, so you always have some extra poly between the body and contacts; just extra stuff in the way. These two extra chunks of leftover poly on either side we will call the resistor **head.** We will have a head region at each end of the resistor. These two head areas have a resistance value that must be included in our equation.

The body could now be 250 ohms-per-square and the original poly left and right could be 2 ohms-per-square. (That's why we changed the middle bit, of course.) Then there's the resistance of the contact as well. These are all different resistances. So now, our total resistance equation includes three parts: the specially treated body, the two extra chunks of original poly (one on each end), and the two metal contacts (also one on each end). Each of these resistances must be calculated separately, then added together in order to get the total resistance value of this resistor.

In simplified form, our total resistance equation now becomes:

$$R = r_b + 2r_h + 2r_c \qquad \textit{ohms}$$

R = *total resistance (ohms)*
r_b = *resistance calculated for the body portion (ohms)*
r_h = *resistance calculated for each head section (ohms)*
r_c = *resistance calculated for each contact (ohms)*

There's a specific ohms-per-square value for the body, plus a specific ohms-per-square value for the head areas, plus another ohm-micron value for the contacts. Not only that, but each region has its own dimensions. So, the equation starts to get quite complicated.

You use the ohms-per-square or ohm-micron values to calculate each resistance portion. If we expand each of our simplified r variables from the above equation, our expanded equation looks like

$$R = \frac{L_b}{W_b} \cdot \rho_b + 2\,\frac{L_h}{W_h} \cdot \rho_h + 2\,\frac{R_c}{W_c} \qquad \textit{ohms}$$

R = *total resistance (ohms)*
L_b = *length of body (microns)*
W_b = *width of body (microns)*
ρ_b = *sheet resistivity of the body (ohms-per-square)*
L_h = *length of head (microns)*
W_h = *width of head (microns)*
ρ_h = *sheet resistivity of head (ohms-per-square)*
R_c = *resistance factor of the contact (ohm-microns)*
W_c = *width of contact (microns)*

Try It

1. Use the equation in the text above to calculate the total resistance of each resistor. Use the given values for R_b, R_h and R_c and width requirements.

Given:
$\rho_b = 300\ \Omega / \square$
$\rho_h = 2\ \Omega / \square$
$R_c = 70\ \Omega$-microns
$W_b = W_h$
$W_c = W_b - 2$ microns

(a) $L_b = 20\ \Omega$
$L_h = 2\ \Omega$
$W_b = 5$ microns
(b) $L_b = 60\ \Omega$
$L_h = 4\ \Omega$
$W_b = 20$ microns
(c) $L_b = 10\ \Omega$
$L_h = 1\ \Omega$
$W_b = 50$ microns

ANSWERS

Substitute variables into the text equation:

$$R = \frac{L_b}{W_b} \cdot \rho_b + 2\,\frac{L_h}{W_h} \cdot \rho_h + 2\,\frac{R_c}{W_c} \qquad \textit{ohms}$$

Solve algebraically. Each term yields one portion of the total resistance—body, head and contact.

$$\text{(1a)}\quad R_{total} = \frac{20}{5} \cdot 300 + 2\left(\frac{2}{5}\right) \cdot 2 + 2\left(\frac{70}{5-2}\right)$$

$$R_{total} = 1200 + 2(.4) \cdot 2 + 2(23.33)$$

$R_{body} = 1200$
$R_{head} = 0.4$ each
$R_{contact} = 23.33$ each
$R_{total} = 1247.46$

$$\text{(1b)}\quad R_{total} = \frac{60}{20} \cdot 300 + 2\left(\frac{4}{20}\right) \cdot 2 + 2\left(\frac{70}{20-2}\right)$$

$$R_{total} = 900 + 2(0.4) + 2(3.89)$$

$R_{body} = 900$
$R_{head} = 0.4$ each
$R_{contact} = 3.89$ each
$R_{total} = 908.67$

$$\text{(1c)}\quad R_{total} = \frac{10}{50} \cdot 300 + 2\left(\frac{1}{50}\right) \cdot 2 + 2\left(\frac{70}{50-2}\right)$$

$$R_{total} = 60 + 2(0.04) + 2(1.46)$$

$R_{body} = 60$
$R_{head} = 0.04$ each
$R_{contact} = 1.46$ each
$R_{total} = 63.00$

Real World Resistor Analysis

We have just seen how to calculate a resistor value from the dimensions you have drawn using your CAD tool. However, in the real world the manufacturing process is not as perfect as your CAD drawing is. The resistor that is fabricated can very often be significantly smaller or larger than the one that you drew. We can compensate for these variations, or **deltas,** in our equations. Let's take a closer look at the various parts of our resistor.

Delta Effects on the Contacts

When the contact holes are etched, there is some real world uncertainty about the actual dimensions. If you over-etch the edges even slightly, that hole will get much bigger. Consequently, the size and width you use for your contact is in reality variable. You need to make sure when you design your resistor you know enough about the way the resistor is built in order to account for this physical discrepancy.

The manufacturers will give you the process variation numbers. They'll measure these for you. You might ask, for example, "If I draw it at 1 micron wide, how big will it end up on the mask?" They might reply, for instance, "If you draw it like this it might oversize by 0.1 micron."

This difference between drawn and real world dimensions is what we call the delta on this width (also called tolerance, error, variation, change in dimension, slop or change). The delta could be positive or negative, over-processed or under-processed. The symbol used to indicate the delta is the Greek letter δ (pronounced delta, which is a good reason why we call it delta). This change in width is denoted δW. Likewise, a change in length is denoted as δL.

For example, W might be 4 microns having a δW of, say, .06 micron. This tells you the actual width could vary from as little as 3.94 microns to as much as 4.06 microns, depending on whether the delta represents over-processing or under-processing.

Delta Effects on the Body

Just like the contacts, the poly could be over- or under-etched, as well; though, typically poly gets smaller during processing. So, you will consider δL and δW for the body resistance calculations as well.

Each delta will be a unique value. One substance and process might have a wiggle error of so much, while another substance or process might end up with a different error altogether. They run a lot of test wafers to determine the delta for each item.

To get all of these numbers you have to make a very big matrix of these tests. Once you extract one value for one item, you substitute it back and hunt for another value. You eventually can figure out each delta.

Delta Effects on the Head

If the body is affected by width variations, it's a good bet that the head of the resistor will be as well. If the body gets longer the head will get shorter. Likewise, if the contact over-etches, the length of the head will get shorter still. All these deltas conspire against us to make our calculations tougher. Adding in these deltas to our equation, we get

$$R = \frac{L_b + \delta L_b}{W_b + \delta W_b} \cdot \rho_b + 2\,\frac{L_h + \delta L_h}{W_h + \delta W_h} \cdot \rho_h + 2\,\frac{R_c}{W_c + \delta W_c} \quad ohms$$

R = total resistance (ohms)
L_b = length of body (microns)
W_b = width of body (microns)
ρ_b = sheet resistivity of the body (ohms-per-square)
L_h = length of head (microns)
W_h = width of head (microns)
ρ_h = sheet resistivity of head (ohms-per-square)
R_c = resistance factor of the contact (ohm-microns)
W_c = width of contact (microns)
δ indicates "change in"

Notice that some of these deltas may be common to more than one element. For example, a resistor length may over-etch by 0.1 micron, making its body longer, but reduces the length of the head by the same amount. The value 0.1, then, is used common to both terms.

Try It

1. A resistor body has the following dimensions as drawn: length = 95 micron, width = 12 microns. The material has a resistivity of 65 ohms-per-square. When fabricated, the above resistor width measures 0.2 microns smaller than drawn. What is its resistance value?

2. A resistor body over-etches, when fabricated, by 0.15 micron. The window that defines its length has an under-etch of 0.2 micron. The material has a resistivity of 150 ohms-per-square. Calculate the resistor values for each of the following sets of dimensions:

Width	Length	$W + \delta W$	$L + \delta L$	R
12	12			
20	60			
16	5			
50	10			
4	75			

3. Tabulate the above resistors again, from their drawn dimensions only. Calculate the percentage difference between the drawn value and the fabricated value:

Width	Length	Drawn Value	Fabricated Value	Percentage Difference
12	12			
20	60			
16	5			
50	10			
4	75			

ANSWERS

1. Determine number of squares. Multiply number of squares by sheet resistivity.

   Calculating as drawn,

   $$\frac{95}{12} \cdot 65 = 514.58 \qquad ohms$$

   Accounting for the 0.2 etching error,

   $$\frac{95}{12 - 0.2} \cdot 65 = 523.30 \qquad ohms$$

2. The resistor body width is a light-field mask, so over-etching results in a smaller dimension. The length is a dark-field mask, so under-etching likewise results in a smaller dimension. Subtract the deltas in both cases.

Width	Length	$W + \delta W$	$L + \delta L$	R
12	12	12 − 0.15 = 11.85	12 − 0.2 = 11.8	149.36
20	60	20 − 0.15 = 19.85	60 − 0.2 = 59.8	451.89
16	5	16 − 0.15 = 15.85	5 − 0.2 = 4.8	45.42
50	10	50 − 0.15 = 41.85	10 − 0.2 = 9.8	29.49
4	75	4 − 0.15 = 3.85	75 − 0.2 = 74.8	2914.28

3.

Width	Length	Drawn Value	Fabricated Value	Percentage Difference
12	12	150	149.36	−0.42%
20	60	450	451.89	+0.42%
16	5	46.87	45.42	−3.099%
50	10	30	29.49	−1.7%
4	75	2812.5	2914.28	+3.62%

Spreading Resistance

Hang on a minute. Look at the overhead view again. Although the width of the poly is measured to the edge of the material, it's fuzzy where the actual electrical flow starts, where it goes, and how it crosses the wider poly area after coming out of the contacts.

Does electricity flow only between metal chunks or does it use the entire available poly width? Where do we measure width? And what about length? Do you measure from the outer corners of the poly? Do you measure only between the metal contacts, or should you really include the metal? Does the width of the poly count more, or just the width of the contacts?

The actual path of the electrons spreads out as the electrons leave the contact, until they eventually use the entire width of the poly. This gradual **spreading** affects resistance, so we need to account for it in our equation.

Spreading resistance is dependent on many things. If you have a wide contact and a wide resistor, the effect is negligible. However, a wide resistor and a small contact means that the current has further to flow before it gets to the full resistor width. This gives more uncertainty as to what we should use as the width of the resistor for our equation.

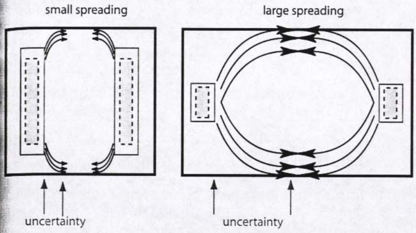


Figure 4–26. Spreading issues can cause varying levels of uncertainty, which must be accounted for in our resistance equation.

If a resistor is very short and badly designed, its value may be very inaccurate due to the inability to appropriately calculate electrical width.

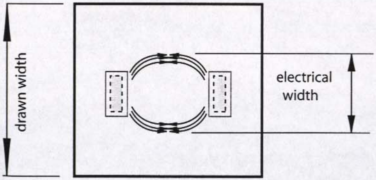


Figure 4–27. The contacts are so small and so close together that there isn't time for the electrons to spread out to the full width of the poly. Our electrical width is smaller than the width drawn for the poly.

Let's take a minute to look at an alternate way of making a resistor that avoids spreading issues.

A resistor can be made with the contacts outside the width of the poly material. There is no spreading to worry about because the poly defines the contact width, effectively making the widths equal.

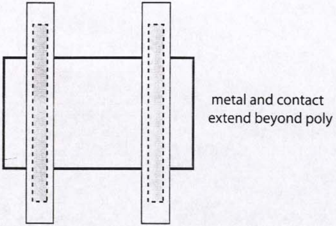


Figure 4–28. Some manufacturers allow the metal and contact to extend beyond the poly. This eliminates resistance due to spreading.

Depending on the technology, you may or may not be able to do that. Some people might not like the idea of having a contact hole out in fresh air. Somebody may make this hole through the oxide differently than somebody else, so they say, "Sorry, mate. This hole has to be all inside the poly because I don't want the etch to get out and get to other stuff." However, other people might be fine with it.

So whether you are able to run your contacts out in fresh air like this, or whether you must run lots of test masks on fully enclosed contacts, depends on your manufacturing process rules.

Some processes only allow one size of contact hole. This is because the process has been fine-tuned to such an extent that nothing other than a square contact will etch correctly. This one-size-fits-all approach is designed to give the process more controllability and get us smaller, more accurate dimensions.

If our process only allows us to use square contacts then we must fill our resistor width with as many contacts as we can to keep the contact resistance low.

Our resistor equation still holds true, but the contact resistance and spreading resistance terms become much more complicated. The exact details of the calculations vary from manufacturer to manufacturer, and are probably proprietary.[1]

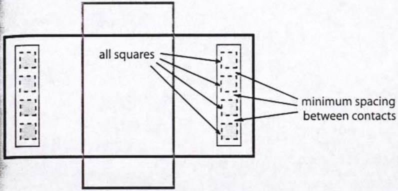


Figure 4–29. Some processes require using squares for contact holes.

Another option to reduce the spreading problem is to make the width of the contact exactly the same as the body.

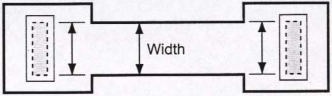


Figure 4–30. Making the width of the contact the same as the width of the center poly also eliminates spreading concerns.

Some people ignore the spreading problem altogether. They just use the ohm-micron approach to lump both the spreading and the contact terms into a single value for a quicker, less accurate calculation. Spreading resistance really should be a separate calculation from the contact resistance since the contact resistance is independent from how the current spreads.

There are various techniques and equations used by IC manufacturers to determine the spreading resistance term. Most of these techniques and equations are proprietary. All we are trying to achieve in this text is to give the

---

[1] Secret spy stuff. I know, but I can't tell you. If I told you I'd have to hire you.

reader an understanding of the physical and electrical processes involved. This should be enough for you to understand and use the equations from your process manual.

The Big, Oh-My-Gosh, Total Resistance Equation

Let's stop to get some perspective. Come up for air, so to speak.

In building your total resistance equation, you start with your basic first order equation using the most prominent variables. Then you build second and third order equations considering spreading results, deltas and other fine adjustments.

That's what the resistance equation is. You see the basic resistance from the body, then you see all these other tiny influences added into the equation. Here's what we've built:

$$R = r_b + 2r_h + 2r_c + 2r_s \qquad \textit{ohms}$$

R = *total resistance (ohms)*
r_b = *resistance from body section (ohms)*
r_h = *resistance from head section (ohms)*
r_c = *resistance from contact (ohms)*
r_s = *resistance from spreading (ohms)*

New term r_s = resistance due to the spreading of electrons (one spreading factor on each end). This number is defined differently by different manufacturers.

Remember that each r term in the equation must be separately calculated. If we write out this equation substituting original units, we can show each individual calculation involved:

THE BIG, OH-MY-GOSH, TOTAL RESISTANCE EQUATION

$$R_{total} = \frac{L_b + \delta L_b}{W_b + \delta W_b} \bullet \rho_b + 2\,\frac{L_h + \delta L_h}{W_h + \delta W_h} \bullet \rho_h + 2\,\frac{R_c}{W_c + \delta W_c} + 2r_s \ \textit{(ohms)}$$

R_{total} = *total resistance (ohms)*
L_b = *length of body (microns)*
W_b = *width of body (microns)*
ρ_b = *sheet resistivity of the body (ohms-per-square)*
L_h = *length of head (microns)*
W_h = *width of head (microns)*
ρ_h = *sheet resistivity of head (ohms-per-square)*
R_c = *resistance factor of the contact (ohm-microns)*
W_c = *width of contact (microns)*
r_s = *spreading factor (see process manual) (ohms)*

NOTE: All the delta values are unique to the process being used. Check the manufacturer's processing book for the δ values you need in your equations.

In the body and the head terms, you divide by width because you are finding the number of squares in the resistive path. Length divided by width gives you number of squares. In these terms, you use ohms-per-square values (sheet rho).

However, in the contact term you divide by width because contact resistance values were designed to be width dependent. So, in this term you use ohm-micron values.

As mentioned earlier, some people combine the contact and the spreading term into a single term.

There you have it. Your total resistance equation. Oh my gosh, it's big. This equation will vary according to the manufacturer, sometimes even according to material. However, you now should be able to interpret any variation of this equation given to you to calculate resistance.

You can see that we've gone from just a simple square to something that is pretty complicated. It's all based on physical dimensions.

The average layout engineer will not routinely get into this depth. But if he has to put a toolkit together, he will have to understand the equations and how to implement them in the CAD tool.

Now we understand our resistor equations. We can use them to help us make layouts that are accurate, tolerance proof and robust. In the next few sections, we will examine a few more considerations when designing resistors.

Practical Minimum Resistor Size

The body resistance should dominate your total resistance. If your resistance is dominated by the contact or head areas, you will not get a very good, controllable resistor.

Remember that manufacturers control the stuff in the middle (body) quite well, but not the outer stuff like the head areas or the contacts.

So, if a circuit designer does not know much about resistor fabrication, he will just say, "I want a resistor that's 2 microns wide and 100 ohms." The layout designer calculates to find that the length of the body is maybe 0.1 micron, which is unrealizable. Even if it were realizable, there would be so much variation in the actual dimensions that the resistance value could go all over the place. It could be out by a factor of two, or more.

Hold to a minimum body length of 10 microns because some of these deltas could be pretty large, say 0.1 of a micron. That keeps your error down to a hundredth of your values. If you want a fairly accurate resistor, make sure it's 10 microns or longer, to minimize the effect of the delta error.

Rule of Thumb: Make sure length of body is at least 10 microns, width is 5 microns.

Likewise, make your minimum width 5 microns. This helps ensure better accuracy and reasonable matching. If you need very highly accurate resistors, just make them wider and longer. The deltas will not change, but you can make their effect proportionally smaller.

Extreme Options for Resistors

High Resistance—Low Precision

Now, there are some cases where you just need a big value. The designer may not care if it is 5K or 10K, as long as it's big. You just need some large resistance in there to protect your circuit.

If you are lucky, you have a high ohms-per-square material you can work with already in your layout. Typically, however, on a CMOS process you are stuck with only a few ohm-per-square values. So, if you have a designer who wants a 10K resistance in the circuit, go ask him, "What's it doing? Why do you want it that high?" and listen carefully to his purpose for the resistor.

If the designer indicates he does not care about the value precisely—he will let it vary up or down by 15%, for example—then you can use existing materials with a minimum process width through the middle, like in Figure 4-31.

The people who make the design rules will set your actual minimum width for you. They will say, for example, "With the layer we've got, the minimum width is 1 micron." So, make the width 1 micron in that case.

Usually the minimum width of a body material can be much smaller than the minimum width of material required for a contact structure. Consequently, a minimum width resistor has a contact that is much larger than the body width. This difference in size gives us a shape that has a very characteristic appearance. The shape looks like a dogbone. That's what they call it, a **dogbone.** For "high value, don't cares" (meaning you need a very high resistance, but you do not care about the accuracy) making your resistor this way is reasonable.

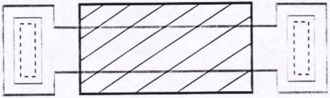


Figure 4–31. Dogbone resistor for higher values.

