Understanding SIGNAL INTEGRITY

Understanding Signal Integrity

Library of Congress Cataloging-in-Publication Data
A catalog record for this book is available from the U.S. Library of Congress.

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library.

ISBN-13: 978-1-59693-981-3

Cover design by Vicki Kane

© 2011 Artech House
685 Canton Street
Norwood, MA 02062

All rights reserved. Printed and bound in the United States of America. No part of this book may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system, without permission in writing from the publisher.

All terms mentioned in this book that are known to be trademarks or service marks have been appropriately capitalized. Artech House cannot attest to the accuracy of this information. Use of a term in this book should not be regarded as affecting the validity of any trademark or service mark.

10 9 8 7 6 5 4 3 2 1

Understanding Signal Integrity

Stephen C. Thierauf

ARTECH
HOUSE

BOSTON | LONDON
artechhouse.com

To Christopher, Kevin, and, as always, Ann

Contents

Preface xv

CHAPTER 1

Introduction to Signal Integrity 1

1.1 What Is Signal Integrity? 1
1.2 The Importance of Signal Integrity: The First Transatlantic Telegraph Cable 1
1.3 What Is a Pulse? 3
1.4 Time and Frequency Domains 4
1.4.1 Line Spectrums 4
1.4.2 Recreating a Pulse with Sine Waves 4
1.4.3 Why Does the Frequency Domain Matter to Signal Integrity Engineers? 5
1.4.4 Upper Bandwidth 7
1.5 Power Integrity and Simultaneous Switching Noise 8
1.5.1 What Is Simultaneous Switching Noise? 8
1.6 What Is the Dielectric Constant and Loss Tangent? 9
1.7 Main Points 10
References 10

CHAPTER 2

The Signal Integrity Process 13

2.1 Introduction 13
2.2 Why Perform a Signal Integrity Analysis? 13
2.3 What Is a Typical Signal Integrity Workflow? 14
2.4 Signal Integrity Worst-Case Analysis 16
2.4.1 Silicon Worst-Case I/O Models 17
2.4.2 What Combination of Environmental Effects Are Worst Case? 17
2.4.3 Using SI Analysis Results During Debug 18
2.4.4 Electrical Overstresses 19
2.5 Main Points 19
References 20

CHAPTER 3
Signal Integrity CAD and Models 21
3.1 Introduction 21
3.2 I/O Models 21
3.2.1 What Are Transistor Level Models? 21
3.2.2 What Are IBIS Models? 22
3.3 Modeling Transmission Lines 23
3.4 S-Parameters 24
3.4.1 What Are S-Parameters? 24
3.4.2 What Is Insertion Loss? 24
3.4.3 What Is Reflection Loss? 26
3.4.4 What Are Touchstone Files? 26
3.5 Field Solvers 26
3.5.1 What Are 2D Field Solvers? 26
3.5.2 What Are 3D Field Solvers? 27
3.6 Main Points 28
References 28

CHAPTER 4
Printed Test and Evaluation Boards 31
4.1 Introduction 31
4.2 Test Boards 31
4.2.1 Test Boards as a Chip Debug Platform 32
4.2.2 Margining Power Supply 33
4.2.3 Testing Decoupling Capacitor Efficacy 33
4.2.4 Testing Sensitivity to Clock Quality 33
4.3 Evaluation Boards 33
4.4 Roll of the Printed Circuit Layout Shop 34
4.5 Fabricating and Assembling the Board 35
4.6 Alternative Methods for Design and Fabrication 35
4.7 Main Points 36
References 36

CHAPTER 5
Printed Circuit Board Construction 37
5.1 Introduction 37
5.2 How Is a Multilayer Circuit Board Constructed? 37
5.2.1 How Are Connections Made Between Layers? 37
5.2.2 What Is the Via Aspect Ratio? 39
5.3 What Is the Significance of Calling a Circuit Board "FR4"? 39
5.3.1 Why Is There Such Variability in *Dk* Between Laminate Manufacturers? 40
5.3.2 Why Is the Same Board Produced Differently by Various Fab Shops? 41
5.3.3 What Are Some Alternatives to FR4? 41

5.4 Circuit Board Traces 42
5.4.1 Microstrip 43
5.4.2 Stripline 43
5.4.3 What Is a Mil and What Does a Trace Thickness in Ounces Mean? 43
5.4.4 What Copper Types Are Used to Form Traces and What Are Their Characteristics? 45
5.4.5 What Is Surface Roughness and What Are Typical Values? 45
5.4.6 Are Traces Rectangular? 46
5.5 Main Points 46
Problems 47
References 48

CHAPTER 6

Transmission Line Fundamentals 49

6.1 Introduction 49
6.2 What Is a Transmission Line? 49
6.2.1 What Is the Signal Return Path? 49
6.3 Circuit Model of a Transmission Line 50
6.4 Impedance and Delay 51
6.4.1 Units and Electrical Length 52
6.5 How Does a Signal Travel Down the Line? 52
6.5.1 What Voltage Gets Launched Down a Transmission Line? 54
6.6 How Does the Current in the Return Path Behave? 55
6.6.1 How Does Current Flow in Return Paths That Are Not Ground? 55
6.6.2 What Is the Behavior When One Return Path Is a Related Power Plane? 56
6.6.3 How Does the Return Current Flow When the Power Supply Is Not the Signaling Voltage? 57
6.7 When Does a Conductor, Via, or Connector Pin Act Like a Transmission Line? 58
6.7.1 Why Does the Signal Rise Time Matter But Not the Pulse Width? 59
6.7.2 Why Does Propagation Delay Time Depend on the Line Length, But Impedance Does Not? 59
6.8 Main Points 59
Problems 60
References 61

CHAPTER 7

Understanding Microstrip and Stripline Transmission Lines 63

7.1 Introduction 63
7.2 DC Resistance and Resistivity 63
7.3 Stripline and Microstrip Capacitance 64
7.3.1 What Is the Effective Dielectric Constant? 65
7.3.2 What Are the Differences Between Microstrip and Stripline Capacitances? 66

7.3.3 How Does Frequency Alter the Loss Tangent? 67
7.3.4 What Is Conductance and How Is It Determined? 67
7.4 Stripline and Microstrip Inductance 68
7.5 How Does the Skin Effect Change the Trace's Resistance? 69
7.5.1 What Is the Proximity Effect? 69
7.6 How Does the Skin Effect Change the Inductance? 71
7.7 Understanding Stripline and Microstrip Impedance 71
7.8 Understanding Stripline and Microstrip Propagation Delay Time 73
7.8.1 How Do the Dielectric and Effective Dielectric Constants Affect Delay Time? 73
7.8.2 How Sensitive Is Delay Time to Changes in the Dielectric Constant? 74
7.9 Main Points 74
Problems 76
References 76

CHAPTER 8

Signal Loss and the Effects of Circuit Board Physical Factors 79

8.1 Introduction 79
8.2 What Are Transmission Line Loss and Attenuation? 79
8.2.1 What Effects Do Losses Have on Signals? 79
8.2.2 How Is Loss Specified? 80
8.2.3 Converting Between Decibels and Percentage Loss 81
8.3 What Creates Transmission Line Loss and How Is It Characterized? 81
8.3.1 Loss Differences Between Microstrip and Stripline Traces 82
8.4 How Do Impedance and Loop Resistance Affect Conductor Loss? 83
8.4.1 How Does Surface Roughness Increase Conductor Loss? 84
8.4.2 Surface Roughness in Perspective 85
8.4.3 What Can Be Done to Reduce Conductor Losses? 85
8.5 Understanding Dielectric Losses 86
8.5.1 Differences in Dielectric Losses Between Stripline and Microstrip 87
8.5.2 Effects of Temperature and Moisture 87
8.5.3 What Can Be Done to Reduce Dielectric Losses? 88
8.6 Summary of Signal Loss and Distortion Characteristics 88
8.7 Effects of Resin Content and Glass Weave on Delay Time and Impedance 89
8.8 Effects of Trace Shape 89
8.9 Main Points 89
Problems 90
References 91

CHAPTER 9

Understanding Trace-to-Trace Coupling 93

9.1 Introduction 93
9.2 Understanding Mutual Capacitance and Inductance 93
9.2.1 What Are the Capacitance and Inductance Matrices? 93

9.2.2 Why Do Some Field Solvers Report Capacitances with Negative Numbers? 95
9.3 How Does Switching Alter the Trace Inductance and Capacitance? 95
9.3.1 How Is the Capacitance Affected? 96
9.3.2 How Is the Inductance Affected? 97
9.4 What Effect Does Coupling Have on Impedance and Delay? 98
9.5 What Are Odd and Even Modes? 99
9.5.1 Can the Odd- and Even-Mode Equations Be Used with More Than Two Traces? 101
9.5.2 Why Is the Even- and Odd-Mode Timing the Same for Stripline But Not Microstrip? 101
9.5.3 By How Much Does the Even- and Odd-Mode Impedance Change? 102
9.6 What Are the Circuit Effects When Switching Causes the Impedance to Change? 103
9.7 How Is Receiver Timing Affected by In- and Out-of-Phase Switching? 103
9.8 Main Points 105
Problems 105
References 106

CHAPTER 10

Understanding Crosstalk 107

10.1 Introduction 107
10.2 How Is Crosstalk Created and What Are Its Characteristics? 107
10.3 Far-End Crosstalk 109
10.3.1 Calculating FEXT 109
10.3.2 What Are Typical *Kf* Values? 111
10.3.3 FEXT Summary 112
10.4 Near-End Crosstalk 112
10.4.1 Calculating NEXT 112
10.4.2 What Are Typical *Kb* Values? 114
10.4.3 NEXT Summary 115
10.5 How Closely Do Calculation and Simulation Agree? 115
10.5.1 How Well Do 2D Field Solvers Agree? 116
10.6 Guard Traces 116
10.6.1 Connecting the Guard Trace 117
10.7 Main Points 117
Problems 119
References 121

CHAPTER 11

Understanding Signal Reflections 123

11.1 Introduction 123
11.2 How Are Reflections Created? 123
11.2.1 A Physical Analogy for Transmission Line Reflections 126
11.3 The Reflection Coefficient 126

11.3.1 What Determines the Reflection Coefficient Value? 126
11.4 How Do the Reflected Waves Combine with the Incident Waves? 127
11.4.1 Difference Between the Response of a Pulse and a Step 127
11.5 What Is the Behavior When There Are Multiple Reflections? 128
11.5.1 What Is the Behavior When There Are Many Pulses? 132
11.6 Reactive Discontinuities 132
11.7 Main Points 134
Problems 135
References 135

CHAPTER 12

Termination Strategies 137

12.1 Introduction 137
12.2 Source Series Termination 137
12.2.1 Selecting the Resistance Value 137
12.2.2 Effect of Driver Resistance 139
12.2.3 Power Savings 139
12.3 Parallel and Thevenin Termination 139
12.3.1 Selecting V_{tt} and V_{term} 141
12.3.2 Thevenin Termination 142
12.3.3 Connecting R_{term} Directly to V_{dd} or Ground 143
12.4 Diode Termination 143
12.4.1 Diode Types 144
12.5 AC Termination 145
12.6 Topologies 146
12.6.1 Single-Load Point-to-Point Termination 146
12.6.2 Multiple Load Point-to-Point Termination 148
12.7 Multidrop Lines 150
12.7.1 Response When Signal Rise Time Is Very Short 151
12.7.2 Response When Signal Rise Time Is Comparable to the Transmission Line Delays 152
12.8 Stubs and Branches 154
12.8.1 Branches 155
12.9 Main Points 157
Problems 158
References 159

CHAPTER 13

Differential Signaling 161

13.1 Introduction 161
13.2 What Are the Electrical Characteristics of Differential Signaling? 162
13.2.1 How Is a Differential Pair Observed? 163
13.2.2 What Is an Eye Diagram? 163
13.2.3 What Is Intersymbol Interference? 164
13.2.4 How Are the Effects of Loss Corrected? 165

13.3 What Are the Electrical Characteristics of a Differential Transmission Line? 165
13.3.1 Why Is the Differential Impedance Twice the Odd-Mode Impedance? 166
13.4 How Are Differential Transmission Lines Terminated? 169
13.4.1 How Should a Diff-Pair Be Terminated When the Propagation Is Not Fully Odd Mode? 170
13.4.2 What Prevents Creating Odd-Mode Signals? 172
13.5 How Are Differential Transmission Lines Created? 172
13.5.1 Manufacturing Trade-Offs Between Loosely and Tightly Coupled Pairs 173
13.5.2 Electrical Trade-Offs Between Loosely and Tightly Coupled Pairs 174
13.6 What Are Some Suggested Diff-Pair Layout and Routing Rules? 174
13.6.1 Obtaining the Proper Differential Impedance 175
13.6.2 Signal Trace Between Diff-Pairs 176
13.6.3 Keeping Signal Traces Far Away from Diff-Pairs 176
13.6.4 Avoid Return Path Splits 177
13.6.5 Route on Same Layers 177
13.6.6 Minimize Layer Hopping 178
13.6.7 Use Differential Vias 178
13.6.8 Match Overall Trace Lengths 178
13.6.9 Match Lengths on Each Layer 178
13.6.10 Ensure That Noise Is Equally Coupled to Both Traces 179
13.7 Main Points 181
Problems 181
References 182

CHAPTER 14

Trace and Via Artwork Considerations for Signal Integrity 185

14.1 Introduction 185
14.2 Mitered Corners 185
14.3 Routing Near the Board Edge 187
14.4 Serpentine Traces 188
14.4.1 How Are Serpentine Delay Lines Modeled? 190
14.4.2 Effects of Corners 190
14.4.3 How Long Should the Segments Be? 190
14.4.4 Guard Traces 191
14.5 Vias 191
14.5.1 Via Circuit Model 191
14.5.2 What Factors Determine Via Capacitance? 192
14.5.3 What Factors Determine Via Inductance? 193
14.5.4 What Are the Effects of Different Antipad Diameters? 193
14.5.5 What Are Nonfunctional Pads? 193
14.5.6 Can the Via Impedance Be Adjusted? 194
14.5.7 What Are Differential Vias? 194

14.6 Main Points 195
References 195

CHAPTER 15
Identifying Common Signal Integrity Problems 199
15.1 Introduction 199
15.2 Questions to Ask When Reviewing PCB Stackup 199
15.3 Questions to Ask When Reviewing Crosstalk Simulations 201
15.4 Questions to Ask When Reviewing Signal Quality Simulations 202
15.5 Questions to Ask When Reviewing Prelayout Simulations 204
15.6 Questions to Ask When Reviewing Postlayout Simulations 205
15.7 Questions to Ask When Reviewing a Termination Strategy 206
15.8 Questions to Ask When Reviewing PCB Layout Design Rules 206
15.9 Questions to Ask When Reviewing ASIC Driver Selection Choices 206
References 207

CHAPTER 16
Solving Common Signal Integrity Problems 209
16.1 Introduction 209
16.2 Reducing Crosstalk 209
16.3 Reducing Reflections 210
16.4 Reducing and Eliminating Simultaneous Switching Noise (SSN) 211
16.5 Improving Inadequate Timing Margins 211
16.6 Correcting Intersymbol Interference (ISI) 212
16.7 Steps to Take When Circuit Board Traces Are Too Lossy 213
16.8 Options to Reduce Circuit Board Thickness 213
16.9 Steps to Take When There Are Not Enough Routing Layers 214
16.10 Steps to Take When a Circuit Board Must Be Cost Reduced 215
16.11 Steps to Take When Data-Dependent Errors Are Causing System Failures 216
References 216

CHAPTER 17
Calculating Trace and Plane Electrical Values 217
17.1 Introduction 217
17.2 How to Convert from Decibel Loss to Signal Swing 217
17.3 Estimating DC Resistance 218
17.4 Finding Inductance and Capacitance When the Physical Dimensions Are Not Known 219
17.5 Finding Stripline Inductance and Capacitance When Impedance Is Known 220
17.6 Finding Stripline Inductance and Capacitance When Trace Geometry Is Known 220
17.7 Finding Microstrip Inductance and Capacitance When Trace Geometry Is Known 221

17.7.1 Microstrip Inductance 221
17.7.2 Microstrip Capacitance 221
17.8 Estimating Inductance and Capacitance of a Plane 222
17.8.1 Inductance 222
17.8.2 Capacitance 223
17.9 Calculating Trace Loss from a Circuit Model 223
17.10 Calculating Stripline Loss from Trace Dimensions 224
17.10.1 Dielectric Loss 224
17.11 Calculating Microstrip Loss from Trace Dimensions 225
17.11.1 Dielectric Loss 225
17.12 Finding Stripline Impedance 226
17.13 Finding Exposed Microstrip Impedance 226
17.14 Finding Solder Mask Covered Microstrip Impedance 227
17.15 Wavelength 228
References 229

About the Author 231

Index 233

Preface

This is an introductory text for engineers, project leaders, and managers wishing to understand the fundamentals of signal integrity.

The electrical characteristics of the various types of transmission lines present on circuit boards are covered in depth, with particular attention to the differences between them. Also covered in detail are reflections, termination strategies, crosstalk, and information regarding circuit board technology. Signal loss and differential signaling are presented in enough detail to provide the reader with a solid background in these important topics.

Because the focus is on those new to signal integrity (including managers), chapters covering the signal integrity process, CAD tools, the value of test and evaluation circuit boards, and troubleshooting strategies are presented. This material is unique and is intended to provide the newcomer with an overview and the manager with an understanding of the tasks that signal integrity engineers routinely carry out. The need to perform some of these tasks (such as model validation or the debugging of individual circuit models) and the uncertainty in the results from simulation are often unknown to management and to those new to the field and are not always properly accounted for in project schedules.

Exercise problems with detailed solutions allow readers to test their knowledge. These problems can be used for practice or as templates to solve similar problems on actual hardware. Several of the problems require information from other chapters and serve to further knit together various fundamental concepts. The solutions are available on the Artech House Web site at http://www.artechhouse.com/static/reslib/thierauf/thierauf1.html.

The reader is assumed to have a basic background in electrical engineering, including an elementary understanding of voltage, current, resistance, capacitance, and inductance. No knowledge of signal integrity, electromagnetics, or transmission line theory is assumed. These subjects are introduced in the text as needed, often with graphs. The formulas are kept simple in those cases where mathematics is used, and to avoid cluttering the text, mathematical derivations have intentionally been omitted. References are provided for those wishing to understand the derivations, or the underlying physics. A catalog of particularly useful formulas has been collected into one chapter at the end of the book for easy reference and to avoid defocusing the earlier chapters. Those uninterested in using formulas can skip that chapter, but those wishing to check the results from field and circuit simulations,

or those wanting to use mathematics to validate their intuition, will find it a helpful resource.

This book has greatly benefited thanks to the generous help from the following:

- Simucad Design Automation for the SmartSpice license necessary to create a number of the circuit simulations appearing in this book;
- Simberian, Inc. for the 3D field solver SIMBEOR used to create and validate all of the 3D circuit models used in the book;
- Polar Instruments, Inc. for the Si9000 2D field solver that generated a number of the resistance, capacitance, and inductance results;
- C&H Technical Sales and TTM Technologies for providing the trace photos appearing in Chapters 5 and 8;
- Tektronix, Inc. for providing the test equipment photograph appearing in Chapter 13;
- The AWR Corporation for the use of AWR® Microwave Office®;
- Laraine Worby of the Framingham Public Library for obtaining many of the obscure references cited;
- The libraries of Wentworth Institute of Technology, Franklin W. Olin College of Engineering, and Worcester Polytechnic Institute.

Once again, the enthusiastic assistance of the Artech House staff and the skillful review provided by the anonymous reviewer are most appreciated in helping to turn a raw manuscript into a finished text.

I also thank my friend and colleague Jeff Cooper for calling to my attention the ageless story of the first transatlantic cable mentioned in Chapter 1 and pointing out its relevance to the problems faced by modern signal engineers and managers.

Most importantly, I am especially grateful to my wife Ann. Without her support, understanding, and encouragement, this book would not have been possible.

CHAPTER 1

Introduction to Signal Integrity

1.1 What Is Signal Integrity?

Signal integrity is the analysis, design, and validation of the interconnect necessary for successful transmission of digital signals.

To do this, signal integrity (SI) engineering borrows ideas and practices from many branches of electrical engineering. To those not having a background in analog or radio frequency electronics, this diversity can make SI seem mysterious, sometimes inconsistent, and occasionally difficult to understand. Logic designers accustomed to working with binary logic can find the apparently imprecise, analog nature of SI particularly frustrating.

Although knowledge in RF and analog electronics is helpful, modern SI CAD allows nearly anyone to create SI simulations and make predictions about circuit operation. However, as Henry Petroski points out [1], in general CAD software vendors do not always advertise or may not fully understand the limitations of their products. When expertly used, CAD is powerful, but its misuse can have dire consequences.

The SI engineer and technical manager working with pulse data transmission can avoid pitfalls like this by understanding the fundamental ideas described in this book.

1.2 The Importance of Signal Integrity: The First Transatlantic Telegraph Cable

Managers and engineers not having a fundamental understanding of the technical principles governing a project are more likely to make serious technical mistakes, especially when pushing the state of the art. Perhaps the first time this became apparent with regards to a significant high-speed signaling project occurred on August 16, 1858. That was when Great Britain's Queen Victoria sent a welcoming message to U.S. President Buchanan over the first successfully laid transatlantic telegraph cable [2–4]. This 3,800-km (about 2,400-mile) length of submarine cable connected

Ireland with Newfoundland and could not be financially viable unless it operated at speeds similar to dry land telegraph cables. Unfortunately, signal integrity problems caused the Queen's 99-word inaugural message to take nearly 17 hours for transmission instead of several minutes as had been expected [2].

The roots of this nineteenth-century disaster are both managerial and technical. The underlying management fault was that the investors and managers did not possess a rudimentary understanding of telegraphy. Because of this they hired a chief electrician lacking a thorough background in electrical engineering or telegraphy theory. This decision was at the root of the cable's technical flaws. Without this knowledge the chief electrician was unable to understand the latest technical developments and instead relied on his intuition to design the cable and instrument set. These intuitive designs were appealing to management but inappropriate for the task.

Some outside experts (most notably Professor William Thompson, later to be Lord Kelvin) had questioned the cable's electrical design and the electrical equipment to be used with it. With his theoretical background, Thompson knew that the electrical characteristics of a long submarine cable were significantly different from the open air cable that had earlier been used in proof-of-concept experiments, and he correctly believed that the submarine cable would distort high-speed Morse code pulses. He preferred another cable design and suggested transmitting and receiving equipment with very different electrical characteristics than the equipment designed by the chief electrician. Because of management's weak technical background, they were unable to ask the kinds of probing, focused questions needed to resolve the technical debate between Thompson and the chief electrician.

After several false starts, the intuitively designed but electrically flawed cable was finally laid, and direct current from a battery in Ireland was detected by instruments in Newfoundland. This low-frequency test seemed to vindicate the design, but it was soon discovered that high-speed pulses could not be properly received. In fact, signaling was only possible at speeds that were far too slow to be profitable.

Management was unable to properly judge the merits of the various approaches to debug suggested by staff. Instead of using the sensitive transmitting and receiving instruments purposed by Thompson (which we now know would have significantly improved the signaling speed), they authorized the chief electrician to increase the voltage of the transmitted pulses. The misguided idea was that the higher voltage would "push" the pulses down the cable with greater force, allowing signaling to occur at high speeds. In fact, this did not improve the signaling rate because the fundamental problem was pulse distortion caused by the cable's large RC delay, not signal strength.

The large RC delay was the critical difference between the dry-land cables used during the proof-of-concept experiments and the submarine cable. The dry-air cables used air as insulation, but the submarine cable used a waterproof insulation that greatly increased the capacitance. We will return to this effect in Section 1.7.

The large RC delay so distorted the pulses that only the slowest of them could be properly received. Modern SI engineers would say that intersymbol interference was created by the way in which the cable attenuated high frequencies (the high-speed Morse pulses) but allowed low frequencies (slow speed pulses, including

the battery's DC) to pass. As we will see in later chapters, this same behavior also occurs when signaling over modern circuit board traces.

During debug the sending voltage was incrementally increased to excessively high values in an unfortunate attempt to improve signaling speed. This destroyed the cable within a month of the Queen's inaugural message. In fact, the cable never sent any commercial telegrams and was a financial failure that set off investigations of financial fraud and technical mismanagement and ruined several careers.

Modern SI engineers and those who manage them can avoid the kinds of systemic problems experienced during the 1858 cable project by learning the ways in which pulses misbehave when traveling along long lines and learning techniques for correcting these types of errors.

We begin by defining a pulse, then examine the time and frequency domains, and describe two of the most important physical constants used in signal integrity work. We conclude by briefly discussing the effects that many drivers simultaneously switching have on power supply noise.

1.3 What Is a Pulse?

A pulse is shown in Figure 1.1. The pulse width in time (*W*, usually nanoseconds or picoseconds) is measured between the two points where the signal crosses half of the signal amplitude ("the 50% points"), while the pulse period (*T*) is measured at the pulse base. The rise time (t_r) is defined as the time required for the pulse to increase from 10% of its steady state value to 90% of that value. The fall time (not shown in the figure) is measured between the same voltage levels on the falling edge. The rise or fall time can be measured in many ways [5, 6], but in this book we will stay with the commonly accepted "10/90" definition shown in the figure.

Under some conditions the pulse's rising or falling edge may exceed and then oscillate around the steady state value. Usually this is undesirable and later chapters describe the transmission line conditions that can cause the overshoot and ringback conditions shown in the figure.

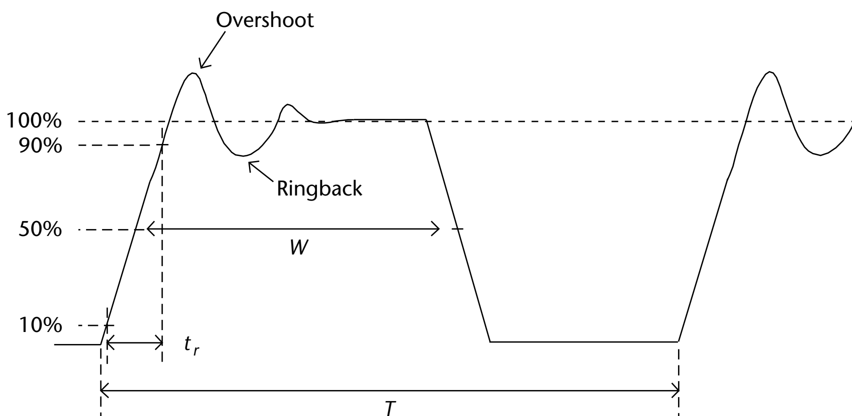


Figure 1.1 Pulse of width *W*, period *T*, and rise time t_r overshoots and then rings back below the steady state value.

1.4 Time and Frequency Domains

Pulses described using time are said to be represented in the time domain. Oscilloscopes are common time-domain instruments, and the SPICE circuit simulator displays its results in the time domain by default.

Most signal integrity work is performed in the time domain, but it is important to understand the way in which the electrical characteristics of circuit board traces change at different frequencies. This frequency domain approach is especially useful when studying how losses on transmission lines affect the shape (distortion) and alignment of pulses (intersymbol interference).

A vector network analyzer (VNA), which displays the impedance of a component or transmission line at various frequencies, and the spectrum analyzer, a device that simultaneously displays the amplitudes across a range of frequencies, are examples of frequency-domain instruments.

1.4.1 Line Spectrums

From signal theory we know that a series of sine waves can be used to recreate a recurring stream of pulses. In fact, a Fourier analysis can be used to find the amplitude and phase of each of the frequencies and determine how many are needed for reconstructing the pulses to the desired level of precision [7, 8].

After doing this analysis, we find that there is a fundamental frequency that has the highest amplitude and possible harmonic frequencies with lesser amplitudes occurring at integer multiples of the fundamental frequency.

The fundamental frequency f_o is found with (1.1) where T is the pulse period.

$$f_o = \frac{1}{T} \tag{1.1}$$

For instance, the fundamental frequency of a 10-ns stream of pulses is 100 MHz. The second harmonic (if present) is twice the fundamental frequency (200 MHz), the third is 300 MHz, and so on indefinitely. The amplitude of each harmonic depends on the duty cycle and rise time of the pulse, and some can have a value of zero (that is, will not be present). None will have a value higher than the fundamental. In fact, (1.1) shows the lowest frequency that will be present in a pulse stream.

For example, a 10-ns square wave (a pulse with the same high and low times) with zero rise time edges has a fundamental frequency of 100 MHz and only the odd harmonics. This is represented in the frequency domain by a discrete line spectrum as shown in Figure 1.2. The amplitude of the fundamental has been set to 1, and the harmonics have been scaled accordingly.

1.4.2 Recreating a Pulse with Sine Waves

The harmonics with the amplitudes shown in the spectrogram can be added together to recreate the pulse stream. In theory an infinite number of harmonics may be necessary, but far fewer suffice in practice.

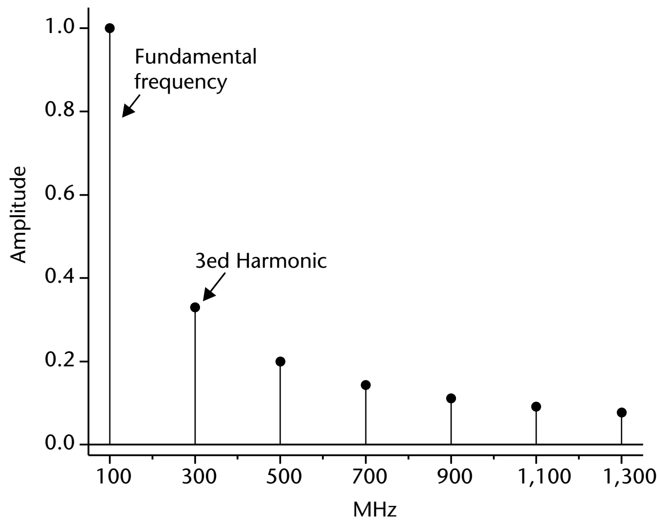


Figure 1.2 Line spectrum of a 10-ns square wave showing the fundamental frequency and the first 13 harmonics. For this pulse only the fundamental frequency and the odd harmonics have values greater than zero.

Figure 1.3 shows how a 100-MHz stream of square waves can be reconstructed by adding together a series of sine waves.

Panel A shows how the pulse looks when only the 100-MHz fundamental frequency is present. The period is 10 ns as expected from (1.1), but the signal is a sine wave rather than a pulse.

Panel B shows the beginnings of a pulse when just the third harmonic is added to the fundamental. From the line spectrum we know the third harmonic is a 300MHz sine wave with an amplitude one-third of the fundamental.

Panel C shows the improvement when the fifth harmonic is added to the fundamental and the third harmonics. The waveform is clearly a pulse, but with large ripple in the flat portion.

Panel D shows the further improvement when the seventh harmonic is added to the previous results. Although the ripple is still present, it is higher in frequency and been reduced in amplitude. This improvement process continues to infinity as more odd harmonics are added.

1.4.3 Why Does the Frequency Domain Matter to Signal Integrity Engineers?

Instead of creating a pulse by adding together sine waves, we get a more powerful insight by observing the effects when harmonics are removed. For example, if the pulse shown in Panel D of the figure is launched though a filter that removes all of the frequencies higher than the third harmonic, the results shown in Panel B are obtained. This shows that by altering harmonics in the frequency domain, the pulse is distorted in the time domain.

This observation is critical because lossy transmission lines act as lowpass filters and gradually remove higher frequencies. As we have just seen (and as the designer of the transatlantic telegraph cable belatedly discovered), this naturally leads to pulse distortion. Losses are discussed in Chapter 9.

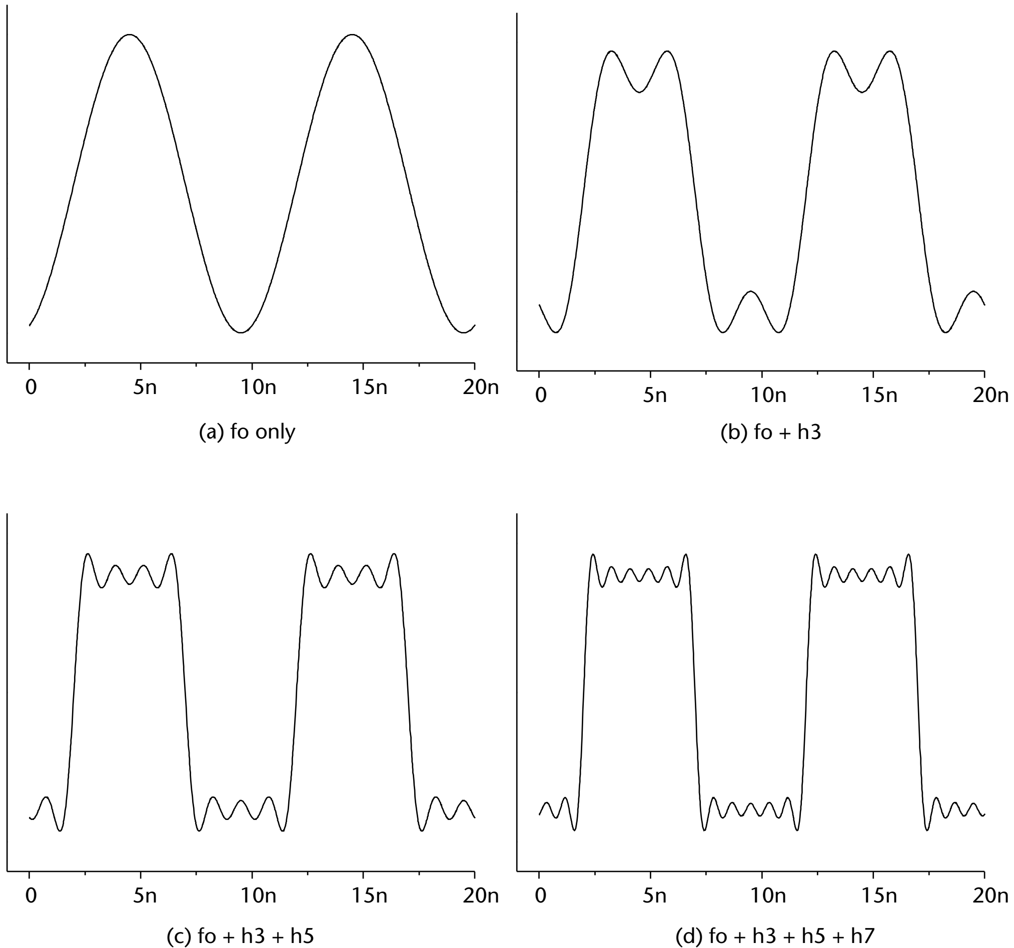


Figure 1.3 Reconstructing a 10-ns pulse stream from the fundamental and the first three odd harmonics. The progressive improvement in the pulse appearance is apparent as more harmonics are added.

A second relationship is shown in Figure 1.4. The solid circles represent the spectra of the 100-MHz square wave when the rise time is zero. This is identical to Figure 1.2 and is included for comparison with the hollow circles, which show the spectra when the same pulse has a rise time of 1.5 ns. The fundamental frequencies are seen to be the same, but the amplitude is lower for the 1.5-ns case, and some even harmonics are now present (for instance, at 200 MHz and 400 MHz). We further notice that for the 1.5-ns rise time case the amplitudes are essentially zero for frequencies of roughly 600 MHz and higher. In contrast, when the rise time is zero, the amplitude is still significant up to 1.3 GHz and beyond.

Apparently the zero rise time pulse requires more harmonics than an identical pulse having a slower rise time. In the frequency domain this means the zero rise time pulse requires more bandwidth than the 1.5-ns rise-time pulse for distortion free transmission. Alternatively, we can say that less bandwidth is needed for distortion free transmission of a long rise time pulse than an identical, sharper pulse.