Another option if you do not care about fine control over value fluctuation, is to snake the path. This is useful especially if you are stuck for space.

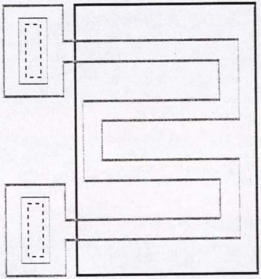


Figure 4–32. Serpentine resistor.

This is called a **serpentine resistor,** named for its serpent-like appearance.

You know the contact, head and spread values that you need. Now the question is, "How do you calculate the number of squares so you can determine the body resistance?"

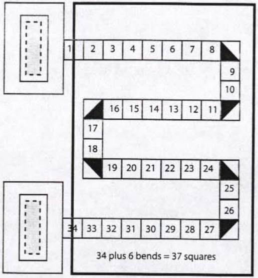


Figure 4–33. *How to count serpentine squares.*

Although it's not perfect, you count squares.

Squares are your friend.

The reason that counting the squares is not perfect is that the outer corner of each bend isn't entirely used.

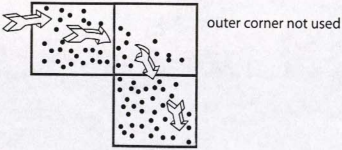


Figure 4–34. *Electrons do not flow through the entire corner square.*

Rule of Thumb: Count the linear squares, but for each bend you only count it as a half square.

The bends are mathematically more than exactly half of a square, but using half is pretty fair. If you look in some of the reference books listed in the bibliography you can investigate the mathematical proofs of how this really works.

You do not see serpentines much anymore. They now have so many variants in resistor types you can choose an option that gives you, say, 2K resistance in a reasonably sized chunk.

Low Resistance—High Precision

If you want a really low value, high precision resistor just use a big chunk of metal. Metal will give you the low resistance you want. The large size will minimize the effect of your deltas, helping your precision.

Checking Current Densities Makes You a Superhero

Your circuit designer may say to you that he has something like this, "I want 10 milli-amps to run through a 200-ohm resistor that has a 200-ohm-per-square sheet rho." How wide do you make that resistor?

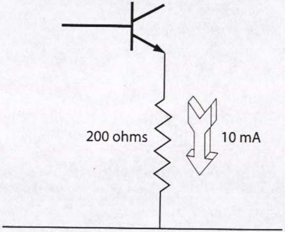


Figure 4–35. *Sometimes you are not given all the information you need. How wide is this resistor?*

Using some simple considerations, we can remove the guesswork. Consider the ohms-per-square of the material required. That should give you an idea of how many squares you require. In this example, we have a 1-square resistor.

Remember, for better accuracy, do not make it any shorter than 10 microns long. In our example, that equates to 10 microns long by 10 microns wide (making it square). 10 by 10, that gives us a square with 200 ohms.

Before you lay out your 10 by 10 resistor, feeling proud that you calculated your own dimensions, stop to consider your answer. For example, you don't use bell wire to wire your house with, do you? You don't use little tiny wire to hook up your washer and drier. If you have lots of current, you want a big, fat, chunky wire. Now, is 10 microns wide enough to handle 10 milli-amps?

You go back to your process book, look in a section called **current densities** for the particular material involved. The manufacturer's number depends on how thick the sheet is. If you have a really thick sheet of material you can get more current through it. (This is not the same as the sheet rho number, which you also will find in the process manual.)

What is a current density? A current density is the amount of current a material can handle reliably. The keyword here is "reliably." For example, you can indeed wire your household clothes drier with thin bell wire, but pretty quickly the wire will overheat. This will melt the insulation, set fire to the house and send your wife to her mother's while you rebuild. Or you to camp on the lake while she rebuilds, which might be the better option considering you just tried to use bell wire for the drier.

A much thicker wire will be able to handle the current of your drier without overheating. It should last a lot longer, perhaps for many years. So any size wire *will* work, but for how long? Thinner wires have shorter lives than thicker wires when connected to your clothes drier. That's why we look at reliability.

Likewise, a narrow resistor will be able to handle various amounts of current—but again, the question is for how long? Consequently, we have to choose a width that is appropriate for the current that is running through it.

Current densities for resistors in an IC are usually very conservative. They are set for a reliability factor in the tens of thousands of hours. You can run a resistor with more current than its rating but you reduce the reliability of the chip. You shorten the life of your product. Then word gets around that so-and-so product only works for awhile then they all die for some reason. The company image is tarnished. Their products are known to be unreliable. Pity. It all could have been avoided with a better layout engineer who knows to check current densities.

*Sometimes you are also given a **fusing current** rating in your process manual as well. This is the current required to blow the resistor up within a certain time. This might be fun, but it's messy.*

Anything in your process that can be used to carry current should have a current density associated with it. If you have a material that you want to use to carry current but you do not have a current density DON'T USE IT. You have no idea how wide to make it so that it is reliable.

If your designer insists on using this undocumented layer, scream blue[2] murder and turn him in to the processing police, or ask the processing police for guidance.

A typical current density is about half a milli-amp per micron of width. It is dependent on width because the wider you make it, the more current it can handle (more stuff in parallel).

■ **Rule of Thumb: .5 mA per micron of width.**

Once we find our milli-amps per micron value, we multiply that value by width. The result gives us the milli-amps the resistor can handle reliably.

$$I_{max} = D \cdot W \qquad \textit{milli-amps}$$

I_{max} = maximum allowable current for reliability (milli-amps)
D = current density of material (milli-amps per micron)
W = width of material (microns)

In the above example, our 10-micron resistor was apparently only designed to handle 5 mA reliably (10 microns multiplied by 0.5 mA). We have a problem.

You go back to your circuit designer and say, "I'm exceeding the current density by a factor of two. I need you to increase the width to at least 20 microns. That will give me 10 milli-amps."

He says, "Doh! I should have known that." He makes it a width of 20 microns.

Current densities are important for choosing the width of a resistor.

If the circuit designer knows his current density numbers in advance for a circuit carrying a lot of current, he will choose his resistor size accordingly. As a layout designer, I always do a double check, however, when I see a big current. I ask myself if this wire or resistor is large enough to handle that kind of current.

Sometimes the circuit designer just feels his width choices are good and he leaves the checking to the layout designer. You often catch a mistake. Some circuit designers do not even think about it. They just throw down something and say, "Oh, 5 microns will do."

---

[2] I learned the expression as "bloody murder" as I grew up. But apparently the Brits feel very strongly about letting their children use such harsh language as bloody, so the UK kids all called it blue murder. Or, so Chris tells me.—*Judy*

Sometimes, if you are lucky, your circuit designer will annotate the schematic to tell you that a wire needs to handle a certain amount of current. However, I always do another double check and say, "Ok, where is this current coming from and where is it going to?" This is why layout people need to know about voltage and current. Pretty simple calculation—you just multiply two numbers together—but you can save a whole project by asking questions.

Remember to check the current densities for everything. Check it all.

"But sometimes chips have thousands of resistors. This would take you forever."—Judy

You can't check every item on a whole chip, but you may come to a block that has, say, a hundred transistors in it. The first thing you ask the circuit designer is, "How much current is flowing in this cell?"

If he says, "10 milli-amp" you ask where it's flowing.

"Through that transistor."

"Is that the only place it's flowing through?"

"No, here and here."

"What about the rest?"

"Oh, that's only a couple hundred micro-amps."

"Ok, we won't care about that." There's only one high current path. We need to do special things with this high current path.

That's where you do your extra calculations.

Try It

1. A resistor needs to carry 100 micro-amps. It has a width of 3 microns. If its current density value is 0.2 milli-amp/micron, will the resistor be reliable?

2. A process has a resistor whose current density value is 0.25 mA/micron. What are the maximum currents for the following widths:

Current Density	Width	Current
0.25	2	
0.25	3	
0.25	5	
0.25	10	
0.25	20	
0.25	60	

3. We need a resistor that is capable of handling 12 mA. The value has to be 50 ohms and needs good immunity to process variation. We have three types from which to choose:

   POLY—current density of 0.27 mA/micron, sheet rho is 225
   NWELL—current density of 0.72 mA/micron, sheet rho is 870
   RDIFF—current density of 0.93 mA/micron, sheet rho is 1290

Which should you use for this application?

ANSWERS

1. Multiply 3 by 0.2 = 0.6 mA. Our maximum allowable current for reliable operation is 600 micro-amps. Yes, that is fine, since we only plan to carry 100 micro-amps.

2.

Current Density	Width	Current
0.25	2	2 × 0.25 = 0.5 mA
0.25	3	3 × 0.25 = 0.75 mA
0.25	5	5 × 0.25 = 1.25 mA
0.25	10	10 × 0.25 = 2.5 mA
0.25	20	20 × 0.25 = 5 mA
0.25	60	60 × 0.25 = 15 mA

3. You can see which material offers a tasty sheet resistivity. However, before you can determine the best resistor type you should look at their respective widths and lengths.

To determine width we use

$$I = D \cdot W \qquad \textit{amps}$$

To determine length we use

$$R = \frac{L}{W} \cdot \rho \qquad \textit{ohms}$$

Resistor Type	Width	Length
POLY	$\frac{12}{0.27} = 44.44$	$50 = \frac{L}{44.44} \cdot 225 \;\; L = 9.87$
NWELL	$\frac{12}{0.72} = 16.66$	$50 = \frac{L}{16.66} \cdot 870 \;\; L = 0.95$
RDIFF	$\frac{12}{0.93} = 12.90$	$50 = \frac{L}{12.90} \cdot 1290 \;\; L = 0.5$

Use POLY with a width of 13 microns or more. Although the other resistors have better current densities, the Nwell and Rdiff need to be so wide in order to get a good length, they become impractical.

Making Your Layout Tolerance-Proof

The body, the head areas and the contacts are the three basic parts of our resistor. We have talked about how real world changes can affect resistor values in each of these areas. Your layout can save the day by being pre-designed to tolerate these inaccuracies.

Effects on Resistor Tolerance

What items affect resistor tolerance?

- Variations in length and width (deltas)
- Variations in sheet resistivity (rho)
- Variations in alignment of various layers (i.e., contact misaligned or over-etched)
- Variations in sheet thickness

What to Do about Tolerances

Processing and etching will cause dimension uncertainties. You may draw something at 5 microns but it hits the silicon at 4.5 microns. Things will vary and there is even some kind of tolerance on that. It could be 4.5 or 5.2. There's real world slop in the actual making of the device.

You have to ask your circuit designer, "What kind of resistor variance can we tolerate? Do you care if two resistors next door to each other vary by 20%?" Or, perhaps you want an amplifier with a gain of 10, and that gain is being set by a resistor value. You would want this resistor to be precise so that your gain is accurate.

In cases like these, when you do care about accuracy, you have to make your resistors very wide or very long. Oversizing helps ensure that the process slop doesn't affect your values in any major way. If you do it right, the only variation will then be the process, not anything due to layout.

Try It

1. You have a 0.1-micron delta on a length of 40 microns. What percentage of the total length is the delta?
2. You have a 0.1-micron delta on a length of only 2 microns. What percentage of the length is the delta?
3. By comparing the answers to problems 1 and 2, should you make resistors smaller or larger in order to decrease the affect of processing errors? Observe a conclusion.

ANSWERS

1. 0.1 divided by 40.1 = 0.25%
2. 0.1 divided by 2.1 = 4.76%
3. The 40-micron resistor has the smaller percentage delta. Therefore, processing errors are less of a problem in longer resistors, relatively speaking. Longer lengths give better tolerance-proof resistors.

- **Rule of Thumb: Make resistors wide, long or both for high accuracy (rule of thumb 10 long × 5 wide).**

So, once you make the process variations negligible the only thing you still worry about is variation in sheet resistivity. You can check into lengths and widths and make sure variations won't bite you, but the basic sheet rho could be plus or minus 10%. If you know that, you design resistors so that even with all the photolithography and processing variations you still only have a combined variation of maybe 15–20%.

Make sure your circuit has been designed with these variations in mind, which means over-design it. If you figure a nominal design then design for the worst case so the resistor at lowest value gives you the performance you want.

Oh, as a last comment I'll mention that you might see feedback loops to compensate for the imperfections of the process. If the resistance might vary by 25%, the circuit designer has to compensate, change design accordingly, include a bunch of matching and servo devices to change the gain back as resistors change. He will design self-compensating loops into the circuitry.

Using Existing Materials—Versatility at No Cost

If we go back to our FET (field effect transistor) theory, the center Gate modulates the depletion region under it. If we were to take off the Gate of an NMOS transistor, what would we be left with?

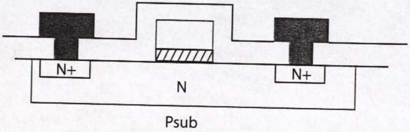


Figure 4–36. *FET (field effect transistor).*

Well, as you can see, we'd essentially be left with an N Type resistor. From a transistor, we get a resistor.

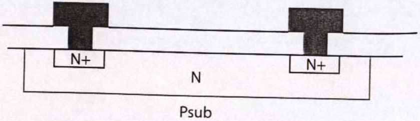


Figure 4–37. *N Type resistor.*

The easiest resistor to build is a poly resistor. In a standard CMOS process, though, you can use anything as a resistor. That's because everything has a resistance of some sort.

You could, if you wanted to, just by using the same type of process steps, have N+ go all the way across. In this case, you wouldn't even need the regular N. This is an N+ resistor. N+ has its own ohms-per-square.

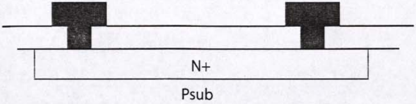


Figure 4–38. *N+ resistor.*

Just by taking the standard stuff that's already being laid down into the device, without building any more masks or layers, we can develop all sorts of resistors. We do not need extra processing steps to build these as long as we are making transistors on the same chip anyway. We can build them for free.

You can do the same trick with a PFET. You are building transistors on the chip anyway, so use the same processes, placed where you need them, and Bob's your uncle,[3] you have free resistors.

Because the substrate is P material, you have to build a PMOS transistor in something that's N. You need N as a separator between the P and the substrate. Isolation, you know.

Remove the Gate.

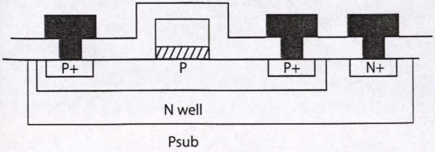


Figure 4–39. *PFET can also give you free resistor materials.*

---

[3] American: "there you are" or "quick as a snap."

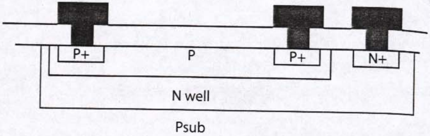


Figure 4–40. PFET without the Gate becomes yet another type of resistor.

On a P Type resistor, because they are in the N well, there is always a third terminal. That's one of the extra issues with a diffusion resistor like this. (It's called a diffusion resistor because it is made out of diffusions in the substrate.) You have to make sure the well is connected to the most positive voltage, called the **VDD.** This is the reason for the third terminal. That goes back to what we covered in basic devices. It stops parasitic activity from turning on.

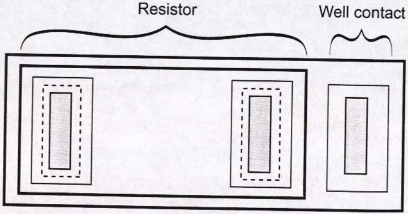


Figure 4–41. Overhead view shows third terminal.

There are all sorts of possibilities here for making resistors.

- Take off the Gate. We have a P Type.
- Take off the Gate and put P+ all across. We get a P+ resistor.
- Take off the Gate, the P, and the P+. You could just end up with N+ in an N well if you wanted.

Building resistors from readily available layers saves money and problems. Experiment. There are many possibilities. You could apply additional layers

but it's more efficient to look at your bare minimum process and say, "Ok, what can I use that is already here?"

Diffusion versus Poly

In the last section, we learned that if you have to build a P Type and an N Type transistor on your chip anyway, you might as well put the existing parts and layers to extra use. However, sometimes you have to build extra layers, and perhaps modify them slightly to get the right kind of resistivity that you need.

If we relate our diffusion resistor back to our poly resistor, they look very similar except for the extra terminal on the right. In contrast to the poly resistors we have used for our previous discussions, this resistor is actually made out of diffusions.

You can still do a similar trick in the middle—put a chunk in the middle to change the resistivity of the body.

We use exactly the same equation for diffusion resistors; all the delta lengths, delta widths, everything that relates to a poly can relate to the diffusion resistor.

- Diffusion resistors. These are diffused down into the substrate, fuzzy around the edges. The diffusions spread out in manufacturing so they are not as well controlled.

- Poly resistors. The Gate is made of polysilicon, a poly layer deposited on the surface. These have exact thickness for better accuracy and well-controlled lengths and widths.

The most noticeable difference with a diffusion resistor in your layout is typically the third terminal to bias the connection.

Why not always use poly? If you had a choice of resistor types, which ones would you choose?

RESISTOR TYPE COMPARISON

POLY	DIFFUSION
• Low power dissipation	• High power dissipation
• Low parasitics	• Higher parasitics
• Good process control	• Worse process control
• Typically low sheet rho	• High and low sheet rho
• 2-terminal device	• 3-terminal device

Double Poly

Now, sometimes the N well and all the diffusions that make up the transistors still do not give us what we need for our resistors. The really picky circuit designers say they want a resistor that does higgly squiggly thus and such.

So, ok, over in the middle of nowhere we can deposit a completely new poly layer. We'll make it really controllable, and that will be the layer we use for our certain higgly squiggly resistors. That's called a **double poly process**—one for the Gates, one for the resistors. You do the same calculations; it's just done using different types of material and values.

Bipolar/BiCMOS

In a **Bipolar process,** they add extra layers to a CMOS process. You can use those extra layers, as well, to make resistors. In this case, you could essentially get a double poly process free. The more steps there are involved in the process, the more choices you have for your resistor material at no extra cost. (See *Bipolar transistors.*)

Closure on Resistance

A good circuit designer studies the sheet resistivities and the whole process as he works. So, layout people will probably be given the right answers to begin with. However, these techniques that we have been learning are widely used by layout engineers everywhere.

As a good layout engineer, you will be the one to catch a mistake if anyone will. You will be the one to simplify a circuit, shrink the size or make it more reliable. Knowing these techniques makes you versatile and valuable.

Here's What We've Learned

Here's what you saw in this chapter:

- Why any size square gives you the same resistance value.
- Why ohm-per-square is used in the resistance equation for body and head sections.
- Why ohm-micron is used in the resistance equation for spreading and contacts.
- How to use existing layers of a chip to increase design flexibility.
- How to use length, width and sheet resistivity in the resistance equation.

- How to identify and calculate the body, head, spreading and contact sections of the resistance equation.
- How and why to compensate for tolerance by using deltas.
- When and how to use a Dogbone, Serpentine or other creative designs.
- Why 10 × 5 microns is used as a minimum resistor size.
- How and why to check current densities and fusion currents to prevent disasters.

And more . . .

Application to Try on Your Own

Process engineering has just issued a report outlining a series of measurements that they have made on a test wafer. Circuits have been coming out of the wafer fab that are not performing as expected.

The accuracy of the resistor model in the design kit is suspected, since the circuit functions but the power supply currents are not as expected. Likewise, the gain of the circuit is also out of spec.

The report, summarized below, outlines a series of measurements taken from the wafer. This circuit only uses two resistor layouts. Calculate the new expected values and feed this information back to the circuit design group so they can re-simulate with the new resistor values.

This problem happened to me, by the way. But not with two resistors, with 120. So I wrote myself a spreadsheet to give me what the new values were. I used that to work out what I needed to do to fix it.

New process information could take six months for fixes to come through in the design kit, but what if you need to do a tapeout next week? If you can't wait six months, then some poor soul has to work out all these calculations.

If you get new information from the lab and you go ahead with your tapeout based on the existing tools, your chip won't work properly. If you rely on rules from your design kit, you'll get caught out.

So you roll up your sleeves, recalculate, modify your layout and you've saved the company a quarter million dollars.

TechnoWizards, Inc.
12 Blastcap Avenue
Germanium Valley, California 95409

MEMO

To: Paul Eghan
From: Circuit Design Group
Re: Info on circuit
Date: 2/Day/Late

	Drawn Dimension	Measured Dimension
Poly	10	10.13
Contact	8	8.21
Resistor Etch Window	25	25.73

	Original	New
ρ body	185 Ω / □	190 Ω / □
ρ head	2 Ω / □	2.2 Ω / □
Contact Factor (*CF*)	100 Ωμ	120 Ωμ
Spreading Resistance Factor (R_{sp})	90	65

The assumption that the body width is not affected by the etch is also incorrect. It has been found that the poly undersizes by a total of 0.2 μ during processing in the region of the resistor etch window.

	L_{bd}	W_{hd}	W_{cd}	L_{hd}	Desired Value
Resistor 1	15	6	5	1	500
Resistor 2	32	15	14	2	400

Assume the above resistor dimensions are from your layout database. There is some uncertainty that the dimensions are correct. The original assumption was that processing had a good handle on their dimension control and that what was drawn was what ended up on silicon.

The original Zoozle Manufacturing Company design manual equation is

$$R = \frac{L_{bd}}{W_{hd}} \cdot \rho_{body} + 2\left(\frac{L_{hd}}{W_{hd}} \cdot \rho_{head} + \frac{CF}{W_{cd}} + \frac{R_{sp}}{W_{hd}}\right) \quad ohms$$

which they show comes from their simplified version:

$$R_{total} = R_b + 2(R_h + R_c + R_s) \quad ohms$$

(I don't know of any real Zoozle company. Any similarity in this application problem to a real company name is unintentional and would be extremely surprising.)

ANSWER

Following are observations drawn from the Report Summary:

- In this process, body width apparently equals head width. (The equation shows the body length divided by *head* width.)
- Poly body undersizes by 0.2μ (stated in middle paragraph)
- The poly head width is affected by the 0.13 error after all.
- Variables in the equations look slightly different than variables used in this textbook. Once you understand the mathematical relationships you can certainly represent a value in any method desired. For example, one source might use L_h in the same place where another source uses L_{hd} or L_{head} or *LH*. Be flexible in how you see the values labeled by different companies. Look for definitions.
- Did you notice all the information you need is located in one spot, right under your nose? This will never happen in the real world.

STEP 1: Determine ohms using original information, just to check that the resistor was laid out correctly in the first place.

Resistor 1

$$R = \frac{15}{6} \cdot 185 + 2\left(\frac{1}{6} \cdot 2 + \frac{100}{5} + \frac{90}{6}\right)$$

$$R = 533.16\ \Omega$$

Resistor 2

$$R = \frac{32}{15} \cdot 185 + 2\left(\frac{2}{15} \cdot 2 + \frac{100}{14} + \frac{90}{15}\right)$$

$$R = 421.48\ \Omega$$

Conclusion from Step 1: Even if processing had not changed their physical values, the resistors were laid out incorrectly anyway. We do not get 500- and 400-ohm values as originally desired.

STEP 2: Answer the concern posed in the Report Summary, "Check the above dimensions in your layout database. There is some uncertainty that the dimensions are correct."

Conclusion from Step 2: Due to so much uncertainty, go back and check everything from the beginning. Include your observations about the dimensions in your report. Let's assume that after checking, the dimensions appear correct to you. (See *Report*, end of Answer section.)

STEP 3: Determine ohms using new information.

Resistor 1

$$R = \frac{15 + 0.73}{6 - 0.2} \cdot 190 +$$

$$2\left(\frac{1 - 0.73 - 0.21}{6 + 0.13} \cdot 2.2 + \frac{120}{5 + 0.21} + \frac{65}{6 + 0.13}\right)$$

$$R = 582.59\ \Omega$$

Resistor 2

$$R = \frac{32 + 0.73}{15 - 0.2} \cdot 190 +$$

$$2\left(\frac{2 - 0.73 - 0.21}{15 + 0.13} \cdot 2.2 + \frac{120}{14 + 0.21} + \frac{65}{15 + 0.13}\right)$$

$$R = 445.94\ \Omega$$

STEP 4: Compare original and new calculations.

	Original	New
Resistor 1	533.16	582.59
Resistor 2	421.48	445.94

Resistor 1 is 49.43 ohms too high.

Resistor 2 is 24.46 ohms too high.

STEP 5: Become a Superhero. Fix it.

Assuming all layers can be modified, we can modify the body length of the resistors as follows:

Resistor 1

We want 500, but we are getting 582.59. Use this information to determine what the new length should be. Use the error amount, 82.59, to find the error in length.

$$82.59 = \frac{L}{5.8} \cdot 190$$

Length error = 2.52 microns

Since we are getting too much resistance, we will shorten our length by that amount.

Resistor 2

We want 400 but we are getting 445.94. Use this information to determine what the new length should be. First, use the error values to find the error in length.

$$45.94 = \frac{L}{14.8} \cdot 190$$

Length error = 3.58 microns

Since we are getting too much resistance, we will shorten our length by that amount.

You are now ready to write your report to the Circuit Design Group.

TechnoWizards, Inc.
12 Blastcap Avenue
Germanium Valley, California 95409

MEMO
REPORT

To: Circuit Design Group
From: Paul Eghan
Re: Info re circuit
Date: 2/Morrow/Morn

The original layout was in error. We desired 500 and 400 ohms, but we laid out resistors that resulted in values of 533 and 421 ohms. Moreover, with the new processing information we find the values are 582.59 and 445.94 ohms.

I have reduced the length of Resistor 1 by 2.52 microns. You will see a total length of 12.48 microns for Resistor 1 on the updated layout.

I have reduced the length of Resistor 2 by 3.58 microns. You will see a total length of 28.42 microns for Resistor 2 on the updated layout.

This should fix the problem.

Extra to Consider

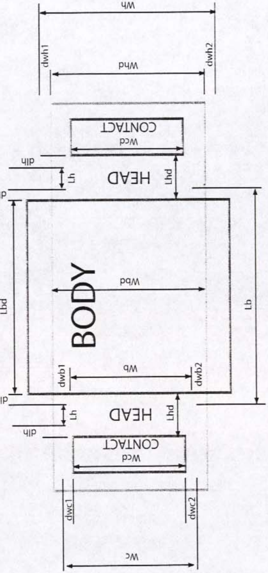
How would you proceed if you were only able to change the body of the resistor? This often is the case.

In the above application problem, you adjusted the length of the overall resistor. You were able to use a simple equation to find the new length. Changing the overall length did not affect any other lengths, so your equation had only a single variable, *L*.

However, if you can only change the body portion of the resistor, this causes two changes in lengths: body as well as head. You will have two variables. You will have to use the full resistance equation.

The quickest method to find these two variables is to hold the minor variable constant (head length), while you solve for the major variable (body length). Then through successive intelligent approximations you can fine-tune the values to determine both the head and body lengths that give you the desired resistance.[4]

Resistor Dimensions and Deltas



Wbd = Width body drawn
Lbd = Length body drawn
Wcd = Width contact drawn
Lhd = Length head drawn
dwc1 = dwc2 = delta on width of contact
dlb1 = dlb2 = delta on length of body

dlh = delta on length of head
Lb = Lbd + dlb1 + dlb2 = True body length
Wc = Wcd + dwc1 + dwc2 = True contact width
Lh = Lhd - dlh - dlb = True head length
Wb = Wbd + dwb1 + dwb2 = True body width
Wh = Whd + dwh1 + dwh2 = True head width

---

[4] I say solving some two-variable equations is faster. To each his own.—*Judy*

CHAPTER 5

Capacitance

Chapter Preview

Here's what you're going to see in this chapter:

- Classic capacitor review
- Hair raising experiment
- Effect of a capacitor on DC
- Effect of a capacitor on AC
- Filtering certain frequencies
- Building a capacitor on your chip
- Determining capacitance values
- Secret capacitance from the edges
- Parasitic capacitance
- Stacking metals

And more . . .

Opening Thoughts on Capacitance

A **capacitor** is a device that stores charge.

You can make a capacitor for yourself. Use two sheets of tin foil. Separate them with a layer of plastic food wrap. There you go. Instant capacitor.

One of the tricks my physics teacher used to do is get some really big capacitors, charge them up, and be very obvious about disconnecting *the battery. Then just before he left the room, he'd say "Now nobody touch these!"*

And, of course, some of the more curious kids would say, "Hmmm, I wonder what that is." They'd pick up these things, start playing with them and get a huge zap out of this capacitor. He'd charge them up to 500 V or so.

That was before the days of lawsuits. But it did demonstrate that a capacitor stores its charge, so it served its purpose.

He could tell who'd done it when he returned to the room by finding the person with their hair standing on end.

Do not try this yourself. It is extremely dangerous

Capacitor Review

A **capacitor** is a device that has the capacity to store a certain amount of **charge,** a certain number of electrons. The ability of a capacitor to store charge is called **capacitance.** The measurement unit for capacitance is **farads.** A capacitor is built from two conducting plates separated by an insulating material called the **dielectric.** It is in this dielectric where the charge is stored.

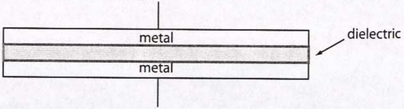


Figure 5–1. *Classic capacitor.*

The value of the capacitor is determined by the thickness of the insulator, the quality of the insulator as valued in the dielectric constant, and the area of the plates that overlap each other.

Capacitors can appear in an integrated circuit with surprising ease. Every time one conducting material crosses over another conducting material a capacitor is formed. This happens even when we do not want it to happen. We will see how to build capacitors that are useful and well controlled.

Capacitor Properties

Connect your capacitor to a battery. Notice the voltage across the capacitor will rise as the capacitor charges. If charged long enough, the voltage across the capacitor will eventually equal the voltage of the battery. The time taken for the capacitor to reach full charge depends on the value of whatever series resistance is in the circuit. In this case, the series resistance comes from the battery. Now open the circuit, or remove the battery. Notice the voltage across the capacitor remains.

The capacitor symbol looks (surprise, surprise) like the actual device. We use two parallel lines to represent the two parallel plates.

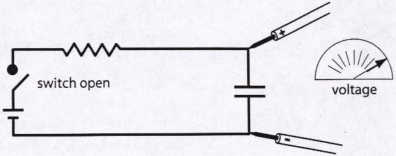


Figure 5–2. *Capacitors retain charge even with removal of the voltage source from the circuit. The capacitor symbol is seen in the circuit at right, between the meter probes.*

If you measure the voltage across the capacitor using an oscilloscope, you can monitor the increasing voltage. The trace on the oscilloscope rises until it approaches the voltage of the source. The trace then remains at this highest level.

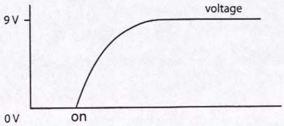


Figure 5–3. *When connected to a voltage source, the voltage across a capacitor increases to a maximum value, then remains constant.*

If the capacitor is removed from the battery, it will hold its charge. The capacitor cannot hold a charge indefinitely, though. The terminals on the oscilloscope have some impedance, for one thing. The charge on the capacitor would bleed gradually away through this impedance.

Capacitors with DC Voltages

If you deliberately add a series resistance into the circuit, you can measure the voltage across the resistor as the capacitor charges. The voltage across the resistor is inversely proportional to the current flowing into the capacitor. The oscilloscope shows a handsome current spike when the switch is first turned on. The current then drops quickly as the capacitor charges. Once the capacitor reaches its maximum charge, your oscilloscope will show no current flowing.

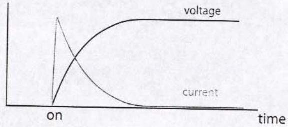


Figure 5–4. *No current flows at maximum charge.*

Due to the insulator between the two plates, a capacitor is very good at blocking static DC voltages.

A capacitor blocks DC voltage.

If a circuit has a certain DC voltage output that you do not want to wander somewhere else in the circuit, you can put a capacitor in the way. It will block that DC voltage. The DC is separated, or decoupled, from that portion of the circuit.

Capacitors with AC Voltages

If we replace our battery with an alternating voltage supply, usually a sinusoidal waveform, a very different story is told. An alternating voltage can sneak through a capacitor, depending on the value of the capacitor, and depending on the frequency. As frequency increases, you will see the amount of AC current passing through increases.

You can think of a capacitor as a frequency-sensitive resistor. If you have a big enough capacitor, once you get past a certain frequency, it's as if you have no capacitor in the circuit at all. It acts like a very low-value resistor.

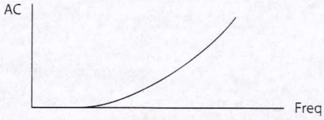


Figure 5–5. *AC current increases across a capacitor as we increase frequency.*

A capacitor is a frequency-sensitive resistor.

You can calculate the effective impedance of a given capacitor at a given frequency.[1] Let's say you want to calculate the gain of an amplifier with a capacitor in the feedback. Calculate the capacitor's impedance at a given frequency, then think of it as a resistor with that Ohm value. Now you can use Ohm's Law. Not too hard. (We are discussing AC current, of course. DC current, you remember, is blocked when the capacitor is charged.)

We have seen that capacitors are used for two types of blocking: blocking DC entirely, and allowing only certain signals through AC. Sometimes you will see capacitors called **blocking capacitors,** or **coupling capacitors.**

Noise Filter

Capacitors also help reduce noise. You might have a DC supply bothered by lots of noise. Noise is high frequency AC. If you connect a capacitor across the voltage supply, the high frequency noise will be shorted to ground, but the DC is blocked from shorting to ground. From the point of the capacitor onward, you are left with this lovely, quiet voltage.

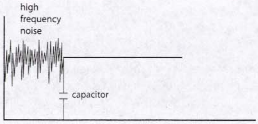


Figure 5–6. *DC cannot short to ground through a capacitor, as high frequency noise can.*

---

[1] They've got equations for everything! There are too many to memorize. The best you can do is know how to look up the right one when you need it.

All the high frequency noise will be shunted down through this low value of resistance because the capacitor is a frequency-sensitive resistor. This type of capacitor is known as a **decoupling capacitor.** It decouples the noise from the supply voltage.

Calculating Capacitance

The amount of area that the top plate overlaps the bottom plate determines the value of a capacitor, for the most part. However, we will make adjustments based on secondary factors.

Area Capacitance

To calculate the value of a capacitor we first find the dimensions of the plate overlap. We then multiply the length of the capacitor plate overlap by the width of the capacitor plate overlap to give us the area of the capacitor.

Each dielectric acts differently. Capacitors of the same area will have different values depending on which dielectric is used.

In an integrated circuit, the thickness of our dielectric is fixed by the process we are using. Therefore we need to multiply the area of the capacitor by a constant we will call C_1. This constant takes into account the process thickness of dielectric as well as the insulator's dielectric constant. The linear measure units for length and width are microns. The unit for the constant is usually femtofarads per square micron.

$$C_{area} = l \cdot w \cdot C_1 \qquad \textit{femtofarads}$$

$$C_1 = \frac{fF}{\mu m^2} \qquad \textit{units for } C_1$$

As we discussed with resistors in Chapter 4, the actual dimensions of finished devices might be slightly larger or slightly smaller than drawn. We call these small measurement variations the *deltas* on the length and width, denoted as the Greek letter delta, δ. Instead of specifying l for length, as read from the schematic, we must also consider the delta on the length after the actual processing. So, we add δl to the length. Likewise, we must consider the delta on width, so we add δw to the width.

$$C_{area} = (l + \delta l) \cdot (w + \delta w) \cdot C_1 \qquad \textit{femtofarads}$$

Our formula becomes a bit more complex, but says the same thing: multiply length by width, then multiply by the appropriate capacitance constant.

Always perform your dimensional analysis. Write out your equations using only the dimensions. Include no numerical values. Then cancel any similar unit factors found in both the numerator and denominator of the same term. Be sure the result is expressed in appropriate units. If the units do not resolve correctly, then we know the equation is not accurate.

$$\textit{femtofarads} = (\textit{microns}) \cdot (\textit{microns}) \cdot \frac{\textit{femtofarads}}{\textit{microns}^2}$$

In this equation, we see that all the micron factors cancel. We are left with femtofarads, the correct unit of capacitance. This equation resolves to the correct units, so it is a correct relationship.

How do people find this magic C number? Are we cheating, multiplying by some imaginary term, just to make sure the units resolve correctly? Well, yes. However, like other constant multipliers in other formulas, the C multiplier is consistent. That should make you feel a little better about it. After all, we do not select a different C number for every instance. Where would we get them all?

If we can determine the value of the constant from a small number of samples, then we can use that constant in the equation to solve for other variables, to predict capacitance in other situations.

To first determine this number, researchers say, "Ok, let's make a huge capacitor a thousand microns by a thousand microns." They obviously know the length and width. They measure the capacitance. From that, they can easily calculate femtofarads per square micron, which is the constant multiplier, C.

Try It

1. A capacitor has drawn dimensions of 50 μm by 35 μm. The process oversizes the defining feature of a capacitor by 0.1 μm. The capacitor has an area constant term of 6.2 fF/μm². What is the value of the capacitor?
2. We need a square capacitor of 5 pF. Assuming perfect processing and given a capacitance area value of 7.1 fF/μm², what are the dimensions of the capacitor?

ANSWERS

1. $C = (50.1)\,(35.1)\,(6.2) \qquad fF$

   $C = 10{,}902.762 \qquad fF$

2. $L = W = \sqrt{\dfrac{5000}{7.1}} \qquad \mu m$

   $L = 26.53\ \mu\text{m}$

   $W = 26.53\ \mu\text{m}$

Periphery Capacitance

These area capacitance constants however, do not scale properly when you have a very small capacitance. By testing many sizes of capacitors, researchers discovered that capacitance was larger than expected on small capacitors.

They found a capacitance hiding along the edges, along the fringe of the plates, so to speak. They call this additional capacitance a **fringing capacitance.** As you move away from the edge of the capacitor, the fringe capacitance becomes negligible, but there is a significant fringing capacitance near the edges, enough to affect calculations. This is the effect of the proximity of the capacitor edges to the bottom plate, emanating from the vertical sides of the top plate.

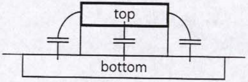


Figure 5–7. Fringing capacitance sneaking out the edges.

The capacitance for the periphery is equal to the total periphery of the capacitor multiplied by a second periphery capacitance constant. In other words, two times the length, plus two times the width, times the periphery constant, C_2. C_2 is measured in femtofarads per micron. Notice, the units are linear in this equation, since the perimeter is linear.