Pulse shaping (including rise time control) is an important aspect of high-speed I/O design [9–12], but is outside the scope of this book. What is relevant to us is the

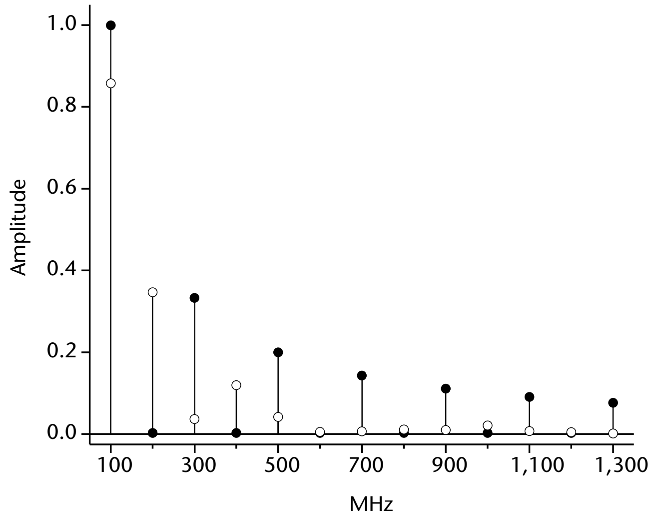


Figure 1.4 Line spectrogram of a 100-MHz square wave pulse train when the rise time is 0 (solid circles) and 1.5 ns (hollow circles).

effect that rise time has on transmission line behavior, and this fundamental topic is discussed in subsequent chapters.

1.4.4 Upper Bandwidth

Equation (1.1) found the lowest frequency present in a pulse stream. The highest frequency of interest is related to the signal rise time.

The rise time (t_r) of a pulse exiting a lowpass filter (such as a circuit board trace) is related to the filter's 3-dB bandwidth (BW) by the relationship shown in (1.2) [13]:

$$BW = \frac{K}{t_r} \tag{1.2}$$

The value of K depends on the shape of the rising or falling edge, and for smooth, Gaussian type pulses when the overshoot is no more than 5%, it ranges from 0.35 to 0.45 (depending on the waveforms characteristics) [14]. A value for K of 0.35 is often used, but more conservative users choose 0.5 [6].

Although this equation shows how a lowpass filter will affect the rise time of the output signal, SI engineers often use it to find the highest frequency of interest in a series of repeating pulses. For instance, to transmit a pulse with a 1.5-ns rise time, the signal integrity engineer would say that the interconnect requires a bandwidth no less than 233 MHz (using 0.35 for K). This is evident in Figure 1.4: Although they have different values, all of the hollow circles for frequencies of 233 MHz and above have amplitudes below 0.6 (which is "3 dB down" from the value of the fundamental). Working with decibels is discussed in Chapters 3, 9, and 16.

The spectrogram of a pulse having glitches or plateaus along its edges, or one that overshoots or oscillates about the steady state value, will have more harmonics

than the well-behaved pulses used here. Equation (1.2) should be used with caution in these situations because those pulses require more bandwidth than the equation predicts.

1.5 Power Integrity and Simultaneous Switching Noise

Power integrity is the analysis and design of power delivery systems. SI engineers are sometimes expected to perform at least some power integrity tasks even though these fields are separate and distinct. For instance, an SI engineer may perform power supply decoupling analysis and power plane assignments within a stackup, especially in small firms such as start-ups.

Although this book does not discuss power integrity in detail, simultaneous switching noise is an important concept that the SI engineer must understand. Those interested in learning more about power integrity are referred to [15, 16].

1.5.1 What Is Simultaneous Switching Noise?

Simultaneous switching noise (SSN) occurs when several output drivers simultaneously switch and draw current from a power supply [12, 17, 18]. The connections for a typical power delivery network (PDN) are shown in Figure 1.5. The noise severity is determined by the PDN resistance, inductance, and decoupling capacitance.

Sometimes SSN is called *deltaI noise* to show that it is the change in current that causes the power supply noise voltage. This term emphasizes that the inductance of the PDN is the root cause of SSN since the changing current through an inductor determines the voltage drop across it. If the switching current cannot be reduced (or the signal rise time lengthened), SSN is improved by lowering the power supply inductance or adding decoupling capacitance (or both). As illustrated

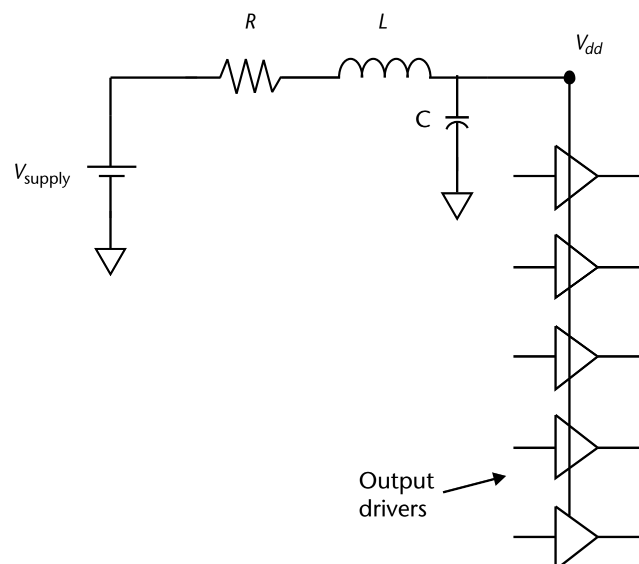


Figure 1.5 Power delivery network resistance, inductance, and capacitance determines the value of V_{dd} when I/O drivers switch.

in Figure 1.6, SSN affects the stability of the power source powering output drivers and input receivers.

The power supply voltage remains reasonably steady when one driver switches, but the increase in current occurring when many drivers switch causes a large voltage drop across the PDN inductance. This reduces the voltage at the transistors of the I/O cell (V_{dd} in the figures). If the decoupling capacitance and resistance are inadequate, the voltage will ring back up and exceed the steady state value when one or more of the drivers stop drawing current. As diagrammed in Figure 1.6 with CMOS I/O drivers this power supply variation changes the driver's timing and often its drive strength. In fact, SSN is one of the major sources of timing jitter (see [19–24]). The SI engineer must consider these factors when selecting I/O models and setting up SI simulations (tasks described in Chapter 3).

1.6 What Is the Dielectric Constant and Loss Tangent?

Signal integrity engineers focus on two electrical parameters when comparing circuit board materials. The capacitance of circuit board traces is determined in part by the value of the laminates relative dielectric constant, and the degree to which the laminate contributes to signal loss is determined in part by the loss tangent.

The relative dielectric constant is abbreviated with the symbol ε_r. Often it is simply called the dielectric constant and is sometimes abbreviated as *Dk*. The two terms are synonymous, but circuit board shops and laminate manufactures generally use *Dk* rather than the more formal ε_r.

The relative dielectric constant of an insulator indicates by how much the capacitance increases when the dielectric (such as the laminate sheets between routing layers), instead of air, is used to separate the plates of a capacitor. For example, a

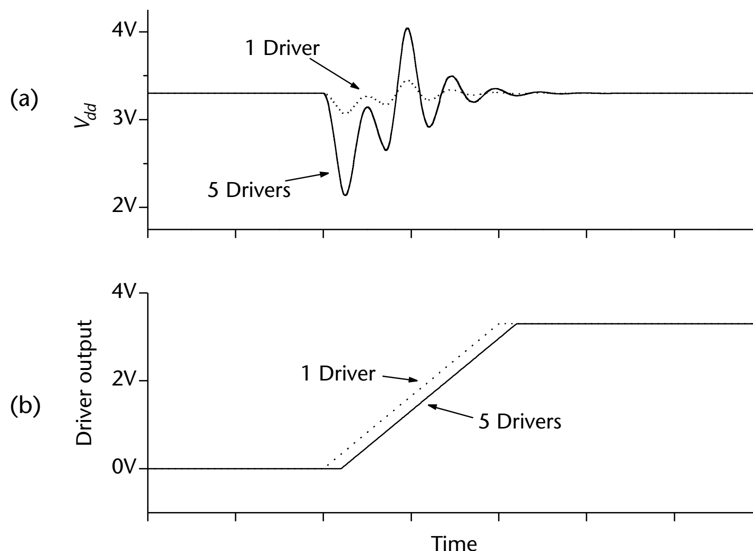


Figure 1.6 PDN inductance causes V_{dd} to drop below the 3.3-V steady state value when I/O drivers switch. The drop is moderate when one driver switches, but is greater when 5 drivers switch simultaneously (a). The changing voltage affects the driver output timing [stylized in (b)].

capacitor formed with a dielectric having $\varepsilon_r = 2.6$ (such as the Gutta-percha insulation used on the transatlantic submarine cable [25]) has 2.6 times the capacitance of an identical capacitor formed in air (the insulation used in the transatlantic cable proof-of-concept experiments). As Chapters 5 and 8 show, ε_r of laminates change with frequency and they do not have a constant value.

The loss tangent (in this book abbreviated as *LT*, but also sometimes called the dissipation factor, *Df*) shows the fraction of signal current lost in the dielectric and so is not available to propagate the signal. This represents a frequency dependent loss, and as we will see in Chapters 7 through 9, is an important factor in distorting pulses when signaling at high frequency.

Loss tangent is often expressed as a decimal fraction, and dissipation factor as a percent. For instance, a typical *LT* value for FR4 is 0.02, which can also be presented as a *Df* of 2%.

Because it is a ratio of currents, the loss tangent has no units. The laminates used in modern printed circuit boards have a low dielectric loss, but moisture uptake (described in Chapter 9) can cause dielectric losses to increase in high-humidity environments.

1.7 Main Points

- Pulses may be examined in the time or frequency domains.
- Pulses can be recreated by combining sine wave harmonics.
- Passing a pulse through a lowpass filter such as a lossy transmission line removes high-frequency harmonics, distorting the pulse.
- Pulses with small rise times require more harmonics (require a higher bandwidth) than a slow rise-time pulse.
- Simultaneously switching output drivers cause power supply noise that can distort the shape and add jitter to the output pulse.
- The loss tangent indicates the amount of frequency dependent loss in the dielectric that a signal will experience.

References

[1] Petroski, H., *Remaking the World: Adventures in Engineering*, New York: Alfred A. Knopp, 1997, pp. 70–71.

[2] Gordon, J. S., *A Thread Across the Ocean*, New York: Walker & Company, 2002.

[3] Standage, T., *The Victorian Internet*, New York: Walker and Company, 1998.

[4] Israel, P., *From Machine Shop to Industrial Laboratory*, Baltimore, MD: The Johns Hopkins University Press, 1992.

[5] Ruston, H., and J. Bordogna, *Electric Networks: Functions, Filters, Analysis*, New York: McGraw-Hill, 1966.

[6] Johnson, H., and M. Graham, *High-Speed Digital Design*, Upper Saddle River, NJ: Prentice-Hall, 1993.

[7] Couch, L., *Digital and Analog Communications Systems*, 5th ed., Upper Saddle River, NJ: Prentice-Hall, 1997

[8] Poularikas, A. D., and S. Seely, *Signals and Systems*, Boston, MA: PWS Publishers, 1985.

[9] Bateman, A., *Digital Communications*, Reading, MA: Addison-Wesley, 1999.

[10] Bissell, C. C., and A. Chapman, *Digital Signal Transmission*, Cambridge, U.K.: Cambridge University Press, 1996.

[11] Kurzweil, J., *An Introduction to Digital Communications*, New York: John Wiley & Sons, 2000.

[12] Dabral, S., et al., *Basic ESD and I/O Design*, New York: John Wiley & Sons, 1998.

[13] Elmore, W. C., "The Transient Response of Damped Linear Networks with Particular Regard to Wide Band Amplifiers," *J. Appl. Physics*, Vol. 19, January 1948, pp. 58–83.

[14] Matick, R. E., *Transmission Lines for Digital and Communications Networks*, New York, IEEE Press, 1969.

[15] Novak, I., and J. R. Miller, *Frequency-Domain Characterization of Power Distribution Networks*, Norwood, MA: Artech House, 2007.

[16] Bogatin, E., *Signal and Power Integrity Simplified*, Upper Saddle River, NJ: Prentice-Hall, 2009.

[17] Hall, S. H., et al., *High-Speed Digital System Design*, New York: John Wiley & Sons, 2000.

[18] Bakoglu, H. B., *Circuits, Interconnections, and Packaging for VLSI*, Reading, MA: Addison-Wesley, 1990.

[19] Li, M. P., *Jitter, Noise and Signal Integrity at High Speed*, Upper Saddle River, NJ: Prentice-Hall, 2008.

[20] Neves, A. P., "Using Measure-Based Modeling to Analyze Backplane Deterministic Jitter," DesignCon 2004, Santa Clara, California, February 2–5, 2004.

[21] Tektronix Corporation, "Understanding and Characterizing Timing Jitter," Tektronix, Inc. White Paper, 2003.

[22] Pandit, V. S., et al., "SSO Noise, Eye Margin, and Jitter Characterization for I/O Power Integrity," *DesignCon 2009*, Santa Clara, CA, February 2–5, 2009.

[23] Miller, M., et al., "A Comparison of Methods for Estimating Total Jitter Concerning Precision, Accuracy and Robustness," *DesignCon 2004*, Santa Clara, CA, January 29–February 1, 2007.

[24] Li, M., "A New Jitter Classification Method Based on Statistical, Physical, and Spectroscopic Mechanisms," *DesignCon 2009*, Santa Clara, CA, February 2–5, 2009.

[25] Federal Telephone and Radio Corp., *Reference Data for Radio Engineers*, 3rd ed., New York: Knickerbocker Printing Corp., 1949.

CHAPTER 2

The Signal Integrity Process

2.1 Introduction

Although modern signal integrity (SI) CAD is sufficiently sophisticated to predict with great accuracy the shape and noise content of signal waveforms, this is not the usual intent of SI analysis (advertisements from SI CAD companies notwithstanding). Instead, an SI analysis defines the boundaries of operation. By creating a group of worst-case simulations, the SI engineer can confidently predict the absolutely best and worst signal behavior and noise. No matter when or where it is manufactured, any given sample of the actual hardware is expected to fall within the bounds defined by the worst-case analysis.

An experienced SI engineer, especially one who in the past has correlated SI simulation results with laboratory measurements on hardware, will instinctively know when a simulation contains too much margin or when the simulation results are physically unreasonable. A less experienced SI engineer, particularly one with no correlation experience, takes the simulation results more literally. These simulations are likely to show excessive margin or unrealistic signal characteristics.

There are special instances when the signaling environment is known in great detail or where the electrical characteristics can be specified to within very small tolerances. These cases often involve very high-performance signaling, or high-reliability systems signaling at more modest rates. In other instances electrical measurements of actual hardware can be used to fine-tune an existing SI model. In these cases modern SI CAD will predict the actual signal wave shape and other attributes with great precision.

2.2 Why Perform a Signal Integrity Analysis?

Perhaps the most common expected outcome of an SI analysis is to insure signal quality and validate system timing. Simulations are performed to quantify the amount of timing margin and to determine that signals have adequate noise margin, do not overshoot by too much, and have manageable crosstalk.

Signal integrity analysis can also determine if I/O drivers are subjected to stresses causing them to operate outside of their safe zone. Some stresses can permanently

damage I/O circuits and shorten the components life. It is often more efficient to set up and perform this type of analysis in simulation than by performing laboratory measurements.

Other types of SI analysis include selecting between components to determine the least expensive device for a particular application and to determine how a product can be cost-reduced. For instance, with SI simulations it is easy to decide with great confidence if an expensive close tolerance termination resistor or expensive type of decoupling capacitor is required in a particular design. It is also straightforward to observe the benefits of various circuit board stackups and materials or, in an ASIC, to select between various I/O drive strengths.

Because well-managed signals generally have lower harmonic content than poorly terminated signals, a common axiom is that good SI yields good electromagnetic interference (EMI) test results. Although good SI is not a guarantee that a product will flawlessly pass EMI testing, experience shows that favorable EMI test results are usually significantly less likely when SI problems are present.

In some instances SI simulations are not performed until after the product has arrived in the laboratory, generally as part of a tardy effort to understand unpredictable system behavior. This can be done on a prototype set aside for debug purposes or on a device already in production that has been identified by manufacturing as having SI problems. Performing an SI analysis this late in the product design cycle is inefficient and expensive, but it can identify the least costly method of improving signal quality.

2.3 What Is a Typical Signal Integrity Workflow?

An idealized workflow performed by one or more signal integrity engineers is illustrated in Figure 2.1. Although eight distinct phases are shown, adjacent phases are sometimes merged, and not all projects need to carry out all of these tasks. Additionally, the interaction between phases is not shown [for instance, the continual feedback between debug and design verification testing (DVT)].

During the product's architectural phase, a senior, experienced SI engineer identifies the limits of technology (such as circuit board size or thickness or ASIC signaling speeds) and broadly defines the circuitry requiring detailed analysis. New CAD tools are also evaluated and purchased at this time, including 2D and 3D field solvers [1]. Decisions are made as to the type of models (transistor level or IBIS circuit models, lossy transmission line, or S-parameter models, for example) that will be used. A rough plan describing the power distribution network is developed.

This phase is often merged with the prelayout phase, especially if the design firm is small (such as a start-up) or has experience with similar technology (and so already has the critical circuit models), or when the product is not aggressively pushing new technology.

The prelayout phase is the time when SI models are obtained and validated and CAD tools are set up. Compatibility between models (verifying that models from different vendors operate correctly when used in the same simulation) is also tested [2].

The validation process individually tests the circuit model of each component by comparing simulation results under simple conditions to laboratory measurements

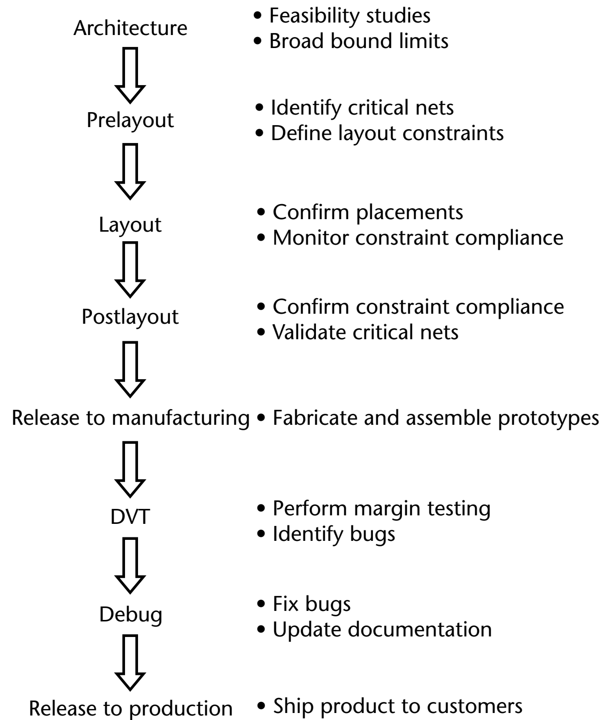


Figure 2.1 Idealized signal integrity workflow showing eight distinct phases. Interaction between phases is not shown.

or to results known to be correct. This time-consuming and often overlooked (and unscheduled) task is vital for proper correlation between SI simulation results and the performance of actual hardware.

Critical nets (and classes of nets) are identified by simulation, and layout constraints (including assumed parts placement) are developed [3]. The need for and the style of termination are determined, and critical layout rules (such as trace widths, spacing, and impedance values) are recorded in an SI layout rules document that will be given to the layout designer. If need be, this is when simulations with advanced laminate systems are carried out to improve signal quality or routing density and the laminate is selected. The circuit board fabrication shop develops a stackup showing the thickness of each layer, the placements of planes, and the trace widths necessary to obtain the specified impedance. The analysis of the power delivery system is begun (sometimes by power delivery specialists or EMI engineers, but often by signal integrity engineers).

During the layout phase, the layout designer creates the artwork necessary to create the circuit board. During this time the SI engineer periodically reviews the circuit board artwork and often continues with prelayout simulations. This is also a time to determine if the manufacturers have updated the circuit models for critical integrated circuits and connectors.

The SI engineer often finds the postlayout phase to be the most hectic of the entire design cycle. This is when the finished artwork is examined for compliance with the SI rules, and when constraints such as spacing between traces, trace widths, and lengths and termination placement and timing are checked. Trade-offs are made to accommodate physical constraints, and simulations are rerun to confirm proper

operation. Some CAD tools support postlayout analysis, so if properly set up during prelayout, the completed layout database can quickly be automatically analyzed for compliance and timing. Although this provides an enormous time savings, especially on large multilayer boards, these tools are no substitute for a detailed visual inspection of the artwork (and the SI engineer should always insist on having enough time in the schedule to do this). The layout designer changes those things identified by the CAD and visual inspections, and the SI engineer reexamines the artwork. This process continues until the board fully meets the SI and timing requirements for the product.

The layout designer formally releases the board to manufacturing once the artwork has been corrected. The SI engineer uses this time to finish up any final simulations and to finalize preparations for design verification testing (DVT).

Generally, DVT is carried out by the engineer responsible for the circuit board, and not the SI engineer. Margin testing where voltage and temperatures are set to extreme values is preformed during this phase, and bugs are uncovered. Test patterns are used to exercise those highest risk nets identified during the prelayout and layout phases. Often this phase is not distinct from debug, but some large firms segregate these tasks and have different specialists for each. As described here, the DVT process includes industry compliance testing, where those signals that must strictly adhere to industry specifications are checked. In some cases, a third-party testing laboratory is used to certify compliance.

Many SI engineers are not directly involved with laboratory debug. Instead, they provide debug support by analyzing measurements taken by others and in identifying critical test patterns and nets. This is especially likely if the SI engineer is a CAD specialist and not one comfortable performing laboratory measurements. However, the most valuable SI engineer is skilled in both CAD and measurement techniques. These engineers have practical experience that is invaluable when interpreting simulation and measurement results.

The system is formally released to production for volume manufacturing once the product has successfully passed through DVT and debug. Generally the SI engineer is not involved in solving production problems, but they may be required to help resolve timing or signal quality problems that arise during volume production and are often required to help with component selection or evaluation when critical parts become unavailable. This is especially so if the SI engineer has laboratory measurement experience, but SI CAD specialists are also enlisted. If properly documented and archived, the prelayout and postlayout design analysis (particularly test patterns and environmental setup assumptions) can be immensely helpful in solving urgent, unexpected production SI problems.

2.4 Signal Integrity Worst-Case Analysis

Often the SI simulations performed during the architectural phase are done under nominal conditions. Circuit models for typical silicon are used, and traces, power supply voltages, and termination resistor values are all assumed to be exactly as specified. As the design progresses, the analysis progressively takes into account process variations in the components, power supplies, and circuit board traces.

Circuit board vendors do not specify the manufacturing tolerances of the trace widths or thicknesses, but on those traces specified as controlled impedance they do indicate the impedance tolerance. This range is used by the SI engineer in simulations to determine the optimum value of termination resistors and to find the highest ringback and overshoot voltages.

2.4.1 Silicon Worst-Case I/O Models

Generally silicon vendors bound the range over which their device will operate by providing nominal, fast (sometimes called best-case), and slow (sometimes called worst-case) I/O circuit models. Because I/O driver performance is strongly dependent on the die temperature (the junction temperature) and the actual power supply voltage, these external factors must be appropriately set along with the proper I/O case for a specific type of analysis. For example, CMOS I/O launches the slowest signal rise time when the lowest power supply voltage and highest temperature is used with the slowest case I/O model. To determine the highest current draw the fast I/O model is used with a low die temperature, highest power supply voltage and lowest transmission line impedance.

The temperature parameter used in the I/O models represents the junction temperature, not the temperature of the ambient air. Sometimes this distinction can be overlooked when the integrated circuit is a low power device, but it becomes critical as the device power dissipation increases. For instance, a commercial product may operate in an ambient air temperature of 85°C or less. The junction temperature will not be much higher than this if the integrated circuit is low power device, but (depending on the packaging) under the same circumstances the junction temperature of a higher power device may exceed 125°C. The junction temperature is estimated during prelayout and this value is used in the evolving SI model.

Although the junction temperature can be set in the simulation, not all CAD tools have the ability to individually set the junction temperatures of various integrated circuits simultaneously coexisting in a simulation. When this is the case, a net containing a device operating at a high junction temperature connected to a device with a lower junction temperature will be simulated at one temperature (often the highest one). The result is that I_{dd} noise (noise on the power supply) and the rise time of the signals launched by the I/O drivers will be underestimated for the device operating at the lower temperature. This can mask signal integrity problems such as the magnitude of overshoots and reflections.

2.4.2 What Combination of Environmental Effects Are Worst Case?

The precise combination of environmental conditions and circuit models that produce a particular worst-case condition depends on the details of a design, and sometimes the combination can be counterintuitive.

Rather than manually running through the various permutations, many SI CAD tools can be set up to vary the parameters automatically and summarize the results. This can be very helpful when designing large and complex circuit boards, but adequate time should be included in the schedule to interpret and react to the results.

For instance, particularly surprising results should be examined (and possibility manually resimulated) to determine if they are valid or if they are the artifact of a defective circuit model. The validity of the stimulus and environmental conditions should also be checked to verify that the conditions presented in the simulation is physically possible in the actual hardware.

Examples of some common worst case configurations are shown in Table 2.1, where L and H represent the lowest and highest values specified by the manufacturer, F is the fast silicon circuit model, and O and E are the odd and even switching modes (described in Chapter 9). The device current is shown as I_{dd}, and the worst signal overshoot is indicated in the highest OS column.

In general, setting the environmental variable to opposite values than shown in the table will produce the opposite effect. For instance, to find the lowest I_{dd}, use the highest temperature, the lowest V_{dd}, the slow silicon process file, and so on.

2.4.3 Using SI Analysis Results During Debug

The results from the SI analysis are valuable when setting up laboratory conditions for debug. For instance, noise between the power and ground planes will generally be the highest when I_{dd} is high and the signal rise time is short. Table 2.1 shows that to make this measurement, integrated circuits with the fastest rise time I/Os are placed on the circuit board, since that implies they are from the fast process corner. The power supply voltage is raised to the highest permissible value, and the temperature is lowered to the lowest possible ambient value. If the testing is done on a single circuit board or small assembly the temperature can be controlled by placing it in a thermal chamber.

The signaling patterns identified during the SI analysis are useful during debug for exercising prototype hardware. By working closely with the diagnostic engineer (a specialist responsible for creating the software and firmware necessary to troubleshoot and diagnose faults in the hardware), the signal integrity engineer can have the most critical SI test patterns included in with the test programs. In this way tests that were run in simulation can be duplicated in prototype hardware and the results compared. Besides validating the design, this correlation exercise provides an excellent opportunity for the SI engineer to understand the limits of simulation and to refine his or her intuition.

Table 2.1 Example Worst-Case Settings

Variable	*Highest* I_{dd}	*Highest Coupling*	*Fastest Rise Time*	*Highest OS*
Temperature	L	L	L	L
V_{dd}	H	H	H	H
Silicon	F	F	F	F
TLINE Zo	L	H	H	H
R_s	L	L	L	L
R_t	L	H	H	H
Pattern	O	E	E	E

2.4.4 Electrical Overstresses

Electrical stresses on an I/O pin can be caused by electrostatic discharge (ESD) events and by signal pulses high enough in amplitude to damage the junctions and gate oxide of the transistors in the I/O cell.

ESD testing is not included in with the normal SI simulations, but signal overvoltages must be analyzed carefully.

Time-dependent dielectric breakdown (TDDB) [4, 5] and hot carrier effect (HCE) [5–8] stresses are present in the transistors forming the receiver and driver portions of an I/O cell in response to high voltages at the I/O pin. Over time these failure mechanisms can cause a gradual degradation of the transistors (changing their behavior), or if the overstress is large enough, it can cause the transistors to fail catastrophically. *High voltage* is a relative term and the actual maximum voltages may only be a few volts higher than the maximum allowed power supply voltage. The specific circuit design of the I/O cell and the details of the process technology determine the highest allowed voltage and the amount of time that a particular voltage can be tolerated [4].

The signal integrity simulations should always be carefully examined to determine the highest voltage present at the I/O pin, and this value must never be allowed to exceed the manufacturer's worst-case specification. Both ends of the net should be examined.

The highest value is usually created by large reflections generated when the transmission line is not properly terminated, but they can also be created by crosstalk on high-impedance nets. Often the worst case occurs when the power supply is at a maximum value, the temperature is low, and the series termination has its lowest value since these conditions encourage the highest reflections. These failure mechanisms are usually not a practical concern in high-speed serial signaling because there the signaling voltages are low and the lines are well terminated.

Often the manufacturer's data sheet will show that an I/O driver is more able to tolerate a high-going voltage than one going below ground, so special care should be taken to design simulations to test the case where signals significantly drop below ground. Once identified, the simulation case creating the worst-case voltage should be noted and that same case duplicated in the laboratory during debug or DVT to verify that the maximum limits are not exceeded.

2.5 Main Points

- SI simulations bound the range of operation and show how the system will behave under all legitimate environmental conditions and manufacturing variations.
- There are eight distinct places in the design cycle where SI analysis can be performed, but not all are performed on a specific product or by an individual engineer.
- Good EMI test results begin with good SI.
- Performing SI analysis late, after the product has been released to production, is costly and time-consuming.

- Patterns and environmental configurations identified during simulation can be used to test actual hardware during DVT and debug.
- SI analysis identifies design vulnerabilities in the design and shows where the debug activities should be targeted.

References

[1] Huang, C. C., "Signal Integrity Modeling and Simulation Tools," *Proceedings of DesignCon*, 2004.

[2] Beal, W., "Use the Right Models for the Simulation of Multi-Gigabit Channels," *Proceedings of DesignCon*, 2004.

[3] Ramakrishnan, P., "CDNLive! Paper—Signal Integrity (SI) for Dual Data Rate (DDR) Interface," *Cadence Designer Network*, Silicon Valley, 2007.

[4] Chandrakasan, A., et al., *Design of High-Performance Microprocessor Circuits*, New York: IEEE Press, 2001.

[5] Pompl, T., and M. Kerber, "Failure Distributions of Successive Dielectric Breakdown Events," *IEEE Transactions on Device and Materials Reliability*, Vol. 4, No. 2, June 2004.

[6] Leblebici, Y., and S. M. Kang, *Hot-Carrier Reliability of MOS VLSI Circuits*, Boston, MA: Kluwer Academic Publishers, 1993.

[7] Dillinger, T. E., *VLSI Engineering*, Upper Saddle River, NJ: Prentice-Hall, 1988.

[8] Groeseneken, G, "Hot Carrier Degradation and ESD in Submicrometer CMOS Technologies: How Do They Interact?" *IEEE Transactions on Device and Materials Reliability*, Vol. 1, No. 1, March 2001.

CHAPTER 3

Signal Integrity CAD and Models

3.1 Introduction

The signal integrity (SI) engineer uses a variety of CAD tools and model types when performing an analysis. This chapter introduces the two most common types of models for representing integrated circuit input and output (I/O) drivers and discusses some of the circuit models used to predict trace and via behavior. Finally, the field solvers necessary to create circuit models from physical structures are described.

3.2 I/O Models

The models of the I/O circuits used in SI simulations are either transistor level circuit models or behavioral type models. The most prevalent circuit models are based on the SPICE circuit simulator, while IBIS is the most common behavioral model.

3.2.1 What Are Transistor Level Models?

Because they come directly from the integrated circuit net list, transistor level models are the most accurate representation of the I/O circuitry. However, the accuracy of the system level simulation also depends on the accuracy of the models used to simulate the transistors (the worst-case I/O models discussed in Chapter 2), and the quality of the models for the interconnect within the micropackage that houses the integrated circuit.

To find the current and voltage at the I/O cell pad, the circuit simulator must solve the equations for each transistor, capacitor, and resistor within the cell at each time point in the simulation. While this can give highly accurate results, these calculations are time-consuming and make transistor level simulations run much more slowly than behavioral model simulations. In some instances the simulator may have difficulty properly calculating the circuit node voltages and currents. The simulation will prematurely terminate if the errors associated with these calculations become large enough. Circuit designers use various optional simulator settings to minimize or avoid these kinds of problems, and occasionally different

I/O models will require that the same simulator option be set to separate values. This makes it impossible to run a signal integrity simulation that simultaneously includes both integrated circuits, since to do so would require the same parameter to simultaneously have different values. Conflicts of this type are more likely to occur when the integrated circuits come from different vendors.

Because a transistor level I/O model is a net list that includes all the connections and values for all the components in the I/O circuit, it is possible to reverse-engineer the model and discover exactly how the circuit works. To prevent this, the I/O circuitry (and sometimes also the micropackage) can be encrypted by the manufacturer. This hides the internal wiring and the transistor parameters, making it impossible for anyone to determine the transistor process information or how the circuitry is wired. The encrypted models are not compatible with all simulators, so the SI engineer must confirm that an encrypted model will properly run on the circuit simulator of their choice (and with the other models in the system).

Although encryption solves the manufacturers' intellectual property dilemma, simulations with encrypted models can be difficult to debug because the SI engineer cannot examine or initialize internal nodes when the simulation goes awry.

3.2.2 What Are IBIS Models?

I/O models based on the ANSI/EIA-656-B standard (called IBIS, for *I/O Buffer Information Specification* [1]) are the most popular behavioral type models. IBIS models present the I/O terminal current and voltage values in look-up tables. The simulator uses this information to determine the response of input and output circuits to different loads and voltages present at the I/O pin. These models include estimates for the package parasitics, so a separate micropackage model is not necessary. The capacitance of the I/O circuitry is also included.

Because the simulator does not perform the complex calculations necessary for transistor level models, simulations using IBIS models run significantly faster than transistor level simulations.

Since no proprietary information is present, IBIS models are not encrypted, and IBIS models can be run on simulators from many different vendors. However, not all IBIS models fully comply with the industry specification, and some simulators may not operate with models from all vendors.

A disadvantage of earlier IBIS models was that current in the power and ground pins is not modeled, thus preventing these models from determining power supply noise. A second problem of importance in high-speed signaling is that complex I/O behavior such as precompensation and equalization (briefly described in Chapters 13 and 16, but more fully described in [2–8]) cannot be properly modeled. However, models conforming to the revision 5 IBIS specification (IBIS-AMI, for *IBIS-Algorithmic Modeling Interface* [1]) remove these deficiencies.

These advanced features are optional and may not be included in all version 5 models. For instance, a model complying with the version 5 specification may deliberately choose not include the power and ground currents but to include the precompensation behavior. Each IBIS model should be checked to insure it has the proper components for a given analysis.

3.3 Modeling Transmission Lines

Short transmission lines are sometimes modeled as a string of resistors, capacitors, and inductors. Models of this type are occasionally provided by vendors for small structures such as connectors, and signal integrity engineers sometimes create these models for vias.

The resistors represent conductor and dielectric signal loss, but because the resistors are fixed values they do not change with frequency. As described in Chapters 6–8, this is undesirable and makes the model most accurate for single frequency sine waves rather than pulses (which contain many frequencies).

Virtually all modern versions of SPICE include lossless and single-frequency lossy transmission line models, and some SPICE implementations have additional models. These are much more computationally efficient and convenient than concatenating strings of RLC circuits. The most common transmission line models are listed in Table 3.1.

The lossless model (the SPICE T element) is a perfect transmission line comprised of only an inductance (L) and capacitance (C) and so its only characteristics are delay and impedance. It can be accurately used with high-frequency pulses when the losses are low (this means that the transmission line is electrically short). In most implementations of SPICE, the T element models only one signal conductor, so coupling between traces (crosstalk) and differential pairs cannot be simulated.