Looking at the dimensional analysis, we see that microns cancel. The peripheral capacitance is resolved in femtofarads, as capacitance should be. The units check.

As usual, we add deltas to the length and width to increase the accuracy of our calculation.

$$C_{periphery} = [2(l + \delta l) + 2(w + \delta w)] \cdot C_2 \qquad \textit{femtofarads}$$

$$C_2 = \frac{fF}{\mu m} \qquad \textit{units}$$

$$C_{periphery} = \frac{(\mu m + \mu m) \cdot fF}{\mu m} = fF \qquad \textit{femtofarads}$$

Our total capacitance for the device, C_{total}, equals the capacitance we found from the general area C_{area}, plus the fringe capacitance we found along the periphery, $C_{periphery}$.

$$C_{total} = C_{area} + C_{periphery} \qquad \textit{femtofarads}$$

Try It

1. A capacitor has drawn dimensions of 50 μm by 35 μm. The process oversizes the defining feature of a capacitor by 0.1 μm. The capacitor has an area constant term of 6.2 $fF/\mu m^2$ and a periphery term of 2.9 fF/μm. What is the value of the capacitor?

2. We need a square capacitor of 5 pF. Assuming perfect processing, a given periphery term of 2.9 fF/μm and a capacitance area value of 7.1 $fF/\mu m^2$, what are the dimensions of the capacitor within 99% of their absolute values?

ANSWERS

1. $C_{area} = 50.1 \cdot 35.1 \cdot 6.2$ fF
$C_{area} = 10{,}902.762$
$C_{periph} = (2 \cdot 50.1 + 2 \cdot 35.1) \cdot 2.9$
$C_{periph} = 494.16$
$C_{total} = 494.16 + 10{,}902.762$
$C_{total} = 11{,}396.922$ fF

2. $L = W = \sqrt{\dfrac{5000}{7.1}}$

$L = 26.53\ \mu m$
$W = 26.53\ \mu m$

But there is periphery capacitance to worry about.

$C_{periph} = (2 \cdot 26.53 + 2 \cdot 26.53) \cdot 2.9$ fF
$C_{periph} = 307.75$

Our capacitor will be too high.

$C_{total} = 5307.75$

Let's remove the periphery capacitance.

$C_{wanted} = 5000 - 307.75$

$L = W = \sqrt{\dfrac{4692.25}{7.1}}$

$L = 25.7\ \mu m$
$W = 25.7\ \mu m$

Let's re-calculate the periphery capacitance to make sure.

$C_{periph} = (2 \cdot 25.7 + 2 \cdot 25.7) \cdot 2.9$ fF

$C_{periph} = 298.12$
$C_{total} = 4692.25 + 298.12$
$C_{total} = 4990.37$ fF

After 10 iterations around this loop the final dimensions are calculated as 25.73 microns. Re-calculate using these dimensions to double check.

You will notice that the periphery term causes many headaches. We need to keep iterating around this loop until we feel that we have as accurate a dimension as we can obtain. In this case, we are better than 99% accurate, so a 25.7-μm square capacitor is good enough.

25.7 μm

N Well Capacitors

So, how do we make a capacitor in an IC? Remember that all the IC processing layers stack on top of each other. Since a capacitor is made of parallel plates, IC layering lends itself very well to making capacitors. How easy is it to make parallel layers if we find ourselves building chips in layers anyway? Fits nicely into our existing processes, doesn't it? We can build some of our capacitors for *free* (no extra steps required during processing).

Between the Gate and the substrate in our FET, we learned about inherent, unwanted capacitance. We called this a nasty parasitic. But, what if we happen to need capacitance? This parasitic is not nasty anymore. Now it's our friend.

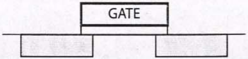


Figure 5–8. *FET Gate is a parallel plate to the source-drain regions. Parasitic capacitance occurs naturally.*

Let's make a really, really, big, huge Gate. I mean much larger than a wimpy little FET Gate. However, let's change our processing ever so slightly. Instead of building a normally Off transistor like we see above, let's build a normally On transistor.

Here's why. If we were to try to make a capacitor out of N+, as we would use for the source-drain of an FET, it would not go all the way under the Gate. It would be blocked by the Gate material above it. Remember, the Gate is laid down first. The N+ is deposited later, unable to creep under the entire Gate.

However, we know the N well is built before the Gate. Therefore, if we make our bottom plate using N well, we can guarantee that we get a complete plate underneath.

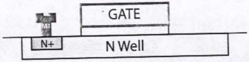


Figure 5–9. *Depletion mode FET, almost. We've lost one contact.*

We can use N+ for the contact to the bottom plate. The N+ will be deposited at the same time as the N+ for the NFET. That saves us some processing steps as well.

■ **Rule of Thumb: Build using existing layers, if you can.**

Examining the above diagram, we can see the regular Gate oxide between the Gate and the N well. This is the capacitor's dielectric. The area formed by the overlap of the polysilicon Gate and the N well defines the area of the capacitor.

The above drawing shows only a single contact. A capacitor built this way will have a significant series resistance in the bottom plate due to the resistance of the N well.

We can reduce this series resistance by placing contacts on both or all sides of the top plate. What people tend to do, however, is make a big block rather than a small strip as they would for an FET. You might see a horseshoe shape for the metal contact.

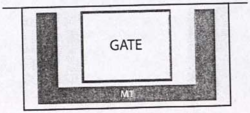


Figure 5–10. *Top down view of a typical diffusion capacitor, with one very large contact horseshoe.*

This pseudo-FET makes a fine capacitor. We call this type of capacitor a **diffusion capacitor.** This is a very simple capacitor, built from existing layers.

Now, you might not have thought of this. Because this capacitor is really a transistor, its capacitance depends on the voltage placed across the plates. If the voltage changes on the top plate, it affects the electrons in the N well underneath it. The capacitance changes as the Gate voltage changes. A capacitor that changes its capacitance. Hmmm. We might not like that.

Diffusion capacitors can work well as fixed voltage decoupling capacitors, though. If the voltage is static, we can rely on the capacitance value.

Parasitic Capacitance

Unfortunately, or fortunately, depending on your point of view, our N well is in a P− region, which provides automatic isolation for us. The fortunate result of this is that we can just plunk chunks of N well down, plunk chunks of poly on top and we get instant capacitors.

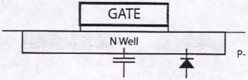


Figure 5–11. N well is layered before the Gate layer.

However, the unfortunate part is that the N well and the P substrate underneath form more parallel plates. Effectively there is another capacitance below the N well between the N and the P. The PN junction is the dielectric barrier.

Therefore, we will always suffer a parasitic capacitance connected to substrate.

Some people model the parasitic as a capacitance. However, since capacitance is built into the equation for a diode already, both are effectively accounted for if modeled as a diode. Therefore, that is what many people do. They model parasitic capacitance as a diode between the N well and substrate.

What if you connect your N well somewhere funny? If you connect your N well where you are not supposed to connect it, you might forward bias the diode between the N well and the substrate.

If you model the parasitic capacitance as a diode, your SPICE simulations could inform you of this possible problem. This is a nice protection. This is another reason why most people model this parasitic capacitance as a diode, with a length- and width-dependent capacitance term.

The thickness of your dielectric determines the capacitance value. The closer the two parallel plates are, the more capacitance you have. So, a thinner dielectric offers more capacitance. Using a good, thin dielectric is covered in our next section.

Metal Capacitors

As we have seen, diffusion capacitors have a parasitic capacitance across the PN junction. Any signal input to the bottom plate of a diffusion capacitor will automatically couple into the substrate.

There are situations in circuit design where we need a capacitor to block a DC voltage but pass the AC signal through to the next circuit block. A capacitor that changes its value depending on the voltage across it is not at all useful for this kind of use.

Consequently, most signal capacitors are made of metal. This eliminates the PN junction, which eliminates the inherent capacitance of that parasitic diode. Also, the voltage dependency is removed.

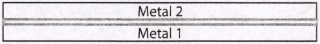


Figure 5–12. Metal capacitors eliminate parasitic capacitance of a PN junction.

In most cases when you run two metals over each other, you want to insure that the top plate does not short to the bottom plate. So, most typical IC processes have a metal insulator that is usually quite thick.

The magic capacitance constant for our equations is slightly different, due to this increased thickness. Other than that, the same equation is used for metal capacitors as for our diffusion capacitors, even with a thick dielectric.

Since the metals are usually kept so far apart, in order to get the same kind of capacitance values as a diffusion capacitor, these have to be much, much bigger. Therefore, a metal-metal capacitor can eat a lot more area than a diffusion capacitor. However, for a good signal capacitor, you might want to live with this tradeoff.

Stacked Metal Capacitors

To recover some of the area used by metal capacitors, we have some tricks. Depending on the number of metals you have, you could make what is called a **stacked capacitor.**

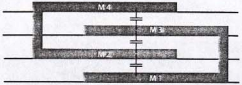


Figure 5–13. *Stacking the metals saves space on the chip.*

Multiple metal plates stack vertically on top of each other like a stack of pancakes.[2] Capacitance occurs between each of the layers, all the way up to the top layer. One strip of metal has been wired into fingers that intertwine with the fingers of another metal piece. Effectively you have interdigitated the metals to get more capacitance per unit of area of chip surface.

Nitride Capacitors

Metal is used for wiring our chip. We normally want to reduce the capacitance between metals, rather than create it. Therefore, metal to metal insulator is usually kept quite thick to reduce the parasitic capacitance.

A thick insulator makes metal capacitors very large, therefore, they are **not** always particularly useful.

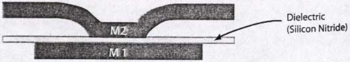


Figure 5–14. *Some dielectrics can be made extra-thin.*

---

[2] I couldn't stop laughing when Chris tried to convince me the UK honors pancakes with a special day! Races? Festivals? He's GOT to be kidding! What's next, a special day for, oh, let's see, say, boxes? "Oh yes, we have boxing day!!" Oh, stop.

You can increase the capacitance value of a metal-metal capacitor, as with any capacitor, by adjusting the thickness of the insulation layer. Rather than use a thick oxide dielectric layer, let's introduce a very thin dielectric layer.

A material with a good dielectric constant, easy to deposit using CVD, is silicon nitride.

By introducing a silicon nitride layer, you create a much more powerful dielectric. However, using a silicon nitride dielectric requires an extra mask and an extra processing step. You will decide if you care to make the tradeoff or not. Is it worth extra processing for bigger metal-metal capacitance values?

Now you have healthy insulation between metals where you use a regular thick dielectric, and you have capacitance between metals where you use a special thin nitride layer. Each is exactly where you placed them. Oxide here. Nitride there. Low capacitance here. Lots of capacitance there. All using the same metal layers. Good job. Your metal layers are working for you just as you wanted.

Closure on Capacitance

Capacitance can be a free byproduct of other layers already being used in your IC. That could be good news. That could be bad news. Your job is to control the capacitance that happens naturally. You might try to eliminate it. You might try to take advantage of it. Or, sometimes you might even cause it to increase substantially.

We try to control capacitance with the fewest number of process steps, and with the smallest amount of chip real estate required for the job.

Maybe you will come up with some tricks of your own. Tell us. We can always add another section to the chapter, and call it by your name.

Go ahead. Invent stuff. It's fun. Besides, this chapter is kind of small.

Here's What We've Learned

Here's what you saw in this chapter:

- Metal-Dielectric-Metal Classic capacitor
- Capacitor functions and uses
- Frequency-sensitive features

- High frequency filter
- Building an N well capacitor
- Capacitance equations
- Periphery capacitance
- Parasitic capacitance
- Stacking interdigitated capacitors
- Nitride capacitors

And more . . .

CHAPTER 6

Bipolar Transistors

Chapter Preview

Here's what you're going to see in this chapter:

- What we can do about that inherent Gate capacitance
- Faster switching of transistors
- How processing limits our choices
- Three parts of a Bipolar switch
- Building switches vertically
- Buried layers brought to the top
- Why we usually don't bother with PNP switches
- The biggest problem for layout people with CMOS experience

And more . . .

Opening Thoughts on Bipolar Transistors

You can build most of the components we have discussed so far using a basic CMOS process.

As we saw in the CMOS layout section, the inherent Gate capacitance in CMOS transistors slows our device. However, in what we call a **Bipolar transistor,** the switching region can be made much smaller. Making regions smaller reduces capacitance.

Bipolar transistors, therefore, help solve the capacitance problem by their size. Also, with their smaller RC time constant, they operate much faster than a CMOS transistor. Fast is good. Bipolar is good. This is a powerful chapter.

We use the name *Bipolar* because these transistors use both electrons and holes at the same time during its operation. It's as if one pole is attracting electrons, while another pole attracts holes. Two (*bi*) poles (*polar*).

You could call a CMOS transistor a *uni*polar device because it only uses one type of carrier during operation. A P Type CMOS transistor uses holes as its main conductor, for instance.

Doping levels of the N and P Type diffusions in a CMOS process are optimized for the CMOS transistor operation, not for Bipolar transistor operation. Extra processing steps, implants and diffusions optimize N and P levels for Bipolar devices. Therefore, you do not produce very good Bipolar devices using plain CMOS.

You will find manufacturers typically offer either pure CMOS processes, or pure Bipolar processes. If you want a mixture of the two types of transistors on the same wafer, the extra processing steps become very expensive and very complicated.

Let's examine the theory behind these Bipolar devices.

Theory of Operation

We do not necessarily need to know how every device works to do basic layout. However, as you understand more about your circuit, you will make better decisions, you will ask better questions. Plus, since we already understand how an FET works, it's a simple step to understand how a Bipolar transistor works, since the concepts are so similar.

We can build two types of Bipolar transistors, NPN transistors and PNP transistors. Let's first look at an NPN device. The PNP is based on the same concepts.

Similarity to FET

Looking at a cross section of the device, we see two PN junctions. That's it. A Bipolar transistor switch is not any fancy sort of new material. It is just two plain, old PN junctions. So far, this still resembles a CMOS transistor.

You can think of a Bipolar transistor as two diodes. (A PN junction is a diode.) The symbol for two diodes would show two arrowed extensions.

The NPN transistor symbol resembles this two-diode drawing. This should help you remember the symbol. However, we use only an arrow coming from the lower extension in our symbol. The arrow denotes conventional current flowing out of the region.

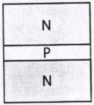


Figure 6–1. *NPN transistor is merely two PN junctions.*

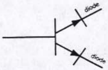


Figure 6–2. *NPN transistor symbol comes from two diode symbols.*

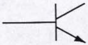


Figure 6–3. *NPN transistor symbol.*

So in an NPN, current comes in the top (which we call the **Collector**), passes through a central area (which we call the **Base**), and goes out the bottom (which we call the **Emitter**). We will see why the terminals are called Collector, Base and Emitter shortly when we see how the device works.

Try It

Answer for yourself why the source of conventional current would be called a Collector, and why the actual Collector of conventional current would be called the Emitter. Doesn't it sound backward to you? Think about it.

ANSWER

It is backward. Conventional current travels in the opposite direction from electron flow.

Did you notice that the NPN symbol resembles the FET symbol?

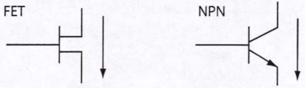


Figure 6–4. *Symbol for FET is similar to symbol for Bipolar transistor.*

An NPN and an FET function similarly. In both devices, current tries to jump across a middle region, from positive to negative, top to bottom. Sometimes it does. Sometimes it does not. An NPN transistor switches current on and off, just like an FET. This switch state depends on the voltage on the controlling terminal. Either the Gate in an FET, or the Base in an NPN. The functions are similar. Therefore the symbols are similar.

How an NPN Works

To understand how a Bipolar switch operates, let's first examine just the lower PN junction.

As you recall, in a regular PN junction we have an abundance of electrons in our N Type region, and an abundance of holes in our P Type region. By placing a positive voltage on the P diffusion and a negative voltage on the N diffusion, we forward bias the PN junction. Electrons start to flow from the N to the P. We effectively have current across a PN junction. (Refer back in this book to remind yourself how a PN junction works.) We call the instigating voltage the **bias** voltage. It forward biases the junction.

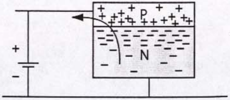


Figure 6–5. *Electrons jump across the PN junction due to the bias voltage applied.*

Let's extend our understanding of a Bipolar switch a bit more. Add another N layer at the top, and another circuit through the entire NPN. We place a much stronger voltage across the entire transistor through this new circuit. We now have two circuits. One voltage is applied across just the bottom PN junction, to get a flow of electrons to jump up into the center P section. A second, more powerful voltage is applied across the entire device.

Imagine what the electrons think that were previously headed toward the bias voltage source off to the left. What would they do now that they see such an attractive positive voltage coming from a just little higher?

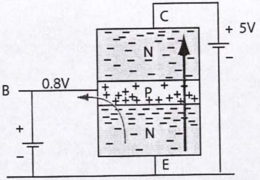


Figure 6–6. *"Oh my, I see a much more attractive voltage up through the other N section. I think I'll just skip straight through this P to get to it."*

For this device to operate, the first PN junction must be forward biased. To forward bias a typical PN junction, we use around 0.8 V. With this minor voltage applied, electrons begin to flow into the P.

If we make the top voltage several volts higher, say 5 V across the whole device, the electrons that have already jumped into the P keep racing upward. They see this much bigger voltage and say "Ah, that's MUCH more interesting." Because the P Base region is so thin, the racing electrons cannot stop. (Slippery socks.) They come pouring over the edge with so much energy that their inertia pushes them right through the thin layer of P and on into the top N layer. If the P region were too thick, the electrons would not have enough energy to get to the top, to see that lovely N waiting with 5 volts. So, the P region has to be very thin.

By adding the more powerful voltage across the entire device, the majority of the current that used to flow in our original forward biased PN junction now flows into the upper N Type region. Some electrons do continue on their original path however. So, there is a very tiny current still flowing through the original circuit.

Does this strike you odd, that our electrons are entering N, rather than leaving N? This is weird. You have a reverse biased diode that is conducting electricity. A reverse biased diode cannot normally conduct electricity. Clever.

The bottom N region is emitting electrons that are being collected by the top N region. Hence the names Emitter and Collector.

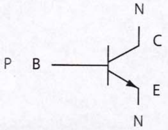


Figure 6–7. NPN sections are the Collector, Base and Emitter. Conventional current direction is shown with an arrow.

So, there they go. Electrons emitted and collected,[1] but only as long as we entice them with our little 0.8 voltage across the first PN junction. Without our little 0.8 V circuit, the electrons would never make it far enough to know anything better existed. Electrons would not leave the bottom chunk of N.

The small current that flows in the forward biased Base/Emitter junction is little more than a nuisance. It can be a factor of a hundred or more smaller than the current across the Collector and Emitter. It acts as the counterpart to the Gate in the FET; as the switch control.

A Gate in an FET only draws a current during the time that the Gate-oxide capacitor is charging or discharging—changing voltage. In contrast, the Base current of a Bipolar transistor is always flowing. Ideally, the Base current of a Bipolar transistor should be zero. If it were zero, then we would have the perfect voltage-controlled switch. However, current must flow for a Bipolar to operate.

Unfortunately, in order to build logic Gates with Bipolar transistors, a constant, static current must flow all the time. Therefore, while Bipolar switches are faster, they burn more current. That is why most microprocessors are CMOS. CMOS uses a lot less power.

Bipolar uses more power.

The ratio between Collector/Emitter current and Base current is called the **beta** of the device. For example, you could have one hundred microamps flowing in the Collector and one microamp as the Base current. This gives you a beta of 100.

However, sometimes any Base current, even small, could be undesirable. It drains current from your circuit. Remember when we discussed building digital Gates, we wanted the Gates normally off, not taking current.

The beta ratio changes depending on how you drive the transistor. The Base current of a Bipolar transistor is variable. For example, at some point your Collector/Emitter current cannot increase further, but the Base/Emitter current can. This variability can cause circuit headaches.

Vertical Processing

For the first time, we will discuss **vertical processing.** Vertical device processing techniques allow more accuracy in the construction of Bipolar transistors. We can make the central P Base region much smaller than using horizontal construction. Therefore, since the P regions are smaller, Bipolar transistors are much faster switches than FETs.

Let's look at the construction of Bipolar devices to better understand this idea.

Switch Area Comparison: FET vs. NPN

Compare a simple FET with a simple Bipolar transistor. On an FET, the Gate length, *L*, determines the speed of device. On the Bipolar, the width of the P region determines the speed of the device. The shorter the distance between N regions, the faster we can switch current flow on and off between these regions. In a speed switching contest, vertical wins.

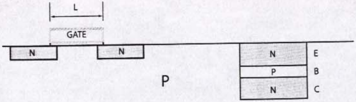


Figure 6–8. Notice the P region of the vertical transistor forms a much shorter distance between N regions than the Gate width in the FET.

The minimum length of the FET Gate depends on how well can you print your pattern on the wafer. However, in a Bipolar transistor you can use just a tiny,

---

[1] Remember that conventional current is the reverse direction from electron flow. Conventional current flows from the Collector to the Emitter.

quick implant to create a very thin layer of P. Therefore, it is easier to create a very thin implant layer than a very thin Gate stripe.

Even if you could build an NPN horizontally, you cannot just place a very small P section between two N's. The problem is that each of the three regions needs a contact. We would have to increase the size of the P region in order to fit it with a contact. That spoils the advantage, doesn't it? The P region becomes too big.

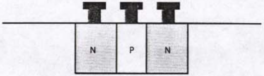


Figure 6–9. Horizontal NPN layout forces the P region larger, so that we may fit a contact to it. This slows the switch.

So Bipolar transistors are built using layers stacked on top of each other. They are known as **vertical devices.**

Construction of Layers

Accessing the Base and the Collector might seem tough with a vertically stacked NPN transistor. These two layers lie below the surface. However, as happens frequently, people were clever. Someone realized that the horizontal length of our layers does not interfere with the speed of the device. That's the key.

Let's build a vertical NPN, step by step, to further understand our device layout. Depending on your technology and processing, you might construct the Base and Emitter very differently. I will just draw a very old-fashioned, diffused version of an NPN transistor as our basic example. The ideas apply to all construction techniques.

Construction of the Base/Emitter junction is much more important than the Base/Collector junction. The electrons in the Emitter barely fall over the barrier ever so gently. However, once the electrons get past the Base, the Collector can be imagined as just a big catcher's mitt. So, the second junction, the Collector, need not be as well controlled.

We want to build our most accurately controlled region at the top, last in the process. Layers placed earlier will suffer more diffusion and stresses than layers placed later. Since we want to control the Base/Emitter junction much more carefully than the Base/Collector junction, we will build the device upside down. Surprisingly, that puts the Emitter on top, the Base in the middle, and the Collector on the bottom.

First, we establish our Collector area with a chunk of N.

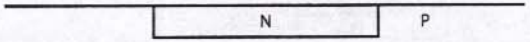


Figure 6–10. First we create the Collector.

Annealing the wafer with a P epitaxial layer over the top diffuses the Collector. Our Collector becomes larger and less distinct.

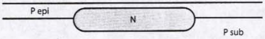


Figure 6–11. Diffusions spread.

Also, when we created the P epitaxial layer, we buried our Collector under it. How can we connect to a material we have just buried? As we discussed briefly in the processing chapter, we can implant some additional N deep enough to make contact with the buried N. This creates an N pathway giving us surface access to the buried Collector. The implanted N we see from the top will become our Collector terminal.

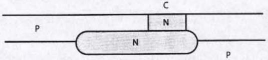


Figure 6–12. Implanting a connection to the buried layer.

Next, we place the Base region, a region of specially doped P, above our buried layer of N. Notice it does not cover the entire buried N region, since the implanted N contact is in the way.

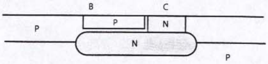


Figure 6–13. *Base region created above the buried layer.*

We already have some P in this area, from the P epi, but we want to control the doping of this P region more carefully. So, our Base is another implant. We keep the P very shallow, of course, to give us faster switching speed.

Our last step in the construction of a Bipolar transistor is to implant some N for the Emitter. Notice how much smaller our Emitter N region is than the Collector N that we buried in earlier steps. That's fine. Since the bottom layer represents our oversized catcher's mitt, we are fine with it being large and diffused. Remember, control matters most in the upper PN junction, not the lower.

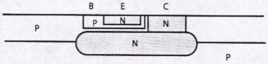


Figure 6–14. *Here comes our Emitter. We now have three horizontal contacts to vertical layers. The contacts are labeled B, E, and C.*

Because we pulled our P Base diffusion out to the side further than needed, we have room to connect to it. Our metal contacts on the surface are in the following order, left to right: Base, Emitter, and Collector.

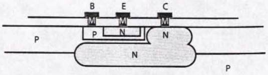


Figure 6–15. *Metal contacts seen in cutaway view.*

The top view of a typical NPN appears as a strip showing the three contacts for Base, Emitter and Collector. Notice the two N contacts are next to each other.

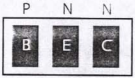


Figure 6–16. *Top view of NPN device. Notice the middle layer contact is located on the far left.*

Find the actual region of NPN. (No, I mean really. Go find it in the above figures before you read on. The area that does the work. Is it the entire device?)

It's not the entire drawing, is it? Everything interesting happens in the center.

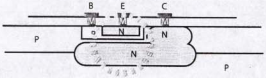


Figure 6–17. *The action happens between N regions, only in the center.*

The construction of a vertical NPN uses excess N and P material, pulled out to the sides, to provide surface access. The actual NPN action is only located where all three layers stack vertically. We use precious chip real estate reaching the buried layers, but the speed of the end device is terrific.

Parasitics of NPN

The Base implant extends out of the sides of the device, beyond the center region of activity. The large size of this diffused implant creates some serious extra resistance getting over and up to the contact.

We also pull our Collector out the side, and up, so there must be substantial resistance through the Collector as well. Moreover, we have our old, faithful PN junction at the bottom, creating a big capacitance to substrate on the Collector.

Of all these parasitics, the two most prominent are the Base resistance and the Collector capacitance. These parasitics slow things considerably. Someday we will have clever solutions to show you, but no one has devised them. Yet. (Let us know when you do.) Until then, just handle the parasitics as well as you can, as you draw your layout.

PNP Transistors

Just as we can build complementary NFET and PFET devices in a CMOS process, the complement to the NPN device is a PNP device. Notice the Collector and Emitter are P, instead of N. The Base is N, instead of P.

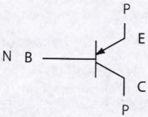


Figure 6–18. Complementary device, PNP.

The arrow indicates the direction of current flow in the Emitter, which, you will notice, is the opposite direction than the NPN arrow.

Lateral PNP

If you use pure Bipolar processes, then your PNP is very easy to build. You can implant at levels you want for Bipolar processes.

A certain process becoming popular combines Bipolar devices and CMOS devices, known as a **BiCMOS** process. BiCMOS offers the speed of an NPN device coupled with the logic functionality of the CMOS technology. We can get the best of both worlds.

However, additional layers are required to adequately isolate the bottom Collector layer in a vertical PNP using BiCMOS. The bottom layer needs an additional layer of isolation that was not necessary when building an NPN. We need one more layer of N underneath, as isolation.

Extra layers mean more processing steps, more money, and more things to go wrong. So, while some Bipolar processes may offer a buried P, most BiCMOS processes just do not bother with a vertical PNP. A BiCMOS vertical PNP adds too much cost.

In most BiCMOS processes, your PNP device, if you have one at all, is a less expensive lateral PNP, built like an FET. A lateral PNP contains a chunk of P in a chunk of N (usually N well) with a chunk of P+ next door. All lateral.

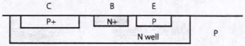


Figure 6–19. Lateral PNP.

If you use lateral PNP's, you can build two in one go, to try to reduce some of the series resistances in the well. A cross section would reveal PNPNP, representing two PNP's sharing the center P region.

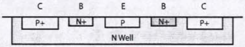


Figure 6–20. If you use a PNP at all, it will likely be lateral. Here we see PNPNP, effectively two PNP's in one.

Looking at the lateral PNP from the top, you might see concentric circles of P, N, and P, rather than simply a one-directional strip as in an FET.[2] You could even build the two-in-one PNPNP in rings.

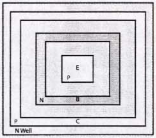


Figure 6–21. PNP built in rings.

If I say they are circles, then they are circles. Never mind the pointy bits at this time.

Instead of having an implanted Emitter, some people dope polysilicon N Type. Doping with N offers lower resistance than a shallow implant.

There you have it. That is how to make our two basic, complementary Bipolar switches, NPN and PNP.

Career Transitioning from CMOS Layout

Unlike CMOS, with all its fancy source drain sharing and flipping, Bipolar is just placed and wired. I find Bipolar layout much simpler in that respect.

Since Bipolar transistors are typically used in either high precision analog or high frequency-high precision analog circuitry, you must learn to deal with these new concerns. Factors such as the wiring and placement of devices with respect to each other, and high frequency cross-talk coupling become very important. Much more so than you would expect. (We cover these and other essential issues in our companion book, which takes you beyond the fundamentals learned in this book.)

Number of Rules

Mask designers with extensive backgrounds in only CMOS technologies usually find it difficult to transition into Bipolar layout. CMOS technologies require many design rules that a mask designer needs to know extensively. However, since the Bipolar devices are pre-built for you, and device sharing techniques are usually not used, you need fewer rules to do your job.

This can sometimes be quite a shock to the human system. Until you have laid out a few cells in a Bipolar technology, it can be very nerve wracking. You expect to need a lot more than you are being told. You will feel like wandering the office looking for all those rules you know must exist. You will wonder if your boss knows what he is talking about.

For example, if you implement source-drain sharing in a CMOS process, you need to know how close contacts can get to Gates, how far Gates can extend past an active diffusion and how often to place a well tie-down. You begin to know your CMOS rules intimately. These rules become your job, as you see it.

With Bipolar technologies you are just given a pre-defined piece of layout and told, "Hey, all the rules are done for you." That can be unnerving until you get used to it. The tendency is to continue searching for more rules to worry over. Layout designers new to Bipolar squirm in their chairs, asking questions like, "What are the design rules?" Or, "What are the numbers I need to worry about?"

In reality, all you need to worry about is how close you can place buried layer to buried layer. Buried layers out-diffuse most. So, they would be the most likely materials to keep apart.

Once you determine how close you can place your buried layers, the rest of your layout is just worrying about the metal wiring rules. So, as you can see, Bipolar layout is more straightforward than CMOS layout.

*From my experience in training others, **CMOS** engineers suffer most from fear of the unknown. They are so accustomed to monitoring 30 to 40 rules. When I say to them, "Well, you just place the transistors down and you wire them up" they get a very scared look on their faces.*

"But I need to know 30 or 40 rules!"

"No, you don't. You need to know 4 or 5."

Once you get past the fear of dealing with just 4 or 5 rules, you can become much more confident about just placing transistors. Your work becomes much more creative. Instead of spending your energy worrying about so many rules, you can devote your creative skills to elegant interconnect, symmetry, matching, or parasitics.[3]

A Bipolar layout person should know more about electricity, electronics and circuit techniques than a CMOS layout designer. Not a lot more, but any electronics you can learn helps a great deal. For CMOS layout, an understanding of circuit function is not as important as understanding the design rules.

In CMOS, we worry more about design rules.

In Bipolar, we worry more about circuit function.

In Bipolar layout, the circuit function is important. You worry about what the circuit is doing and how it is doing it, rather than whether a diffusion is too close to a Gate stripe. You worry about whether you have your Emitter strapping matched from this device to some other device on the other side of the circuit, or whether you have a good signal flow through your layout.

---

[3] See authors' companion book on essential techniques for mask design. We're trying to keep each book affordable, so that you can have both on your shelf. Remember to write your name in each book to discourage permanent borrowing.

Because you are dealing with much higher frequencies, the circuit function demands more of your attention than, say, the rules for the diffusions.

Black Box Placement

Most Bipolar devices are not as easy to model as an FET. Consequently, unlike stretching and multiplying a device to your desired length and width in CMOS, Bipolar processes give you a selection of 4, 5, maybe 10 fixed devices, whose models have been well determined. You cannot alter these 4, 5 or 10 given boxes. You are told, "This is how they work, period." You never touch the diffusions. Never stretch the devices. You treat the Bipolar transistor like a black box, a magic, untouchable rectangle.

It is useful to understand what is within the boxes to know what you are laying out. Once you understand the box, your layout becomes just a join-the-dot exercise. Of course, someone certainly designed and tested the box in the first place, understood the places, diffusions, how it all worked. It is nice to know how a Bipolar device works in case you are told to originate some Bipolar devices.

Closure on Bipolar Transistors

> *I've done mainly Bipolar layout in my life. Consequently, I hate doing CMOS layout . There are all these rules you have to follow. They're a pain in the neck.*
>
> *You'll find your circuit designer will be much more critical of your layout with Bipolar technologies because of the circuit functionality and frequencies involved. Things you can get away with in CMOS you can't get away with in Bipolar, like thinking "Oh, there's no room to place this transistor close so I'll just move it out of the way."*
>
> *You'll get spanked by the circuit guy. He will say, "I don't care about your having no room. This transistor has to be next to this transistor because the wire that runs from the Collector of transistor 1 to the Emitter of transistor 2 has to be as short as we can possibly make it. That piece of wire is critical to the operation of the circuit."*
>
> *Certainly as frequencies start to rise in CMOS, layout is more critical, but usually CMOS is just "make it as compact as possible." Worry about some current density rules and away you go. Bipolar is much more creative.*

Employers diligently seek people offering analog Bipolar layout skills. They receive more money than people with just plain CMOS skill. It is rare to find people who have done a lot of Bipolar layout. When employers do find them, especially with high frequency experience, they are well looked after.[4]

Here's What We've Learned

Here's what you saw in this chapter:

- Reducing capacitance with vertical layering
- Faster switching using thinner layers
- PNP requires extra processing for isolation
- How an NPN transistor operates
- Emitters, Bases and Collectors
- Vertical transistor construction
- Implanting contacts to buried layers
- PNP drawbacks for both lateral methods
- Advice for the CMOS-experienced learner

And more . . .

---

[4] I've seen ridiculous bidding wars between companies trying to hire a Bipolar high frequency experienced design team. In fact, if you have this experience, call me. I have some nice incentive plans for you.

CHAPTER 7

Diodes

Chapter Preview

Here's what you're going to see in this chapter:

- A diode is a PN junction
- How several types of diodes are built
- A look at some different uses of diodes
- Why we would want to use a transistor as a diode
- How to prevent voltage zaps from ruining our chips
- Some variations on the typical shapes of diodes

And more . . .

Opening Thoughts on Diodes

This is a rather simple chapter. As we have seen, a diode is just a PN junction.

Integrated circuits have P Type implants and N Type implants all over the place. Within reason, you can make a diode out of any pair of them. Some of those PN junction combinations are more useful than others. The usefulness depends on the doping and thickness of the implants, and other factors. Not all PN junctions are the same.

As we learned earlier, current through a diode may only pass in one direction. We use that feature of diodes to isolate devices from each other. In addition, we will also see a few creative uses for this unidirectional feature.

Types of Diodes

"Let's see what happens if we implant some N into a P substrate." I don't think so. People typically want to control their diodes better than by just random implantation. In fact, people have developed several sophisticated methods of making a diode just so they can improve their control.

What do you want the diode to do? How do you want it to work? How should it react with other devices in your process? Will you need Bipolar transistors or CMOS transistors? The answers to these questions determine how you will construct your diode—which P and which N to use, how to configure the layout, whether to include additional transistor components, and so on.

Regular Diode

One of the simplest ways to make a PN junction is to implant some N into the P Type substrate. It is not very controllable, though, because you are at the whim of whatever the P Type substrate's doping level happens to be. However, if the dopant is the right level, you will create a useful diode.

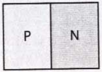


Figure 7–1. A diode is a PN junction. Implanting N into a P substrate is the simplest form.

Applications for Regular Diode

Many circuits need diodes, particularly analog circuits. Diodes in a CMOS process are very useful for providing voltage references, temperature compensation, or temperature measurement. Diodes which provide feedback based on the temperature of the chip, for example, could raise or lower the power within a particular circuit depending upon how hot the circuit becomes when operating.

As another example, you can make a logarithmic amplifier using regular diodes. The diode is inserted into the feedback path of an op-amp. The amplifier response now becomes logarithmic instead of the linear response you would get with a resistor.

Bipolar Transistor Diode

In Bipolar transistor circuits, the choice of diode depends on circuit technique. Instead of having a chunk of P and a chunk of N, which form our regular diode, you could use a Bipolar transistor as a diode.

Remember from the last chapter, a Bipolar transistor is NPN; Collector, Base, Emitter. The Base/Emitter PN junction is a diode. So, you could use this junction from a transistor as your diode.

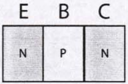


Figure 7–2. Transistors are made of PN junctions. Yes, indeed. You can use parts of a transistor as a diode.

Sometimes we need the electrical characteristics of a diode to react the same way as a Bipolar transistor, to track the transistor. The electrical characteristics of a regular, plain diode in substrate will not track the same as the characteristics of a Bipolar transistor, since their fabrication steps are very different. So, when tracking a Bipolar transistor, use another transistor for the tracking. However, you only use the Base and Emitter portions of the Bipolar transistor for the diode tracking properties. Since you created an entire Base/Emitter/Collector transistor, but are only using the Base and Emitter, you have to do something with the Collector. Most people just short it to the Base connection.

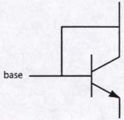


Figure 7–3. Shorting the Collector leaves us with a Base/Emitter diode.

Application for Bipolar Transistor Diode

Let's look at that tracking example a little more closely. We have a very small circuit—a Bipolar transistor with a resistor connected to the Emitter of the transistor. We want to measure the current going through the right resistor, R_2. However, any measurement we do across R_2 will affect our circuit. If we can emulate the characteristics elsewhere, we can take measurement there, without bothering the original circuit.

If we connect a diode, D_1, and a similar resistor, R_1, across the Base to the ground, we can measure the voltage across R_1. The characteristics of the diode are the same as the characteristics as the Base/Emitter junction in Q_2.