The lossy line (usually referred to as an LTRA line and often referenced as the O element) is defined by specifying values for the capacitance, inductance, series resistance (representing the loss in the conductor), and in some implementations a shunt conductance (representing the dielectric losses). The resistance (R) and conductance (G) are fixed values that do not change with frequency, making this line most accurate at the one frequency for which R and G were specified. As such, it is best used with single frequency sine waves and will not be accurate when used with digital pulses. Like the T element, in most implementations the O element models one signal and cannot be used to simulate crosstalk or differential pairs.

At least two vendors [9, 10] provide a multiconductor lossy transmission line element (often called the W element) where the circuit values automatically change with frequency. The parameters in this model can be specified in many ways, but the most common usage at low to moderately high frequencies is for the SI engineer to provide values for L, C, R, and G. This model differs from the O element in that the simulator computes new values for the conductor and dielectric losses (R and G) and optionally for L and C at the various frequencies present in the signal. This allows the model to properly account for losses in pulse streams where the signal rise time and pulse width change (as can be caused by signal reflections). This

Table 3.1 Common SPICE Transmission Line Elements

SPICE Element	*Type*	*Notes*
T	LC Lossless	Only delay and impedance; no losses
O	RLGC Lossy	Losses correct at only one frequency
W	RLGC Lossy	R, L, G, C can be made to vary with frequency

model can simulate a single conductor or a multiconductor bus and so can be used to model crosstalk or differential pairs.

The Exercises for Chapter 7 show how to compute R and G for use in O and W element circuit models.

3.4 S-Parameters

Interconnect (including traces, vias, or connector pins) can be modeled in the frequency domain with *scattering parameters* (S-parameters). Although they originated in RF engineering, S-parameters are now commonly used in high-speed digital signaling work, especially in serial signaling applications.

Because of their importance to advanced SI applications, an overview is presented here, but since S-parameters are not used in this book, we will not go into detail. Additional introductory information on S-parameters is presented in [11–13]. More theoretical explanations are presented in [14–17].

3.4.1 What Are S-Parameters?

S-parameters show the degree to which sine waves are reflected (scattered) from a signal source of known impedance (the reference impedance) by each pair of terminals (ports), and the amount of the signal appearing between ports. This means that S-parameter models show loss, impedance mismatch (reflections), and the coupling between ports (crosstalk between traces).

A trace over a ground plane has two pairs of terminals and so is an example of a simple two-port network, as shown in Figure 3.1.

Subscripts are used to indicate the terminal pair being measured (identified by the subscript m) and the terminal pair injecting the signal (the subscript i). This is shown in (3.1).

$$S_{mi} \tag{3.1}$$

For instance, measuring the signal at port 2 and injecting a signal into port 1 gives S_{21} (the *insertion loss,* since it shows how much the signal is attenuated between ports 1 and 2).

Simultaneously injecting a wave and observing the reflection at port 1 yields S_{11} (the *reflection loss* because it shows how much of the signal is reflected back to the generator). Interconnects such as transmission lines and vias have the same behavior in both directions, so for traces $S_{21} = S_{12}$ and $S_{11} = S_{22}$.

In S-parameter models the *S-parameter coefficients* are tabulated for a range of frequencies and are represented by two numbers, one for the amplitude and the other for the phase. A sample S-Parameter listing of just the amplitude for a 4-mil (0.11-mm)-wide half-ounce 50 stripline on FR4 is shown in Table 3.2.

3.4.2 What Is Insertion Loss?

For a lossless transmission line, the insertion loss (S_{21}) and its twin (S_{12}) are 0, since that indicates that no signal energy is lost when the transmission line is inserted

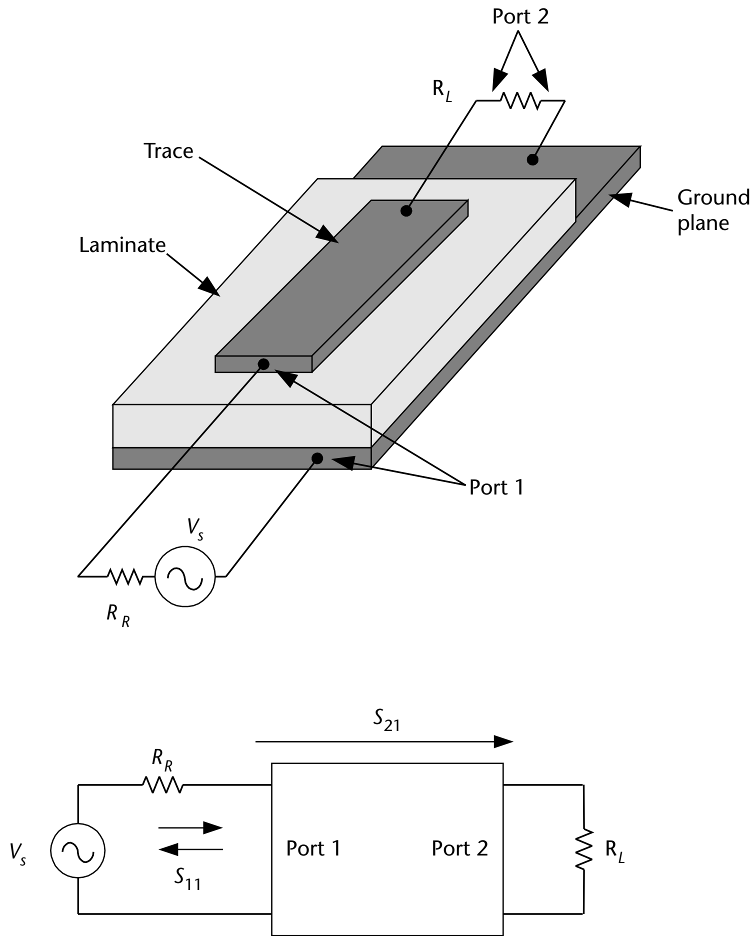


Figure 3.1 Single trace over a ground plane two-port network. For S_{11} port 1 is simultaneously driven and measured at port 1. S_{21} is measured at port 2 when port 1 is driven. In both instances port 2 is terminated and driven with a reference impedance (R_R).

Table 3.2 S-Parameter Coefficients in Decibels for a 15-cm (6-Inch)-Long 50Ω Stripline as Calculated by Linpar [18]

GHz	S_{11}	S_{21}	S_{12}	S_{22}
1.0	–56	–1.3	–1.3	–56
1.5	–51	–2.2	–2.2	–51
2.0	–46	–3.2	–3.2	–46
2.5	–42	–4.7	–4.7	–42

between the transmitter and the load. From the table we see that the insertion loss for the 4-mil-wide 50Ω trace is actually –1.3 dB at 1 GHz and increases to –4.7 dB at 2.5 GHz. The use of decibels is described in Chapters 8 and 17, where we will see that a negative value indicates a loss and larger negative values represent larger loss.

Here we note that a loss of 1.3 dB means that the output is only about 86% of the input value. This means that about 14% of the signal has been attenuated. About 41% of the signal is lost at 4.7 dB.

Chapter 1 showed that the 3-dB bandwidth is a convenient way to estimate the highest frequency of interest in a pulse stream. From Table 3.2 we see that frequencies of 2 GHz and higher attenuate the signal by more than 3 dB. Therefore, the bandwidth for this trace is just under 2 GHz, and signals (including the signal harmonics) with frequencies higher than this will be severely attenuated.

3.4.3 What Is Reflection Loss?

The reflection loss (S_{22} and S_{11}) indirectly shows by how much the transmission line impedance differs from the reference impedance. Chapter 11 shows that a perfect match results in no reflections, which would make the reflection loss zero. Since on a decibel scale a large negative number represents a small quantity, a large negative reflection loss value in decibels shows that the reference impedance and transmission line impedances are very similar. For instance, a reflection loss of –26 dB represents a 5% impedance mismatch, which is acceptable in most low and moderate speed applications. From Table 3.2 we see that at 2 GHz (the 3-dB bandwidth of this trace) the reflection loss is much smaller than this (–46 dB). That represents about a 0.5% mismatch in the impedance.

3.4.4 What Are Touchstone Files?

Those versions of SPICE capable of running S-parameter models nearly universally use the Touchstone File format [19] to present the S-parameter data, although other formats also may be supported. Touchstone format model files can come from test equipment (measurements) or field solvers (simulations).

Often the data is presented in Touchstone files using a magnitude-angle format with the magnitude in decibels (dB) and the phase angle measured in degrees, but the specification also defines other formats.

Usually S-parameters are specified at 50Ω, but any reference impedance may be used and its value recorded in the file.

3.5 Field Solvers

Electromagnetic field solvers analyze a physical structure such as a circuit board trace or the complex geometry of a connector and produce an electrical circuit model for use in signal integrity simulations.

The most common (and often the easiest to use) field solvers are those that analyze two-dimensional structures (2D solvers). Three-dimensional structures are analyzed with 3D simulators.

3.5.1 What Are 2D Field Solvers?

The underlying assumption behind 2D solvers is that the structures shape is uniform and so can be properly described with only two dimensions. A group of straight

circuit board traces that have a constant width, spacing, height above the ground plane, and thickness is a good example of the type of structure appropriate for 2D analysis.

The field solver calculates the resistance, capacitance, inductance, and conductance of the 2D structure and tabulates them across a user-defined frequency range. The frequency range must be carefully selected because, as we saw in Chapter 1, digital pulses contain high-frequency harmonics and (as we will see in Chapters 6 through 8) losses increase with frequency.

Since the shape is uniform, the electrical parameters are specified per unit length. The value for a specific trace is found by multiplying the field solver results by the actual length of the structure. For example, if the capacitance of a trace is reported as 140 pF/m, the capacitance for a 1-cm-long trace is 1.4 pF.

The best solvers allow the user to choose the output formant from a variety of circuit model types, so a library of circuit models can be created that are usable by a number of different circuit simulators. At a minimum the field solver that you choose should produce both an RLGC output file and an S-parameter model (preferably in Touchstone format). The RLGC data of single traces can be validated by comparing it with the results from the approximate formulas presented in Chapter 17. This is a good way to verify that the model has been setup correctly. As described in Chapter 7's Problems, the RLGC information can be used to create a W line model in those cases where the field solver does not directly create one. Alternatively, it can be used with the SPICE LTRA model (noting the limitations discussed in Section 3.3).

3.5.2 What Are 3D Field Solvers?

Circuit models for physical structures that are not uniform in shape (such as vias or traces that vary in width or spacing or that bend around objects) are created with a 3D field solver. Some of these solvers require knowledge of radio frequency terms and concepts to produce valid models, but other simulators are specifically designed to be operated by signal integrity engineers having little RF knowledge.

Because the structure is simulated in three dimensions, setting up the problem space is usually more time-consuming and error-prone than with 2D simulators. Before purchasing a 3D simulator, check that entering the circuit board stackup and defining the material properties are intuitive. Furthermore, it should be straightforward to verify that the values have been assigned properly and assigned to the correct layer in the stackup. Surface roughness (as shown in Chapter 8, an important factor when calculating conductor loss) should be easy to apply to traces and planes. When working with high-speed signals, the ability of the simulator to include the way in which frequency changes the dielectrics electrical properties is also very important.

As part of the selection process, the drawing interface should be thoroughly tested to determine the ease in creating and editing complex 3D structures. Although some 3D simulators can directly import the 3D mechanical drawings created by CAD drafting software, or the stackup from circuit board artwork tools, it is routinely necessary for the SI engineer to "clean up" the drawing and to crop regions once it has been imported.

Some 3D solvers have *wizards* or template files that make it easy to setup commonly found problems such as vias or differential pair traces (or differential vias). This valuable feature can save significant time and is especially important if the SI engineer is not using the simulator on a daily basis.

Generally S-parameters are the default output model type, but RLGC files or W line models can also sometimes be created. The electrical parameters calculated are of the complete geometry. For instance, a structure might have a capacitance of 3 pF, but it would not have a value of 3 pF/inch.

3.6 Main Points

- I/O models are either transistor based or behavioral.
- Transistor level models are most accurate but run significantly slower than IBIS models.
- Transistor level models may be encrypted to protect intellectual property. This can make the SI debug more difficult.
- IBIS models are the most common behavioral type model and run significantly faster than transistor level models.
- Field solvers calculate the resistance, capacitance, inductance, and conductance of physical structures.
- Two-dimensional field solvers are used to create circuit models of regular structures such as uniform traces.
- Three-dimensional field solvers are used to create circuit models such as vias and irregular structures.
- Some 3D simulators are difficult to set up and to properly run and should be thoroughly tested to verify that the operation is intuitive.
- The output from 2D and 3D field solvers can be tabular data, circuit models of various forms, or S-parameters.

References

[1] IBIS version 5.0, ratified August 2008 by the IBIS Open Forum, http://www.eigroup.org/ibis/specs.htm.

[2] Thierauf, S. C., *High-Speed Circuit Board Signal Integrity*, Norwood, MA: Artech House, 2004.

[3] Proakis, J. G., *Digital Communications*, New York: McGraw-Hill, 2000.

[4] Bateman, A., *Digital Communications*, Reading, MA: Addison-Wesley, 1998.

[5] Bissell, C. C., et al., *Digital Signal Transmission*, Cambridge: Cambridge University Press, 1996.

[6] Horowitz, M., et al., “High-Speed Electrical Signaling: Overview and Limitations,” *IEEE Micro*., Vol. 18, No. 1, January/February 1998.

[7] Dally, W., et al., “Transmitter Equalization for 4-Gbps Signaling,” *IEEE Micro*, Vol. 17, No. 1, January/February 1997.

[8] Dally, W., et al., *Digital Systems Engineering*, New York: Cambridge University Press, 1998.

[9] SmartSpice, Simucad Design Automation, Santa Clara, CA.

[10] HSPICE, Synopsys, Inc., Mountain View, CA.

[11] Quint, D., and K. Bois, "The Digital Designer's Complete Lossy Transmission Line Model," *Proceedings of the 52nd Electronic Components and Technology Conference*, May 28–31, 2002, pp. 1073–1079.

[12] Hewlett Packard Test & Measurement Application Note 95-1, "S-Parameter Techniques for Faster, More Accurate Network Design," November 1996.

[13] Bogatin, E., "How Do You Spell S Parameters?" *Printed Circuit Design Magazine*, February 2003.

[14] Hall, S. H., et al., *High-Speed Digital System Design*, New York: John Wiley & Sons, 2000.

[15] Young, B., *Digital Signal Integrity*, Upper Saddle River, NJ: Prentice-Hall, 2001.

[16] Agilent AN 154, "S-Parameter Design," Application Note, Agilent Technologies, 2000.

[17] De Geest, J., et al., "Electrical Modeling of High Speed Interconnection Links," *Proceedings of the 51st Electronic Components and Technology Conference*, May 29–June 1, 2001, pp. 1245–1249.

[18] Djordjevic, A. R., et al., *LINPAR for Windows*, Norwood, MA: Artech House, 1999.

[19] Touchstone Specification, Version 2.0, ratified April 24 2009 by the IBIS Forum, http://www.eigroup.org/ibis/specs.htm.

CHAPTER 4

Printed Test and Evaluation Boards

4.1 Introduction

Usually debug is performed in the laboratory on prototype versions of the actual product board. This allows hardware, software, and firmware developers to perform debug on hardware that is essentially identical to what will become the final product.

However, this approach is not always the most efficient, especially when working with new signaling or integrated circuit technologies and new devices (such as connectors or sockets) or when the communication between subassemblies must be tested (such as with a backplane or daughter card situation). This is particularly so when signaling at high speeds, but as we will see in later chapters, signal integrity problems are not limited to high-speed circuits.

In any of these cases it can be more cost effective to design and build circuit boards with limited functionality specifically designed to test the high-risk portions of the actual design. These printed circuit boards are called *test boards* to distinguish them from *evaluation boards*. As discussed in Section 4.3, evaluation boards are often provided to customers for evaluation of one or more ASICs in a well-defined environment.

4.2 Test Boards

The signal integrity engineer should consider test boards as opportunities to reduce risk and to further their own understanding of the way simulation portrays physical effects.

For instance, laboratory measurements can be made on the test board to:

- Validate I/O and connector circuit simulation models.
- Validate a routing strategy in a high-density pin or connector field.
- Validate a stackup.
- Validate the performance of a split plane.
- Test power supply decoupling strategies, including component values and types, and layout rules.

- Validate circuit models of various routing topologies, and to decide between them.
- Determine if a given trace width has adequately low loss and compare to the simulation.
- Determine crosstalk of critical routes, especially in dual stripline situations that are difficult to simulate properly.
- Validate I/O cell circuit models.
- Determine the ability of a manufacturer to meet impedance specifications.
- Determine the characteristics of traces routed near the board edge or other cutouts in the board.

However, it is important for these boards not to become too intricate, especially if the experiments require excessive support from others. For instance, a dual stripline crosstalk test would require that many lines be driven and some number of them monitored. Support from other engineers will be required unless the signal integrity engineer can design the logic or create the firmware necessary to do this. In fact, it is critical not to lose sight that test boards are intended to reduce risk and not to create it by siphoning off engineering resources that might be better used in the development of the actual product.

4.2.1 Test Boards as a Chip Debug Platform

An important use of test boards is as a platform to facilitate the test and debug of a new ASIC. For instance, a fabless semiconductor company would create such a board so their engineers can validate and demonstrate the features of their chip in an electrically controlled environment. This can be particularly helpful when providing customer support if the applications engineer can configure the ASIC in the same way as the customer and demonstrate the same problem.

A test board is an important engineering tool because it is far less costly to debug a new chip design on a test card than it is on the manufacturer's automatic test equipment (ATE). Although deep, high-quality test patterns running on ATE are critical for the manufacturer to determine if the ASIC works the way the designers had intended, tracking an elusive bug on the ATE is inefficient and costly.

Instead, a test board can be developed that provides an electrically sound environment for the ASIC that includes jumpers, LEDS, test points, and connectors that allow the designer to configure and test the chip under normal and extreme conditions.

Test boards often have connections to visibility points that allow the designer to observe or set logic deep within the ASIC. This is particularly helpful when the chip can be placed in one or more test modes. Such modes are often not intended for customer use and frequently redefine the normal function of ASIC I/O pins. With a properly designed test card, the designer can use these pins in a test mode without interference from the logic to which they would normally be connected.

A well-coordinated debug strategy includes a test plan that describes how to use the test card with whatever test logic has been designed into the chip. In this way test patterns can be written and any special fixtures or long lead time test

equipment and other parts can all be in place before the test card arrives in the laboratory.

4.2.2 Margining Power Supply

With the proper laboratory equipment, the power supply to the chip can be varied (margined) to see how low or high the voltage can be set before the chip begins to malfunction. Margining can also be used to determine which circuits are the first to fail. This technique is especially powerful when the temperature of the air surrounding the integrated circuit is simultaneously margined. As shown in Chapter 2, in general, low voltage and high temperature represent the extreme slow case for CMOS integrated circuits, while low temperature and high voltage generally represent the fast operation, high-current draw case.

4.2.3 Testing Decoupling Capacitor Efficacy

The frequency content of the current drawn by an integrated circuit or a group of chips under various switching and operating conditions can be determined by measuring the current and performing a Fourier analysis on the waveform to obtain the frequency spectra. Many modern oscilloscopes will directly perform a fast Fourier transform (FFT), or the readings can be saved and externally processed by mathematical CAD programs such as MathCAD [1] or MATLAB [2].

By observing the spectra under various conditions, including with various I/O switching patterns, the signal integrity engineer can determine the frequencies over which the power supply decoupling network is operating. Specific frequencies with particularly high amplitudes can be identified, and if necessary, decoupling can be specifically designed to eliminate them.

4.2.4 Testing Sensitivity to Clock Quality

The chips sensitivity to clock amplitude, frequency, duty cycle and jitter can be tested by replacing the expected clock source (often a crystal oscillator or PLL/fanout circuit) with a laboratory function generator or test set. In this way the sensitivity of the chip to clock quality, especially when tested over extremes of temperature and power supply voltage, can be determined.

4.3 Evaluation Boards

Ideally, the test board used for debug can be given to prospective customers for them to evaluate the operation of the ASIC, but in practice it is usually better to have a different board for this purpose. In the past, the strategy of creating a single test and evaluation board made economic sense when circuit board artwork could only be created by layout specialists and when fabricating and assembly shops were reluctant to take on small sized orders.

While it is true that the routing of high-speed, high-performance circuit boards is best left to experienced specialists, the proliferation of layout design CAD tools (some of which include auto routing software that is driven from the schematic to

route traces automatically) makes it possible for the skilled signal integrity engineer to lay out simple, small, multilayer circuit boards for customer use.

Often these “eval boards” are small and contain only the minimum amount of peripheral circuitry necessary to demonstrate operation of the ASIC. Connectors are provided as necessary so that the user can connect the evaluation board to his or her application.

Although the interconnect present on these boards is often simple, careful thought must be given to the stackup and to decoupling and filtering. These must provide the electrical environment necessary to demonstrate the best possible operation of the ASIC. However, unnecessarily complex decoupling, or a complex stackup (for instance, using many power and ground planes or planes that have multiple splits) can be seen by the potential customer as an undesirable requirement for their design. In fact, often the customer will request (and sometimes will require) copies of the evaluation board artwork files so that the layout of the end product will exactly match the evaluation board. In this way the ASIC will operate in the same electrical environment in the end product and so should behave in the same way. An overly complex stackup in the evaluation board can lead to higher costs in the end-user board, which could discourage potential customers.

4.4 Roll of the Printed Circuit Layout Shop

A layout designer is the specialist who creates the artwork files required by a circuit board manufacturing shop to actually build the circuit boards. They are employed in firms ranging in size from a single proprietor operation to a multiperson office to a company that has staff located throughout the world.

Layout designers understand the manufacturing requirements and technological limitations imposed by manufacturers, but usually they are not expert in circuit design or signal integrity. Instead, in the absence of specific rules specified by the customer, they create the artwork by using previous experience of apparently similar designs and general rules of thumb.

For this reason it is important that the designer or firm has experience laying out boards in the same technological field as the test or evaluation board. For instance, a firm that customarily designs back planes and high-end server boards may not have meaningful experience in the automotive or medical electronics fields. In a similar way, a layout designer experienced with aerospace electronics circuit boards may not be the best choice to design a low-cost ASIC evaluation card.

Although the layout shop may not have direct experience in signal integrity, often they work closely with signal integrity experts, and commonly the large layout firms have signal integrity specialists on staff or under contract. This can be a helpful resource when designing evaluation and test cards, but in general the layout designer will not seek help from these specialists without prior approval.

It is customary for the layout shop to send the incomplete layout database at regular intervals to the customers so they can track progress and to inspect the artwork. Usually the project manager and the circuit designer will review these snapshots for schedule and design compliance, but the signal integrity engineer should also examine them to ensure that the layout is in compliance with whatever

signal integrity rules have been specified and that it in general conforms to good engineering practice.

4.5 Fabricating and Assembling the Board

The printed circuit fabrication shop (the fab shop) is the manufacture that uses the circuit board artwork designed by the layout shop to create the actual printed circuit board. Generally, the layout design shop will directly send the layout artwork files to the board shop selected by the customer. Often the layout shop can recommend a fab shop with which it has worked that is most suitable for a particular design.

The boards delivered from the fab shop are drilled and punched to accept components such as connectors and some electronic components, but the fab shop does not install components. That task is the responsibility of an assembly shop or of a contract manufacturer.

If requested, for controlled impedance boards the fab shop will supply the customer with the test coupons they used to verify that they had achieved the target impedance value. If the customer has access to a time domain reflectometry (TDR), these coupons can be helpful in determining the actual impedance of a given lot of boards. If requested, the fab shop can also provide test reports showing the trace impedance and line widths.

Like layout firms and circuit board manufacturers, contract manufacturers range in size and in capabilities. Some specialize in specific markets (such as aerospace, automotive, or telecommunications), but the larger firms tend to work in all markets. The largest contract manufacturers provide design and consulting services, which can include schematic capture, signal integrity analysis, circuit board artwork generation, circuit board fabrication, circuit board assembly (attaching components to the circuit board), and electrical testing once all the components have been attached (“final test”).

4.6 Alternative Methods for Design and Fabrication

Using an experienced layout designer specialist and a circuit board fab shop with moderate to high-end processing capabilities remains the lowest-risk way to produce complex or high-performance test and evaluation boards. However, the creation of small test and evaluation circuit boards is now straightforward with the proliferation of easy-to-use circuit board layout software and the increasing number of fab shops willing to manufacture circuit boards in prototype quantities. This provides an opportunity for the signal integrity engineer to design these boards without using a specialty layout shop. Small start-up companies and small fabless semiconductor firms often find this approach particularly attractive.

The cost per board will be reasonably low if the boards have only a few layers (usually 4 or less) and if the board area is no more than roughly 25 square inches (160 square cm). The signal integrity engineer can still lay out and have manufactured boards that are larger than this and that have more layers, but the price of the CAD software necessary to layout these boards will increase, and fewer shops

will be willing to manufacture the board at reasonable cost in small (prototype) quantities.

Although prototype fab shops routinely manufacture circuit boards suitable for surface mount technology (in some cases including BGA technology), usually the lowest-cost shops do not support the most advanced technologies. In general, the low-end shops will require larger line width and wiring pitch design rules than higher-cost shops, and impedance control may not be available as an option.

It is common for a low-end fab shop to specify the available drill and via sizes (including the pad size) from which to choose. In virtually all cases the antipad sizes are automatically determined by the via and pad size, and cannot be changed. This means that by using these shops the signal integrity engineer is constrained to using only those specific sizes and custom via/pad/antipad design intended to produce a specific impedance (discussed in Chapter 14) cannot be designed.

Before selecting a fab shop to make a test or evaluation board, the signal integrity engineer should determine the trace width and pitch requirements for the board. Since not all shops have the same capabilities, selecting the width and pitch will eliminate many shops form consideration, especially if traces and spaces narrower than 5 mils (0.13 mm) are required. Many shops now routinely etch traces with a width of 4 mils (0.11 mm) or less, but 5 mils remain the most narrow that the least expensive shops will fabricate.

4.7 Main Points

- Test boards are a good way to engineering debug and evaluate custom integrated circuits and signal integrity structures.
- Evaluation boards for customer use are an effective way to demonstrate new products.
- Moderately complex test and evaluation boards can be created by the signal integrity engineer without using layout specialists.

References

[1] Parametric Technology Corp., Needham, MA, http://www.ptc.com.

[2] The MathWorks, Inc., Natick, MA, http://www.mathworks.com.

CHAPTER 5

Printed Circuit Board Construction

5.1 Introduction

Understanding circuit board construction and terminology helps the signal integrity engineer discuss trade-offs with board designers and board fabrication shops. Understanding the materials used to form circuit boards (and the way in which they are specified) is crucial when performing signal integrity analysis.

In this chapter we describe characteristics of circuit boards common to all multilayer circuit boards (even though the construction details often differ between fabrication shops), and the terms commonly used in the circuit board industry are introduced and defined.

5.2 How Is a Multilayer Circuit Board Constructed?

Multilayer circuit boards are made by applying heat and pressure to a stackup of fully cured laminate cores and partially cured prepreg mats. Circuit traces are photolithographically patterned and then etched on the copper sheets. A six-layer board stackup is shown in Figure 5.1.

The prepreg is a mat of woven glass fibers saturated with partially cured resin. FR4 grade epoxy resin is most common, but other chemistries are available. Pressing together the stackup under heat fully cures the prepreg. In contrast, a core is a fully cured prepreg mat that has copper sheets (foil) bonded to one or both sides.

The heat and pressure cause the soft prepreg to flow and form around the copper traces patterned on the core. Once cured, the prepreg hardens and forms a tight bond between cores, securing them.

Manufacturers can create the circuit board outer layers on a core (core construction as shown in Figure 5.1) or on prepreg (foil construction as illustrated in Figure 5.2). Because it offers better control and requires less processing, many manufacturers prefer to use foil construction [1].

5.2.1 How Are Connections Made Between Layers?

Holes drilled through the circuit board dielectric materials form tube-like structures called vias (more descriptively, plated-through holes). Although the process varies,

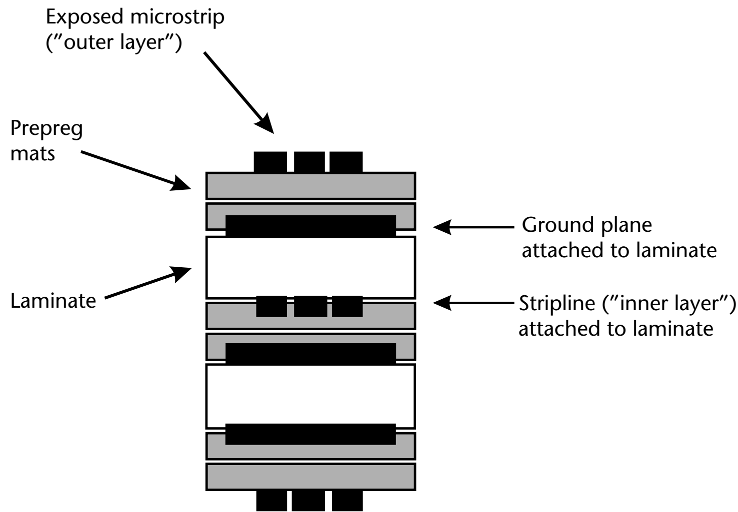


Figure 5.1 A six-layer stackup using core construction. Copper traces or planes are shown in black, and cores are shown in white. Prepreg mats are shaded.

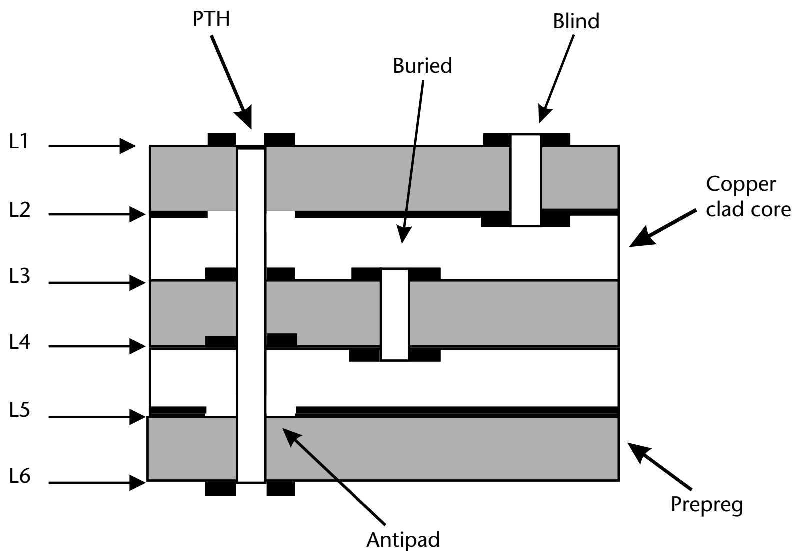


Figure 5.2 Circuit board vias in a six-layer foil construction board. The PTH connects L1 and L4. Note the nonfunctional pad on L3. The buried via connects L3 to L4. The blind via connects layer L1 to L2.

commonly the plating process deposits copper inside the tube, making the plated-through hole (PTH) a conductor.

The PTHs have circular pads (rings of copper) that connect the via to signal traces or power/ground planes. Three types of vias commonly found on circuit boards are shown in Figure 5.2.

This stackup has six copper layers (L1 through L6). Signals are routed on layers L1, L3, L4, and L6. Layer L2 is a ground plane, while L5 is a power plane.

The PTH passes completely through the board but only connects L1 to L4.

Antipads (holes precisely created in the copper planes) allow the PTH to pass through (L2 and L5 in the figure) without connecting.

The nonfunctional pad on L3 does not provide an electrical connection to traces on L3, but it does help anchor the PTH in the stackup. The parasitic electrical effects of these pads can be detrimental to very high-speed signals. Depending on the board thickness, high-end fab shops can selectively remove nonfunctional pads on selected vias.

Blind vias (shown connecting L1 to L2) are PTHs that start on either of the outer layers but do not pass entirely through the board. Blind vias can connect to more than just the two layers suggested in the figure.

The buried via (sometimes mistakenly called a blind via) is shown connecting together signals on L3 and L4. It is a PTH that does not pass completely through the board and is not visible from the outer layers.

5.2.2 What Is the Via Aspect Ratio?

A via that is much longer than its outer diameter can be unreliable and is difficult to manufacture [2]. The via aspect ratio calculated in (5.1) is a way to gauge this difficulty:

$$\text{Via aspect ratio} = \frac{PWB_thickness}{O.D.} \tag{5.1}$$

Some shops call the via aspect ratio the drilling ratio, or simply the aspect ratio. A via aspect ratio of 5 means that the fab shop can reliably produce boards that can be five times thicker than the outer diameter (O.D.) of the smallest via (the pad size is not a factor). For a 10-mil diameter via, this means that the board can be no more than 50 mils thick.

Presently, boards with aspect ratios of 5 or less are considered as having mainstream volume production worldwide. Aspect ratios in the 6–8 range are usually considered *standard production*, ratios in the 9–14 range are considered *advanced* or *low volume*, and ratios greater than about 14 are considered *prototype* or *experimental*.

The signal integrity engineer uses the via aspect ratio provided by the manufacturer and the estimated board thickness to determine the minimum sized via that can be used in a design (for example, when escaping from a small pitch BGA).

5.3 What Is the Significance of Calling a Circuit Board "FR4"?

Although often used to describe the electrical characteristics of a typical circuit board, the term FR4 has little electrical significance. It actually refers to the National Manufacturers Association (NEMA) level of fire retardation.

Ratings FR1 through FR3 are paper-reinforced epoxy resin systems, while ratings FR4 and 5 are glass fabric–reinforced epoxy resin systems with "reduced burning rate" [3].

Table 5.1 shows two important electrical parameters for a small number of the FR4 class laminate systems available to fabrication shops. As discussed in Chapter 7, *Dk* refers to the dielectric constant, and *LT* refers to the loss tangent.

The dielectric constant is different for each vendor's version of FR4, and all the values are lower than the 5.4 worst-case ASTM specification. Because trace width, impedance, and stackup thickness are all related to the dielectric constant, Table 5.1 illustrates why different fabrication shops will produce FR4 boards having very different electrical characteristics.

5.3.1 Why Is There Such Variability in *Dk* Between Laminate Manufacturers?

A material called E glass is most often used for the reinforcing fibers, but some higher-performance laminates use S glass or NE glass because of their lower *Dk* and *LT* values.

Various resins can be used to bind the fibers. As indicated in Table 5.2, the glasses and resins do not have the same electrical characteristics. In fact, the relative amount of glass fibers to resin (the glass-to-resin ratio) determines the *Dk* and *LT* values for a fabricated core or laminate mat.

The relationship between core thickness and resin content for FR4 with E glass as defined in IPC-CC-110 [9] is illustrated in Figure 5.3.

Industry has not universally agreed on the glass-to-resin ratio to use when publishing *Dk* values. This lack of standardization complicates comparing data sheet values; the comparison is only valid when the resin content (and the glass weave) is identical.

Table 5.1 Electrical Characteristics for a Sampling of FR4 Laminate Systems

	Dk	*LT*	*Dk*	*LT*	
Name	*(100 MHz)*		*(1 GHz)*		*Reference*
ASTM D 1867	5.4	0.035	—	—	[3]
FR402	4.6	0.016	4.25	0.015	[4]
PCL226	4.5	0.015	4.3	0.015	[5]
IS410	4.0	0.015	3.9	0.019	[6]

Table 5.2 Electrical Properties at 1 MHz of Circuit Board Glass and Resins [7, 8]

Material	*Dk*	*LT*
E glass	6.2	0.004
S glass	5.2	0.003
NE glass	4.4	0.0012
FR4 epoxy resin	3.6	0.032
BT resin	3.1	0.003
Polyamide resin	3.2	0.02
Cyanate ester resin	2.8	0.002

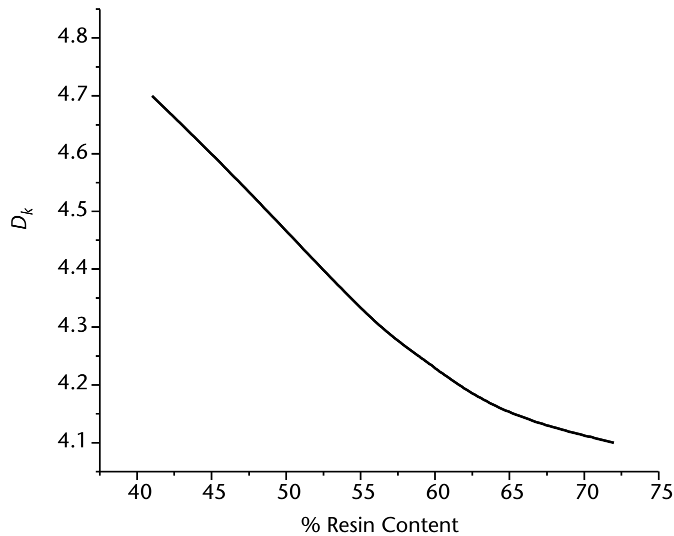


Figure 5.3 FR4 dielectric constant depends on the glass-to-resin ratio. Because the resin has a much lower *Dk* than E glass, *Dk* falls as the amount of resin increases. Raw data from [9].

5.3.2 Why Is the Same Board Produced Differently by Various Fab Shops?

Sending the same artwork to multiple circuit board fabricators does not guarantee that each shop will produce boards having the same electrical characteristics. In fact, unless the board is very simple, it is nearly certain each shop will use a unique combination of prepreg and core thicknesses.

For example, one shop might choose to use two thick prepregs with a 42% resin content to form a layer, while another shop might use four thinner sheets each having a 53% resin content to obtain the same overall layer thickness. As illustrated in Figure 5.3, this difference in resin content changes the *Dk* from 4.65 to 4.35. As explored in later chapters, this difference affects the signal time of flight and the trace impedance (and possibly the trace width).

5.3.3 What Are Some Alternatives to FR4?

Manufacturers have created a variety of laminate systems, which are electrically and mechanically (but sometimes only mechanically) superior to common FR4 resin systems. Table 5.3 shows a specific product offering from various vendors for the chemistries appearing in Table 5.2.

Megtron 6 is a PPO (polyphynoleneoxide) chemistry offered by Panasonic [10]. Generally, the *Dk* and *LT* values for PPO-based laminates are less affected by frequency than standard FR4 systems. G200 is a BT (bismaleimide/triazine blend)–based chemistry from Isola [11]. The higher temperature capabilities and lower moisture uptake make these laminates attractive to the micropackaging industry. Nelco N7000-1 [12] is a polyimide-based system. Polyamide circuit boards can withstand high temperatures and are often used in commercial test equipment and in aerospace applications. Nelco N8000 is a cyanate ester [13] system offered by Park Electrochemical Corporation. Laminates based on these chemistries are often

Table 5.3 A Sampling of Alternate Laminate Systems

Chemistry	*Product Name*	*Dk/LT at 10 GHz*
PPO	Megtron 6	3.4/0.004
BT	G200	3.65/0.015
Polyamide	N7000-1	3.8/0.016
Cyanate Ester	N8000	3.5/0.011
Ceramic	RO4003	3.4/0.0022

used in RF applications. The RO4000 series of materials from the Rogers Corporation [14] are reinforced hydrocarbon/ceramics that have low and stable *Dk* values. Theses are popular in high-performance digital and RF applications.

5.4 Circuit Board Traces

The copper conductors used to interconnect circuits (the traces) are formed either on the boards surface (the outer layers) or buried within the circuit board material (inner layers). Outer layer traces are called microstrip (or microstripline); inner layer traces are called stripline.

The four basic topologies used to interconnect digital circuits are shown in Figure 5.4. Microwave and RF designers use additional topologies that are not shown.

The outer layer traces can either be exposed microstrip or embedded microstrip, and the inner layers can be either stripline or dual stripline.

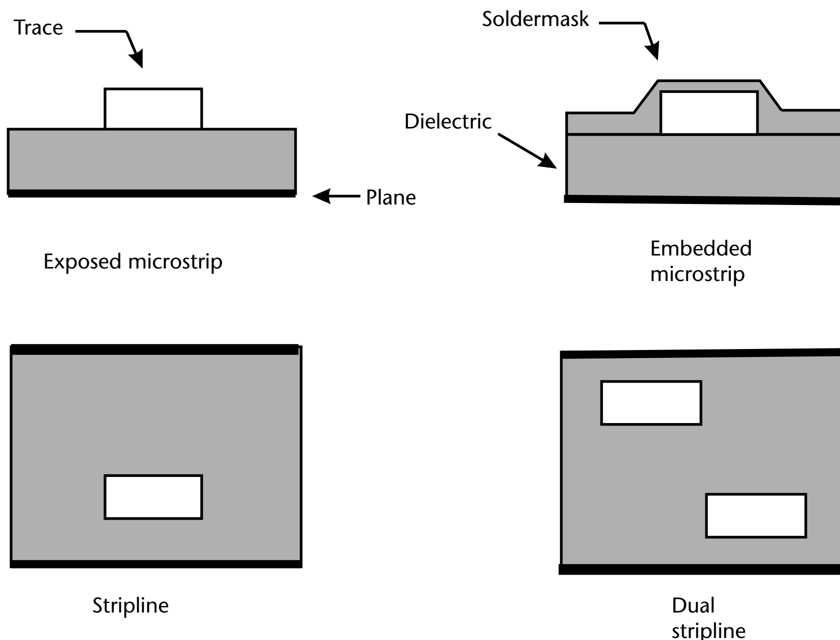


Figure 5.4 Types of circuit traces used to connect digital circuits. Microstrips are formed on the outer layers, while inner layers create striplines.

5.4.1 Microstrip

Exposed microstrip is formed when a copper trace is separated from a metal plane by a block of dielectric material. The trace forms a transmission line with the plane (the return path), which is connected to either power or ground. In exposed microstrip the copper trace is directly exposed to the environment. It is not covered by an encapsulate or other nonconducting sealant materials.

This contrasts with *embedded microstrip*, which is a microstrip covered by an insulating material such as solder mask. On commercial digital logic circuit boards, virtually all microstrips are covered with solder mask, making them embedded. Although commonly called *microstrips*, Figure 5.4 shows that they are actually embedded microstrips.

Solder mask is an insulating material used to prevent solder bridges forming between traces. Coating the traces with this dielectric material lowers the trace impedance and increases the trace capacitance and delay time. To obtain a specific impedance, the manufacturer will adjust the trace parameters (generally by making the trace less wide) to compensate for the solder mask.

Microstrips are made thicker than the base copper layer as part of the plating process used to create vias. The amount of plating necessary is determined by factors associated with forming the via, and the final microstrip thickness is not usually the primary consideration for the manufacturer. An example appears in Figure 5.5.

5.4.2 Stripline

Stripline traces are sandwiched between two metal planes. A block of dielectric material separates the planes and trace. Often striplines are thought of as being centered between the planes, but usually this is not the case in practice. Instead, because of the way in which circuit boards are manufactured, offset stripline (where the trace is located closer to one of the planes) is the norm.

It is possible to sandwich two traces between the planes rather than just one. Such dual stripline increases routing density, but can produce excessive noise coupling (crosstalk, discussed in Chapter 10) between the traces.

By routing the traces parallel to and directly underneath one another, the coupling can be intentionally accentuated, and dual stripline can form a broadside coupled differential pair (discussed in Chapter 13).

5.4.3 What Is a Mil and What Does a Trace Thickness in Ounces Mean?

It is common practice in North America to define the copper trace width in mils and the thickness in terms of copper weight, which is measured in ounces per square foot. A mil is 0.001 inch and is equal to 25.4 μm (0.0254 mm).

Common copper thickness values are 1/2-, 1-, and 2-ounce copper, but others are available. Table 5.4 shows the minimum copper thickness in mils and µm required by IPC-2221 [15] for inner and outer layer traces.

Table 5.4 shows that a half-ounce trace is nominally 0.67-mil (17 μm) thick. It must be at least 0.47-mil (12 μm) thick when on an outer layer (such as microstrip) and 1.3-mils (33 μm) thick when on an inner layer (such as stripline).

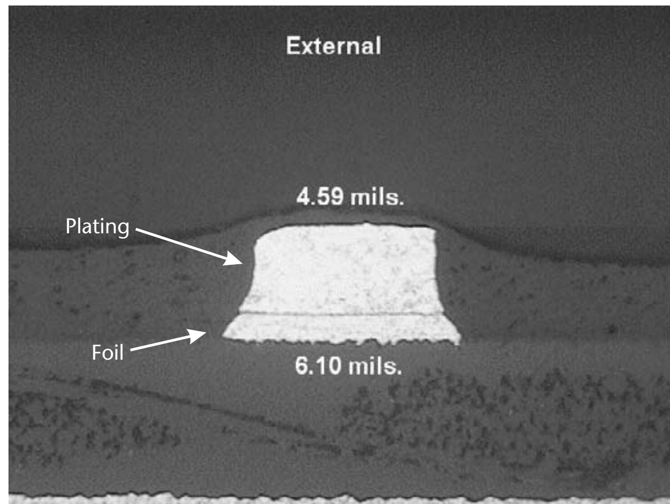


Figure 5.5 Outer layer trace is made thicker than the base metal by plating. (Courtesy of TTM Technologies, Inc. Used with permission.)

Table 5.4 Minimum Copper Thickness by Weight Per IPC-2221 [15]

Copper Foil Weight	*Nominal Thickness Mils (μm)*	*Internal Layer Mils (μm)*	*External Layer Mils (μm)*
1/8		0.14 (3.5)	0.79 (20)
1/4	0.35 (9)	0.24 (6)	0.87 (22)
3/8		0.31 (8)	0.98 (25)
1/2	0.67 (17)	0.47 (12)	1.30 (33)
1	1.38 (35)	0.98 (25)	1.81 (46)
2	2.76 (70)	2.20 (56)	3.00 (76)
3		3.58 (91)	4.21 (107)
4		4.80 (122)	5.39 (137)

Not all weights have the nominal thicknesses specified.

Although a maximum thickness is not defined, to manage material costs and to improve etching, the internal layer copper is usually not much thicker than the IPC nominal values.

Typically after processing, the half-ounce copper used in internal layers are 0.6 ± 0.1-mils thick and are 1.24 ± 0.25-mils thick for 1-ounce copper [16].

From experience, outer layer half-ounce copper thickness ranges from 1.8 mils to 2.7 mils, depending on the board thickness. Other studies [16] have found the average to be between 2.1 mils and about 2.6 mils. That same study showed that 1-ounce copper had an even greater variation, ranging from 1.3 to 3.5 mils, depending on the vendor.

5.4.4 What Copper Types Are Used to Form Traces and What Are Their Characteristics?

The copper foil used on circuit boards is created by either an electrodeposition or rolling process. The end result from either process is very pure copper foil, but the foil's mechanical and high-frequency electrical characteristics are quite different.

Electrodeposited (ED) foils are formed on a drum rotating in a solution containing copper [3]. This leaves one side of the foil smooth and the other side rough. The rough side of these foils attach to the laminate. This foil is the most commonly used in commercial circuit boards. *Double treat* foils roughen both sides. *Reverse treat* foils roughen the normally smooth side of the foil. With these foils the smooth side attaches to the laminate. These configurations are illustrated in Figure 5.6.

In contrast, rolled (also called wrought) foils are created by squeezing copper sheets through closely spaced rollers. This hardens the copper and crushes the copper crystals, making the copper smooth on both sides. In fact, rolled foils are smoother and denser than ED foils. For this reason the bulk resistivity of rolled foils is lower than that of ED foils [17], as is apparent in Table 5.5. In later chapters this information is used to determine the DC resistance and AC loss of traces.

Although the majority of ridged circuit boards for the commercial market use ED foils, rolled foils are often used in flex circuitry [10] and in those situations where there is a large temperature swing (such as space and avionics applications) [17]. As shown in Chapter 8, high-frequency loss depends in part on the smoothness of the copper surface. For this reason, rolled foils are often used in RF work.

5.4.5 What Is Surface Roughness and What Are Typical Values?

To promote adhesion, the copper foil is mechanically roughened and various adhesion, barrier, and stabilization layers are added [19]. These steps are necessary

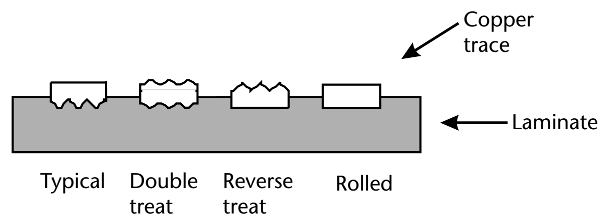


Figure 5.6 Roughness of various copper foils, exaggerated for clarity. Typical foils are rougher on the laminate side, while double treat is rough on both sides. Reverse treat attaches on the smooth side, and rolled copper is smooth on both sides.

Table 5.5 Resistivity in $\mu\Omega$-cm for Copper Traces [18]

Weight Ounces (μm)	*ED*	*Rolled*
1/2	1.82	1.78
1	1.78	1.74
2	1.78	1.74

because a trace (especially if it is narrow) would easily peel away from the substrate without peaks and valleys in the copper to grab hold and anchor it to the dielectric.

To improve peel strength on commercial circuit boards, ED foils with a more pronounced tooth profile are usually used on outer layer traces, while smoother ED foils are usually used on inner layers [20]. However, to promote adhesion with the prepreg, the top surface of inner layer copper is sometimes "micro-roughened" [21] as well.

Surface roughness is specified by foil manufactures in terms of the maximum valley to peak height of the ridges, or as R_a, their RMS average sampled over a region. Their amplitude and shape depend on how the copper is formed and on the surface treatments used to further roughen the surface.

As shown in Table 5.6, average surface roughness values for electrodeposited copper range from 0.7-μm RMS to as much as 3-μm RMS [18–22]. This wide range reflects the diversity of copper foil types and roughening techniques used in the industry. Copper foil adheres better to some laminates than others, so the copper used with alternate laminate systems is sometimes made rougher than that used in conventional FR4 systems [9].

5.4.6 Are Traces Rectangular?

Etched traces are not usually rectangular in shape because the etching solution attacks the copper both vertically and horizontally. This results in the traces being more trapezoidal than rectangular. This was first shown in Figure 5.5 for an outer layer trace, but as shown in Figure 5.7, the effect also happens on inner layer traces.

Figures 5.5 and 5.7 show the trace as wider at the bottom than the top, but it is possible for the trace to be wider at the top than the bottom. Because etching times are shorter, thinner copper generally tends to produce more rectangular traces than thicker copper. Similarly, rough copper requires more etching time than smooth copper, so smooth copper traces tend to be more rectangular than rougher traces.

5.5 Main Points

- Multilayer circuit boards are created by fusing together prepreg mats and laminate cores.
- The glass-to-resin ratio is an important factor in determining trace impedance and delay.
- Microstrips are formed on the circuit board outer layers.

Table 5.6 Typical Average Roughness Values for Treated Copper Foil [18–22]

Copper Weight (Ounces)	*R_a ED (μm, RMS)*	*R_a Rolled (μm, RMS)*
1/2	0.7–1.9	0.3–1.4
1	1.0–2.4	1.2–1.4
2	1.1–3.0	1.4

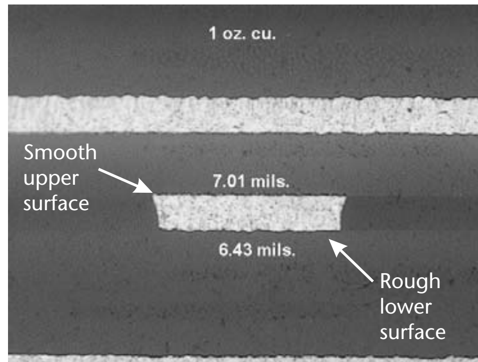


Figure 5.7 Cross-section of a 1-ounce trace. Etching causes traces to be trapezoidal in shape. The roughness of the copper surface is also apparent. (Courtesy of TTM Technologies, Inc. Used with permission.)

- Striplines are formed on inner layers.
- The traces are generally trapezoidal rather than rectangular in shape.
- FR4 is a level of fire retardant and does not define the electrical characteristics of a board system.
- Advanced laminate systems can use resins other than the traditional FR4 systems.
- The plating process creating vias makes the shape and thickness of microstrip traces less predictable than stripline traces.
- In North America, trace thickness is specified in terms of copper weight in ounces.
- Traces are usually formed with electrodeposited copper foils. Generally, these foils are rougher than the rolled foils used in microwave work.

Problems

Answers to these problems are available on the Artech House Web site at http://www.artechhouse.com/static/reslib/thierauf/thierauf1.html.

5.1 The signal integrity engineer reviewing your completed FR4 circuit board informs you that the conductor loss is too high on some high-speed nets. What manufacturing options should you consider?

5.2 In reviewing signal integrity models, you notice that the half-ounce stripline and microstrip traces were assumed to be 0.65-mil thick. Is this valid?

5.3 Prototype quantities of your circuit board have been manufactured by a high-end production shop, but once production begins, the design will be transferred to a manufacturer more accustomed to high-volume production. What are the signal integrity risks?

5.4 Given the scenario described in Problem 5.3, what signal integrity tests should be performed on the volume production boards?

5.5 Based only on the via aspect ratio, what is the smallest sized through-hole via that you should expect in standard production if your circuit board is 0.092-inches (2.3 mm) thick?

References

[1] Merix Corp., *Design for Manufacturability of Rigid Multi-Layer Boards*, Revision 7/99, Forest Grove, OR, July 1999.

[2] Goval, D., "Reliability of High Aspect Ratio Plated Through Holes (PTH) for Advanced Printed Circuit Board (PCB) Packages," *Annual Proceedings, IEEE International Reliability Physics Symposium*, Denver, CO, April 8–10, 1997, pp. 129–135.

[3] Shugg, W. T., *Handbook of Electrical and Electronic Insulating Materials*, 2nd ed., New York: IEEE Press, 1995, pp. 270–271.

[4] "Isola FR402," Isola Group, Chandler, AZ, data sheet #5023/5/02, Rev 3/16/07.

[5] Polyclad Laminates, Inc.," PCL-FR-226/PCL-FRP-226," data sheet PCL-226, Isola Group, Chandler AZ, October 2001.

[6] "Isola IS410," Isola Group, Chandler, AZ, data sheet IS410 08-13-08.

[7] Thierauf, S. C., *High-Speed Circuit Board Signal Integrity*, Norwood, MA: Artech House, 2004.

[8] Coombs, C. F., Jr., *Printed Circuits Handbook*, 6th ed., New York, McGraw-Hill, 2008.

[9] Jawitz, M. W., *Printed Circuit Board Materials Handbook*, New York, 1997.

[10] "DATA SHEET Megtron 6," Panasonic PCB Materials, Santa Ana, CA, undated.

[11] "Isola G200," Isola Group, Chandler, AZ, data sheet, Rev 5-16-08, dated 2006.

[12] "Nelco® 7000-1," Nelco Advanced Circuitry Materials, Park Electrochemical Corp., data sheet Rev 2-07 (undated).

[13] "Nelco® 8000, Nelco N8000Q," Nelco Advanced Circuitry Materials, Park Electrochemical Corp., data sheet Rev 2-07 (undated).

[14] Rogers Corp., "RO4000 Series High Frequency Circuit Material Data Sheet," No. 92-004, Advanced Circuit Materials, Chandler, AZ, 1992.

[15] Institute for Interconnection and Packaging, *Electronic Circuits Generic Standards on Printed Board Design*, IPC-2221 2215, Northbrook, IL, February 1998.

[16] Brist, G., "Design Optimization of Single-Ended and Differential Impedance PCB Transmission Lines," *PCB West Conference Proceedings*, 2004.

[17] Merix Corporation, "Electrodeposited vs. Rolled Copper," Merix Corporation, www.merix.com.

[18] Rogers Corp., "Copper Foils for Microwave Circuits," Publication #92-243, dated 2/00, Microwave Materials Division, Chandler, AZ.

[19] Hilburn, R., et al., "The State of Copper," *Printed Circuit Design & Manufacturer*, May 2005, pp. 24–29.

[20] Brist, G. et al., "Non-Classical Conductor Losses Due to Copper Foil Roughness and Treatment," *ECWC 10/IPC/APEX Conf.*, Fall 2005.

[21] Lee, B., "The Impact of Inner Layer Copper Foil Roughness on Signal Integrity," *Printed Circuit Design & Manufacturing*, May 2007, pp. 27–29.

[22] Normyle, S. J., et al., "The Impact of Conductor Surface Profile (Rrms) on Total Circuit Attenuation in Microstrip and Stripline Transmission Lines," White Paper, Taconic Advanced Dielectric Division, Petersburg, NY, undated.

CHAPTER 6

Transmission Line Fundamentals

6.1 Introduction

In this chapter we use a circuit analysis approach to introduce fundamental transmission line concepts. This approach is more intuitive than field theory, but [1–3] should be helpful for those interested in exploring transmission lines in that way.

We begin by defining a transmission line and its return path and developing a transmission line circuit model. We will restrict our study here to transmission lines having low losses, because this lets us develop an intuitive understanding without the mathematics necessary when losses cannot be ignored. Lossy transmission lines are discussed in Chapter 8. We discuss return paths in some detail because return path management is an important aspect of signal integrity analysis. We conclude the chapter by developing rules of thumb that define when a conductor acts as a simple lumped circuit or when the circuit elements are distributed.

6.2 What Is a Transmission Line?

As illustrated in Figure 6.1, a transmission line is formed with two or more conductors. The signal conductor carries the signal energy from a generator to a load, and the second conductor (the return) completes the circuit by returning the signal current back to the generator. The signal conductor could be a microstrip trace and the return path the underlying ground plane. The generator is any signal launching device, such as an ASIC I/O driver.

6.2.1 What Is the Signal Return Path?

Frequently, signal integrity engineers speak of the "signal return path," but often it is taken for granted and not fully understood. Studying the return path is important because it determines the trace's capacitance, inductance, resistance, and noise coupling.

Because capacitance is often more intuitive than inductance, we will ignore inductance for now and focus on the importance of the return path in determining the capacitance. Resistance is also affected by the return path, but we will wait until Chapters 7 and 8 to discuss this important topic.

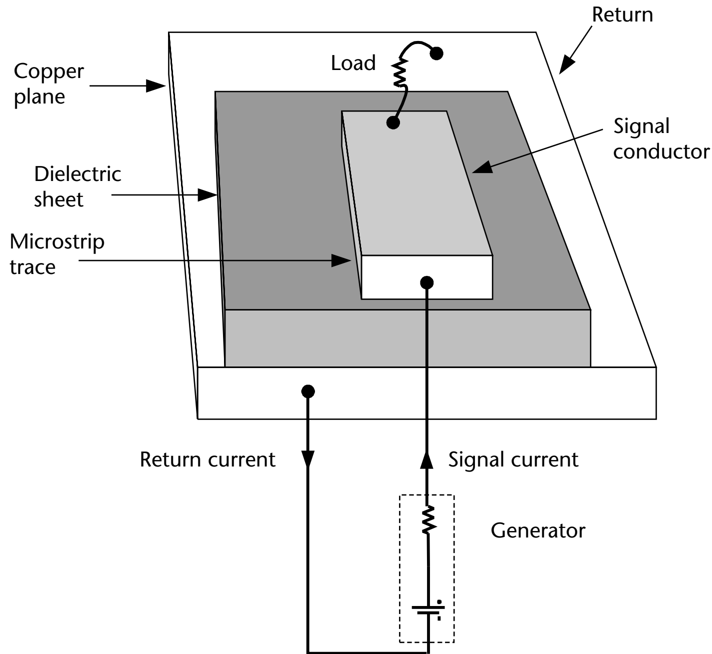


Figure 6.1 A transmission line created by a signal trace and its return.

The side view of a trace over a ground plane is shown in Figure 6.2. We can see that the trace is one plate of a capacitor with the second plate being the return path (such as a ground plane). The distributed inductance is formed by the current loop area between the trace and ground plane. The capacitance and inductance values are determined by the width of the trace and the distance to the ground plane.

As we will see, the inductance and capacitance values determine the trace impedance. This is why an impedance value is really only meaningful when it is quoted with respect to a specific return plane. We will return to this later in this chapter and somewhat more indirectly in Chapter 7.

6.3 Circuit Model of a Transmission Line

A circuit model for a small segment of trace is shown in Figure 6.3. The trace has a resistance (R) and an inductance (L). The capacitance formed between the trace and ground plane is modeled by capacitor C, and the capacitor losses are modeled

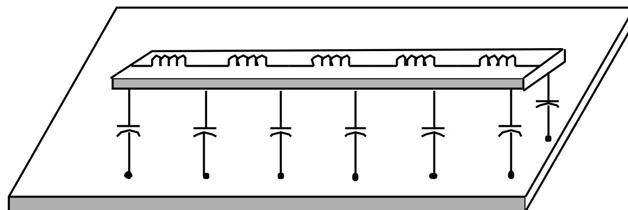


Figure 6.2 Distributed parallel plate capacitors and loop inductance formed between the signal trace and the ground plane return path.

by conductance G. In Chapter 7 we discuss typical values for these parameters, and in Chapter 17 we show how to calculate them. We will avoid this level of detail here and note that the entire trace is made up of an infinite number of identical copies of these tiny segments. In fact, the segments are so small they blend together smoothly. The resistors, inductors, and capacitors are distributed evenly along the length of line.

The circuit in Figure 6.3 is a constant-k lowpass filter, and circuit theory can be used to analyze its behavior. Taking this approach lets us develop a circuits-based intuitive understanding, but in the following chapters we will use some concepts from field theory when explaining the different characteristics of various types of circuit board transmission lines.

The circuit has the following characteristics:

- *It imposes a delay.* Time is required for the signal to travel from one end of the line to the other. This is called the transmission lines propagation delay time, tpd.
- *The line possesses an impedance.* The current launched down a such a line is limited by the transmission lines characteristic impedance, Z_o.
- *The line exhibits loss.* The signal is attenuated and, if the loss is severe enough, the signal will become distorted. This is discussed in Chapter 8.

6.4 Impedance and Delay

By using circuit analysis [4–6] and ignoring the resistance and conductance (R and G), the characteristic impedance and the delay of a long chain of RLCG circuits such as Figure 6.3 are found to be:

$$Z_o = \sqrt{\frac{L}{C}} \tag{6.1}$$

$$tpd = \sqrt{LC} \tag{6.2}$$

Impedance (Z_o) has units of ohms and it has that value for any length of line. For instance, a transmission line may have an impedance of 75Ω, but it does not

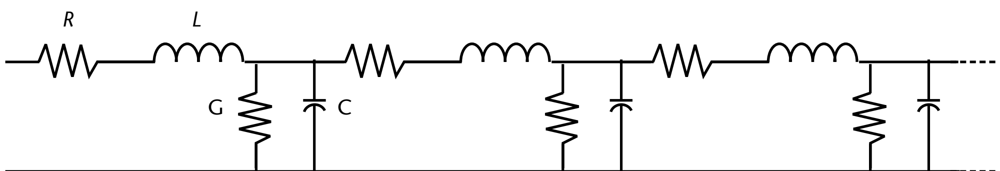


Figure 6.3 Trace circuit model. The trace's distributed circuit elements are represented by a string of closely spaced lumped elements.

have an impedance of 75Ω per meter. This is different from the resistance, where the transmission line length does matter. For example, a trace might have a resistance of 0.1 Ω/inch.

The delay (*tpd*) has units of time (typically picoseconds or nanoseconds per unit of length. A 10-inch-long trace having a delay of 180 ps per inch has a total delay of 1.8 ns.

In Section 6.8 and the Problems we explore why the impedance does not depend on the lines length but the time delay does.

Equations (6.1) and (6.2) are only true when the effect of the resistance (R) is very small compared to the effect of the inductance (L), and when the capacitor losses (G) are very small compared to the effects of the capacitance (C). In fact, these equations calculate the lossless impedance and lossless delay time because the losses in the conductor (represented by R) and in the capacitor (represented by G) are assumed to be small enough to drop out of the general equations for impedance and delay. This is a good assumption for circuit board transmission lines carrying signals with rise times of about 10 nS or faster [7].

For instance, the effect of the inductance is some 25 times higher than the effect of the series resistor for a pulse having a 1-ns rise time traveling down a 50Ω, 5-mil (127 μm)-wide half-ounce stripline trace on FR4. In contrast, the effect of the inductance is only about eight times larger than the resistance when the rise time increases to 10 ns.

6.4.1 Units and Electrical Length

The units of length associated with the values for L and C are important when using these equations. For instance, field solvers usually give these values per unit length (sometimes referred to as PUL), such as picofarads per centimeter or nanohenry per inch. Total values for the entire length of the line are also sometimes given.

As an example, a 50Ω transmission line might have a PUL capacitance of 3 pF per inch and an inductance of 6.5 nH/inch. A 10-inch-long line would have a total capacitance and inductance equal to the PUL value times the total length: 30 pF and 65 nH.

The electrical length of a transmission line having PUL values of 3.5 pF and 6.4 nH is 150 ps. It increases to 1.5 ns when the PUL values are 10 times as large. We will see the advantage of looking at transmission lines in this way later on in this chapter and in Chapters 11 and 12 when we discuss reflections and terminations.

6.5 How Does a Signal Travel Down the Line?

Because we took results from circuit theory rather than field theory, we have not seen that energy travels as waves along transmission lines. We can develop an intuitive understanding of this by thinking about the transmission line circuit shown in Figure 6.4 in the following way [8].

The circuit shows the resistance, inductance, and capacitance of a microstrip over a return path such as a ground plane. For simplicity, the capacitor losses (the G element in Figure 6.3) are not included.

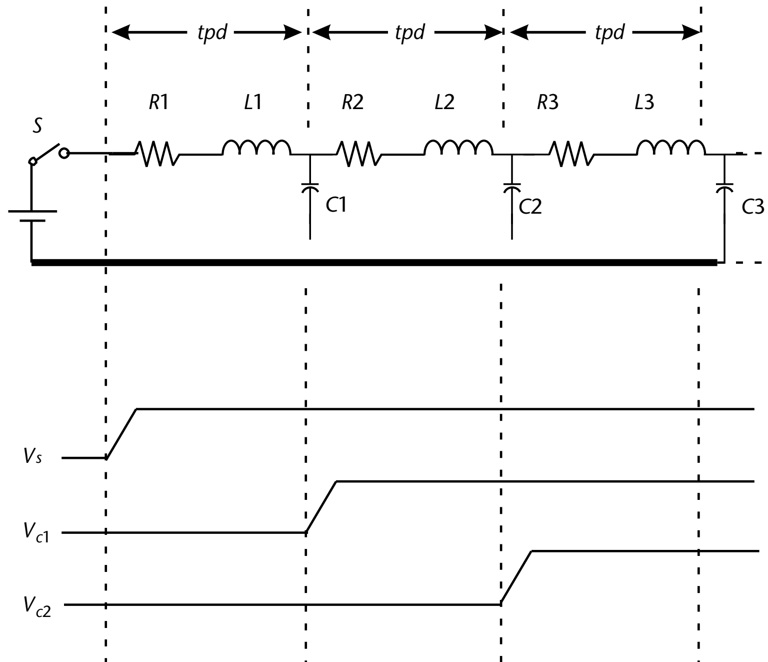


Figure 6.4 Applying voltage to transmission line causes inductors and capacitors to charge in a wave-like manner.

We will start off by first assuming that the switch is off, that there is no current flowing in the circuit, and that the capacitors have all been discharged. We will also assume that the line is infinitely long and so there are no reflections.

Closing the switch causes current to flow through $L1$ and $R1$, charging capacitor $C1$. However, the inductor and resistor prevent the capacitor from charging up instantaneously. There is a time delay equal to *tpd* between the switch closing and $C1$ being charged.

Once the delay is over, $C1$ is fully charged and current can now pass through $L2$ and $R2$ to charge $C2$. After a delay of *tpd* seconds, $C2$ is fully charged and then $C3$ starts to charge. This process continues until all of the capacitors along the line have been charged.

By measuring the voltage across each capacitor, it appears as if a voltage wave is traveling from left to right in the figure. By monitoring the current through the inductors, it appears that a current wave travels along with the voltage wave.

This description of line behavior is simplistic. Rather than waiting for $C1$ to become fully charged, current begins to flow in $L2$ while $C1$ is still charging, and $C2$ can start to charge $C3$ before $C2$ is fully charged, and so on. This means that the waveform gradually works its way down the line rather than doing so in steps as we have described. However, if the segments are made small enough, the steps become very small, and the observed behavior appears as a smooth wave traveling down the line.

6.5.1 What Voltage Gets Launched Down a Transmission Line?

Provided the line has low losses (which means that R and G in Figure 6.3 can be ignored), the transmission line impedance appears as a simple resistance. The voltage and current waves at the input of an infinite length of line obey Ohm's law, and will have a ratio equal to Z_o:

$$Z_o = \frac{V_i}{I_i} \tag{6.3}$$

where the launched voltage is called V_i, the incident voltage, and the launched current is called I_i, the incident current.

Equation (6.3) cannot be used when the impedance changes along the line or at its end, which gives rise to signal reflections. Signal reflections are waves that travel along the line independently of the launched wave. These waves combine in complex ways, but for lines with no reflections, (6.3) shows that the characteristic impedance determines the relationship between the voltage and current along the line.

Because the impedance of a lossless transmission line acts like a simple resistance, the resistance voltage divider principle can be used to find the launched voltage if the power supply voltage (V_{cc}) is known. This is illustrated in Figure 6.5 and in (6.4), where the impedance of the driver (such as an ASIC I/O driver) is Z_g, and the transmission line characteristic impedance is Z_o.

$$V_i = V_{cc}\frac{Z_o}{Z_o + Z_g} \tag{6.4}$$

This model shows how the power supply voltage, line impedance, and I/O driver impedance interact:

- For a given driver impedance and V_{cc} value, a higher voltage will be launched if the transmission line characteristic impedance increases. This can be used to improve the received noise margins by increasing the voltage at the receiver, but care must be taken to prevent excessive reflections and overshoots.

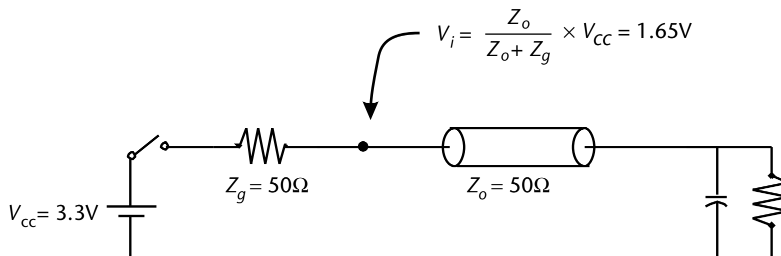


Figure 6.5 Voltage divider between the generator and transmission line impedances determines the launched voltage (V_i). The line is long enough that the load is not a factor.

- The actual value of V_{cc} is important in determining the launched voltage. Sags on V_{cc} caused by many drivers switching all at once (simultaneous switching noise, described in Chapter 1) result in lower launched voltages. This can lead to reduced receiver noise margins, and because the amount of sag depends on the number of lines switching, the reduction in noise margin becomes data dependent.
- A higher voltage will be launched when low output impedance output drivers are used. This may initially increase the received voltage, but care must be taken because drivers with too low of an impedance will cause unacceptably high reflection and overshoot voltages.