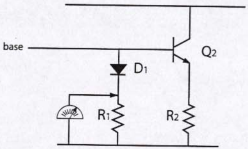


Figure 7–4. Challenge: Measure across R$_2$ *without affecting the circuit.*

Now let's change the circuit to include a Bipolar transistor diode in place of the regular diode. Because this diode is composed of the same material made during the same processing as the transistor Q_2, this diode will have the same characteristics as that transistor, provided they are fairly close to each other on the chip.

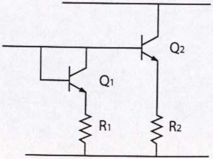


Figure 7–5. Example of tracking using a Bipolar transistor diode.

Voila, that's what we wanted. We have emulated the characteristics of R_2 elsewhere on the chip. We can take measurements across these new components, which we know match the characteristics of R_2. Therefore, we know what the measurements would be if we were to measure the actual circuit in question. All that, and we have not bothered our primary circuit. We're happy campers.

From a layout standpoint, you can make this tracking diode by either of two methods. One method uses a regular transistor, connecting the Collector to the Base.

Alternatively, you can use a Bipolar transistor layout, eliminating the buried layer, the Collector, and its contact.

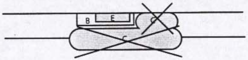


Figure 7–6. Tracking diode, made by eliminating Collector and contact from Bipolar transistor layout.

This method could use a little more explaining. Let's take a closer look.

Effectively, using a Bipolar transistor gives us a PN junction made of the same processed materials as the device we need to track, though we have lost the Collector. Although it provides a tracking diode, the device does not match exactly without all its parts.

Figure 7–7. We have lost the Collector. Saves space, but can we trust the characteristics of a tracking device that is missing an element?

So, although it uses more space, most people leave the Collector in place, and use the shorting technique, in order to ensure better matching. By using two identical transistors, you are guaranteed the devices are the same. You will see bigger parasitics than had you used a regular diode, due to the presence of the Collector.

From a layout perspective, you do not draw that much to accomplish this. You place the transistor and simply connect it as required.

Varactor Diode

Another of the various types of diodes available is the **varactor diode.** The varactor diode has a highly variable junction capacitance that depends on the voltage you put across it. All diodes exhibit this characteristic, but in a varactor diode, the dopants have been specially chosen to enhance this variable capacitance feature.

The capacitance of the diode depends on the number of electrons and holes present. As you increase or decrease the applied voltage, the electrons and holes in the semiconductor are either attracted to, or repelled by the voltage we apply. The potential barrier increases or decreases in size and the parallel plates of the semiconductor get further apart or closer together. As we saw in the chapter on capacitance, the capacitance value is proportional to the gap between the plates. Hence, the capacitance changes as the applied voltage changes.

Application for Varactor Diode

Varactor diodes are very good for making voltage-controlled oscillators. The variable capacitance properties of a varactor diode can be used with an on-chip inductor to form a series or parallel resonant circuit. If we can vary the capacitance of the diode with an external tuning voltage, then we can vary the frequency at which the circuit resonates.

ESD Protection

One helpful property of diodes in an IC is protection against what we call **ESD, electrostatic discharge.** Try wearing a 90% nylon jumper[1] all day. You pick up a chip and feel a sharp electrical zap in your finger. Those zaps are in the thousands of volts. Due to the very thin oxides involved, you can blow things up with those kinds of voltage levels.

Remember we talked about reverse breakdown voltage of a diode? That can actually help us. I'll show you how.

In Figure 7-8 is a single transistor circuit. If we zap the transistor without protection, the oxide at the Gate can die.

What you can do is place two diodes before the transistor, as in Figure 7-9.

Even if you only have a pure CMOS process, another way you can make an EDS diode is from the CMOS transistors themselves. They appear as transistors in Figure 7-10, but since we are only using them for their diode properties, we will refer to them as diodes.

---

[1] American: sweater.

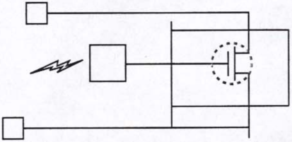


Figure 7–8. *ZAP! Electrostatic discharge kills our thin little oxide Gate when you pick up the chip.*

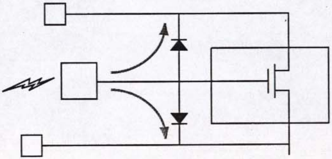


Figure 7–9. *Diodes will conduct under high reverse bias, allowing easy escape for all that zap current. No damage to the Gate this time.*

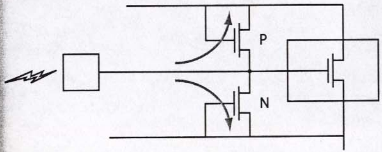


Figure 7–10. *Using transistors as ESD protection diodes.*

If you pick up a chip, zapping enough static discharge from the chip pin to negative rail in a positive direction, you will see that all the energy flows through the bottom diode. None of the discharge voltage gets through to the transistor Gate.

ESD diodes protect the circuit by reaching their reverse breakdown voltage before damage is done further down the line. The reverse breakdown voltage of a diode can be 12 V or so. Once the junction starts to break down, it just conducts current freely like a wire. When the circuit receives a static discharge, it is easier for the current to flow the wrong way through the diode, at this high voltage, than to flow into the circuit.

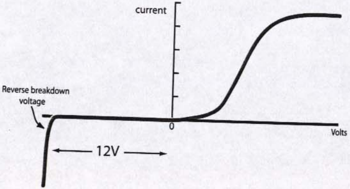


Figure 7–11. *When we have reverse breakdown voltage through a diode, the diode conducts very well, though in the other direction.*

In addition, having diode protection clamps the transistor Gate at a maximum of 12 V. The transistor, or any device next in line, cannot receive any bigger voltage than the reverse breakdown voltage of the diodes. When the voltage resumes normal levels, the diode again functions as a diode instead of as a wire.

Because of the very high voltages involved, ESD diodes need to be laid out very carefully. Missing one opportunity to prevent that critical zap could ruin a good project.

Substrate ESD Diodes

Good ESD diode layout is all about energy flow. You could have a layout for a diode comprised of just an N and a P in the substrate.

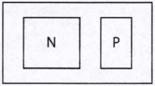


Figure 7–12. *Diode formed by having N and P in substrate.*

We want to get as much energy in and out of this diode as we possibly can. Therefore, some people draw the diode as a ring structure. Surrounding the N contact with a ring of P contact ensures that the escaping energy is picked up in all directions as soon as possible. Nice, easy exits, located in all directions.

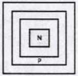


Figure 7–13. *Ring structure maximizing PN junction frontage.*

From a side view, we would see P, N and P sections.

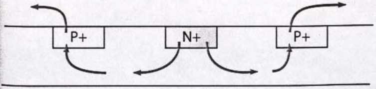


Figure 7–14. *Cross section of substrate ring diode.*

When the ring diode is zapped, the energy comes in through the N contact. That's done quite easily. The difficult part is trying to dissipate the energy as quickly as you can. Having P surrounding our input gives our zap of energy many directions to travel. This is the advantage of the ring structure—lots of places for the energy to leave the chip. This type of ESD diode, an N Type in P substrate, is typically used for most ESD protection.

In a CMOS process, these substrate diodes are free, requiring no additional processing costs. Use the diffusions you already have. Draw these diodes on your existing layers.

N Well ESD Diodes

We also typically have an N well on our chips. Using the N well, you can implant a center of P, surrounded by a ring of N. This is the same idea as the substrate ring diode.

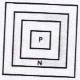


Figure 7–15. *Taking the zap in an N well, we can ring the P with N.*

The side view of this diode resembles the cross section of the substrate diode. However, the N and P are reversed. Notice now the flow direction is also reversed.

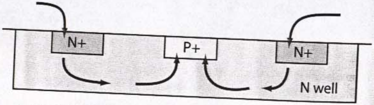


Figure 7–16. *Cross section of N well ring diode.*

This type of diode is called a **well diode.** Well diodes are typically used to protect an input to the most positive rail, the power supply. Substrate diodes, from the previous section, are typically used to protect the most negative supply.

Some people in the Bipolar world use the well diode extensively for both the positive and negative supplies. They do not like the idea of coupling all their chip signals into the substrate. They would rather confine it within a known space, namely the N well. Your use of N well diodes depends on the circuit, what you are trying to do, the frequencies involved, and so on.

Every input and output pad should have some ESD protection on it. Who knows where you will grasp the chip to pick it up—the corner, the middle, or the side? Therefore, a perfectly protected chip will have some sort of ESD protection on every pin.

Placing ESD diodes on every signal pin has a drawback. It is possible that you can ruin a perfectly good chip with ESD diodes. Imagine that your chip has a very sensitive input pin and that there are some very noisy output pins. The ESD diodes will connect the outputs to the inputs using the substrate and the ESD diode capacitance. On very high frequency circuits, ESD diodes can be a very big issue.

As frequencies increase, capacitances become lower impedances. As you raise the frequency in your circuit, the capacitance from well to substrate can almost connect all your inputs and all your outputs to each other. So in some really high frequency circuits you will see that people do not even put ESD diodes on their layouts at all. However, certainly on big CMOS microprocessors, ESD protection is a major concern.

Special Layout Considerations

Here are two unique designs to keep in mind as you layout your diodes.

Circular Layout

If you have ever played with a high voltage generator, like a lightning conductor, for instance, you are aware that the high voltage is seen as a sudden spike that forms in a concentrated spot. If we use the square layouts as shown above, we run the risk of causing these voltage spikes to leap from the corners, where the charge is concentrated.

The perfect structure for an ESD diode is a circle (a real circle). Lack of sharp corners stops the high voltage and current from damaging the diode.

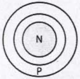


Figure 7–17. *Some layout tools will not draw real circles.*

You can see the circular pattern would contain a voltage zap quite well. Current flow cannot concentrate at any one given point, since there are no corners.

CAD tools do not like circles. The mathematical calculations slow the functioning of the software. Sometimes when you want a circle, you may have to draw a square, since that might be as close to representing a circle as you are allowed.

Multi-faceted polygons, like the square, are used to approximate circles. The more facets, the closer the approximation. Some CAD tools will not allow angles other than 45 or 90, restricting you to much rougher circle approximations. Others will draw perfect circles for you.

Multi-Finger Diode

Sometimes on special diodes, such as ESD and varactor diodes, for instance, you will see a multi-fingered layout like the one we needed earlier when working with our long, thin CMOS transistors. Likewise, we can change a long, thin diode into separated chunks, place them side by side, and wire them in parallel, as well.

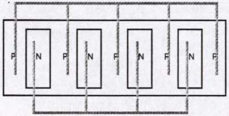


Figure 7–18. Splitting a long, thin diode into fingers, layout view.

Splitting the diode helps reduce resistance, conserves chip real estate, and provides us with a more easily managed, compact design.

Closure on Diodes

Usually you just select a standardized diode off the shelf and use it. However, now you understand why and how they are built. As with other devices, you might be asked to generate a new diode in a new process, based on certain criteria. If you haven't a clue what things look like from a processing point of view, then you are stuck. Now you won't be stuck.

That's diodes. No magic. Go hook 'em up.

I told you it was a simple chapter.

Here's What We've Learned

Here's what you saw in this chapter:

- Using your PN junctions as diodes
- Regular diodes, feedback loops and logarithmic amps
- Voltage matching and tracking using transistors as diodes
- Oscillation and varactor diodes
- ESD protection for both positive and negative power supplies
- Ring construction
- Approximation of circular layout
- Multi-finger layout

And more . . .

CHAPTER 8

Inductance

Chapter Preview

Here's what you're going to see in this chapter:

- Relationship between current and magnetic fields
- Reluctance of inductors to conduct high frequency
- Properties of high frequency transmission
- Helping signals bounce around corners
- Effect of pointy corners on math modeling
- Inductance that spirals back on itself
- Device parasitics can sync in harmony
- High frequency chokes
- Building transformers
- Placement issues

And more . . .

Opening Thoughts on Inductance

With the increasing prominence of high frequency integrated circuits, wiring properties such as inductance require special consideration. Many of these considerations are new to the digital world. Therefore, you are starting to see people looking at analog techniques for tools to deal with these frequencies. The popular adage, "The world has gone digital, analog is dead" is not so. Far from it.

Analog skills will always be needed. In fact, my recommendation to people studying electronics is "Forget the digital world, study analog and you will have a job for life." Sure, somebody has to know digital, but if you want to earn the real big bucks, study analog.

With analog, you might spend three months trying to design 10 transistors, as opposed to three months to design 12 million transistors in digital. It's too easy to design the 12 million transistors. Just a lot of copying. Let someone else do it. Go analog.

Let's get to work.

Basic Inductance

Whenever a current flows in a wire, a magnetic field is generated around the wire. Likewise, if we generate a magnetic field near a wire, current flows inside the wire. Current and magnetic field occur together. Always.[1]

The **Right Hand Rule** tells you the direction of the generated magnetic field. Grab the conductor with your right hand, your thumb extended in the direction of current flow. Your curled fingers mimic the lines of magnetism. This is also called the **Hitchhiker Rule.** If the current flowed in the opposite direction, you would see reverse effects in the nearby wires.

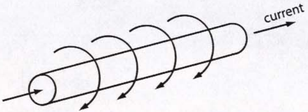


Figure 8–1. Current and magnetic fields induce each other.

If one wire is conducting current, the magnetic field it generates influences any nearby wires to also conduct current. The second wire has been recruited, or induced, to produce current. Hence the name, the **inductor.**

This induced transfer of current occurs even among the tiny wires in an integrated circuit. Wherever current flows in an IC, it generates a magnetic field, which in turn induces current in nearby wires.

---

[1] Have you ever really stopped to think about how weird this is? It bugs me. I want to know how magnetism pushes electrons from a distance. How does it know where they are? What is it that the electrons *see*? Do they feel it when they get pushed? Are we just little electrons in a bigger universe, being pushed to eat cake, ride motorcycles and make babies? Is that what electrons feel?—*Judy*

The magnetic field not only interacts with surrounding IC components, but the magnetic field interacts with the current flowing through its own wire as well. An inductor produces self-inductance. As the current flows, the magnetic field generates a back voltage, which tries to stop the current flow that is producing it.

A constant DC current produces a static magnetic field. The static field will interact with other conductors but not generate any current.

Try It

Make a coil of wire. Connect it to a voltmeter and measure the voltage across the coil. You should see zero voltage. Now lay a permanent magnet next to the coil and measure the voltage across the coil. You should still see zero voltage.

Now wave the magnet around slowly, close to the coil. Provided you have the voltmeter set to a sensitive enough range, and the coil of wire is big enough, you should see the reading on the voltmeter vary wildly. This is the basic principle a power generator uses.

Referring only to AC, in the last chapter, we noticed that our capacitor was frequency-conscious. As the applied voltage frequency increased, the capacitor's ability to conduct current increased. An inductor is the other way around.

As the frequency of the applied voltage across the inductor increases, the frequency of the current flowing also increases. However, as we have seen in our coil experiment above, a changing magnetic field induces voltage and current. The induced voltage and current are generated in the opposite direction of the voltage and current that is applied, thus canceling some of it. The higher the frequency the bigger is this effect, and the smaller is the overall current through the inductor.

Inductors block high frequency.

A capacitance lets high frequencies through. An inductor stops high frequencies.[2]

Particularly in high frequency circuits, you sometimes want to block your frequencies from certain locations. Inductors can block those high frequencies for you.

---

[2] Running DC through an inductor has no effect. All the DC current passes through since there is no change in current.

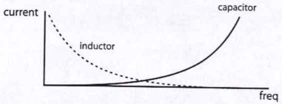


Figure 8–2. As frequency increases, capacitors pass more current, inductors restrict more current.

Transmission Lines

Picture a block of layout connected with a wire to another block of layout. Send a high frequency signal from one block to the other, along the wire. If you measure the amplitude of the original signal, it will be bigger than the amplitude measured at the opposite end of the wire where it is being received.

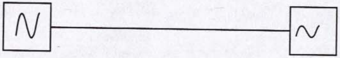


Figure 8–3. Signal received with same frequency, but smaller amplitude.

The signal frequency will be the same, but the amplitude is smaller. Why is this?

Straight Segment Characterization

Well, to start with, the wire has some resistance. This resistance reduces some of our signal due to a simple potential divider effect.

As we increase our signal frequency, however, the loss in amplitude becomes greater than can be accounted for by just the resistance of the wire alone. Of course. You guessed it. The wire also has some capacitance and inductance that become significant as the signal frequency increases. This is just what we were discussing in the previous section.

If our signal contains many different frequencies, like in music for instance, then some frequencies will be reduced in amplitude differently than others. If our signal generator is a HiFi amplifier, the wiring to the loudspeakers can have an effect on the sound we hear. The higher frequencies will be blocked differently in one speaker than in the other.

The wiring of an integrated circuit has exactly the same problem. The parasitic capacitance, inductance, and resistance of the wiring produce a noticeable effect at high frequencies. We can characterize our wiring, however, and adjust our circuitry to compensate for the losses. These characterized wires are known as **transmission lines.** We *transmit* our signal from one end of the wire to the other.

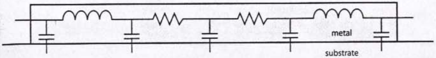


Figure 8–4. Parasitics, all over a transmission line, like fleas on a dog.

An IC transmission line is just like the heavy lines that transmit power around the country in the electricity grid. You transmit power from one end to the other. Same thing.

If you want to run a signal from one end of a chip to the other, use a transmission line characterized for that purpose. If the frequency of our signal is 1800 MHz and they have only characterized the transmission line up to 300 MHz, you have no idea what will happen if you use it. Make sure that the transmission line model is adequate for what you are trying to achieve.

So what does a transmission line look like in an integrated circuit? Well, to be honest, not much. In its simplest form, it is just a regular wire like any other. However, as we have seen, the parasitics of the wire make a big difference.

A typical transmission line will be wider than minimum and exists on the lowest capacitance metal layer, usually the last metal in the process. Some IC processes also insist that other structures or metals lie beneath the main wire, to give the transmission line more reproducible characteristics. Most transmission lines are characterized as unbent wires of a given width. The length is variable.

Corner Characterization

Unfortunately, an IC cannot be wired entirely with unbent lines. We must introduce corners to avoid other chunks of circuitry in the layout. Every time we turn a corner to avoid circuitry, we disrupt our nice, perfectly characterized transmission line.

Luckily, high frequency circuits have few sets of wires that must be treated as transmission lines, but the bends do pose a problem, nonetheless.

One way to approximate our wiring in our circuit simulators is to represent the wire as a series of individual, straight transmission lines. So, each section of a staggered path could be represented individually. This would yield the same mathematical model regardless of the number of bends.

Another approach to help mathematically account for these bends is to measure the centerline of our bent wire and model it as a straight line of that length. Again, the corners themselves disappear mathematically.

Measurements have found that these two methods of approximation are not always perfect. Why would you suppose that is?

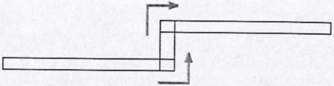


Figure 8–5. *Ever tried running fast around a corner wearing clean fluffy socks on a newly waxed linoleum floor? Electrons bump their heads on the walls, too.*

We are trying to transmit energy down a wire. As the energy propagates along the wire, it soon hits a bend. Some energy can actually reflect back up the wire against the oncoming flow. The signal energy reflects off the bend wall like light reflecting off a mirror.