6.6 How Does the Current in the Return Path Behave?

We have seen how current flows down the transmission lines' signal conductor, but we have not seen how the current flows in the return.

We might think that the signal must first reach the load before current begins to flow in the return, but in actuality current starts flowing in the return path as soon as the signal is launched down the signal line.

We can see this by examining the microstrip shown Figure 6.6 and following a current pulse (i) as it travels down the signal line. The current pulse reaches the first capacitor, causing it to charge. To complete the circuit, the capacitor's displacement current flows in the return path, back to the source. This occurs immediately, before the signal has reached the load. Displacement current flows again when the signal current reaches the next capacitor, and the process continues as the pulse makes its way down the line. By monitoring the return path we observe that return current continues to flow as each successive capacitor is charged. Because the distance between each capacitor is so small, a steady, uninterrupted current flows in the return path. As illustrated by the arrows, this current moves in lockstep with the signal current, but because it is flowing in the opposite direction it appears to be a mirror image of it.

6.6.1 How Does Current Flow in Return Paths That Are Not Ground?

In the previous example the return path was a single plane, tied to ground, but the return could have been a power plane rather than ground. The analysis for that situation is not presented here, but is similar to the analysis presented next for striplines when one of the return paths is ground and the other is a voltage plane.

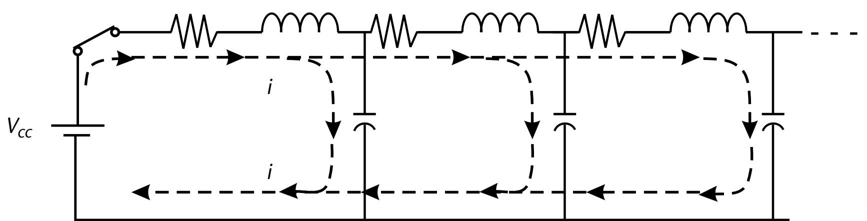


Figure 6.6 Current in the return path is coincident with current i flowing in the signal trace.

The power plane can either be connected to the same supply powering the I/O driver, or it could be an unrelated supply that powers the ASIC core or other logic on the board. As we will see, the return paths are very different in these situations.

6.6.2 What Is the Behavior When One Return Path Is a Related Power Plane?

An ASIC driving a stripline is shown in Figure 6.7. The I/O driver drives the line to a logic high by closing the switch and connecting the resistance (which represents the on resistance of the driver) to the ASIC I/O power supply.

The stripline is perfectly centered between a ground plane and a power plane that connects to the ASIC I/O driver. This makes the capacitance (and inductance not shown) from the trace to the ground plane identical to the capacitance (and inductance) from the trace to the power plane.

This is not necessarily the case in practice. For example, with offset stripline the trace is closer to one plane than the other, increasing the coupling to that plane and decreasing it to the other plane. We can develop our explanation with that configuration, but to simplify things, we will assume the capacitances (and the inductances) to each plane are equal.

We start by assuming that the driver output has been at 0V long enough to fully discharge all the capacitors between the trace and ground. Because the line is at 0V, the capacitors formed with the power plane are all charged to the plane voltage.

From Figure 6.7 we see that once the driver switches to a logic high it launches 1.5V down the 50Ω transmission line. Solving for I_i with (6.3), we find that the launched current is 30 mA.

Since the capacitances are equal, the same amount of return current flows in the two planes. As expected, 15 mA flows in the ground plane back to the driver. Figure 6.7 shows that the power plane current is actually the result of the driver discharging the upper capacitors, which were charged when the line was held at 0V. The discharge path is through the ASIC power rail to the pull-up in the driver and out to the line.

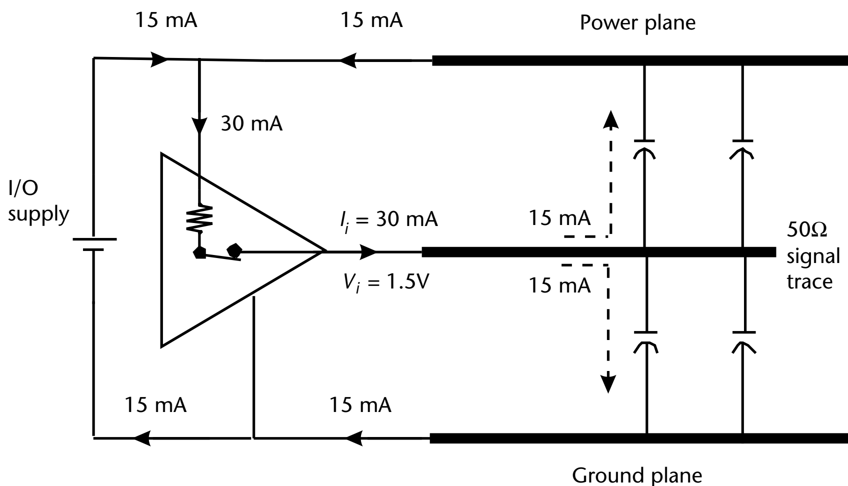


Figure 6.7 Return path current when stripline is formed between a power plane connected to the I/O driver and a ground plane.

This analysis suggests that to properly manage the return currents it is not necessary to decouple the I/O supply. In fact, to achieve the proper trace impedance, enough capacitance must be placed between the power and ground planes so that the pair has an impedance very much less than the parallel impedance of all the traces sharing the return. For instance, across frequency the decoupling should have an impedance well under 0.5Ω if ten 50Ω traces were routed as in Figure 6.6. This is explored in the Problems.

6.6.3 How Does the Return Current Flow When the Power Supply Is Not the Signaling Voltage?

In Figure 6.8 we see a 50Ω signal trace sandwiched between a ground plane and a power plane. In this situation the power plane does not connect to the pins on the ASIC powering the I/O drivers. Instead, the stripline is formed with a plane carrying an "unrelated voltage," such as the power supply for the ASIC's internal logic. For instance, the internal ASIC logic supply might be 1.8V and the I/O supply is 3.3V.

Forming the stripline with the 1.8-V internal logic plane causes the capacitors (and the inductors, not shown) to be created between that power plane and ground. Once again, we will assume that the trace is perfectly centered between the two planes so that the capacitance to each is identical.

Figure 6.8 illustrates that, as in the previous case, the launched current evenly divides between the two return planes. However, in this case, to discharge the upper capacitors, the 15 mA flowing in the power plane must flow through both power systems before it can make its way back to the driver.

In general, the power supplies are resistive, and for frequencies above a few tens of hertz, they are also inductive. The inductance prevents them from being an effective path for high-frequency return current. In practice, the current mostly returns by way of the circuit board's interlayer and decoupling capacitors. Interlayer capacitance is the parallel plate capacitance naturally formed between a power and ground plane. Each of the planes forms one of the capacitor plates, and the laminate is the dielectric.

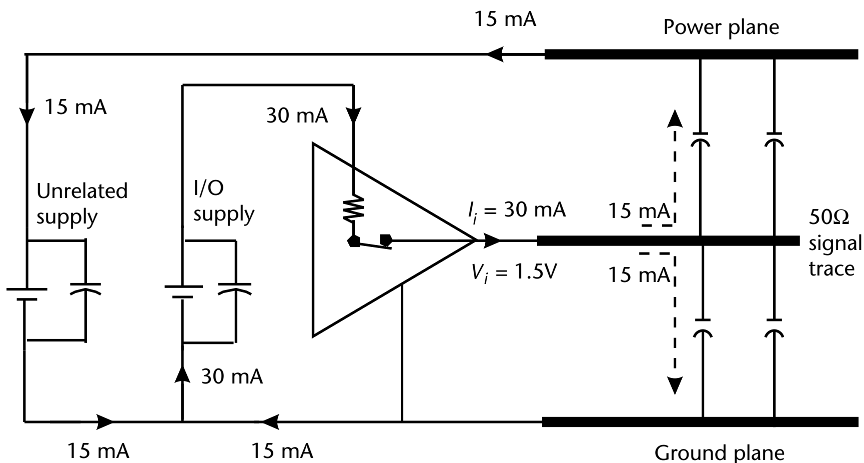


Figure 6.8 Return path current when stripline is formed when using a power plane not connected to the I/O supply. Adequate decoupling capacitance is necessary to properly steer the return current.

Generally, the path through the interlayer capacitance is low inductance, but the capacitance is small (often in the tens of nanofarads range). This means that in most cases the signal return currents flow through the decoupling capacitors explicitly placed between the power and ground planes. Unfortunately, the path through the decoupling capacitors is inductive, which raises their impedance and makes them increasingly less effective at high frequency. This is an important reason why using an unrelated power supply to form a transmission line is poor practice and should be avoided, especially when working with fast rise time signals since these signals have high frequency harmonics that make achieving a low impedance AC short difficult. As illustrated in the Problems, this means that the decoupling requirements for Figure 6.8 are much more stringent than those of Figure 6.7.

6.7 When Does a Conductor, Via, or Connector Pin Act Like a Transmission Line?

As a rule of thumb, interconnect will act as a transmission line when the electrical length of the line [as given in (6.2)] is longer than half the signal rise time [9–11]. Here we refer to the signal rise time to mean either the rise or the fall time, whichever is the fastest. The relationship is shown in:

$$\sqrt{LC} \geq 0.5t_r \tag{6.5}$$

For instance, a line having a delay of 150 ps would be considered electrically long (and so would exhibit transmission line behavior) when carrying signals having rise times of 300 pS or less.

More conservative estimates are limits of 0.2 [6] to 0.125 [12] rather than 0.5. From (6.5) we see that:

- It is the signal's rise time relative to the interconnect length that determines if a line is lumped or acts as a transmission line.
- The signal pulse width and duty cycle are not factors in determining if a line exhibits transmission line behavior such as reflections. Slow repetition pulses with sharp edges are just as likely to trigger transmission line effects as are narrow pulses with fast edges. The pulse width is important in determining what the signal looks like as it travels up and down the line, but it does not determine if reflections will occur.

A common signal integrity mistake is to assume that test logic or other low-speed diagnostic interconnect used during debug or manufacturing test and configuration need not be analyzed for signal integrity. Unfortunately, although these signals may be the slowest in the system, they are often generated by the same circuit technology that creates the high-performance signals. The result is a sharp rise time pulse with a slow repetition rate. Unless managed, this creates signal reflections and can lead to false triggering.

As explored in the Problems, this rule of thumb can also be used to decide if an interconnect should be modeled as a lumped RLC circuit or as a transmission line.

6.7.1 Why Does the Signal Rise Time Matter But Not the Pulse Width?

We have seen how the relationship between the length of line and the signal rise time determines if the interconnect acts as a transmission line, but we have not explained why the pulse width is not a factor.

Assuming that the transmission line resistance R is very small, circuit theory shows that the voltage drop across the inductor is determined by the amount of time required by the current to change. This is illustrated in:

$$Vl = L \times \frac{change\ in\ current}{change\ in\ time} = L \times \frac{di}{dt} \tag{6.6}$$

For a given change in current (di), the voltage drop Vl is affected by the total inductance L and the change in time (dt). As the voltage drop across the inductors becomes smaller, the individual capacitors in Figure 6.4 increasingly act as if they are connected in parallel. In this situation the line is no longer a distributed RLC circuit, but is a single lumped RLC (or, if the voltage drop across the inductor and resistor is very small, simply a lumped C circuit).

However, the capacitors become separated from one another by the inductor's voltage drop when it is high enough to no longer be negligible. In this situation the inductors and capacitors are distributed along the line, and the line exhibits distributed RLC (transmission line) behavior.

This means that, because a short piece of interconnect does not have much inductance, the signal must have a very fast rise time before the voltage across the inductors becomes noticeable and the line acts like a transmission line. On the other hand, if the interconnect is long, the inductance will be high, so even a slow rise time signal can cause a noticeable voltage drop across the total inductance, causing transmission line behavior. This matches our intuition. A short trace will not act like a transmission line unless the signal rise time is very fast, but a long trace will act like one even to slow rise time signals.

6.7.2 Why Does Propagation Delay Time Depend on the Line Length, But Impedance Does Not?

It may seem natural to assume that the impedance and the propagation delay time should depend on the line's length, but, in fact, the impedance does not depend on the length of line, while the propagation delay does. This comes about because (as illustrated in the Problems) in (6.1) the impedance is the ratio of the inductance to capacitance, while the propagation delay in (6.2) is the product of them. Because of this, the length units cancel when calculating the impedance, but they do not in the propagation delay calculations.

6.8 Main Points

- A transmission line is comprised of two or more conductors: the signal and its return.

- The return path determines the signal trace capacitance, inductance, and resistance.
- A transmission line can be modeled as a chain of very small RLCG elements.
- The transmission line circuit elements are usually given as per unit length.
- The transmission line capacitance and inductance cause the line to possess impedance and impose a delay.
- The transmission line impedance is independent of length, but the length determines the line's propagation delay time.
- Current flows in the return path simultaneously with the signal current as a mirror image.
- The signal rise time determines if a line will behave as a transmission line or as a lumped circuit.
- A line will act as a transmission line if its electrical length is greater than half the signal rise time. In some circumstances even shorter lines will act as transmission lines.
- Voltage divider action between the driver impedance and the transmission line impedance determines the value of the launched signal voltage.

Problems

Answers to these problems are available on the Artech House Web site at http://www.artechhouse.com/static/reslib/thierauf/thierauf1.html.

6.1 An ASIC drives a 50Ω transmission line with a driver having an output impedance of 30Ω. The driver is connected to a 2.2-V supply. What current is drawn from the supply when the driver switches from a logic low to a logic high?

6.2 In the above setup, what voltage gets launched?

6.3 An ASIC has 30Ω, 50Ω, and 75Ω drivers in its library, and they can be connected to either a 3.3-V or a 2.2-V supply. Which driver impedance and supply voltage should be specified if the goal is to launch no less than 1.2V but no more than 1.7V down a 50Ω transmission line?

6.4 An ASIC drives a 32-bit bus from a 3.3-V supply. The driver impedance can be set to either 50Ω or 75Ω, and the circuit board traces can be created with either of these impedances. When the drivers switch, it is necessary to limit the total current drawn from the I/O power supply to under 750 mA. Using only the ASIC total I/O current as the design consideration, what driver and transmission line impedance should be chosen?

6.5 Using the results from Problem 6.3, which combination of driver and power supply creates the least amount of on-die heat?

6.6 Find the propagation delay and impedance for a half-meter-long transmission line having L = 280 nH/meter and C = 118 pF/meter.

6.7 A circuit board trace has a total capacitance of 20 pF and a total inductance of 50 nH. What is the propagation delay for this interconnect, and what is the fastest rise time that can be sent down this trace before transmission line effects occur?

6.8 Show how the propagation time but not the impedance scales with distance by computing the delay and impedance for a 1-meter and a 5-cm-long transmission line. Assume that L is 280 nH/m and C is 110 pF/m.

6.9 A 50Ω transmission line has a propagation delay of 174 ps. What are the line's capacitance and inductance?

6.10 The trace routed to a termination device creates a half-inch stub. Assuming that the signal rise time is 350 ps and the line has a propagation delay of 166 ps/inch, should the stub be modeled as a transmission line or as a lumped capacitor?

6.11 Ten 50Ω stripline traces are routed as in Figure 6.6. Determine the minimum number of 10-nF decoupling capacitors that must be placed between the power and ground planes when the signal has a 1-ns rise time. Assume that when soldered to the circuit board each capacitor has an equivalent series inductance of 700 pH.

6.12 Apply the results from Problem 6.11 to the configuration shown in Figure 6.8 and determine the number of needed capacitors.

References

[1] Young, B., *Digital Signal Integrity*, Upper Saddle River, NJ: Prentice-Hall, 2001, p. 326.

[2] Miner, G. F., *Lines and Electromagnetic Fields for Engineers*, New York: Oxford University Press, 1996.

[3] Ramo, S., et al., *Fields and Waves in Communication Electronics*, 3rd ed., New York: John Wiley & Sons, 1994.

[4] True, K. M., "Data Transmission Lines and Their Characteristics," National Semiconductor Application Note 806, February 1996.

[5] Matick, R. E., *Transmission Lines for Digital and Communication Networks*, New York: IEEE Press, 1969, p. 7.

[6] Baloglu, H. B., *Circuits, Interconnections, and Packaging for VLSI*, Reading, MA: Addison-Wesley, 1990.

[7] Johnson, H. W., and M. Graham, *High-Speed Digital Design*, Upper Saddle River, NJ: Prentice-Hall, 1993.

[8] Skilling, H. H., *Transient Electric Currents*, New York: McGraw-Hill, 1952, p. 262.

[9] Blood, W. R., *MECL System Design Handbook*, Motorola Semiconductor Products, 1988.

[10] Buchanan, J. E., *Signal and Power Integrity in Digital Systems*, New York: McGraw-Hill, 1996.

[11] Hart, B. L., *Digital Signal Transmission Line Circuit Technology*, Berkshire, U.K.: Van Nostrand Reinhold, 1988.

[12] Royle, D., "Rules Tell Whether Interconnections Act Like Transmission Lines," *EDN*, June 23, 1988, pp. 131–136.

CHAPTER 7

Understanding Microstrip and Stripline Transmission Lines

7.1 Introduction

We saw in Chapter 5 how circuit board traces are arraigned as either microstrip or stripline transmission lines. Figure 7.1 shows that microstrips are suspended over a single plane at a height *h*, while striplines are sandwiched between two reference planes that are separated by a distance *b*. From Chapter 6 we know that these planes act as signal returns and can either be ground or power planes. As shown Figure 7.1, the microstrips are usually covered with a solder mask, and the circuit board may be hermetically sealed with an encapsulate (not shown).

By thinking of transmission lines as circuit elements rather than explaining them with field theory, previous chapters did not show how the physical differences in the construction of stripline and microstrips affect their electrical characteristics. These factors are addressed in this chapter.

7.2 DC Resistance and Resistivity

When measured with an ohm-meter, the resistance of a circuit board trace (the trace's DC resistance) will usually be a few ohms or less, but it climbs to much higher values when measured at high frequency. As is discussed in Chapter 8, this resistance increase means that conductor losses are much higher than at low frequencies.

The trace DC resistance depends on the trace resistivity (with units of ohm-meter), cross-sectional area (thickness and width), length, and temperature.

In Chapter 5 Table 5.5 showed that the resistivity of copper depends on how it is formed. A typical value is 1.75×10^{-8} Ω-meters at room temperature.

The trace resistance increases with temperature. Although the exact amount depends on the copper alloy, typically the resistance increases by just over 4% for every 10°C rise in temperature [1]. This is a greater than 25% increase in DC resistance when the temperature increases from 25°C to 85°C. Because larger resistance increases conductor losses, the signal integrity engineer should account for the effects of temperature when using field solvers to create worst-case trace models.

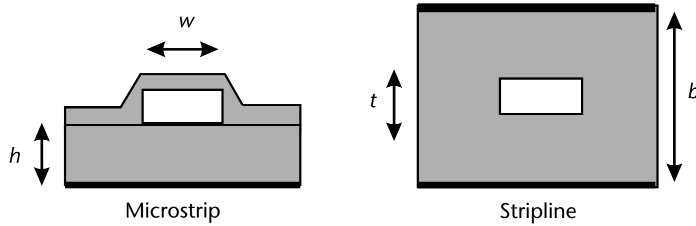


Figure 7.1 Microstrip and stripline traces of width *w*, thickness *t*, and distance *h* or *b*.

Field solvers sometimes use conductivity (the reciprocal of resistivity, with units of Siemens-meter) rather than resistivity to specify the metal characteristics. Depending on the alloy, at room temperature the conductivity of copper is about 5.7×10^7 S-m.

Figure 7.2 shows the room temperature resistance for half-ounce and 1-ounce perfectly rectangular copper traces. The resistance increases dramatically for half-ounce traces under 10 mils and 1-ounce traces less than 15 mils. The resistance is higher than shown if the trace is trapezoidal. Equations for calculating trace resistance appear in Chapter 17.

7.3 Stripline and Microstrip Capacitance

As illustrated in Figure 7.3, a parallel plate capacitor is formed between the microstrip trace (as the upper plate of the capacitor) and the ground plane (as the

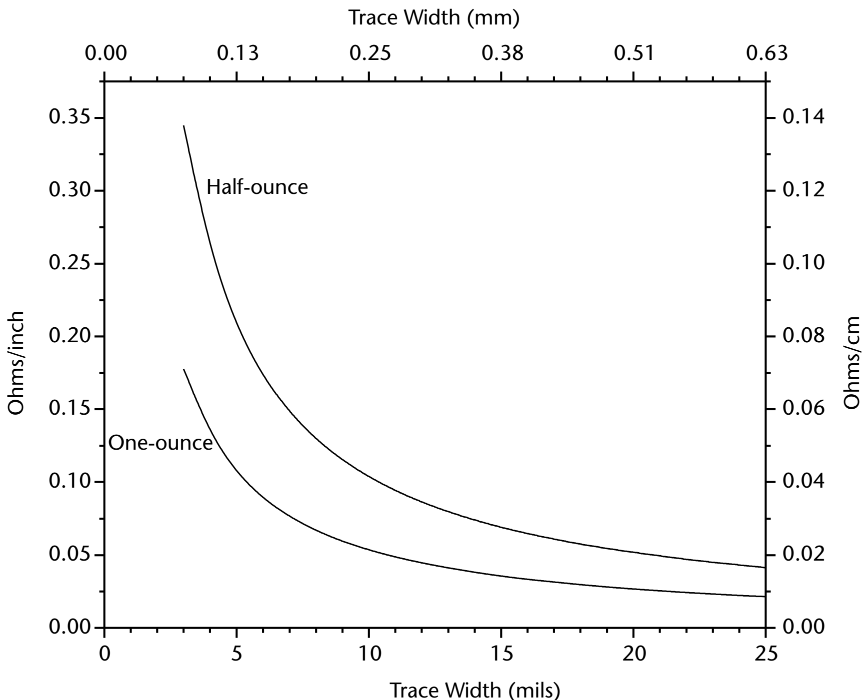


Figure 7.2 Room temperature (25°C) DC resistance of half-ounce (0.65-mil-thick) and 1-ounce (1.2-mil-thick) rectangular copper traces.

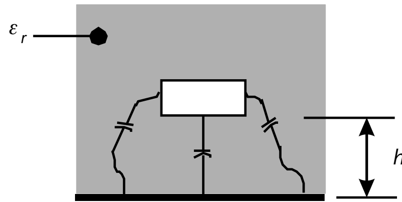


Figure 7.3 Trace over a ground plane creates a parallel plate capacitor.

bottom plate). In stripline the coupling appears to both planes, even if one is a power rather than a ground plane.

The fringing capacitance from the trace edges along with the capacitance between the trace bottom and the ground plane is not shown.

Increasing the distance h from the trace to the ground plane or reducing the dielectric constant (ε_r, but, as mentioned in Chapter 1, sometimes called Dk) decreases the capacitance.

A typical dielectric constant (and the way in which they change with frequency) for three FR4 laminate systems is shown in Table 7.1.

Because the polarization of the molecules in the dielectric is affected by frequency and temperature [5], the dielectric constant generally falls as frequency increases, but in some laminates the dielectric constant will increase again after reaching a minimum value at moderately high frequencies. FR406 exhibits this behavior: ε_r falls from 4.6 at 1 MHz to 4.28 at 500 MHz and then climbs slightly to 4.29 at 1 GHz. The rate of the change depends on the resin composition and the amount of glass (the glass-to-resin ratio, introduced in Chapter 5).

7.3.1 What Is the Effective Dielectric Constant?

Because they are formed in one dielectric (or two that are very similar), striplines are said to have a homogeneous dielectric and have a dielectric constant equal to ε_r.

The solder mask covering a microstrip has a dielectric constant close to typical values for FR4, so microstrips are often thought of has having only two distinct dielectrics: the laminate/solder mask and air. Because these two dielectrics have different dielectric constants, the microstrip dielectric is not homogeneous, but it is possible to think of the blend as having one aggregate value. The microstrip effective dielectric constant (ε_{r_eff}) is an equivalent value that takes into account the value of ε_r of the laminate and the microstrip width and thickness and its height above the return path. These physical factors are important because they determine

Table 7.1 Dielectric Constant and (Loss Tangent) for Three FR4 Type Laminate Systems for a Glass-to-Resin Ratio of 44%

Laminate	*1 MHz*	*500 MHz*	*1 GHz*
FR404	4.6 (0.025)	4.26 (0.014)	4.25 (0.014)
FR406	4.6 (0.023)	4.28 (0.014)	4.29 (0.014)
FR408	3.8 (0.010)	—	3.7 (0.010)

Source: [2–4].

how much of the signal energy travels in air and how much travels in the laminate and how much each influences the capacitance and (as we will see in Section 7.8) the signal delay.

The effective dielectric constant will have a value between 1 (for those fields traveling only in the air) and the value of the laminate itself. This means that the effective dielectric constant will always be lower than the dielectric constant of the underlying laminate, so signals propagating along microstrip experience a lower dielectric constant than striplines. This is true even when they are fabricated with the same laminate materials. As we will see, this makes the electrical characteristics of microstrips quite different from striplines.

7.3.2 What Are the Differences Between Microstrip and Stripline Capacitances?

Figure 7.4 plots the capacitance for several microstrip and stripline traces for three impedances and traces of various widths. The traces are perfectly rectangular, ε_r is 4.2, and the microstrips are covered with a 1-mil-thick solder mask that also has an ε_r of 4.2. Chapter 17 shows how to calculate capacitance using trace dimensions or trace impedance.

The graph shows that high impedance traces have less capacitance than low impedance traces, and for a given impedance, the capacitance of stripline traces is higher than microstrip traces.

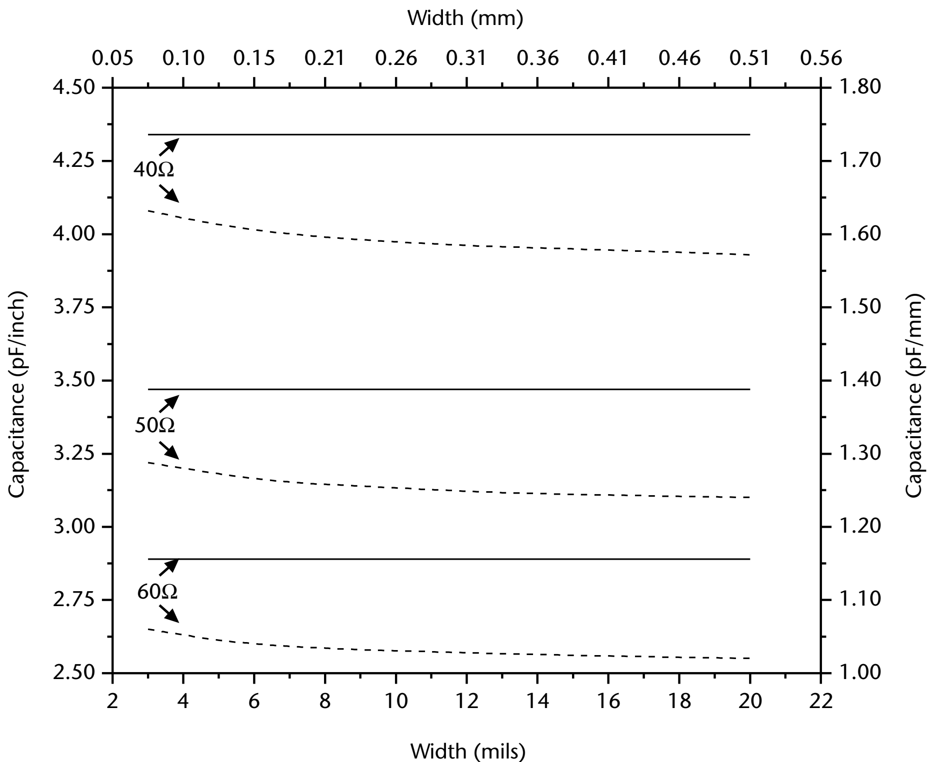


Figure 7.4 Capacitance of rectangular solder mask microstrip (plotted with dotted lines) and stripline (solid lines) traces of various widths. Raw data from the Polar Si9000 2D field solver [6].

For instance, the capacitance of a 5-mil (0.13-mm)-wide, 60Ω microstrip is about 2.6 pF/inch, but it increases to about 4 pF/inch for a 40Ω microstrip of the same width. A 40Ω stripline of any width has a capacitance of about 4.3 pF/inch, which is about 10% higher than the microstrip capacitance.

7.3.3 How Does Frequency Alter the Loss Tangent?

In general, the loss tangent (introduced in Chapter 1) will have its highest value at low frequencies, and the value of the loss tangent falls as frequency increases. This trend is evident for three FR4 laminate systems in Table 7.1. For instance, the loss tangent of the first laminate is 0.025 at 1 MHz, but falls to 0.014 for higher frequencies. As in this case, once reaching a high enough frequency, the loss tangent will often remain nearly constant and in some cases it may begin to climb again. Moisture uptake (described in Chapter 8) can cause the loss tangent to increase in high-humidity environments.

7.3.4 What Is Conductance and How Is It Determined?

Conductance is used in many SPICE lossy transmission line models to account for stripline and microstrip dielectric losses, and this chapter's Problems show how to create SPICE lossy transmission line models. In Chapter 8 we will see how to directly determine loss from the loss tangent without using the conductance. However, because conductance is so widely used in simulators, it is worthwhile to understand how its value is determined.

From Chapter 6 we recall that the conductance (G, units of Siemens/length) is modeled as a resistor in parallel with the trace capacitance. Higher values of conductance represent higher dielectric losses. As shown in (7.1), the conductance value is determined by the trace capacitance (C), the frequency (f), and the value of the loss tangent (LT).

$$G = 6.28 \times C \times f \times LT \tag{7.1}$$

Evidently traces with larger capacitance and loss tangent values have correspondingly larger values for G (and so, higher dielectric losses). This means that dielectric loss is highest for low impedance traces (because these have larger capacitance), and the loss increases with higher loss tangent values.

As we just saw, for a given impedance the capacitance of a stripline is the same for all trace widths, but it changes for microstrip widths. This means that for a given impedance, the dielectric loss is the same for striplines regardless of its width, but the dielectric loss will be different for various widths of microstrip. This assumes the loss tangent has the same value everywhere on the circuit board.

The effects with microstrips are small, especially if the microstrip is covered with solder mask, and although not evident from the equation, because the loss tangent for air is lower than that of the laminate, for any given impedance G is lower for microstrip than for stripline.

For instance, on FR4 ($\varepsilon_r = 4.2$, $LT = 0.02$), G is just over 17m S/m for a 50Ω, half-ounce stripline of any width at 1 GHz. At that same frequency it is slightly more than 15m S/m for a 5-mil-wide microstrip that is covered by 1 mil of solder

mask. It falls to about 14.75m S/m when the 50Ω solder mask–covered microstrip is 10 mils wide. It is even less (just over 9m S/m) when the 10-mil microstrip has an impedance of 75Ω.

This example and (7.1) show the advantage of using wide, high-impedance microstrips to route high-speed signals, especially when an inexpensive circuit board having a high loss tangent is used. We will return to this when discussing dielectric losses in Chapter 8.

7.4 Stripline and Microstrip Inductance

Circuit board traces exhibit parallel plate inductance with the ground plane. The inductance is determined by the height and width of the trace: Increasing the height (h and b in Figure 7.1) causes the inductance to become larger, while increasing the trace width (w) lowers the inductance. The skin effect causes the inductance to fall slightly at high frequency as the internal inductance is reduced. In fact, at very high frequencies the inductance is mostly determined by the shape and dimensions of the trace. These things are described in more detail in Section 7.5.

The high-frequency inductance for several microstrip and stripline traces is presented in Figure 7.5. The traces are perfectly rectangular, the dielectric constant is 4.2, and the microstrips are covered with a 1-mil-thick solder mask having a dielectric constant of 4.2.

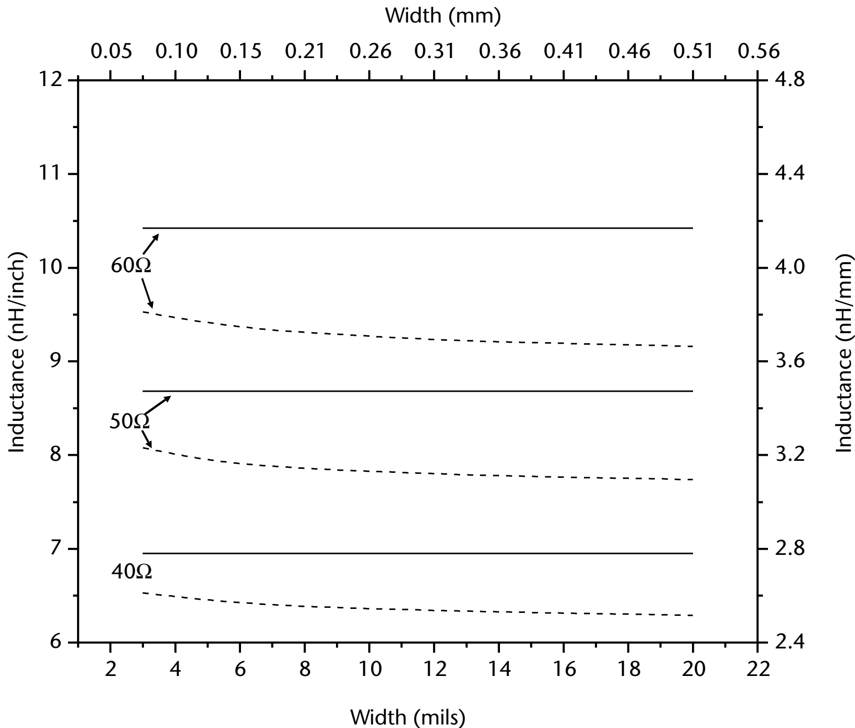


Figure 7.5 Inductance at a very high frequency of rectangular traces of various widths. Striplines are shown with solid lines; the dotted lines are microstrips. Raw data from the Polar Si9000 2D field solver [6].

Low-impedance traces are seen to have lower inductance than high impedance traces. For example, the 40Ω stripline inductance is about 7 nH/inch for any width versus about 10.5 nH/inch for the 60Ω stripline.

Also evident is that for a given impedance and width, the stripline has a higher inductance than microstrip. For instance, a 5-mil-wide 50Ω stripline has an inductance of 8.7 nH/inch, while the microstrip is 8 nH/inch.

7.5 How Does the Skin Effect Change the Trace's Resistance?

At low frequencies the signal current flows through the entire cross-section of a trace. However, at high frequencies the skin effect causes current to only flow near the trace surface in a region defined by the skin depth. The skin depth (or the depth of penetration) is large for low frequencies, which is why the current flows through the entire trace cross section at DC.

However, at a high frequency the resistance of the trace and the return path (the loop resistance, R_{ac}) increases as the square root of the frequency. This is shown in:

$$R_{ac_f2} = R_{ac_f1}\sqrt{\frac{f_2}{f_1}} \tag{7.2}$$

For instance, increasing the frequency by a factor of 4 from 500 MHz to 2 GHz causes the resistance to increase by 2.

The skin effect becomes apparent in 1-ounce traces at roughly 14 MHz; it is evident at about 65 MHz for half-ounce traces [7].

The loop resistance at 350 MHz (corresponding to the 3-dB bandwidth of a 1-ns pulse) is plotted for various widths of striplines and solder mask covered microstrips in Figure 7.6. For the laminate and the solder mask, the dielectric constant is 4.2, and the loss tangent is 0.02. The solder mask is 1-mil thick.