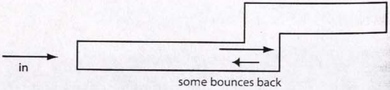


Figure 8–6. *Walls reflect current.*

In very high frequency systems, it is often useful to think of the signal energy as electromagnetic radiation. A good analogy is to think of the signal energy as light bouncing around inside a fiber optic cable.[3]

---

[3] Chris prefers the bouncing light imagery. I prefer seeing current as water flowing through pipes. My physics teacher had us do a lot of wave tank work. It stuck. I get very thirsty drawing these circuits now.—*Judy*

Based on what we know from light waves, we can help the energy get around corners. If we angle the corners of our bends to 45 degrees, the angled corners help the signal energy bounce into the next wire segment. The same principle helps light bounce around the mirrors in a periscope.[4]

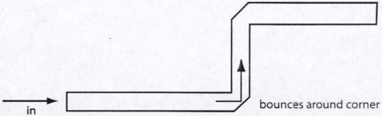


Figure 8–7. *We can help energy around those bumpy corners.*

This is a useful layout technique for all very high frequency IC's. You will use this technique almost exclusively in microwave-frequency integrated circuits.

Some designers go even further. Besides modeling the straight sections as transmission lines, they also model the corners as completely separate components.

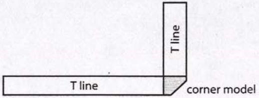


Figure 8–8. *Modeling corners separately makes calculations more accurate.*

In the figure, we would model the entire wire using three modeled components; the first transmission line, the corner, and the last transmission line. This modeling method is the most accurate.

Spiral Inductors

An inductor can be a very useful circuit component. However, the length of a piece of wire must be fairly long in order to get a reasonable inductance value.

---

[4] Get some small mirrors, some paper towel tubes, and teach your kids how to make their own spy scope.

We can save some space by making a spiral inductor. A **spiral inductor** is literally wire in a spiral shape.

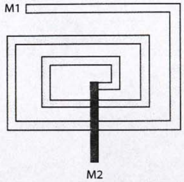


Figure 8–9. *Spiral inductor. First spiral inductor layouts discovered in prehistoric cave drawings, Central America.*[5]

Winding our wire in a spiral like this also has one other advantage. The magnetic field from each turn of the spiral interacts with every other turn in the spiral, making the total inductance greater than that of a single wire of the same effective length. This interaction is known as **mutual inductance.**

Spiral inductors are not modeled as transmission lines with bends. They are modeled as a unique element.

Spiral inductors do eat up a lot of area. Watch your space.

The metal layers of a spiral inductor can seriously affect the component's performance. The parasitic resistance of an inductor made of very thin, resistive metals will affect what is known as the **Q** of the inductor. We will discuss the Q of the inductor next.

Inductor Quality

An ideal inductor at all frequencies is impossible to fabricate. Parasitic resistance and capacitance handicap the way the inductor functions. At low frequencies, the series resistance pulls the inductor frequency response away from ideal. At high frequencies, the parasitic capacitance pulls the inductor frequency response away from ideal. For any given inductor, performance can be perfect only at one frequency. The ideal inductor would match across all frequencies.

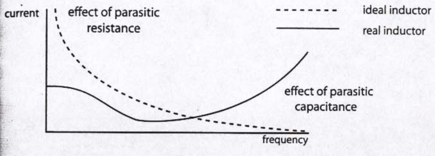


Figure 8–10. *Either the resistance is a problem, or the capacitance is a problem. No inductor is perfect across all frequencies.*

The inductor is said to have a Q at a certain frequency. The Q of an inductor measures how good it is, telling us how low the parasitics are at that frequency.

A Q of 40 is good—very low parasitics.

A Q of 5 is bad—very high parasitics.

How can we reduce parasitics and thereby improve Q?

- Series resistance of the spiral reduces easily. As we mentioned earlier, we use the thickest, lowest resistance metal to make our spiral inductor.
- A wide metal trace also helps improve the Q. Unfortunately, wide metal increases our parasitic capacitance.
- Depending on the process you are using, there may be structures you can place under the spiral to reduce capacitance.

We make many compromises when laying out an inductor. A good knowledge of your process is essential.

Stacked Inductors

If you have enough metals, you can make what they call a **stacked inductor.** Coil your turns in one metal, as discussed above. Bring it out from the center to another metal layer, and spiral another inductor stacked on top of the first, continuing the spiral in the same direction.

Stacking inductors produces more inductance per unit area. However, the layers create complex interactions. You see some very odd frequency responses.

---

[5] Some claim as proof of higher intelligence: alien visitations. This argument was dismissed when archeologists located a cave drawing with a shorted circuit.

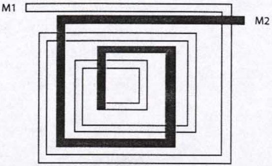


Figure 8–11. Continuing the spiral in another layer stacked on top.[6]

So, unless they are modeled and characterized very well, you may not be able to actually use stacked inductors. Just stick with the existing, accurately measured spirals, rather than experimenting with new stack designs, flying on hope.

Inductor modeling has been a very difficult job. In the past, the only way to get accurate information about inductors was to build them and measure them. Once you knew what you had built and how it reacted over various frequency ranges, you could use those results in your next design.

However, each inductor is used differently than how it was used when characterized. Circuits using stacked inductors typically do not work as well as expected. For example, just laying a wire slightly too close to an inductor can affect the inductor's value.

Try laying out some wiring right up against an inductor and see the color drain from the face of your circuit designer. (Buy the book "101 Fun Tricks to Play on Your Circuit Designer.")

Some companies might give you a library of inductors, and say, "This is what you use. Don't change them." In that case, that is all you use.

---

[6] The Q of one spiral continues to the next. You have found the Q continuum.

Other companies might give you some parameterizable inductor cells. You enter the inductance. The program calculates what the inductance and the Q could be, based on some model someone has put together.[7]

RF Choke

Inductors are mainly used in very high frequency circuits, either as matching circuits or as an **RF choke.** If you want to connect a very high frequency circuit with a very low frequency circuit without allowing the high frequency through, you may place an inductor in the way. The low frequency signals will pass through the inductor, but the high frequency will be stopped, or choked off. Hence the name RF choke.

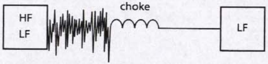


Figure 8–12. Low frequency seems to have an intergalactic Multi-Pass. It is allowed through any inductor at any airport. The high frequency is choked at the inductor.

You can also use spiral inductors to make on-chip transformers. To make on-chip transformers, wind two parallel wires into a single spiral.

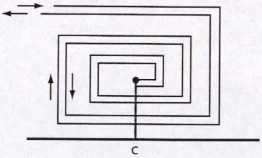


Figure 8–13. Making a transformer from connected spiral inductor coils.

---

[7] Notice, I said *could* be. The science of engineering is not that far along yet, to know for sure in every case. More work is yet to be done here, as in many areas of electronics. But that's why you're here.

This type of transformer uses the following schematic. Notice the connection point at the bottom of the schematic.

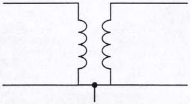


Figure 8–14. Schematic for our connected spirals.

Another transformer option is two totally independent spiral inductors, interwound with each other. This isolates one piece of circuitry from the other.

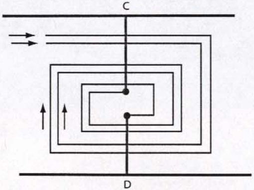


Figure 8–15. Inductors spiraling in the same direction, not connected.

Notice in this schematic we do not see any connection between the two inductor circuits.

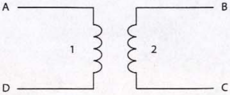


Figure 8–16. Schematic for same direction double spiral.

There you have a variety of spiral designs and their layouts. Next let's examine why your circuit designer's face went pale when you laid those wires right up next to your inductor.

Proximity Effect

Keep all wires away from inductors. A wire running close to an inductor will affect the value of the inductor. All of your circuit designer's excellent planning could be laid to waste by some nearby wires.

Many rules of thumb dictate how close you can place wires to inductors. Some people say to keep them 5 line widths apart. Let's use that as a fair starting point.

■ **Rule of Thumb: Place wires at least five line widths from inductors.**

For example, if the line of an inductor is 10 microns, then put wires no closer to the inductor than 50 microns.

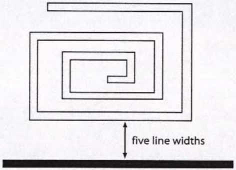


Figure 8–17. Distance your wires from your inductors.

However, distancing our wires from our inductors eats up even more room on the chip. So, discuss this with your circuit designer. Ask how nervous he is putting the particular wire in question close to that certain inductor. Ask what he is doing with the circuit, whether you should care about the distance or not in this case. Check to see whether the inductor has been characterized well enough.

Always communicate with your circuit designer. Do not just use a rule of thumb, like 5 widths, just because you read it in the Saint Book. Everything depends on chip size, what you have been told to do, how experienced your designer is, and other factors.

In general, try to keep wires as far away from inductors as possible. With companies wanting to make more profit, you will feel the big push to keep chip size down. Companies keep driving chip elements closer, and closer, and closer. Sometimes they are prepared to take the hit in case the line affects the conductor. They build in a contingency. They will just see what happens. But they put in other compensating devices, knowing the wire will affect the circuit.

As a layout designer, you can begin with rules of thumb, but remember to work with your circuit people from the beginning of the layout process with these types of questions.

Inductance is everywhere on an integrated circuit. Every wire has some inductance associated with it.

■ **Rule of Thumb: Inductance is everywhere in an IC. Worry most about power supply wires.**

If you are not careful with inductance, your layout choices may cause the chip to fail. Pay particular attention to power supply wires. There is usually some reasonably high current flowing in them. At high signal frequencies, the parasitic inductance can particularly hurt, so worry most about your power supply wires.

Closure on Inductance

Microprocessor people are starting to worry about all these effects. Design is more crucial in high frequency than low or audio frequency. They have not had to consider these effects before. The wireless people and the microwave people have known about all these effects for years. The digital world is starting to become more analog, and the analog world is starting to take on more digital functions.

Digital designers did not have to know anything about analog techniques before. Now they have to know a bit more about analog circuits, propagation delays, frequency responses, and trying to transmit power. In a CMOS digital chip, the clock signal runs everything. We transmit that clock frequency energy into the heart of the microprocessor at higher and higher frequencies. So, we see the fields start to converge.

The trick with really high frequency layout is to keep it smooth. Keep it flowing. Understand your circuit. Imagine where the energies move around. Make it easy for the energy to flow. Do not draw sharp corners causing signals to hit walls dead on. Place nice, slow bends where you must have bends. Anything abrupt will disrupt the energy flow.

■ **Rule of Thumb: Keep it smooth.**

Sensing high frequency flow is one of those things you just have to be exposed to for a long time before you develop a good, intuitive grasp.

Here's What We've Learned

Here's what you saw in this chapter:

- Electromagnetic force, EMF
- Frequency sensitivity of inductors
- Properties of high frequency transmission
- Corner effects on high frequency current
- Using transmission line and corner-specific mathematical models
- Spiral and stacked inductors
- The Q (circuit resonance)
- High frequency chokes
- Building transformers
- Proximity effect

And more . . .

Glossary

active (diffusion)—A region of an integrated circuit where the thick field oxide has been thinned to allow the creation of semiconductor devices.

active device—A semiconductor device whose electrical response can be tailored to specific characteristics, typically transistors and diodes.

alternating current (AC)—Current that reverses direction in a circuit at a regular rate.

alternating waveform—The cyclic voltage waveform that swings both positively and negatively.

amperes (amps)—Unit of measure for current in the meter-kilogram-second system.

analog—From the word *analogous,* meaning *representation.* The signal remains unchanged, not digitized, as it passes through a circuit.

AND—Logic function defined as one that requires all inputs to be at a logic one in order for the output to be at a logic one. If any input is not at logic one, the function returns a logic zero.

annealing—Process of heating a wafer during semiconductor fabrication to repair the crystal lattice. This helps all atoms loosen and settle with each other, forming a more consistent structure. Byproducts can be a new oxide layer and diffused areas of previously concentrated material.

antenna effect—The phenomenon of charge buildup on polysilicon gates during the polysilicon etch process. If you have too much voltage on polysilicon that is used as a transistor Gate, the Gate oxide could become damaged and kill the transistor.

applied voltage—The voltage that has been applied across two points.

atom—Fundamental component of matter containing a nucleus surrounded by swirling electrons.

Base—Terminal of a bipolar transistor. Used to entice preliminary current from the Emitter. See also Collector, Emitter, transistor.

decoupling—The process of noise reduction on DC voltage supplies, using capacitors, to prevent unwanted signals from passing between circuits.

decoupling capacitor—A capacitor used to eliminate high frequency noise from a supply voltage.

delta—The difference between a drawn dimension in a CAD tool and the real dimension on silicon. In the real world the manufacturing process is not as perfect as your CAD drawing is. A device that is fabricated can very often be significantly smaller or larger than the one you drew. The difference is represented by the Greek letter, delta δ. So, δL, for instance, could be the difference between the dimension of the CAD drawing for the length and the actual processed length.

depletion mode FET—A normally On FET.

design manual—A book containing all the physical and electrical information required to design integrated circuits for a given process. Also called the Process Manual or Rulebook. Typical examples are sheet resistivities, current handling capabilities, reliability information, transistor model information, layout rules, and so on.

dielectric—An insulating material between two conductors.

dielectric constant—A reference number pertaining to the characteristic of a material's ability to store charge. Air has a constant of one.

diffusion—A region of semiconductor containing implanted impurities which has been heat-treated (annealed). The heat treatment causes the impurities to spread in all directions within the semiconductor.

diffusion capacitor—A capacitor whose bottom plate is made from diffusion. Typically a large area, depletion mode, FET. The capacitor is formed using the gate as the top plate, and utilizes the thin gate oxide as the dielectric.

digital—Conversion of signal into a binary logic. A series of pulses.

diode—The junction between P and N materials that restricts current flow chiefly to one direction, used extensively in integrated circuits to isolate devices from each other.

direct current (DC)—An electric current flowing in one direction only. See also alternating current.

dogbone—A term used to describe a resistor that has a contact much larger than the main resistor body. The shape looks like a dogbone. Good shape for "high value, don't care" resistors.

dopants—The materials used in implantation to create N and P Type regions.

doping—The process of introducing dopants into a semiconductor.

double poly process—A semiconductor process using multiple polysilicon layers to create different types of devices. One poly layer might be used to create CMOS transistors while another poly layer might be used specifically for

resistive components. Each poly layer will be deposited at a different stage fabrication. Each poly layer requires a unique mask.

drain—One terminal of a CMOS transistor. May be thought of as int changeable with the source, depending on the polarity of the applied volta across the transistor.

effective Gate width—The total combined width of parallel CMOS transi tors.

electric charge—See charge.

electron—Fundamental particle of an atom. Considered negatively charged. moving electron is current. See current.

electron flow—Opposite direction of conventional current. Reflects actua movement of electrons.

electrostatic discharge (ESD)—A sharp electrical zap in the thousands o volts, usually felt in your finger as you touch a conductive material after build ing up a large number of free electrons in your body. Can destroy a chip. Als called ESD; see also diode, reverse biased.

Electromotive force (EMF)—Electric potential; voltage.

Emitter—A terminal of a Bipolar transistor. This area emits either holes or electrons which are collected at the Collector terminal. See also Base, Emitter, transistor.

enhancement mode FET—A normally Off FET.

epitaxial deposition—The process of growing a new layer of silicon on top of another layer of silicon while maintaining the lattice structure. This is typically performed using a mixture of gases, a specialized form of Chemical Vapor Deposition.

epitaxial layer—A thin semiconductor layer grown on a thicker substrate layer using chemical vapor deposition. Extremely well-controlled, maintains regular crystal structure.

etching—The process of removing material from a semiconductor wafer using reactive chemicals.

evaporative deposition—A process of depositing metal onto a semiconductor wafer. The metal is vaporized then condensed onto the wafer surface in a very low pressure chamber.

Exclusive OR (XOR)—Logic function defined as the OR function with the exclusion of all inputs being at logic one. If no input is at logic one, the function returns a logic zero. If all inputs are at logic one, the function returns a logic zero. Any other state returns a logic one.

extrinsic semiconductor—Doped silicon.

fan out—The number of digital logic Gates that can be driven by the output of a single logic Gate.

farad—The measurement unit for capacitance.

femto—A quantity multiplied by ten raised to the power of −15.

FET—See Field Effect Transistor.

field effect—The phenomenon that occurs in semiconducting material when an applied voltage is placed nearby. Electrons in the material are either attracted or repelled by the voltage, which results in an inversion of the semiconductor material type. For example, N Type silicon can be changed to P Type silicon.

Field Effect Transistor (FET)—A voltage-operated switch built from either N or P Type semiconductor material. Nearby voltage turns on or off electron flow through the semiconductor.

force/sense—Circuit technique that removes any uncertainties due to wiring resistance by providing a nonconducting measurement path. Can be used to determine ohms-per-square for a material, for example. Also known as 4-point measurement method.

forward biased—When a junction is in conduction, it is said to be forward biased. Holes are moving forward across the junction, electrons are moving backward across the junction.

4-point measurement—Circuit technique that removes any uncertainties due to wiring resistance by providing a nonconducting measurement path. Can be used to determine ohms-per-square for a material, for example. Also known as the force/sense method.

frequency—The number of times in one second that an alternating voltage or current completes a cycle, measured in hertz (cycles per second).

fringing capacitance—Extra capacitance created between the edges (fringe areas) of the top plate of a parallel plate capacitor and its bottom plate.

full wave rectification—Method of inverting the negative portion of an alternating waveform into its positive quadrant.

fusing current—The current required to destroy a material within a certain time.

gallium arsenide (GaAs)—Semiconductor material made from the elements gallium and arsenic. Useful for microwave circuits because of its high speed of operation.

Gate—One terminal of a CMOS transistor. The terminal we use to apply our nearby controlling voltage. The voltage on this Gate generates the field effect.

Gate area—The area of the Gate that crosses the active diffusion. Gate length multiplied by Gate width. The device parasitic capacitance is mainly dependent on this area.

Gate tie-down—A small diode in the substrate attached to the poly connection of a transistor Gate. Used as a protection against the antenna effect. The diode will limit how high the voltage can get. Also known as NAC diode.

giga—A quantity multiplied by ten raised to the ninth power.

GND—A signal name for the most negative voltage in a circuit, usually in analog circuitry.

ground—The lowest voltage potential in any circuit. To be *grounded* mea be connected to the ground voltage.

harmonic—A sinusoidal component of a waveform that is a whole mul of the waveform's fundamental frequency, meaning the wave has been m plied by an integer.

head—A portion of a resistor between the resistor body and the resistor tact. Usually made from low ohm-per-square material, whose resistanc sometimes considered as part of the contact resistance.

henry—Unit of inductance. One henry is the inductance required to prod a back EMF of one volt when the current flow through the inductor change the rate of one amp per second.

high (circuit voltage state)—Voltage level representing a logic one state.