The loop resistance for frequencies other than 350 MHz can be found by using (7.2). For instance, Figure 7.6 shows the loop resistance of a 4-mil (0.11-mm)-wide 50 stripline to be about 0.8 Ω/inch. Increasing the frequency to 1.4 GHz causes the resistance to increase by a factor of 2 (f_1 = 350 MHz, f_2 = 1.4 GHz, R_{ac_f1} = 0.8Ω), to 1.6 Ω/inch.

7.5.1 What Is the Proximity Effect?

One property of the skin effect (the proximity effect) is that as frequency increases, current in the return path gradually collects underneath the trace, rather than spreading out and using the entire ground plane as it does at lower frequencies [9–12]. For a given trace width, the spreading is greatest for traces further away from the ground plane, which is why higher-impedance traces (which as we will see in Section 7.7 are further away from the return plane) use more of the return path metal than lower-impedance ones. This is shown in Figure 7.7.

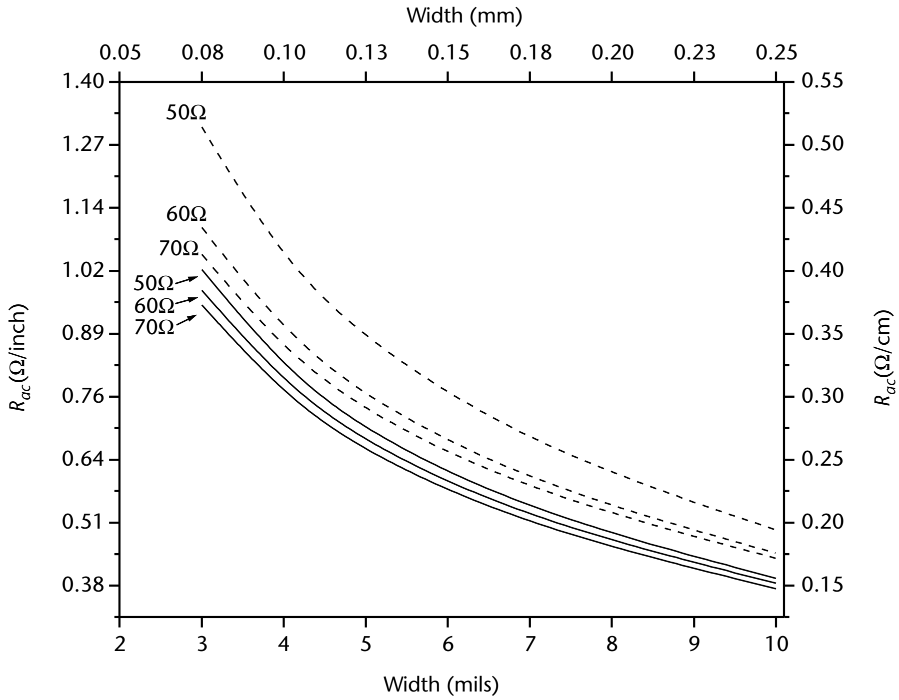


Figure 7.6 Loop resistance as calculated by Linpar [8] at 350 MHz for perfectly rectangular 0.65-mil (0.017-mm)-thick stripline (solid curve) and solder mask–covered microstrip (dotted curve).

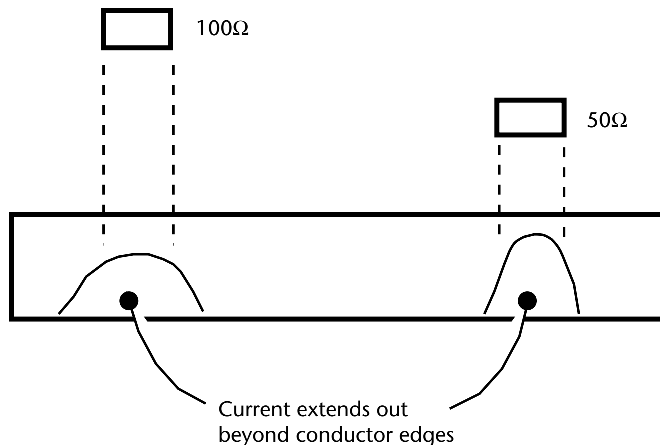


Figure 7.7 Cross-section of two microstrips showing the current in the ground plane gathering underneath the traces at high frequency. The proximity effect cause more spreading for traces of the same width that are a greater distance from the return.

Confining the return current to a small channel in the vicinity of the trace raises the loop resistance. Conversely, the resistance will be lower if the current can spread out along a larger section of the return plane (as it does for traces higher above the plane). The net effect is that, for a given dielectric constant and trace width, the loop resistance is lower for high-impedance traces at any given high frequency.

7.6 How Does the Skin Effect Change the Inductance?

The total trace inductance of a trace or round wire is the sum of the inductance internal to the trace (the internal inductance) and the inductance due to the trace geometry (the external inductance) [5, 13].

The internal inductance is small and its value is reduced as the skin effect causes the current to move away from the center of the trace and migrate outward, toward the trace surface. This leaves fewer lines of magnetic flux present inside the conductor, lowering the internal inductance. Because this behavior is driven by the skin effect, the internal inductance decreases as the square root of the increase in frequency (the same rate at which the resistance increases).

For instance, at 500 MHz the internal inductance of a 4-mil-wide 50Ω stripline trace is about 5% of the total inductance. It falls to less than 1% at 20 GHz. Assuming that the capacitance does not change, this reduction in inductance causes the trace impedance to be about 2% lower at 20 GHz than it is at 500 MHz.

7.7 Understanding Stripline and Microstrip Impedance

A trace's dimensions (its width, thickness, and distance to the ground plane or planes) determine its impedance. To understand how these interact, (7.3) (simplified from [14]) calculates stripline impedance, and (7.4) calculates microstrip impedance. For clarity, (7.4a) is slightly modified from [15], and uses a more comprehensive equation for the effective dielectric constant (7.4b) [16].

More complex (and accurate) equations appear in Chapter 17, but these equations are simple enough to show how the trace dimensions interact and determine the impedance. The traces are assumed to be perfectly rectangular, and the stripline is assumed to be centered between the return planes.

$$Z_o = \frac{94}{\sqrt{\varepsilon_r}} \ln\left(\frac{b+w}{w+t}\right) \tag{7.3}$$

$$Z_o = \frac{60}{\sqrt{\varepsilon_{r_eff}}} \ln\left(\frac{7.5h}{w+1.25t}\right) \tag{7.4a}$$

$$\varepsilon_{r_eff} = \frac{\varepsilon_r + 1}{2} + \frac{\varepsilon_r - 1}{2}\left(\sqrt{\frac{w}{w+10h}}\right) \tag{7.4b}$$

From these equations we see that:

- When holding the distance to the trace (b or h) constant, the impedance gradually goes down as the trace gets wider (w) or as it gets thicker (t).

- When holding the width (w) and thickness (t) fixed, the impedance gradually goes down when the distance to the ground plane (b or h) is made smaller.
- The impedance goes up as the dielectric constant (ε_r for stripline, or ε_{r_eff} for microstrip) goes down, and it does so as the square root of the change.

For instance, doubling the trace width from 5 mils to 10 mils but keeping the spacing b constant at 13.5 mils causes the stripline impedance to fall by less than half from 50Ω to 31Ω on FR4. In a similar way, reducing the dielectric constant in half from 4.2 to 2.1 causes the stripline impedance to increase from 50 to nearly 71Ω (rather than doubling).

Because the units cancel, either inch-based or metric units may be used with these equations to obtain the impedance. For impedances of 40Ω or higher, (7.3) is accurate for striplines to better than 5%. Equation (7.4a) is accurate to within 10% for microstrips without solder mask across that same impedance range when used with (7.4b) to calculate ε_{r_eff}. Chapter 17 shows how to calculate the impedance of solder mask–covered microstrip.

This chapter's Problems further illustrate uses for these equations, but they are plotted for FR4 ($\varepsilon_r = 4.2$) in Figure 7.8.

For stripline the trace bottom is assumed to be centered between the planes at the distance shown. For the microstrip traces the distance is from the ground plane to the trace bottom. The solder mask covers the microstrip by 1 mil and has a dielectric constant of 4.2.

Figure 7.8 is for illustration only because manufacturing details such as the actual shape of the trace and the composition of the laminate are not accounted for.

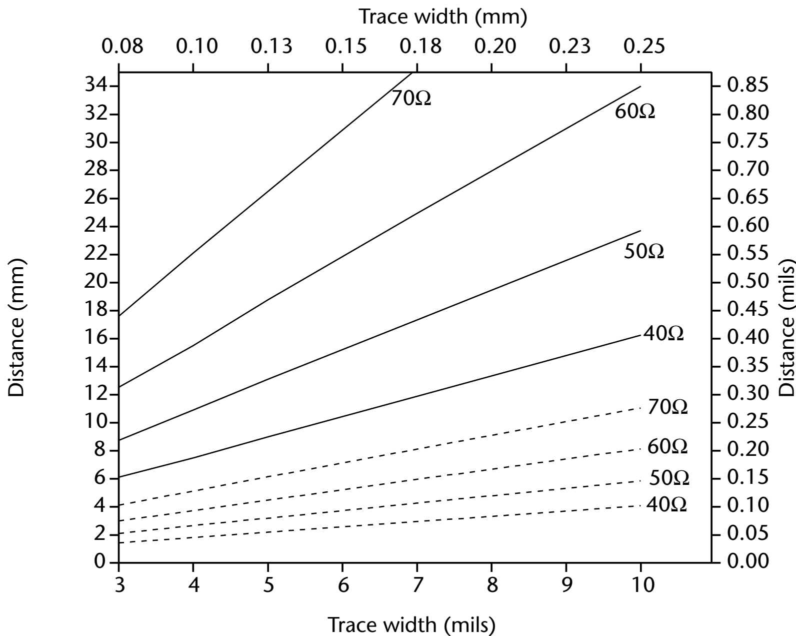


Figure 7.8 The distance (b or h) and trace width (w) necessary to obtain various impedances for stripline (solid curve) and solder mask–covered microstrip (dotted curve). Data from the Polar Si9000 2D field solver.

7.8 Understanding Stripline and Microstrip Propagation Delay Time

In air, electromagnetic energy requires about 3.3 ns to travel the distance of 1 meter (about 85 ps per inch), but as shown in (7.5), it takes longer to travel through a dielectric such as a circuit board laminate. The physics behind this is widely available (for just one example, see [5]), but we will note here that as shown in Table 7.1, ε_r for circuit board laminates is higher than the value for air (which is about 1).

$$tpd = 84.7\sqrt{\varepsilon_r}\ \text{ps/inch} \tag{7.5a}$$

$$tpd = 33.5\sqrt{\varepsilon_r}\ \text{ps/cm} \tag{7.5b}$$

In (7.5) *tpd* is the propagation delay time [in picoseconds/inch for (7.5a), and picoseconds/centimeter for (7.5b)], and ε_r is the dielectric constant of the laminate for stripline or the effective dielectric constant (ε_{r_eff}) for microstrip.

Because only one dielectric is involved with stripline (the dielectric is homogeneous), (7.5) shows that the delay is not affected by the trace dimensions or impedance. For example, the delay of a 4-mil-wide, 25Ω stripline will be identical to the delay of a 10-mil-wide, 100Ω stripline formed on the same layer (where we assume ε_r is identical for both traces). In fact, the delay for stripline traces is independent of the trace geometry.

However, because the two dielectrics (air and the laminate) have such different dielectric constants, this is not so for microstrips. In this case, the dielectric is not homogeneous, and the trace width and impedance do influence microstrip delay.

To understand this, we notice from Figure 7.8 that obtaining a given impedance while making w larger requires the trace to be further away from the ground plane. This means that more of the field lines travel in the air as compared to a trace that is closer to the laminate. Since signals travel faster in air than in a laminate, this results in the delay being smaller. In fact, ε_{r_eff} becomes closer to 1 as the trace is moved further away from the ground plane (that is, as h in Figure 7.1 is made larger). From (7.5), lower values of ε_{r_eff} cause *tpd* to be less. Microstrip is said to be faster than stripline for this reason.

Similarly, as is evident in Figure 7.8, obtaining a particular impedance for a given trace width requires that h change. This means that, unlike stripline, the delay of a 25Ω microstrip will not be the same as the delay of a 100Ω microstrip. In fact, the 25Ω microstrip will be closer to the laminate than the 100Ω microstrip, making its delay longer.

7.8.1 How Do the Dielectric and Effective Dielectric Constants Affect Delay Time?

Equations such as (7.4b) do not clearly show how the delay changes with trace width and impedance. The delay for various trace widths is plotted in Figure 7.9 to show this relationship.

Signal delay on microstrip is reduced slightly with increases in the trace width and impedance, and microstrips are faster than stripline. For instance, on FR4

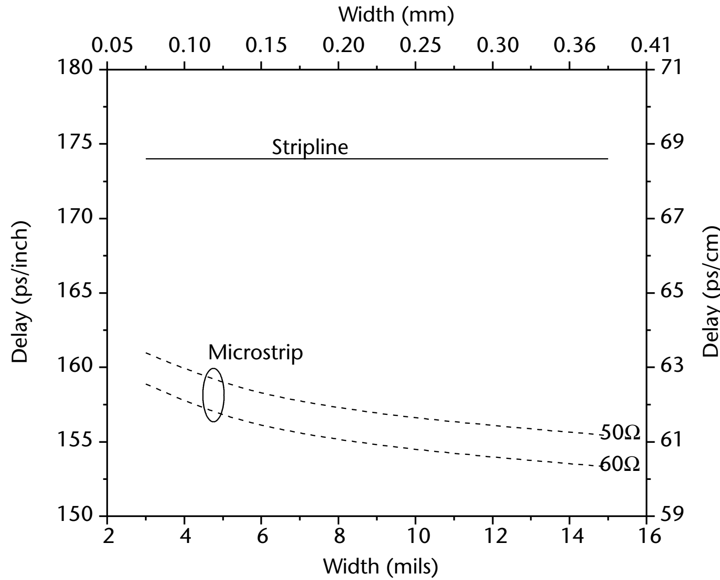


Figure 7.9 Approximate delay of half-ounce-thick stripline and microstrip on FR4 (ε_r = 4.2). The stripline of all widths and impedances has a delay of 174 ps/inch (68.5 ps/cm), but the microstrip trace width and impedance factor into the microstrip delay. The solder mask thickness is 1 mil (25.4 μm).

the 3-mil-wide 50Ω microstrip is about 8% faster than any of the shown stripline traces, and is about 12% faster for a 15-mil-wide trace.

Wide, high-impedance microstrips have smaller delay because they are further away from the ground plane and more of the field lines travel in air. However, the effect is small. Changing the impedance by 20% from 50Ω to 60Ω reduces the delay by just over 1%.

The data shown is for a 1-mil (25.4-μm)-thick solder mask, which has a dielectric constant of 4.2. A thicker coating of solder mask causes the trace to be slower (doubling the solder mask thickness from 1 to 2 mils causes the propagation time to increase by about 3% on FR4). Increasing the microstrip thickness to 1-ounce copper increases the delay by about 1%.

7.8.2 How Sensitive Is Delay Time to Changes in the Dielectric Constant?

The delay for a 3-mil (0.08-mm)-wide and 10-mil (0.25-mm)-wide 50Ω microstrip and a stripline is shown in Figure 7.10 plotted against different values for the circuit board dielectric constant (ε_r). The traces are perfectly rectangular half-ounce copper with a 1-mil (25.4-μm) solder mask, whose ε_r is 4.2.

The delay difference between stripline and microstrip traces grows as the dielectric constant gets larger. Large dielectric constants also reduce the difference in time of flight between wide and narrow microstrip traces.

7.9 Main Points

- The trace and return path form parallel plate capacitors and inductors.

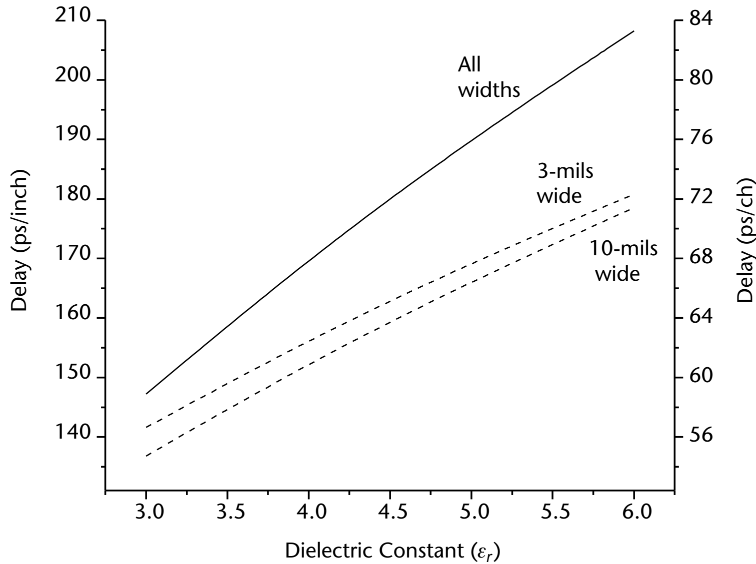


Figure 7.10 Approximate delay for stripline (solid curve) and solder mask covered microstrips (dotted curve) fabricated on circuit boards of various dielectric constants. Stripline delay is not affected by trace width, but microstrip delay is.

- The distance between the trace and the return path determines the trace resistance, inductance, capacitance, and conductance.
- The dielectric systems are not the same for microstrip and stripline traces, making them have different electrical characteristics.
- Transmission line loss is caused by trace resistance and leakage in the dielectric.
- The skin effect causes the AC loop resistance to increase and becomes apparent for half-ounce traces at about 65 MHz (lower for thicker traces).
- For a given dielectric constant, the trace thickness, width, and distance from the return path determine the trace impedance (Z_o).
- In striplines the trace propagation delay time (*tpd*) is only determined by the dielectric constant, giving all striplines on a routing layer the same delay per length.
- Trace geometry factors in the microstrip delay time.
- For striplines of a given impedance, the trace capacitance (C), inductance (L), and conductance (G) are the same for all trace widths.
- For microstrips of a given impedance, the trace capacitance, inductance, and conductance are affected by the trace width, especially when the dielectric constant is low or when the solder mask is thin or has a low *Dk*.
- The skin effect causes a traces inductance to decrease as frequency increases.

Problems

Answers to these problems are available on the Artech House Web site at http://www.artechhouse.com/static/reslib/thierauf/thierauf1.html.

7.1 What is the approximate impedance of a half-ounce 5-mil-wide stripline on FR4 when the plane separation is 16 mils?

7.2 What is the approximate impedance of a half-ounce 8-mil-wide solder mask microstrip on FR4 when the separation from the trace to the return plane is 5.2 mils?

7.3 How will the board thickness change when a half-ounce, 50Ω stripline on FR4 is changed from 5 mils wide to 8 mils wide?

7.4 How long will it take a signal to travel down a 12-inch (30.5 cm) long, 75Ω, 5-mil (0.127-mm)-wide, half-ounce stripline fabricated on FR4 ($Dk = 4.2$)?

7.5 A certain circuit board cannot be made too thick, yet it is necessary to obtain the maximum number or routing channels possible. What influence does this requirement have on selecting the trace width and impedance?

7.6 A 50Ω stripline is fabricated on FR4 ($\varepsilon_r = 4.2$). What will be the impedance and delay time if the laminate is changed to one having $\varepsilon_r = 3.75$?

7.7 A 10-mil-wide, 50Ω stripline runs parallel to and on the same layer as a 75Ω, 5-mil-wide stripline. Assuming that the laminate is FR4 and both traces are a half-ounce thick, which trace will delay the signal the least?

7.8 The circuit board manufacturer has decided to reduce the thickness of each stripline layer in a multilayer circuit board from 15 mils (0.38 mm) to10 mils (0.25 mm) as a way to meet the overall board thickness goal. What effect will this have on the trace impedance and time of flight?

7.9 A stripline is known to be precisely 31.5 inches (80 cm) long. A signal takes 5.4 nS to travel the entire trace length. What is the dielectric constant of the laminate?

7.10 From a 2D field solver, R_{ac} is found to be 7 Ω/meter (0.18 Ω/inch) at 100 MHz. What is the resistance at 1 GHz?

7.11 What is the conductance at (a) 100 MHz and (b) at 1 GHz for a 5-mil (0.13-mm)-wide half-ounce 50 stripline on FR4 ($Dk = 4.2$ and $LT = 0.02$)?

7.12 Build a W line model for a 5-mil-wide, half-ounce 50Ω stripline on FR4.

References

[1] The Chemical Rubber Company, *CRC Handbook of Chemistry and Physics (1986–1987)*, 67th ed., Cleveland, OH: CRC Press, 1987, pp. E-9.

[2] Polyclad Laminates, Inc.,” PCL-FR-226/PCL-FRP-226,” data sheet PCL-226, dated 10/01, Isola Group, Chandler AZ.

[3] Isola Laminate Systems, Inc., “FR408 Product Data Sheet,” No. 5035/9/99, LaCrosse, WI, 1999.

[4] Isola Group, “Isola IS620,” data sheet IS620 rev 07-08-08, Chandler, AZ.

[5] Miner, G. F., *Lines and Electromagnetic Fields for Engineers*, New York: Oxford University Press, 1996.

[6] Polar Si9000 Field Solver, Polar Instruments, Inc., http://www.polarisinstruments.com.

[7] Thierauf, S. C., *High-Speed Circuit Board Signal Integrity*, Norwood, MA: Artech House, 2004.

[8] Djordjevic, A. R., et al., *LINPAR for Windows*, Norwood, MA: Artech House, 1999.

[9] Bertin, C. L., "Transmission-Line Response Using Frequency Techniques," *IBM Journal of Research and Development*, January 1964, pp. 52–63.

[10] Faraji-Dana, R., and Y. L. Chow, "The Current Distribution and AC Resistance of a Microstrip Structure," *IEEE Journal of Microwave Theory and Techniques*, Vol. 38, No. 9, September 1990, pp. 1268–1277.

[11] Pucel, R. A., D. J. Masse, and C. P. Hartwig, "Losses in Microstrip," *IEEE Transactions on Microwave Theory and Techniques*, MTT-16, No. 6, June 1968, pp. 324–350.

[12] Pucel, R. A., D. J. Masse, and C. P. Hartwig, "Correction to Losses in Microstrip," *IEEE Transactions on Microwave Theory and Techniques*, Vol. MTT-16, No. 12, December 1968, p. 1064.

[13] Matick, R. E., *Transmission Lines for Digital and Communication Networks*, New York: IEEE Press, 1969.

[14] Bakoglu, H. B., *Circuits, Interconnections, and Packaging for VLSI*, Reading, MA: Addison-Wesley, 1990.

[15] Kaupp, H. R., "Characteristics of Microstrip Transmission Lines," *IEEE Transactions on Electronic Computers*, Vol. EC-16, No. 2, April 1967, pp. 185–193.

[16] Scheider, M. V., "Microstrip Lines for Microwave Integrated Circuits," *Bell System Technical Journal*, May-June 1969, pp. 1421–1444.

CHAPTER 8

Signal Loss and the Effects of Circuit Board Physical Factors

8.1 Introduction

In previous chapters we assumed that the copper traces were rectangular and had smooth surfaces and that the stripline traces were embedded in a single block of dielectric material (the dielectric was homogeneous).

These ideal assumptions are not valid in practice. Copper surfaces are rough, often traces are trapezoidal in shape, and the glass-to-resin ratio (see Chapter 5) determines the actual dielectric constant. All of these can alter the trace impedance, delay, and loss characteristics, and it is important for the signal integrity engineer to determine when it is necessary to account for them in simulation or when to discuss them with circuit board fabrication shops.

8.2 What Are Transmission Line Loss and Attenuation?

Loss is the mechanism that reduces (attenuates) a signal and in circuit board traces is caused by resistance in the conductor and leakage in the circuit board laminate. Distortion is the change in pulse shape caused by the loss unequally affecting each of the harmonics making up the signal.

8.2.1 What Effects Do Losses Have on Signals?

To understand loss, we must think about signals in the frequency domain. As described in Chapter 1, a pulse is made up of many frequency components (harmonics). Because the amount of loss depends on the frequency, each harmonic making up a pulse will be attenuated differently as the pulse travels down the line. In fact, each successively higher harmonic will be attenuated more than lower-frequency ones. This means that when the harmonics recombine at the load their amplitudes and phase are not in the same proportion as when they were launched. The pulse has become distorted.

As shown in Figure 8.1, the attenuation of the high-frequency components cause the received pulse to become rounded and reduced in amplitude, and the base of the pulse smears and becomes wider than the base of the transmitted pulse.

The progressive distortion of a 1-V, 400-ps pulse stream as it progresses down a 1-meter (39-inch)-long transmission line is illustrated in Figure 8.2. The effect on the pulse is evident when compared to the undistorted pulse. In fact, the pulse base has become wide enough to collide with the rising edge of the succeeding pulse. This effect (dispersion) can severely distort a pulse stream, which is of particular concern in serial signaling. For instance, dispersion severe enough to cause a pair of back-to-back pulses (such as a 0110 pattern) to interfere with one another may not be great enough to cause significant interference when the pulses are widely separated (for example, a 1001 pattern). This leads to “data-dependent” signaling errors.

8.2.2 How Is Loss Specified?

In signal integrity work, loss is defined as the ratio of the received voltage to the transmitted voltage. As shown in (8.1), the ratio is given in terms of the common logarithm and has units of decibels (dB).

$$Gain\ in\ dB = 20 \times \log_{10}\left(\frac{V_{received}}{V_{transmitted}}\right) \tag{8.1}$$

When calculating loss, decibels are used rather than a simple percent because it is only necessary to add together the losses in decibels of the various parts to find the total loss. As we will see, microstrip and stripline losses come in two parts, and the total loss in dB is simply the sum of the two.

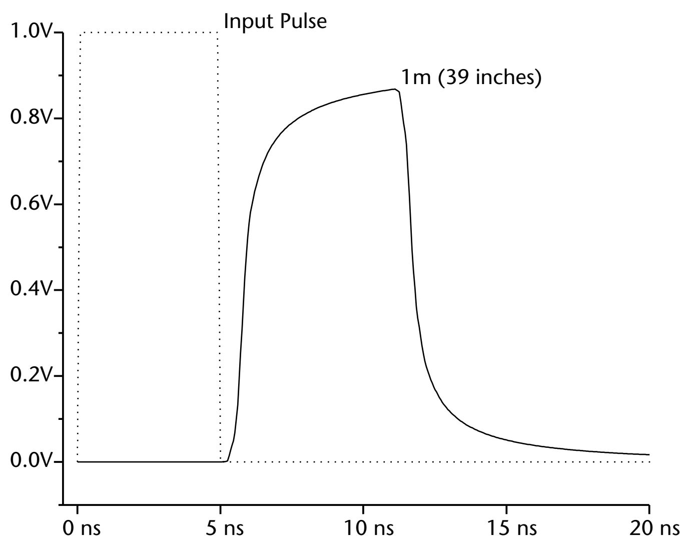


Figure 8.1 Effects of loss created by a 1-meter (39-inch)-long microstrip on a sharp edged, 5-nS-wide pulse (dotted curve). The pulse exiting the line (solid curve) is smaller, rounded, and wider at its base.

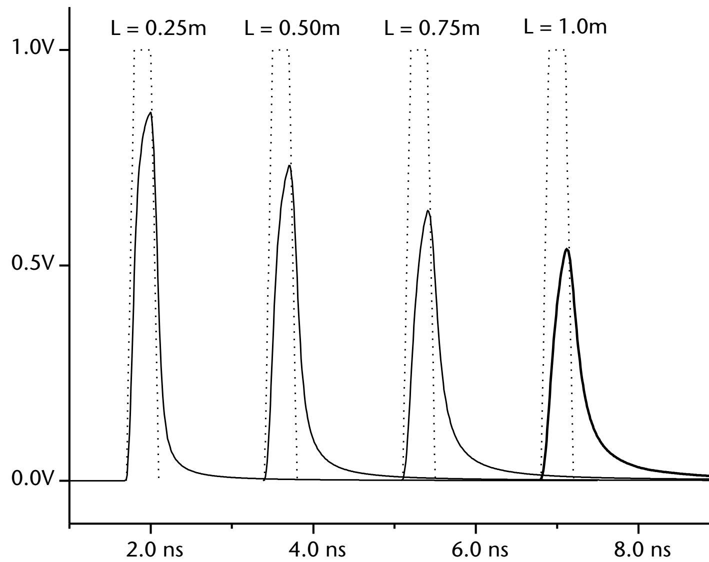


Figure 8.2 Transmission line loss causes pulse distortion and spreading of its base. The position and shape of the pulse without distortion are shown for reference.

Because the logarithm of a fraction has a negative value, the gain will be negative when $V_{received}$ is less than $V_{transmitted}$. For instance, the gain is –14 dB when $V_{received}$ is 0.2V and $V_{transmitted}$ is 1V. However, it is customary to drop the negative sign and say that the circuit has a loss (or the signal has experienced an attenuation) of 14 dB.

This equation is further demonstrated in the Problems, and Chapter 17 shows how to convert from decibels to the ratio of the received and transmitted voltage.

The equation cannot be used to represent the power lost in dB, but by using the received and transmitted currents ($I_{received}$ and $I_{transmitted}$), it can show the current attenuation.

8.2.3 Converting Between Decibels and Percentage Loss

The relationship between the loss in percent and in decibels is shown graphically in Figure 8.3. The negative value is implied and the attenuation is shown as a positive number.

A 3-dB loss is shown to mean that the signal voltage is reduced by just over 29%, leaving nearly 71% of the original signal intact at the receiver. The nonlinear nature of the decibel scale is evident: increasing the loss from 1 to 10 dB reduces the received amplitude by 2.8 times (89% to 32%) rather than 10 times.

8.3 What Creates Transmission Line Loss and How Is It Characterized?

Signals lose energy by heating the conductor resistance (represented by R in Figure 8.4) and by leakage currents heating the dielectric (G). As we saw in Chapter 7, both of these change with frequency. For a circuit board, the capacitor dielectric is the laminate or prepreg, and the conductor is a trace and the ground plane(s).

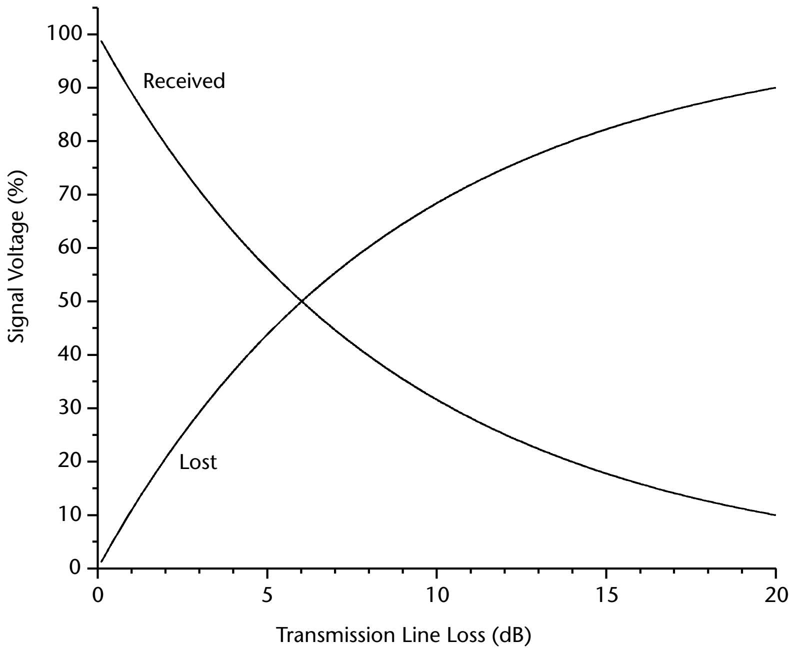


Figure 8.3 Amplitude (in percent) to decibel conversion. A 3-dB loss is seen to represent a signal swing of 70.8% at the receiver.

Alpha (α) is used to indicate loss, with a subscript showing that it is either the dielectric (α_d) or the conductor (α_c) part. As shown in (8.2), the total loss in dB is the sum of the two and is represented by α_t.

$$\alpha_t = \alpha_c + \alpha_d \tag{8.2}$$

As Figure 8.5 shows for FR4 (ε_r = 4.2, LT = 0.02), at low frequencies losses in the conductor contribute more to the total loss than dielectric losses. However, the two losses increase at different rates, so at some frequency losses in the dielectric overtake and eventually contribute more to the total than the conductor loss. In this example the crossover point occurs at about 1 GHz.

This type of analysis is used to decide if a laminate system having dielectric constants and loss tangent values lower than typical FR4 should be considered. Using Figure 8.5 as a specific example, a laminate having very low loss tangent values (which lowers the dielectric loss) is not warranted if most of the signal harmonics are below roughly 1.5 GHz.

8.3.1 Loss Differences Between Microstrip and Stripline Traces

The difference in loss of microstrips and striplines is evident in Figure 8.6. By picking one frequency (the 3-dB bandwidth of a 1-ns rise time pulse), the graph clearly shows how the two physical parameters most easily adjusted by the signal integrity engineer (trace width and impedance) influence loss.

The loss for narrow traces is higher in microstrips than stripline, but the situation reverses for wider traces, where stripline losses are higher than microstrip

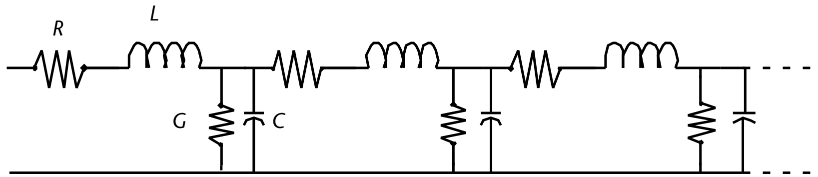


Figure 8.4 Transmission line circuit model. Resistance in the conductor and losses in the dielectric cause signal loss.

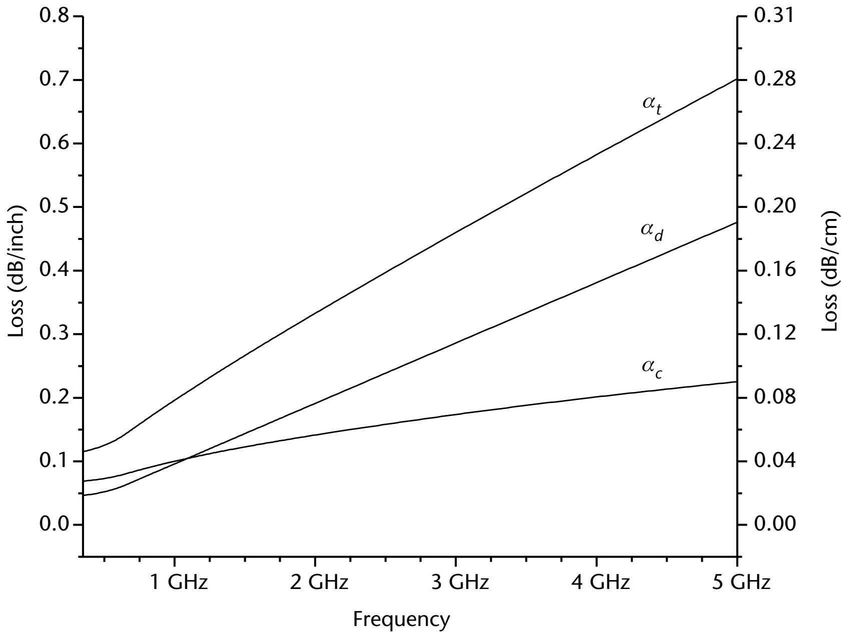


Figure 8.5 Losses for a 5-mil-wide (0.13-mm), 50Ω half-ounce copper stripline on FR4 as computed by Linpar [1]. Total loss in dB (α_t) is the sum of the dielectric (α_d) and conductor (α_c) losses.

losses. This comes about because at a given frequency the dielectric loss in stripline traces is the same for all widths, but it falls slightly as the width of microstrips increase.

8.4 How Do Impedance and Loop Resistance Affect Conductor Loss?

The conductor loss (α_c) in decibels (dB) is related to the AC loop resistance (R_{ac}, introduced in Chapter 7) and the trace impedance (Z_o) as shown in (8.3), which is valid for either stripline or microstrip traces:

$$\alpha_c = \frac{4.3 \times R_{ac}}{Z_o} \tag{8.3}$$

Because of the way in which field lines become concentrated at sharp corners, calculating R_{ac} by hand is not easy for rectangular traces. Fortunately, there are

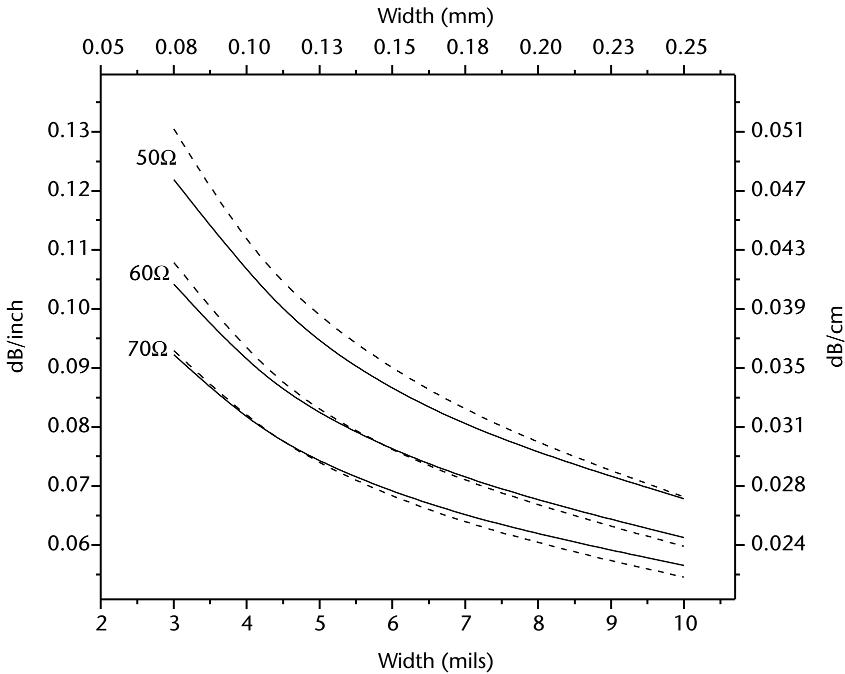


Figure 8.6 Total signal loss from Linpar [1] at 350 MHz for half-ounce rectangular stripline (solid curve) and microstrip (dotted curve) on FR4.

ways to accurately calculate microstrip and stripline loss without first calculating R_{ac}, and we will see how in Chapter 17. The purpose of introducing (8.3) here is to show how the loop resistance and impedance interrelate to affect conductor loss.

Equation (8.3) shows how increasing the trace impedance lowers the conductor loss and how increasing the AC loop resistance increases the conductor loss. This makes sense; long, narrow, thin, heavily undercut traces should have a higher resistance (and so conductor loss) than short, wide, thick, perfectly rectangular traces (which have a lower resistance).

Chapter 7 showed how the skin effect causes the loop resistance to increase as the square root of frequency. From (8.3) this means the conductor loss will also.

Equation (8.3) has another implication. For a given resistance (and therefore trace width), high-impedance traces will have a lower conductor loss than lower-impedance traces. For example, a 65Ω, 5-mil-wide trace will have a lower conductor loss than a 50Ω, 5-mil-wide trace.

The Problems show how to use (8.3) and R_{ac} to estimate the conductor loss.

8.4.1 How Does Surface Roughness Increase Conductor Loss?

Manufacturers roughen one or both sides of the copper foil sheet so it better adheres to the epoxy substrate material (see Chapter 5). An example of surface roughness is shown in Figure 8.7.

Chapter 7 showed how the skin effect causes high-frequency currents to flow on the trace surface, increasing its resistance. Any irregularities or imperfections present on the copper surface forces the highest frequency currents to flow through the peaks and valleys of the rough surface, further raising the resistance to those

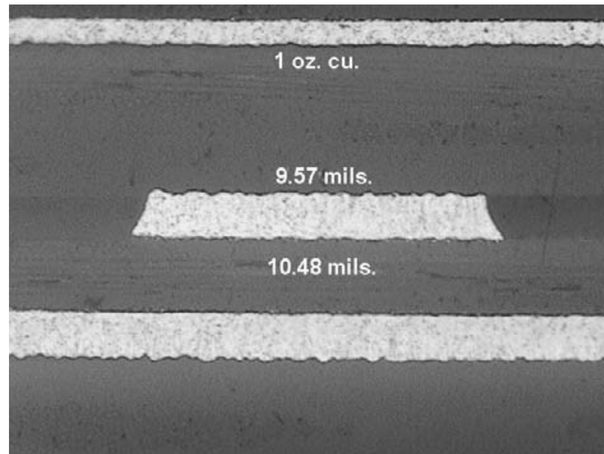


Figure 8.7 Cross-section of a 1-ounce stripline. The roughness of the trace surface is clearly evident. (Courtesy of TTM Technologies, Inc. Used with permission.)

currents. This added resistance, created by the need to roughen the copper for adhesion, becomes pronounced for frequencies where the skin depth is comparable to the average roughness (R_a) of the copper foil [2]. This is explored in the Problems, where we find this typically occurs for frequencies of a few gigahertz and more. For those currents the conductor loss is larger than predicted by the square root of f skin effect model.

The impact of surface roughness (R_a) is seen in Figure 8.8 for a 50Ω, 5-mil-wide solder mask covered microstrip to be small at frequencies of a few hundred megahertz and below, but its effects steadily increase at higher frequencies.

8.4.2 Surface Roughness in Perspective

In general, it is important to include surface roughness effects in simulations when high frequencies are present, especially if the circuit board has low dielectric losses. Dielectric losses are lowest when the loss tangent is small and, in general, when using wide microstrips. As we have seen, stripline dielectric losses are usually higher than microstrips for given impedance and trace width.

However, the impact of surface roughness on the total loss depends on the relative magnitude of the dielectric loss. For instance, at 5 GHz (which represents the fundamental frequency of a 10-Gbps binary NRZ signal stream), moderate amounts of surface roughness ($R_a = 1\ \mu$m) increase the microstrip conductor loss by 8% more than a smooth trace, when but when dielectric loss for FR4 is included, the overall total loss is only increased by about 5%.

8.4.3 What Can Be Done to Reduce Conductor Losses?

To lower conductor losses:

- Use wide traces since this lowers the loop resistance (R_{ac}).

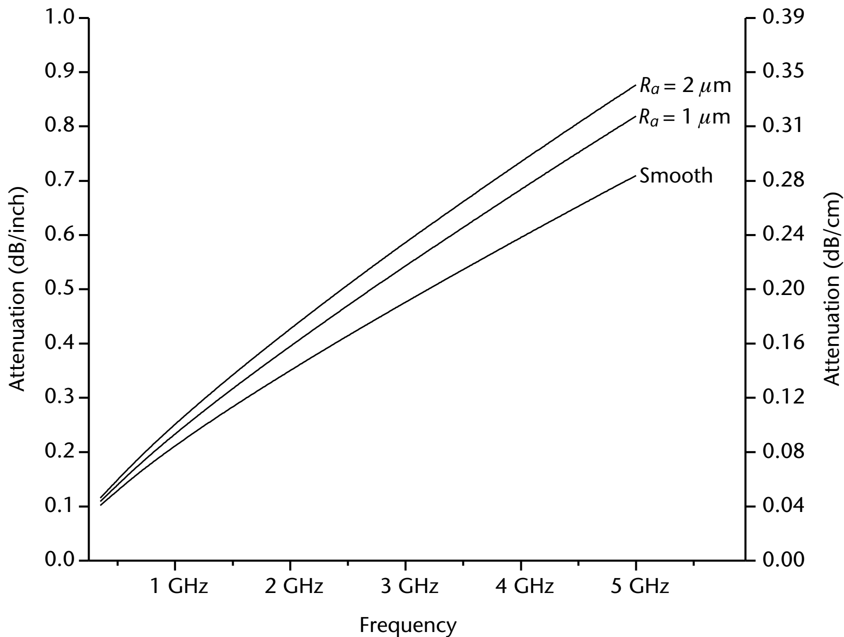


Figure 8.8 Conductor loss of a 5-mil-wide, 50Ω, half-ounce solder mask covered microstrip with and without surface roughness. A 60° tooth angle is applied to both trace and return plane on adjacent surfaces. Resistance increases slightly for smaller angles. (Data from the Simbeor 3D electromagnetic field solver [3].)

- When the traces are controlled impedance and the board is a fixed thickness, select a laminate system with a lower dielectric constant. The Problems show how this can sometimes improve conductor loss.
- Specify high-impedance traces. This increases the distance to the return plane (especially for wide traces), lowering the loop resistance.
- Use copper with low surface roughness.

Often 50Ω traces are required for high-speed signaling, which can make it impractical to lower conductor losses by signaling with higher impedance traces.

Note that because of the skin effect, much of the high-frequency current flows near the trace surface, so increasing the copper thickness will not significantly reduce high-frequency conductor losses.

8.5 Understanding Dielectric Losses

In Chapter 7 we saw how conductance (the circuit element used to represent dielectric loss in many circuit simulators) is affected by the impedance and trace width for microstrips and striplines. Here we examine dielectric losses directly, without using conductance.

The amount of signal energy lost to the dielectric is affected by frequency, the dielectric constant, and the loss tangent as shown in (8.4), which is slightly modified from that appearing in [4, 5].

$$\alpha_d = K \times f \times \sqrt{\varepsilon_r} \times LT \quad (8.4)$$

In (8.4) α_d is the dielectric loss at the frequency f (in gigahertz), and LT is the loss tangent. To get the loss in dB per inch, set K to 2.32; use 0.91 to get the loss in dB per cm. For stripline, ε_r is the dielectric constant of the laminate and prepreg, but it can be replaced with the effective dielectric constant (ε_{r_eff}) for microstrip.

Equation (8.4) shows how the dielectric loss increases with frequency and the loss tangent. The loss at 3.5 GHz will be 10 times the loss at 350 MHz (assuming that the loss tangent remains the same). It also shows how small changes to the loss tangent only slightly affect the dielectric loss. This is important because the loss tangent for many high-performance laminate systems is only moderately lower than typical FR4 values. This means that switching from FR4 to a higher-cost alternate laminate system will not reduce the total loss unless the signaling speeds are very high.

8.5.1 Differences in Dielectric Losses Between Stripline and Microstrip

Because in stripline the dielectric constant is the same for all trace widths and impedances, (8.4) shows that losses in the dielectric are not affected by stripline geometry. For example, the dielectric loss for a 5-mil-wide, 50Ω stripline will be the same as the losses for a 10-mil-wide 65Ω stripline fabricated on the same routing layer.

However, we have already seen how in microstrips the width and height above the ground plane strongly affect the effective dielectric constant (ε_{r_eff}). It is evident by using the effective dielectric constant in (8.4) that dielectric losses are affected by the geometry of a microstrip. This means that because the geometry also affects the impedance, microstrips of different impedances will have different dielectric losses. In stripline the dielectric constant is the same of all impedance values, so the dielectric losses are the same for all impedances.

This significant observation is illustrated in Figure 8.9.

Since air has a loss tangent lower than the laminate, dielectric losses will be smaller for those microstrips that have proportionally more of the signal energy propagating in air. This corresponds to wide, high-impedance microstrips because they are further away from the ground plane than narrow, low-impedance ones. It is for this reason that the dielectric losses are lower for wide, high-impedance microstrips than narrow, low-impedance ones.

8.5.2 Effects of Temperature and Moisture

The dielectric constant and loss tangent values for laminate systems are affected by temperature, humidity (moisture uptake), and the glass-to-resin ratio.

For instance, increasing the temperature from 25°C to 80°C causes the dielectric constant of typical FR4 epoxies to increase by under 10%, and the loss tangent to fall by more than 40% [6]. From (8.4) this change is significant and the actual operating temperature of the system should be factored into the loss model.

Exposing the FR4 laminate to 90% relative humidity for 200 hours can cause the loss tangent to increase by up to 35% compared to its value when dry, and, depending on the glass-to-resin ratio, the dielectric constant can increase by more than 10% [5–7].

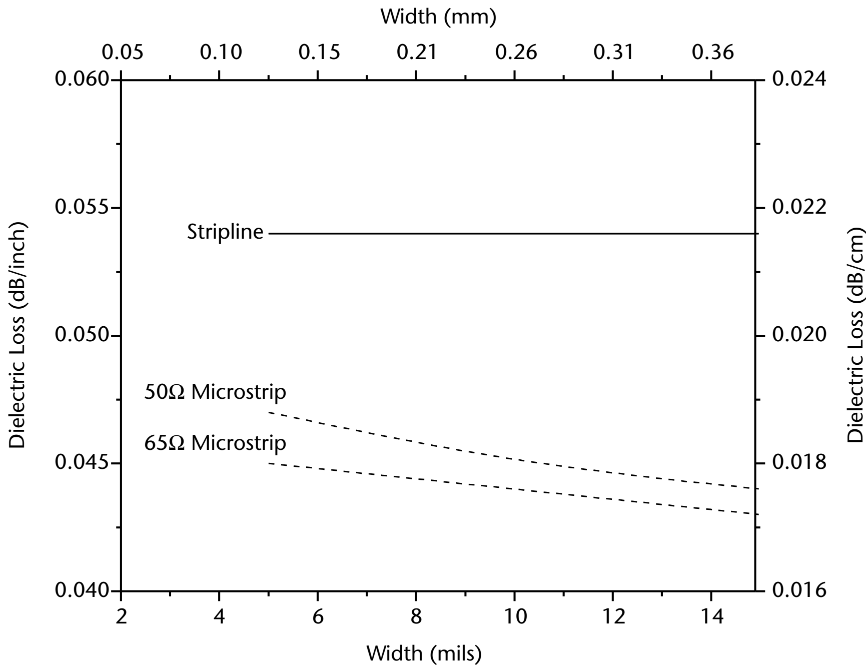


Figure 8.9 Dielectric loss of stripline is the same for all trace widths and impedances, but changes for solder mask–covered microstrip. (Data for FR4 ($\varepsilon_r = 4.2$, $LT = 0.02$) from Linpar [1].)

8.5.3 What Can Be Done to Reduce Dielectric Losses?

To lower dielectric losses:

- Use a low loss tangent laminate system.
- Use a low dielectric constant laminate system.
- Use wide, high-impedance traces when using microstrip since that lowers ε_{r_eff}.

8.6 Summary of Signal Loss and Distortion Characteristics

- A truly lossless line is distortionless. Provided the line is properly terminated, pulses sent down these transmission lines do not change shape.
- Signal energy is lost to the conductor resistance and to the dielectric.
- A lossy line distorts the pulse by attenuating the high-frequency harmonics making up the pulse.
- Loss reduces the amplitude and rounds the pulse.
- Conductor losses increase as the square root of frequency, but dielectric losses increase directly with frequency.
- The crossover frequency where the dielectric loss outweighs the conductor losses depends on the trace width, its thickness, and its impedance.

- For FR4 systems higher temperatures increase conductor resistance (and therefore loss), while the temperature lowers loss tangent (and therefore dielectric losses).

8.7 Effects of Resin Content and Glass Weave on Delay Time and Impedance

Researchers [8] have observed a variation in the dielectric constant of nearly 9% occurring across a wide bus even when all its members are routed on the same layer.

This occurs because a trace running directly on top of a glass yarn bundle in the laminate has a higher dielectric constant than one running along a resin filled trough where there is no glass. Since the pitch of the glass yarn making up the mats is typically between 16.7 and 23 mils (0.4–0.9 mm) [9], this effect is most likely to be observed with wide (multitrace) buses, but it also can be significant for high-speed serial signaling involving only two traces [10].

For instance, a 16-bit bus routed with 5-mil-wide traces and 5-mil space between traces will occupy 155 mils (0.4 cm) in width, which is several times the pitch of the glass yarn. It is very likely that at least one of the traces will route over a glass bundle while another in the same bus routes over an epoxy trough.

This effect can be ignored when timing skew between various signals does not have to be tightly controlled. When signaling at high speeds, the two extremes are addressed by creating two transmission line models: one when the trace is routed over glass and the other when routed over epoxy.

8.8 Effects of Trace Shape

Figures 5.5 and 5.7 show that actual circuit board traces are trapezoidal rather than perfectly rectangular in shape.

Because the surface area exposed to the return path will be less, a trapezoidal trace will have lower capacitance and higher inductance than rectangular traces, which increases the trace impedance. Trapezoidal traces also have less copper to support current flow, which increases their resistance and the amount of conductor loss.

Using a perfectly rectangular 5-mil-wide, 50Ω stripline as an example, R_{ac} increases by some 22%, but the inductance only increases by about 4% and the capacitance falls by about 2.5% when the top of the trace is reduced to 3.7 mils (corresponding to a heavily overetched trace). Because the impedance changes as the square root, this only results in the impedance increasing to 51Ω.

8.9 Main Points

- Loss grows with frequency, causing a pulse to be distorted in shape and reduced in amplitude as each of its harmonics is attenuated differently.

- Signal energy is lost to the trace resistance and to leakage currents in the dielectric.
- At low frequencies the conductor loss outweighs dielectric loss, but this reverses at high frequency.
- High-impedance traces have lower conductor loss than low-impedance traces.
- Surface roughness of the trace and the return planes increase the conductor loss, especially at high frequency.
- Dielectric loss increases with frequency and is determined by the laminate's loss tangent value.
- Changes in temperature alter the conductor and dielectric losses, and the dielectric loss is also affected by moisture.
- Traces are usually not rectangular in shape. This alters the transmission line's electrical characteristics.

Problems

Answers to these problems are available on the Artech House Web site at http://www.artechhouse.com/static/reslib/thierauf/thierauf1.html.

8.1 A serial interface requires a minimum signal of 175 mV peak to peak at the receiver. Assuming that the transmitter launches 800 mV peak to peak, what is the maximum loss allowed in the interconnect?

8.2 Transmission line loss is a concern in a design linking two custom-designed integrated circuits. What impedance and trace width should be selected to minimize loss?

8.3 Estimate the conductor loss of a 15.75-inch (40-cm)-long, 8-mil (0.20-mm)-wide, 50Ω half-ounce stripline on FR4 at 1.25 GHz.

8.4 In a particular application, 50Ω controlled impedance stripline traces on FR4 have too much loss. The signal integrity engineer states that although it will cost more to manufacture, both the dielectric and conductor losses can be reduced by changing to a low *Dk* laminate. Is this true, and if so, why?

8.5 To reduce cost and current consumption, a gate array manufacture has moved an old product line to a new process technology. A side effect is that the I/O signal rise time that was 1 nS becomes 700 pS. According to the manufacturer, this is advantageous because it results in more setup time for downstream logic. Considering only signal integrity, what must be considered before you can decide if it is safe to use this new version of the chip in place of the older chip in existing designs?

8.6 A 50Ω stripline has a dielectric loss of 3 dB at 2 GHz. What is the dielectric loss at 1 GHz for a 100Ω stripline that is fabricated on the same circuit board?

8.7 What the dielectric loss for a 40-cm-long stripline trace at 1 GHz when fabricated on IS410?

8.8 A 4-ns pulse has harmonics every 250 MHz. Calculate the total loss for a 5-mil-wide, 50Ω stripline on FR4 for the pulse fundamental frequency and the first five harmonics.

8.9 Determine the frequency where surface roughness is important to conductor loss.

References

[1] Djordjevic, A. R., et al., *LINPAR for Windows*, Norwood, MA: Artech House, 1999.

[2] Brist, G., et al., "Non-Classical Conductor Losses Due to Copper Foil Roughness and Treatment," *ECWC 10/IPC/APEX Conf.*, Fall 2005.

[3] Simbeor Electromagnetic Simulation Environment, Simberian Corporation.

[4] Cohn, S. B., "Problems in Strip Transmission Lines," *IEEE Transactions on Microwave Theory and Techniques*, Vol. MTT 3, No. 2, March 1955, pp. 119–126.

[5] Thierauf, S. C., *High-Speed Circuit Board Signal Integrity*, Norwood, MA: Artech House, 2004.

[6] Khan, S., "Comparison of the Dielectric Constant and Dissipation Factors of Non-Woven Aramid/FR4 and Glass/FR4 Laminates," Technical Note, DuPont Advanced Fibers Systems Division, Richmond, VA, undated.

[7] Matsushita Electronic Materials, Inc. "Megtron R5715 Technical Brochure," Forest Grove, OR, Number 199707-1.5Y, 1997.

[8] Heck, H., et al., "Impact of FR4 Dielectric Non-Uniformity on the Performance of Multi-Gb/s Differential Signals," *IEEE Conf. Electrical Performance of Electronic Packaging*, October 27–29, 2003, pp. 243–246.

[9] Brist, G., et al., "Woven Glass Reinforcement Patterns," *Printed Circuit Design & Manufacture*, November 2004, pp. 28–33.

[10] McMorrow, S., and C. Heard, "The Impact of PCB Laminate Weave on the Electrical Performance of Differential Signaling at Multi-Gigabit Data Rates," *DesignCon 2005*, Santa Clara, CA, 2005.

CHAPTER 9

Understanding Trace-to-Trace Coupling

9.1 Introduction

This chapter discusses how coupling caused by parasitic capacitance and inductance linking traces changes their impedance and propagation time. We will use this understanding in the following chapters, especially when discussing crosstalk, serial signaling, and transmission line terminations and reflections.

9.2 Understanding Mutual Capacitance and Inductance

Chapter 7 showed how capacitance is created between a trace and the ground plane (the return path). In fact, as shown in Figure 9.1(a), besides this self-capacitance (and implied but not shown, self-inductance), parasitic coupling capacitance and inductance are also present between traces.

We know from Chapter 7 that a transmission line can be thought of as a chain of very small, identical distributed circuit elements. It is apparent from Figure 9.1(b) that this is also true for the parasitic circuit elements between traces (the mutual terms). The line itself is made from an infinite number of these tiny segments, each with their own identical self and mutual elements. This means that similar to the self terms, the mutuals can be described per unit length (PUL), introduced in Chapter 6).

9.2.1 What Are the Capacitance and Inductance Matrices?

It is convenient (and mathematically desirable) to record all of the self and mutual elements as matrices.

As shown in (9.1a) and (9.1b), and in Figure 9.1, a two-digit numbering scheme is used to identify the terms.

$$L = \begin{matrix} L_{11} & L_{12} & L_{13} \\ L_{21} & L_{22} & L_{23} \\ L_{31} & L_{32} & L_{33} \end{matrix} \tag{9.1a}$$

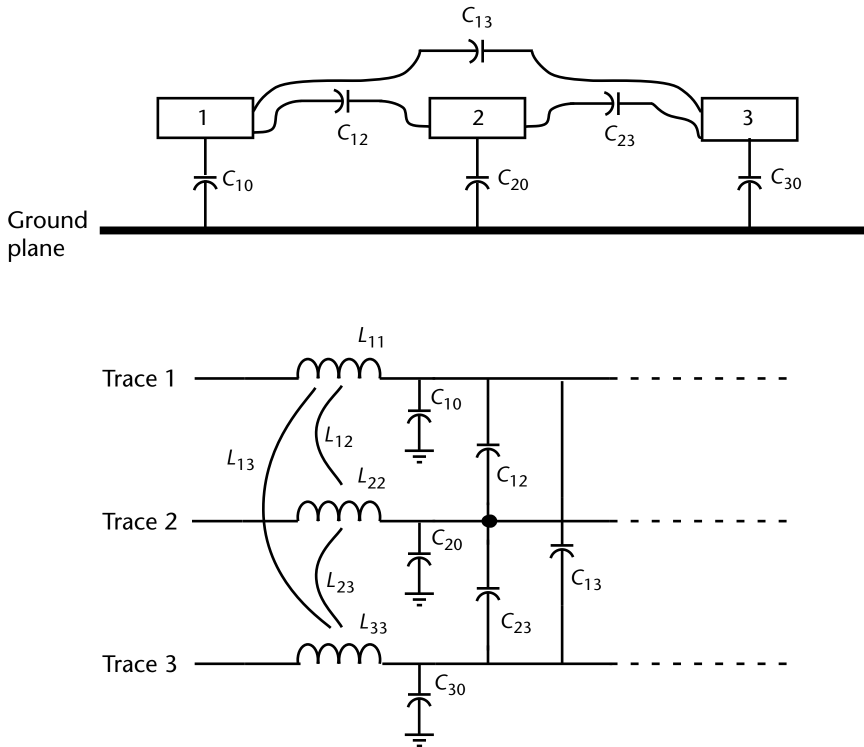


Figure 9.1 (a) Trace cross-section and (b) circuit model of a small segment of line showing the capacitances and inductances for three traces.

$$C = \begin{matrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{matrix} \tag{9.1b}$$

The self-inductance (L_s) of trace 1 is L_{11} and is measured with no current flowing in the other traces. The mutual inductance (L_m) linking trace 1 to 2 is L_{12}.

The self-capacitance (C_s) and mutual capacitance (C_m) are similar, but as shown in (9.2), C_{11} and the other diagonal terms such as C_{22} in the matrix is the capacitance to ground (C_{10} in Figure 9.1) plus the sum of all the mutual capacitances when they are grounded.

$$C_{11} = C_{10} + C_{12} + C_{13} \tag{9.2}$$

By rearranging (9.2), we find that the capacitance to ground of just trace 1 (C_{10}) is the self-capacitance (C_{11}) minus all of the mutual capacitance.

To illustrate, the inductance and capacitance matrices per meter for a pair of traces are found by a 2D field solver to be:

$$L = \begin{matrix} 2.876\text{E-}07 & 3.564\text{E-}08 & 9.859\text{E-}09 \\ 3.564\text{E-}08 & 2.863\text{E-}07 & 3.564\text{E-}08 \\ 9.859\text{E-}09 & 3.564\text{E-}08 & 2.876\text{E-}07 \end{matrix} \tag{9.3a}$$

$$C = \begin{matrix} 1.176E\text{-}10 & -5.306E\text{-}12 & -6.049E\text{-}13 \\ -5.306E\text{-}12 & 1.180E\text{-}10 & -5.306E\text{-}12 \\ -6.049E\text{-}13 & -5.36E\text{-}12 & 1.176E\text{-}10 \end{matrix} \tag{9.3b}$$

In this example, the self-inductance (L_{11}) is 287.6 nH/m, the mutual inductance from trace 1 to 2 (L_{12}) is 35.64 nH/m, and the mutual inductance from trace 1 to 3 (L_{13}) is 9.859 nH/m.

The total capacitance to ground on trace 1 (C_{11}) is 117.6 pF/m when traces 2 and 3 are grounded. The mutual capacitance from trace 1 to 2 (C_{12}) is –5.306 pF/m; it is –0.6049 pF/m from trace 1 to 3 (C_{13}). In the following sections we will see the reasoning and benefits of showing the mutual capacitances as negative numbers. Here we notice that the total mutual capacitance is the sum of the two mutual terms: –5.911 pF.

This makes intuitive sense. Because traces 1 and 2 are closer together than traces 1 and 3, the mutual capacitance between traces 1 and 2 should be higher than between traces 1 and 3, and since the self-capacitance includes all the mutual capacitance, C_{11} must be higher than the value of the mutual terms. This is further explored in the Problems.

By examining these matrices, we notice that the mutual terms on either side of the diagonal are the same. For instance, L_{12} has the same value as L_{21} (35.64 nH/m). This is reasonable because we would expect the coupling between traces 1 and 2 to be the same whether it is measured from traces 1 to 2, or from traces 2 to 1. SPICE-type simulators take advantage of this symmetry and require that only half of the matrix be entered when creating transmission line models.

9.2.2 Why Do Some Field Solvers Report Capacitances with Negative Numbers?

Field solvers commonly report the mutual capacitances as negative numbers, but this is not a universal practice. The negative value for the mutual capacitance is mathematically correct and is an artifact of the way in which charge is calculated when the field solver computes the capacitance of a structure. Those interested in these details are referred to [1–3].

Although the negative values are proper, their presence is a common source of confusion because all of the hand calculations treat the mutual capacitance values as positive numbers.

As shown in the Problems, calculations throughout this book always take the mutual capacitance as positive, even if the field solver reports them as negative.

9.3 How Does Switching Alter the Trace Inductance and Capacitance?

Figure 9.2 shows a trace flanked by neighboring traces. The aggressors switch at the same time as the victim trace, and all of them have the same signal amplitude. This implies that the voltage on each trace has the same rate of change. In practice loading differences may make the signal rise times unequal, but to simplify the analysis we will assume that they are the same.

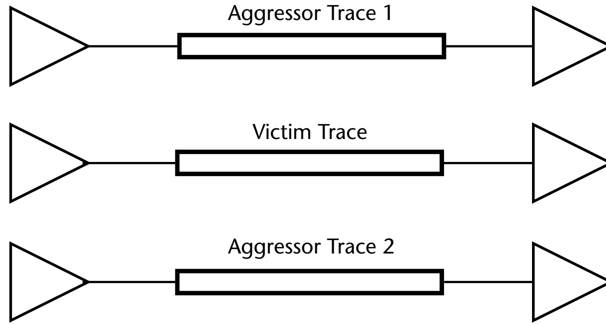


Figure 9.2 Top view showing a victim trace flanked by neighboring aggressor traces.

9.3.1 How Is the Capacitance Affected?

The capacitance for a tiny segment of the three traces illustrated in Figure 9.1 and 9.2 appears in Figure 9.3.

Figure 9.3 shows that when the victim and aggressors simultaneously switch in phase with identical rise times, the voltage at each terminal simultaneously reaches 1V. This makes the voltage difference across either of the mutual capacitors (C_{12} or C_{23}) 0, which means no current flows and these capacitors can be removed without altering the circuit behavior. The net result is that the total switched capacitance of trace 2 is only C_{20}: its capacitance to ground.

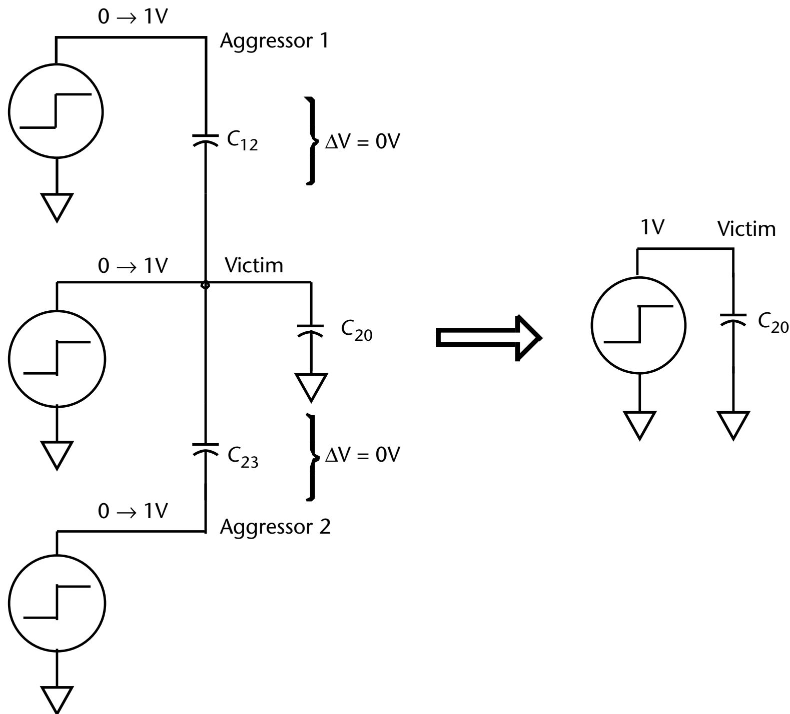


Figure 9.3 When the victim and aggressor traces switch in phase, the victim capacitance is the self-capacitance to ground (C_{20}).

The situation is quite different when then signals switch out of phase. Figure 9.4 shows the victim switching low to high while the aggressors surrounding it switch in the opposite direction.

In this case the aggressors switch from 1V to 0V while the victim switches from 0V to 1V. The mutual capacitances C_{12} and C_{23} experience a 2V change in voltage, effectively multiplying their value by a factor of 2. As shown in Figure 9.4, this makes the total capacitance on the victim trace equal to twice the mutual capacitance plus C_{20} (the capacitance of the victim trace to ground).

From (9.2) the victim's self-capacitance (in this case, C_{22}) already includes the mutual capacitance. As shown in Figure 9.4, this allows us to simplify the total capacitance to be the sum of the self-capacitance (C_{22}) plus the mutual capacitances.

9.3.2 How Is the Inductance Affected?

From circuit theory (for example, see [4]), the mutual inductance occurring between side-by-side wires can either increase or decrease the total circuit inductance depending on the direction of the current flow. In particular, when the signals switch out of phase, the currents in the aggressor lines are flowing in the opposite direction to the current in the victim, and the mutual inductance subtracts from the self-inductance.

However, the currents flow in the same direction when the signals switch in phase. In this situation the mutual inductances add to the self inductances. This useful observation will be used later in this chapter when we analyze switching modes.

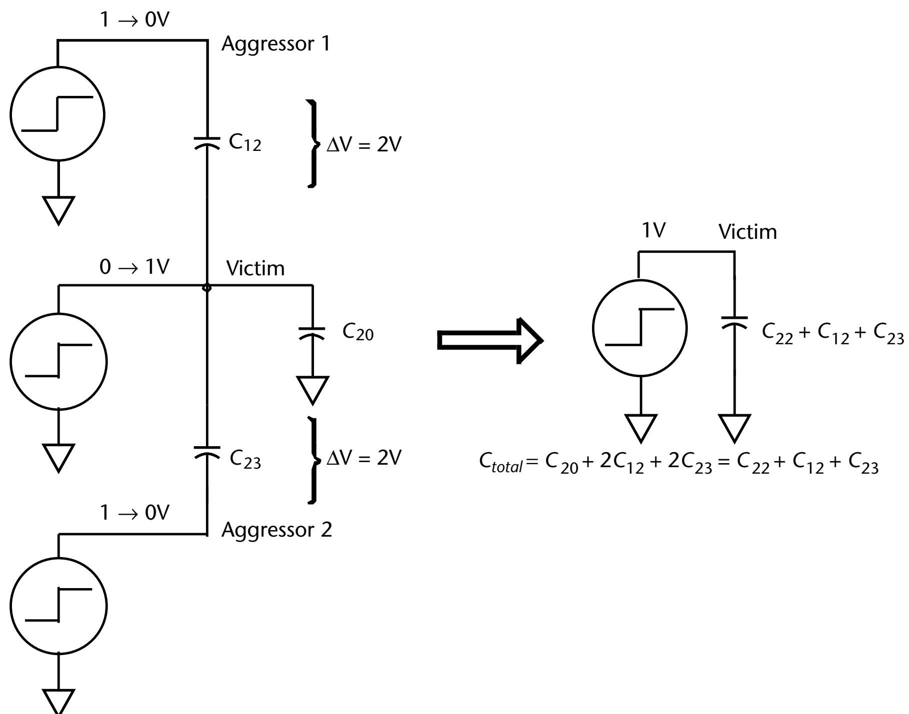


Figure 9.4 When the victim and aggressor traces switch out of phase, the victim capacitance is the capacitance of C_2 plus the mutual capacitances ($C_{22} + C_{12} + C_{23}$).

9.4 What Effect Does Coupling Have on Impedance and Delay?

The circuit setup shown in Figure 9.5 shows four aggressor microstrip traces (A4 through A1) surrounding a victim (V) microstrip. They are designed to have an impedance of 50Ω, and are driven from ASIC 1 with identical I/O drivers.