Hitchhiker Rule—Method of determining direction of generated magn field in a conductor. Grab the conductor with your right hand, your thu extended in the direction of current flow. Your curled fingers mimic the li of magnetism. This is also called the Right Hand Rule.

hole—A vacancy in the electron structure of a semiconductor crystal latti Can be thought of as carrying a positive charge.

impendance—Effective resistance of a capacitor or inductor at a given f quency, measured in ohms.

implantation—Method of introducing impurities into a semiconductor cry tal lattice. Usually achieved by bombarding a silicon-based wafer with hig energy ions.

impurity—Implanted atoms in a silicon semiconductor crystal lattice that c either increase or decrease the number of available free electrons.

inductance—The ability of a conductor to resist change in current flow, mea ured in henrys.

induction—The process that causes current to flow in one conductor by usin the magnetic field generated by current flowing in a nearby second conducto

inductor—A component whose impedance increases as the frequency of a alternating current flowing through it increases.

insulator—A material that is a poor conductor of electricity or heat.

integrated circuit—An electronic circuit containing active and passiv devices connected together on a semiconductor base wafer.

interdigitated—A method of device placement that alternates elements.

intrinsic semiconductor—Undoped silicon.

invert a region—Changing N Type semiconductor to P Type semiconductor or vice versa, using the field effect.

Inverter—A logical NOT Gate, whose output is the opposite binary state from its input.

I/O—Input and output of a system.

ion—A charged atom that has acquired a charge by gaining or losing one or more electrons.

iterative process—A mathematical technique for calculating a desired result by successively closer approximations.

jumper—1. Connection joining two points electrically. 2. Pullover wooly clothing, used to demonstrate electrostatic discharge.

Junkyard Wars—Cathy Rogers' innovative use of trash in team construction competition, idea originated as result of movie *Apollo 13* throwing box of junk on the table with imminent need for creativity.

kilo—A quantity multiplied by ten to the third power.

Kirchhoff's Current Law—All the various currents leaving a junction should add up to the total current entering the junction.

Kirchhoff's Voltage Law—All voltage drops in a closed circuit should add up to the total voltage applied to the circuit.

latch-up—A phenomenon that occurs when parasitic transistors in the substrate of a semiconductor wafer turn on, causing large destructive currents to flow.

layout—The process of creating the two-dimensional representations of the fabrication layers, which are then used to manufacture an IC chip.

light field—A primarily clear mask, with only small areas of chrome, used in photolithography.

logic function—An expression of the relationship between binary inputs and outputs of a circuit; including AND, OR, NOT, NAND, NOR and XOR (exclusive OR).

logic one—The output or input of a logic gate, usually represented by a high voltage in the circuit.

logic zero—The output or input of a logic gate, usually represented by a low voltage in the circuit.

low (circuit voltage state)—Voltage level representing a logic zero state.

magnetic field—A really spooky phenomenon surrounding a material conducting a current, which is capable of influencing certain other nearby materials. It makes iron filings move in pretty patterns at the Exploratorium.

mask—An extremely flat quartz glass plate containing an image of layer detail etched in chrome.

masking—Coating selected areas of a semiconductor wafer, leaving other areas exposed for diffusion, etching, or metallization.

micro—A quantity multiplied by ten to the −6 power.

micron—Metric unit of linear measure equal to one meter multiplied by ten to the −6 power. That is, one millionth of a meter.

milli—A quantity multiplied by ten to the −3 power.

models—Mathematical representations of the circuit element behaviors. Accurately reflect the physics of the device and its electrical and physical characteristics. A basic device could consume months to model.

multimeter—Handheld device used to measure voltages, currents, and resistances.

mutual inductance—Extra inductance due to the proximal interaction of inductors with each other.

N Type—A semiconductor material doped with impurities that have extra electrons compared to the silicon atom. *N* stands for *Negative,* denoting the resultant charge on the doped material.

N well—A deep N Type diffusion used to build PMOS transistors.

N well contacts—Areas of N+ doped diffusion inside an N well that form Ohmic contacts. Usually tied to the most positive voltage in the circuit to prevent latch-up.

Net area check (NAC) diode—A small diode in the substrate attached to the poly connection of a transistor Gate. Used as a protection against the antenna effect. The diode will limit how high the voltage can get. Also called a Gate tie-down.

NAND—A function which outputs the inverted results of the AND function.

nano—A quantity multiplied by ten to the −9 power.

negative charge—A surplus of electrons.

negative resist—Resist that remains unaffected by exposure to light. Exposed regions remain hard and cannot be removed by the resist developer. So, the shaded regions are what naturally dissolve in our developer. Also see photolithography, photoresist, positive resist.

network—Circuit.

node—A junction of two or more circuit components.

NOR—A function which outputs the inverted results of the OR function.

normally Off—Term used to describe a transistor that does not conduct in its unbiased condition. Also called an Enhancement Mode transistor.

normally On—Term used to describe a transistor that is conductive in its unbiased condition. Also called a Depletion Mode transistor.

off (switch state)—A transistor state in which the transistor is unable to pass current.

ohms—Unit of measure for resistance, *R,* denoted as symbol omega Ω.

Ohm's Law—The relationship between voltage, current and resistance. Ohm's Law states that the current flowing in a resistor is directly proportional to the applied voltage and inversely proportional to the value of the resistor. $V = IR$.

ohms-per-square—Resistive value for one square of material. All squares of the same material in the same process have the same value.

on (switch state)—A transistor state in which the transistor is able to pass current.

OR—Logic function defined as one that requires any of the inputs to be at a logic one in order for the output to return a logic one. If no input is at logic one, the function returns a logic zero.

P Type—A semiconductor material doped with impurities that have fewer electrons compared to the silicon atom. *P* stands for *Positive,* denoting the resultant charge on the doped material.

P well—A deep P Type diffusion used to build NMOS transistors.

parallel—A circuit that can conduct current from a common entry point along two or more paths to a common exit point.

parasitic element—Unwanted capacitance, resistance or inductance that is inherent in all conducting materials and devices.

passive device—An integrated circuit component whose electrical response cannot be changed by external influences, typically resistors, capacitors and inductors.

PECVD—Acronym for Plasma Enhanced Chemical Vapor Deposition.

photolithography—Phrase literally meaning *printing with light.* The process of transferring an image to the surface of a semiconductor base wafer using light shone through a blocking mask. See mask, masking, photoresist.

photoresist—Light-sensitive liquid that is dropped onto a spinning semiconductor base wafer to provide a uniformly thin film that protects the base wafer from subsequent processing steps.

pico—A quantity multiplied by ten to the −12 power.

pinched off—Used to describe the state of a CMOS transistor where the field effect has inverted enough of the conducting region to prevent electron flow.

planarization—Technique to make an uneven surface flat. Can be by etching, grinding or polishing, for example. Allows subsequent processing steps to be more accurate.

plasma—A high-energy, ionized gas. A fourth state of matter, other than solid, liquid or gaseous.

Plasma Enhanced Chemical Vapor Deposition (PECVD)—Very similar to CVD, but instead of high temperatures to start the chemical reaction, a plasma is used instead.

poly-crystalline silicon (poly)—Silicon made from many randomly orier
small crystals as opposed to a silicon ingot, for example, which is one v
large crystal.

polysilicon—Same as poly-crystalline silicon.

positive charge—A lack of electrons.

positive resist—Resist that is affected by exposure to light. Exposed regi
can be dissolved and removed by the resist developer. See also negative res
photolithography, photoresist.

potential—The ability of a charge to do work.

potential difference—Electromotive force. Voltage. The difference betw
two potentials.

potential barrier—A region formed when N Type and P Type semiconduc
materials are joined. The region forms a barrier that inhibits electron flow. S
diode.

power rail—A wide metal wire that delivers voltage and current to regions
circuitry.

process manual—A book containing all the physical and electrical inform
tion required to design integrated circuits for a given process. Also called t
Design Manual or Rulebook. Typical examples are sheet resistivities, curr
handling capabilities, reliability information, transistor model informatio
layout rules, and so on.

propogation time—The time required for a signal to reach one point fro
another, be it through free space, a transmission line, an amplifier, or a logic Ga

proprietary—We can't tell you.

publishing contract—Totally confusing pieces of paper.

Q—A measurement of inductor quality. The higher the number the closer t
inductor is to ideal.

rat's nest—A useful technique used in CAD tools to graphically demonstra
the interconnections between various circuit blocks. Each circuit block h
direct lines drawn from each connection point to every other connection poi
in the circuit, drawn straight, *as the crow flies,* so to speak. The more tangl
the diagram, the harder it will be to wire.

RC time constant—A figure of merit reflecting how quickly a capacitor w
charge to a given voltage through a given resistor, defined as T = RC, resis
ance multiplied by capacitance.

reactive ion etching (RIE)—An etching technique using a plasma that cher
ically reacts with the material to be etched.

real estate—Chip area.

rectifier—A device that converts an alternating waveform into one that
entirely positive.

reliability—An indication of an integrated circuit's expected lifetime.

resist—Same as photoresist.

resistance (value)—A measurement of a material's opposition to current flow. The unit is expressed in ohms. One ohm is defined as the resistance value that will produce one amp of current flow when one volt is applied.

resistivity—The amount of resistance in a given volume of conducting material.

resistor—A device used to control current in an electric circuit by providing resistance.

reverse biased—A description of a PN junction that has a voltage applied in such a polarity that no current flows.

reverse breakdown voltage—A voltage at which a reverse biased junction starts to conduct current.

RF choke—A term used to describe an inductor that is being used to inhibit radio frequency signals from entering unwanted regions of circuitry.

RIE—Short for Reactive Ion Etching.

Right Hand Rule—Method of determining direction of generated magnetic field in a conductor. Grab the conductor with your right hand, your thumb extended in the direction of current flow. Your curled fingers mimic the lines of magnetism. This is also called the Hitchhiker Rule.

rulebook—A book containing all the physical and electrical information required to design integrated circuits for a given process. Also called the Process Manual or Design Manual. Typical examples are sheet resistivities, current handling capabilities, reliability information, transistor model information, layout rules, and so on.

saturation—The level at which increasing one quantity no longer alters a second quantity.

saturation current—As the voltage across a device is increased, so the current flowing through it increases. Once the current ceases to increase further, as the applied voltage continues to increase, the device is said to be in saturation, and we have reached the saturation current for that device. Typically used in conjunction with bipolar and CMOS transistors.

schematic—Diagram containing symbols that represent the electrical components of a circuit and their connectivity. Used for documentation, simulation and layout reference.

sea of electrons—A term used to describe the abundance of conduction band electrons available in a conductor or semiconductor.

seed crystal—A small starter crystal, used in a crystal puller, whose alignment determines the alignment of the entire silicon ingot produced.

self-aligned Gate—A phrase used to describe the technique of aligning source drain regions to the Gate of a CMOS transistor. The Gate stripe of the transistor acts as a mask for the source drain implant process. Subsequent annealing pushes the implant underneath the Gate stripe, reducing the transistor Ga length, thus producing a faster device.

semiconductor—A crystalline material that has a very small differen between its valence band electrons and its conduction band electron energie Electrons can be easily coaxed into conduction. The introduction of impuriti can allow a semiconductor to conduct more easily and more controllably.

series—A circuit containing only one possible path for current flow.

serpentine resistor—A resistor which has been laid out in a snake-like pat in order to facilitate a more compact layout of a lengthy resistor.

sheet resistivity—The amount of resistance in a given surface area of fixe depth conducting material. Also known as sheet rho.

sheet rho—Same as sheet resistivity.

shells—Description of the discrete energy levels in which electrons orbit nucleus.

short (circuit)—An undesired low-resistance path.

simulator—A computer program that predicts how a circuit might functio once built, which saves time and effort otherwise spent making test chip Requires circuit models, schematics, circuit specifications and operatin parameters.

sinusoidal waveform—A signal voltage having the shape of a sine wave.

source—One terminal of a CMOS transistor. May be thought of as interchangeable with the drain, depending on the polarity of the applied voltag across the transistor.

source-drain sharing—A technique used to reduce chip layout area that combines CMOS transistors that have terminals with identical circuit connection

specs—Short for circuit specifications.

SPICE—Simulation Program for Integrated Circuits Emphasis (o Engineering). Circuit simulation programs. The original SPICE was develope by the University of California, Berkeley, in the 1970's. A popular family o computer-aided circuit-analysis programs.

spiral inductor—An inductive wire wound into a square spiral to reduce chi area.

spreading—The phenomenon that becomes problematic when a contact i much smaller than the conductive path. As electrons leave the contact, the spread out toward the entire width of the conductor. This leads to ambiguity i the determination of the length of a resistor, for example.

spurious signal—Unwanted signals of chance or questionable origin.

sputterer—A machine used for depositing metal using high energy plasma t erode a metal target. The eroded metal is then deposited onto our semiconductor base target.

stacked capacitor—Two or more capacitors stacked vertically on top of each other. Typically, this technique is used to obtain higher capacitance values in a small area. Metal capacitors are usually the type used with this approach. See interdigitated.

stacked inductor—Two or more inductors stacked vertically on top of each other.

static (electricity)—The buildup of electric charge on nonconducting material, usually by stripping electrons away from it.

stick diagram—A very simple representation of CMOS devices and their interconnections. The diagram is a halfway sketch between the schematic and the final layout. Once the stick diagram is complete, it gives you the device placement and connections.

stochastic—Relying on trial and error or probability for a solution; as opposed to algorithmic.

substrate—The bottom layer of raw silicon. The base upon which a transistor or IC is fabricated.

substrate contacts—P+ connections to the silicon base wafer, used to tie the substrate to the lowest circuit voltage in order to ensure device operation and prevent latch-up.

substrate material—Same as substrate. Usually P−.

substrate ties—Same as substrate contacts.

superhero—Person with blue uniform, red cape and a big S on the chest.

supply—The highest voltage potential in any circuit.

time constant—A figure of merit reflecting how quickly a capacitor will charge to a given voltage through a given resistor, defined as $T = \text{RC}$. Same as RC time constant.

transistor—Solid state switch built using semiconductor materials. Comes from the words transfer and resistor (*TRANS*fer res*ISTOR*). An active semiconductor device having three electrodes.

transmission line—Characterized wire for parasitic capacitance, inductance and resistance, used as a channel to carry signal from one end to the other.

tub—Same as N well.

underpass—The use of polysilicon to enable wiring connections that are not possible using metal.

valence band—Region of an energy diagram that designates electrons without enough energy to conduct. Electrons in the valence band hold a substance together.

varactor diode—A diode that has a highly variable junction capacitance, that depends on the voltage you put across it. The dopants have been specially chosen to enhance this variable capacitance feature.

VCC—A signal name for the most positive voltage in your circuit, usually used in analog circuitry.

VDD—A signal name for the most positive voltage in your circuit, usually used in digital circuitry.

vertical device—A semiconductor device built using diffusion layers stacked on top of each other.

vertical processing—Steps used to make a vertical device. Vertical device processing techniques allow more accuracy in the construction of Bipolar transistors.

via—A hole in an insulating dielectric that allows connection from one metal layer of an integrated circuit to another.

voltage—The potential difference between charges across two points.

volt—Unit of measure for voltage, denoted as symbol V. One volt is defined as the potential difference that will cause a one-amp current to flow in a one-ohm resistor.

VSS—A signal name for the most negative voltage in your circuit, usually used in digital circuitry.

wafer—Thin, round slice of semiconductor base material cut from a single crystal ingot, sort of like a dinner plate.

well diode—A PN junction formed between an N well and the P− substrate. Can be used as protection against ESD.

well contact—N+ connection used to tie the N well to the most positive circuit voltage in order to ensure device operation and prevent latch-up.

well ties—Same as well contacts.

x-ray lithography—Photolithography using x-rays instead of visible light. X-rays are smaller, and therefore can be used to create finer detail.

Yterbium—Strangely named element. We needed a y word.

zymurgy—A good last word. Look it up.

Suggested Readings and Resources

Electronics Pocket Handbook
by Daniel L. Metzger
Paperback—329 pages, 3rd edition (1998)
Prentice Hall PTR; ISBN 0137841906

CMOS Circuit Design, Layout, and Simulation
IEEE Press Series on Microelectronic Systems
by R. Jacob Baker, Harry W. Li, David E. Boyce
Hardcover—928 pages (January 1998)
IEEE; ISBN 0780334167

CMOS IC Layout: Concepts, Methodologies, and Tools
by Dan Clein
Paperback—304 pages, book & CD-ROM edition (December 1999)
Butterworth-Heinemann; ISBN 0750671947

Design of Analog CMOS Integrated Circuits
McGraw-Hill Series in Electrical and Computer Engineering.
Electronics and VLSI Circuits
by Behzad Razavi
Hardcover—600 pages, 1st edition (August 1999)
McGraw-Hill; ISBN 0072372710

The Art of Analog Layout
by Alan Hastings
Hardcover—576 pages, 1st edition (December 15, 2000)
Prentice Hall; ISBN 0130870617

The Art of Electronics
by Paul Horowitz, Winfield Hill
Hardcover—1125 pages, 2nd edition 1989)
Cambridge University Press; ISBN 0521370957

Microchip Fabrication: A Practical Guide to Semiconductor Processing
by Peter Van Zant
Hardcover—640 pages, 4th edition (April 3, 2000)
McGraw-Hill; ISBN 0071356363

Analysis and Design of Analog Integrated Circuits
by Paul R. Gray, Robert G. Meyer
Hardcover—792 pages, 3rd edition (January 15, 1993)
John Wiley & Sons; ISBN 0471574953

Physical Design of CMOS Integrated Circuits Using L-Edit™
by John P. Uyemura
Hardcover—256 pages, 1st edition (October 5, 1994)
ITP Educational Division; ISBN 0534943271

Schaum's Outlines: Basic Electricity
by Milton Gussow
Paperback—454 pages (1983)
McGraw-Hill; ISBN 0070252408

Educational Programs

Eastfield College
Eastfield College is located in Mesquite, Texas, just outside of Dallas. They are a two-year college and offer an Associate Degree in Computer Aided Design and Drafting— Integrated Circuit Design.

Austin Community College
Austin Community College is a two-year college located in Austin, Texas. They offer an Associate of Applied Science degree with an IC Certificate. The two IC layout courses in the certificate are Integrated Circuit Layout and Design I, and Integrated Circuit Layout and Design II.

Mesa Community College
Mesa Community College is located in Mesa, Arizona. They have two classes for Integrated Circuit Design. The first is CMOS 1A , and the second class is CMOS 1B.

San Jose City College
San Jose City College is located in San Jose, California, also known as Silicon Valley. They offer CMOS Mask Design I and II courses.

American River College
American River College is a two-year college located in Sacramento, California. Courses include CMOS Design I and II.

Chemeketa Community College
Chemeketa is a dynamic, comprehensive educational institution located in the heart of the Willamette Valley in Salem, Oregon. They offer an Integrated Circuit Mask Design program.

Silicon Artists
Silicon Artists offers courses in the field of IC Mask design— basic and advanced level classes. Each course is taught by an experienced industry professional, in such a manner that does not require an Electrical

Engineering degree. Silicon Artists offers both IC Mask Design and Advanced IC Mask Design.

Feel free to inform us of other educational programs, so that we may include them in future editions.

Index

About the Authors

Christopher Saint is currently Manager of the IBM West Coast Physical Design Group. He served as lead layout engineer for the Commquest GSM, AMPS, and CDMA chipsets, as well as spent many years with the Analog Devices, LSI Logic, and GEC/Plessey semiconductor layout design groups.

Judy Saint is a professional writer and illustrator, with 11 years of successful teaching experience. She provided and directed the motivational aspects of this book, keeping the information easy to read and understand.