By systematically switching each aggressor in or out of phase with the victim trace, we can determine how the mutual capacitance and inductance affect the victim impedance and delay for each switching mode. This is shown in Table 9.1, where a "+" indicates that the signal transition is in the same direction, while a "–" means the signals swing in opposite directions. A signal that does not switch (is static) is shown with an S.

Evidently, even though the traces are designed to be 50Ω, they only have that value when their neighbors are static. The impedance is lowest when all of the aggressors are out of phase with the victim, and it is highest when they are all in phase.

The switching activity also affects the microstrip (but not stripline) propagation time. Microstrip delay is shortest when the signals are all out of phase, and it is the longest when they are all in phase. As we saw in Chapter 7, stripline propagation delay is determined by the dielectric constant, which is why it is the same for all switching modes and so is not shown in Table 9.1.

However, data-dependent timing jitter will occur at the output of receivers connected to either stripline or microstrip traces. This behavior is a property of the circuit rather than the transmission line and is an artifact of the changing impedance

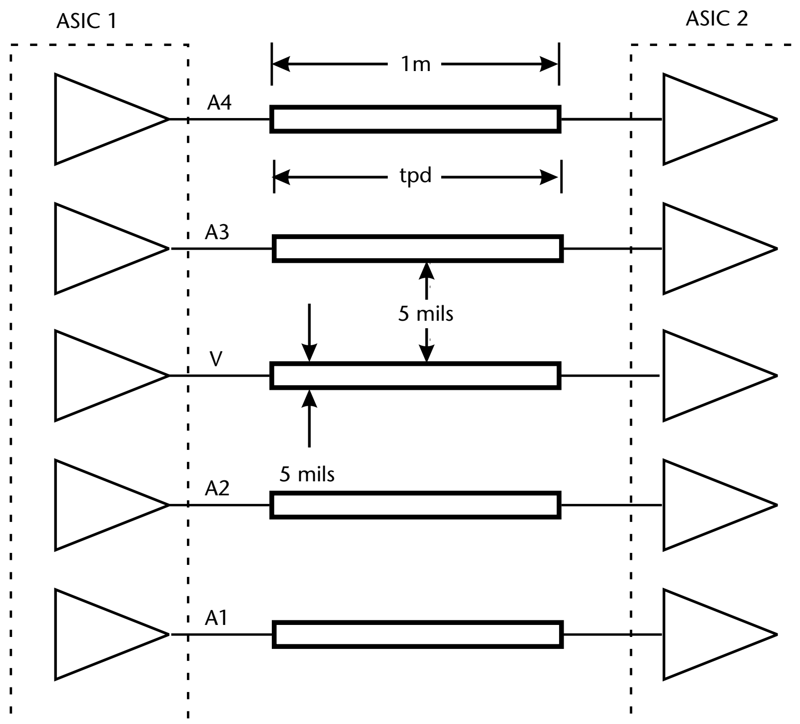


Figure 9.5 Top view of four aggressor microstrip traces surrounding a victim trace connecting ASIC 1 to ASIC 2 and driven by identical drivers.

Table 9.1 Impedance and propagation delay of the victim 50Ω microstrip in Figure 9.5.

A4	*A3*	*V*	*A2*	*A1*	*Z*	*tpd*
–	–	+	–	–	42	6.31
+	–	+	–	+	43	6.43
S	S	+	S	S	50	6.46
–	–	+	+	+	52	6.54
–	+	+	+	–	63	6.65
+	+	+	+	+	69	6.92

Transmission line impedance (*Z*) is in ohms and the delay (*tpd*) is in nanoseconds/meter.

of the victim trace causing different voltages to be launched when neighboring traces switch.

To summarize:

- Stripline and microstrip victim traces have the nominal impedance when all neighbors are static.
- The lowest impedance for stripline or microstrip victims occurs when all neighbors switch out of phase.
- The highest impedance of stripline or microstrip victims occurs when all neighbors switch in phase.
- Stripline propagation delay is not affected by the switching activity of its neighbors.
- Microstrip has the nominal delay when all its neighbors are static.
- Microstrip has the lowest propagation time when all its neighbors switch out of phase and has the largest propagation time when all its neighbors switch in phase.

9.5 What Are Odd and Even Modes?

In a two-trace system the switching mode is said to be even when the traces switch in phase; it is odd when they are out of phase [5, 6].

Differential signaling (as discussed in Chapter 13, where one trace carries a positive going signal and the other carries its complement) is an example of odd-mode switching since the two signals are always intended to be out of phase.

From Figure 9.6 with even-mode switching, the inductive coupling moves in the same direction and encompasses both traces, while the capacitive coupling does not include the neighboring traces. From this we would expect the total inductance to be equal to the self plus the mutual inductance and the total capacitance to be only the capacitance to ground. Since by its definition the self-capacitance includes all of the mutual capacitance, to obtain the total even-mode capacitance, the mutual capacitance must be subtracted from the self-capacitance.

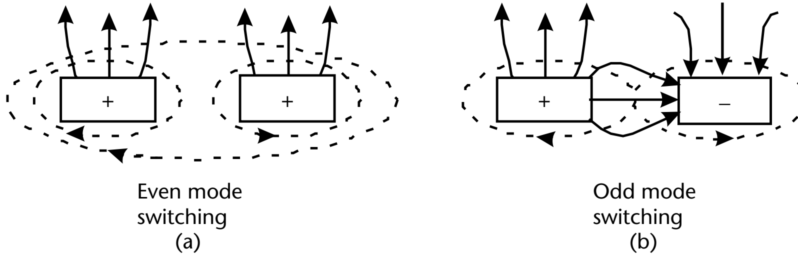


Figure 9.6 Edge view of two pairs of traces switching in (a) even mode and (b) odd mode. Electric fields (capacitive coupling) are shown as solid lines and magnetic fields (inductive coupling) are shown as dotted lines.

This contrasts with odd mode where the magnetic field lines move in opposite directions and the electric fields extend from one trace to the other. From this we would expect the inductance to only include the self-inductance and the capacitance to include both the self- and mutual capacitances.

This reasoning is confirmed in (9.4) and (9.5) for even-mode switching and in (9.6) and (9.7) for odd-mode switching [5]. These equations may be used for either microstrip or striplines.

$$Z_{oe} = \sqrt{\frac{L_s + L_m}{C_s - C_m}} \tag{9.4}$$

$$tpd_{even} = \sqrt{(L_s + L_m)(C_s - C_m)} \tag{9.5}$$

$$Z_{oo} = \sqrt{\frac{L_s - L_m}{C_s + C_m}} \tag{9.6}$$

$$tpd_{odd} = \sqrt{(L_s - L_m)(C_s + C_m)} \tag{9.7}$$

The even-mode impedance is called Z_{oe} and odd-mode impedance is called Z_{oo} to distinguish them from the characteristic impedance (Z_o). Similarly, t_{pd_even} and t_{pd_odd} distinguish the even- and odd-mode delays from the propagation time of a single trace.

In the equations L_s and C_s are the self-inductance and capacitance as displayed by a field solver, and L_m and C_m are the sum of all the mutual terms. All are per unit length. For instance, from the capacitance matrix in (9.3b), C_s is 117.6 pF/m and C_m is −5.91 pF/m.

We will return to these equations in future chapters. Here we again note that the mutual capacitance is taken as positive in the equations, even if they are reported as negative. As shown in the Problems, the even-mode impedance for trace 1 using the information in (9.3) is 54.6Ω, and not 51.9Ω because +5.9 pF is used for C_m, even though it is reported as −5.9 pF.

9.5.1 Can the Odd- and Even-Mode Equations Be Used with More Than Two Traces?

As demonstrated in the Problems, by summing all of the mutuals and using them as L_m and C_m, (9.4) through (9.7) can be used to find the impedance and delay of a victim trace when many neighbors switch.

9.5.2 Why Is the Even- and Odd-Mode Timing the Same for Stripline But Not for Microstrip?

From Chapter 7, we know that in stripline the propagation delay is determined by the laminate dielectric constant and not by the trace shape or impedance. This appears to conflict with (9.5) and (9.7) because they show that, in general, the timing for even and odd modes is different.

We must turn to field theory to resolve this apparent paradox. It can be shown [7, 8] that for the types of transmission lines of interest to us (lines where the electric and magnetic fields travel crosswise to the direction in which the signal is moving, called transverse electromagnetic, or TEM, transmission), the even-mode and odd-mode propagation delays are the same. The significance of this is that because striplines with low loss are true TEM lines, the even-mode propagation delay given by (9.5) will be the same as the odd-mode propagation delay given in (9.7). In fact, when used with striplines, they will give the same result as (7.5) where we used the dielectric constant to find the delay. This is summarized in Figure 9.7 and explored in the Problems.

However, the situation is very different in microstrips because even when they are lossless, the two or more dielectrics involved means the signal propagation is not truly TEM. In fact, microstrip delay is determined by how much of the signal energy propagates in air and how much propagates in the laminate. This makes the delay shorter when the victim and aggressors are out of phase because under those conditions more of the signal energy travels in the air than in the laminate. The Problems illustrate this further by showing that the odd-mode microstrip propagation time is smaller than the even-mode propagation times.

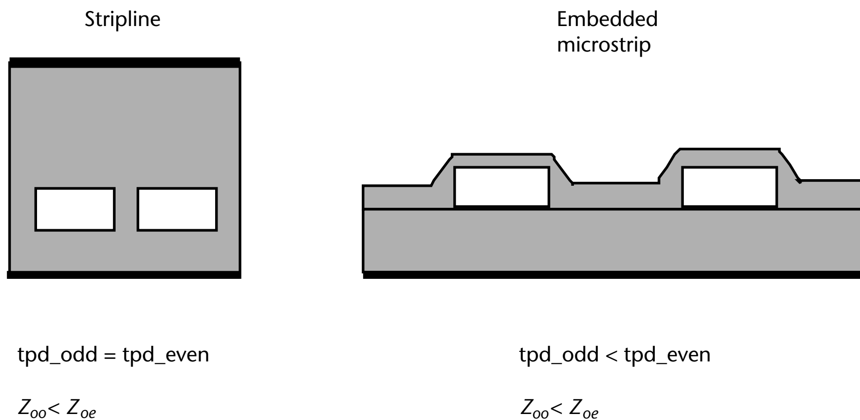


Figure 9.7 Odd- and even-mode delays are the same in stripline traces but not in solder mask–covered microstrips.

9.5.3 By How Much Does the Even- and Odd-Mode Impedance Change?

We saw in Table 9.1 how coupling between traces during switching changes the victim trace impedance. The coupling can be reduced by increasing the spacing, which makes the even- and odd-mode impedances equal. This is confirmed in Figure 9.8, which shows the odd- and even-mode impedances with respect to the edge-to-edge spacing for a pair of 5-mil-wide, half-ounce 50Ω and 65Ω stripline and microstrip without solder mask traces on FR4.

The difference between the even- and odd-mode impedances is greatest when the separation is small and the impedance is high, since those conditions make coupling the highest. Increasing the trace-to-trace spacing reduces the coupling and causes the even-mode impedance to fall and the odd-mode impedance to increase. Eventually the traces are far enough apart to where they both equal the value the traces would have in isolation.

Graphs such as Figure 9.8 are used to judge the relative coupling between a pair of traces. For instance, we see the 50Ω traces have a reasonably loose coupling (indicated by the small difference between the even- and odd-mode impedances) with an edge-to-edge separation of 20 mils. However, the impedance difference is quite a bit larger for the 65 traces. This shows that even though the traces are the same width, for a 20-mil spacing the 50Ω traces are loosely coupled, but the 65Ω traces are not.

This observation highlights the danger of applying simple rules of thumb to determine if coupling is severe between traces. For instance, the 3-W rule is sometimes used to determine how close together traces can be routed. This well-known rule [10] is illustrated in Figure 9.9 and holds that to minimize coupling the center distance between a pair of traces should be three or more times the trace width.

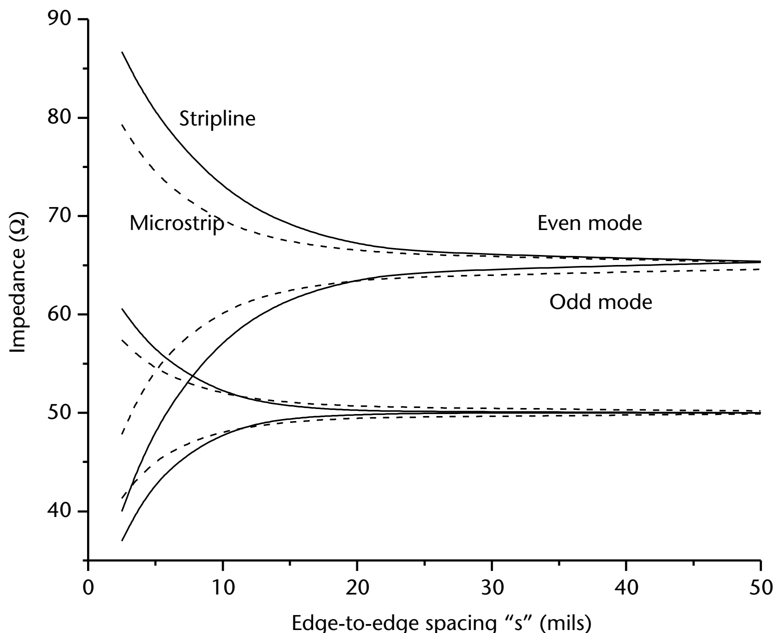


Figure 9.8 Odd- and even-mode impedance for a 5-mil-wide, half-ounce 50Ω and 65Ω stripline (solid curve) and microstrip (dotted curve) on FR4. (*From:* [9]. Used with permission.)

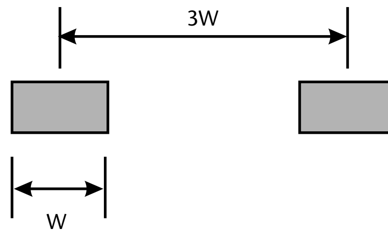


Figure 9.9 Side view of two traces showing the 3W rule of thumb.

Using this rule with 5-mil-wide traces would result in traces being routed as close together as 10 mils. Judging from Figure 9.8, for the 65Ω trace the coupling is quite high at that distance and it must be increased to over 25 mils before the coupling is reduced to a reasonable value.

9.6 What Are the Circuit Effects When Switching Causes the Impedance to Change?

Chapter 6 showed how voltage divider action between the impedances of the trace and driver determines the voltage launched down the victim trace. Because switching activity of neighboring traces causes the victim impedance to change, neighbors switching causes the launched voltage to change even if the power supply voltage is held constant.

This is illustrated in Figure 9.10, which shows the voltage launched down the transmission line circuit appearing in Figure 9.5, which was used to create Table 9.1. The drivers have a drive strength of 32 mA.

Figure 9.10 illustrates the two worst-case situations that bound the results. Even though the power supply is 3.3V, the voltage launched by a 32-mA driver is 3.1V when the victim and neighbors are all switching in phase (corresponding to a victim impedance of 69Ω in Table 9.1), but it falls to 2.9V when the neighbors switch out of phase (the 42Ω impedance case). As explored in the Problems, a weaker driver would launch a lower voltage and the difference in launched voltages would be greater then the 200 mV shown in Figure 9.10.

9.7 How Is Receiver Timing Affected by In- and Out-of-Phase Switching?

Table 9.1 shows how the switching activity of neighboring microstrips causes the trace delay to vary from a low of 6.31 ns to a high of 6.92 ns, for a total variation of 610 ps. However, as Figure 9.11 illustrates, the variation is 650 ps when the measurement is made at the output of the victim receiver.

The added delay is caused by the difference in launched rise times between the in-phase and out-of-phase cases. With an impedance of 69Ω, the in-phase signaling case represents less capacitive loading to the driver and so has a sharper rise time than the out-of-phase case, which has an impedance of 42Ω. The result is that more time is required for the signal to reach the receiver switch point. This loading

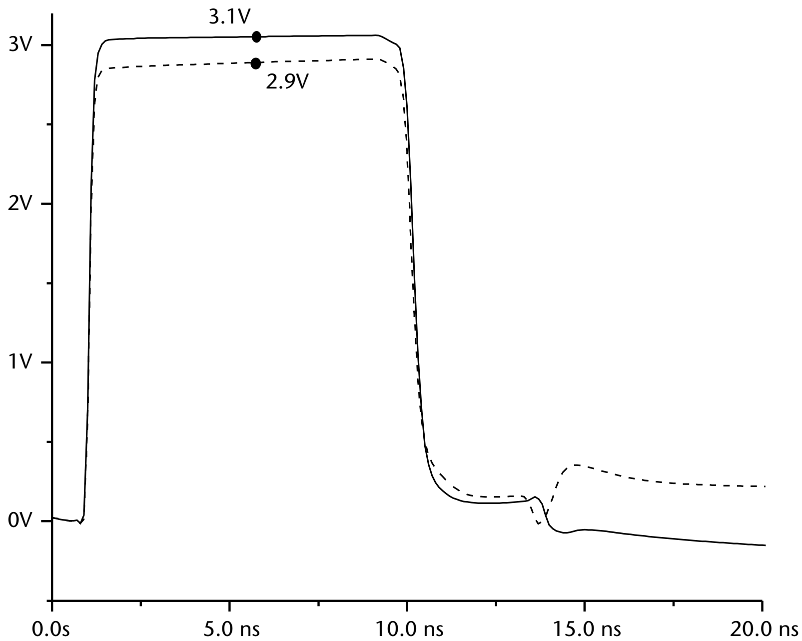


Figure 9.10 Launched voltage is lower when the signals are out of phase (dotted curve) than when in phase (solid curve).

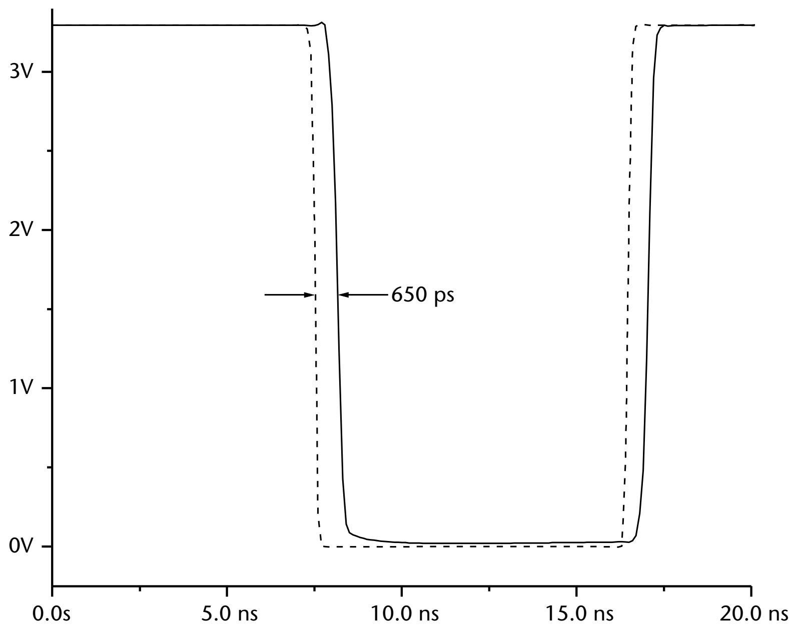


Figure 9.11 Victim receiver responds sooner when the traces switch out of phases (dotted curve) than when they are in phase (solid curve).

behavior is characteristic of CMOS I/O drivers and is independent of any parasitic inductance in the power or signal leads.

9.8 Main Points

- Inductive and capacitive coupling is present between traces.
- The switching activity of neighboring aggressor traces determines the victim trace impedance and for microstrip (but not stripline) the propagation delay.
- When only two traces are involved, the switching patters are called even mode and odd mode.
- Odd mode occurs when the victim and aggressor switch out of phase. This gives the lowest impedance and for microstrip the shortest propagation delay.
- Even mode occurs when the victim and aggressor switch in phase. This gives the highest impedance and for microstrip the longest propagation delay.
- The even- and odd-mode delay is the same for stripline.
- The change in impedance caused by switching modes affects the launched voltage for both microstrip and stripline traces.

Problems

Answers to these problems are available on the Artech House Web site at http://www.artechhouse.com/static/reslib/thierauf/thierauf1.html.

9.1 Under switching conditions a victim trace in a two-trace system has the following impedance values: 36Ω, 60Ω, and 97Ω. Identify the odd, even, and characteristic impedances.

9.2 Under switching conditions a victim trace in a two-trace system has the following propagation times: 154 pS/inch and 142 ps/inch. Identify the odd and even propagation delays.

9.3 A field solver produced the following output file for a five-trace system. Identify L_s, L_m, C_s, and C_m for trace 3.

L MATRIX [H/M]

2.875E-07	3.557E-08	9.705E-09	4.228E-09	2.416E-09
3.557E-08	2.862E-07	3.530E-08	9.620E-09	4.228E-09
9.705E-09	3.530E-08	2.861E-07	3.530E-08	9.705E-09
4.228E-09	9.620E-09	3.530E-08	2.862E-07	3.557E-08
2.416E-09	4.228E-09	9.705E-09	3.557E-08	2.875E-07

C MATRIX [F/M]

1.176E-10	–5.300E-12	–5.833E-13	–2.531E-13	–1.521E-13
–5.300E-12	1.180E-10	–5.261E-12	–5.677E-13	–2.531E-13
–5.833E-13	–5.261E-12	1.180E-10	–5.261E-12	–5.833E-13
–2.531E-13	–5.677E-13	–5.261E-12	1.180E-10	–5.300E-12
–1.521E-13	–2.531E-13	–5.833E-13	–5.300E-12	1.176E-10

9.4 Use the inductance and capacitance matrix from Problem 9.3 to find the two worst-case impedance values for trace 3.

9.5 Use the inductance and capacitance matrix from Problem 9.3 to find the two worst-case propagation times.

9.6 Does the matrix in Problem 9.3 represent a microstrip or stripline?

9.7 A field solver produced the following output file for a two-trace system. Is this the matrix for a stripline or microstrip?

L MATRIX [H/INCH]	
1.052e-8	2.040e-9
2.040e-9	1.052e-8
C MATRIX [F/INCH]	
2.835e-12	–5.490e-13
–5.490e-13	2.835e-12

9.8 Compare the results from (9.5) and (9.7) with (7.5) for the two-trace system given in Problem 9.7. Assume that the dielectric constant is 4.0.

9.9 Use the matrices in Problem 9.3 to show how the coupling falls as the distance from a given trace increases. Also show that the traces are symmetrical.

9.10 What voltage will be launched down the transmission line described by the matrix appearing in Problem 9.7, assuming that the ASIC I/O driver has an output impedance of 25Ω and is connected to a 2.5-V supply? How does the voltage change when the driver impedance is reduced to 10Ω?

9.11 Use the matrices in Problem 9.7 to show that incorrect results are obtained when negative values are used for mutual capacitance when calculating the even- and odd-mode delays and impedances.

References

[1] Young, B., *Digital Signal Integrity*, Upper Saddle River, NJ: Prentice-Hall, 2001.

[2] Paul, C. R., *Analysis of Multiconductor Transmission Lines*, New York: John Wiley & Sons, 1994.

[3] Baloglu, H. B., *Circuits, Interconnections, and Packaging for VLSI*, Reading, MA: Addison-Wesley, 1990.

[4] Johnson, D., et al., *Basic Electric Circuit Analysis*, 5th ed., Upper Saddle River, NJ: Prentice-Hall, 1995.

[5] Marx, K. D., "Propagation Modes, Equivalent Circuits, and Characteristic Terminations for Multiconductor Transmission Lines with Inhomogeneous Dielectrics," *IEEE Transactions on Microwave Theory and Techniques*, Vol. MTT-21, No. 7, July 1973, pp. 450–457.

[6] Trinogga, G., et al., *Practical Microstrip Circuit Design*, Upper Saddle River, NJ: Prentice-Hall, 1991.

[7] Matick, R., *Transmission Line for Digital and Communication Networks*, New York: McGraw-Hill, 1969.

[8] Ramo, S., et al., *Fields and Waves in Communication Electronics*, 3rd ed., New York: John Wiley & Sons, 1994.

[9] Thierauf, S. C., *High-Speed Circuit Board Signal Integrity*, Norwood, MA: Artech House, 2004.

[10] Montrose, M. I., *Printed Circuit Board Design Techniques for EMC Compliance*, New York: IEEE Press, 1996.

CHAPTER 10

Understanding Crosstalk

10.1 Introduction

Crosstalk is the coupling of signal energy from one or more aggressor signals to one or more victim signals. This causes unintended currents to flow in the victim, which creates noise voltages.

When designers think of crosstalk, they often only consider how the noise voltage causes false triggering at a receiver and so disturbs system timing or reduces noise margin. Signal integrity engineers take a wider view and recognize that in addition to those types of problems, crosstalk is capable of causing pulse distortion and (as described in Chapter 2) damage to the I/O cell if the transient voltages are large enough.

10.2 How Is Crosstalk Created and What Are Its Characteristics?

The mutual capacitance and inductance appearing between the aggressor and victim traces [1–6] create the coupling responsible for crosstalk. This coupling was introduced in Chapter 9 and is shown for two traces in Figure 10.1.

Although only two traces are shown, the mutual capacitance (C_m) and mutual inductance (L_m) represent the sum of the electric and magnetic coupling from any number of aggressors to the victim.

The electric and magnetic coupling causes current to flow in the victim [8], as shown in Figure 10.2.

By making the loads resistors equal in value to the impedance of the line, Ohm's law can be used to find the load voltages, and reflections (discussed in Chapter 11) are eliminated.

The current injected by the mutual capacitance (I_c) flows in the victim trace with half flowing toward the near end and the other half flowing toward the far end. The current induced by the mutual inductance (I_l) does not behave in this way because Lenz's law causes the victim current to flow in the direction opposite to the aggressor current (I_s). For these things to be so, the victim and aggressors must be far enough apart that the presence of the victim does not change the aggressors' impedance (and so the voltages coupled to the victim cannot influence the aggressor).

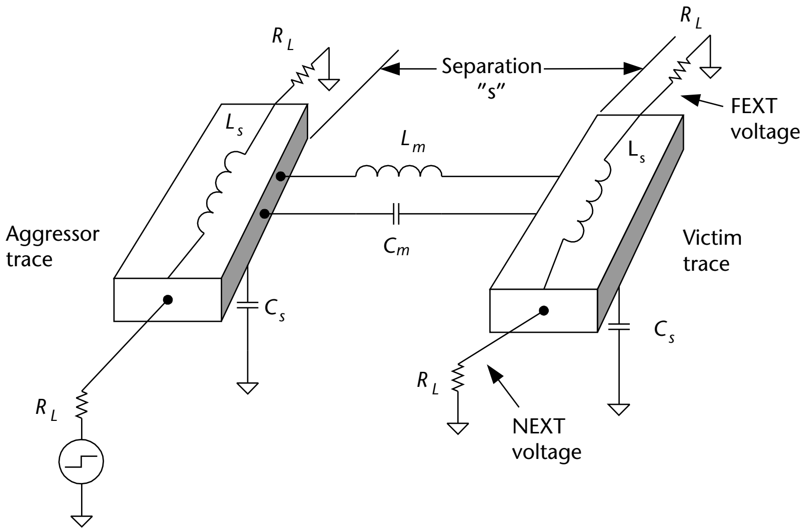


Figure 10.1 Electric and magnetic coupling between traces transfer energy from the aggressor to victim trace. The mutual capacitance (C_m) represents the electric coupling, and mutual inductance (L_m) represents the magnetic coupling. (*From:* [7]. Used with permission.)

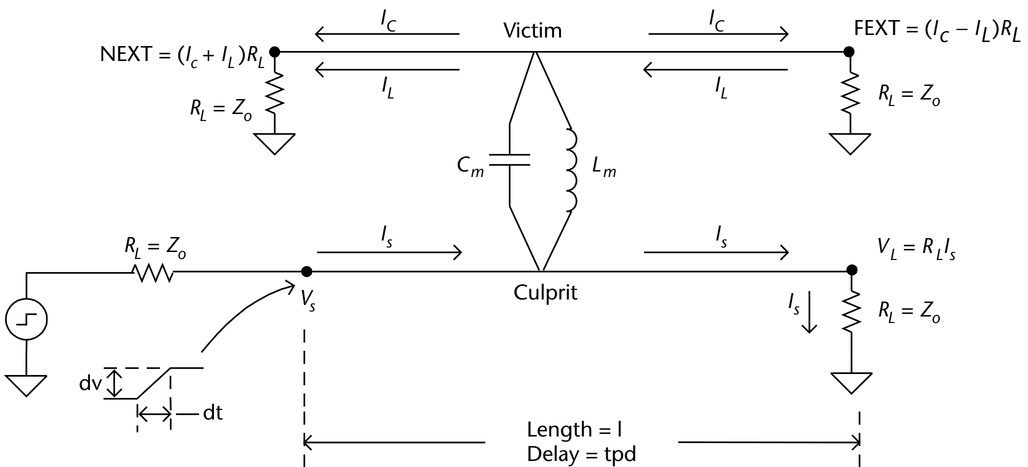


Figure 10.2 Mutual capacitance and inductance cause current to flow in the tiny portion of the victim trace shown. (*From:* [7]. Used with permission.)

The lines are considered to be loosely coupled in this way when the coupling is under 25% [3].

The result is that the current from the mutual inductance flows back toward the source, while the current from the mutual capacitance flows in both directions.

The currents combine to create a total noise current at each end of the victim. At the near end the noise current is the sum of the two and creates the near-end crosstalk voltage (NEXT). At the far end the two noise currents flow in opposite directions, making the total the difference between the current coupled by the mutual capacitance and the current coupled by the mutual inductance. The combined

current creates far-end crosstalk (FEXT) when it flows through the far-end load resistance.

It is apparent from this discussion that FEXT and NEXT have very different electrical properties, and in fact they must be analyzed separately.

10.3 Far-End Crosstalk

If the current from magnetic coupling is greater than the capacitive coupling current, the victim FEXT pulse will have a polarity opposite of the aggressor pulse. The polarity will be the same if the magnetic coupling is less than the capacitive coupling. If the magnetic and capacitive couplings are equal, the currents cancel and there is no FEXT pulse. This is the case with stripline, provided that the losses are low. However, because the currents never fully cancel, FEXT is always possible in microstrips.

On circuit boards the coupling is usually mostly capacitive to nearby traces, but it becomes increasingly inductive to traces progressively further away [1]. This is why on circuit boards FEXT polarity is usually opposite of the aggressor pulse polarity, but it can have the same polarity if the coupled distances are great. This means that both the amplitude and the polarity of the FEXT pulse depend on the switching behavior of neighboring traces.

As illustrated in Figure 10.3, the FEXT pulse has a width equal to the rise time of the aggressor signal (*dt*, the change in time). This is because there is no DC connection between the traces, so coupling only occurs while the aggressor pulse is transitioning. This AC connection means that the FEXT pulse is no longer created once the aggressor pulse has reached its steady state value. In fact, FEXT is determined by the change in voltage (*dv*) rather than the magnitude of the voltage, and because the aggressor pulse has two edges, two FEXT pulses are created from a single aggressor pulse.

10.3.1 Calculating FEXT

The forward crosstalk coupling factor (*Kf*) presented in (10.1a) shows the amount and polarity of the coupling between traces. It has units of time per unit length (for example, picoseconds/inch) and is used with (10.1b) and (10.1c) to find the actual FEXT voltage. [4, 5, 8–10].

$$Kf = -0.5 \times \left(\frac{L_m}{Z_o} - \left(Z_o \times C_m \right) \right) \tag{10.1a}$$

$$FEXT = Kf \times length \times \frac{dV}{dt} \tag{10.1b}$$

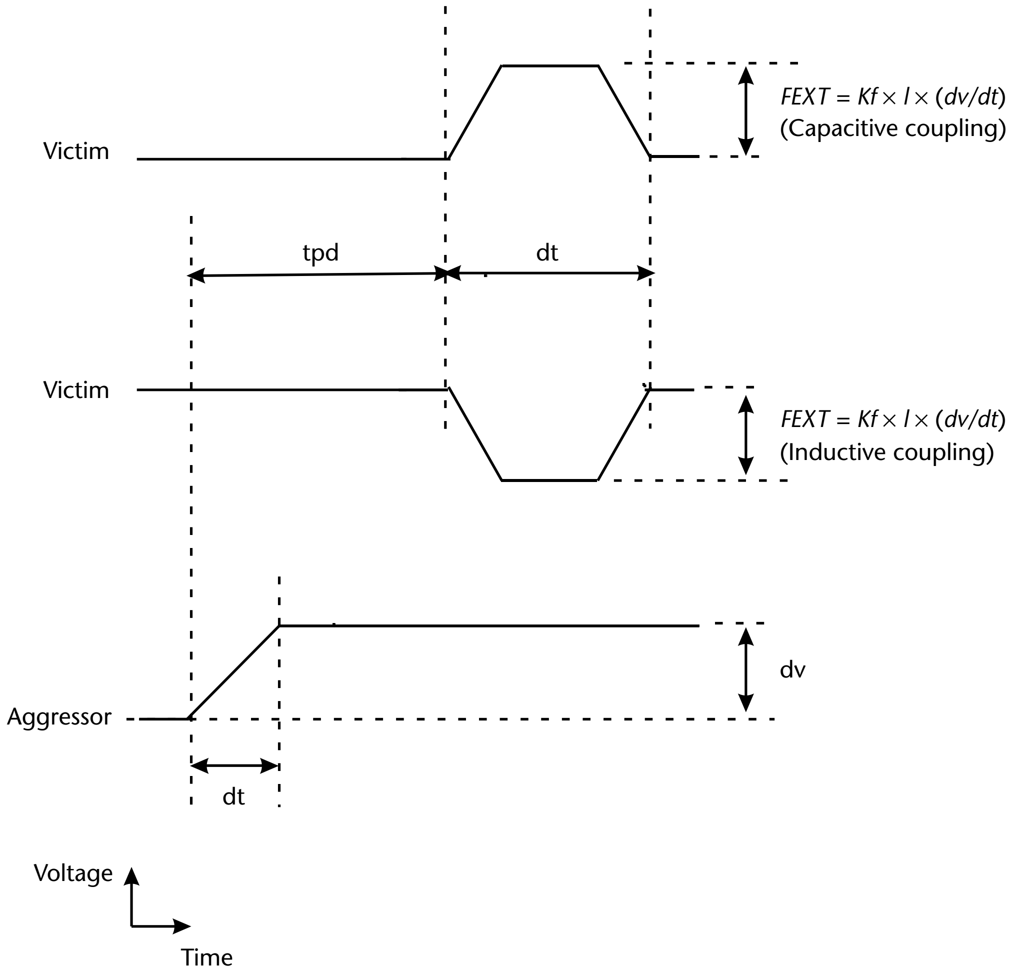


Figure 10.3 FEXT victim pulse width equals the aggressor rise time (*dt*). It has the same polarity as the aggressor when capacitive coupling dominates but the opposite when inductive coupling dominates.

$$Z_o = \sqrt{\frac{L_s}{C_s}} \tag{10.1c}$$

The sum of all the mutuals can be used in these equations to find *Kf* when more than one trace aggresses a victim. The Problems include a worked example of how to do this.

Because the mutual capacitance (C_m) term subtracts from the mutual inductance term (L_m), it is apparent from (10.1a) how the inductive and capacitive effects can cancel one another as happens with stripline. This is explored in the Problems. It is also apparent how either the mutual inductance or capacitance can become dominant to determine the polarity of *Kf*.

The change in voltage (*dv*) and time (*dt*) emphasize that it is the difference between voltage levels (rather than the magnitude) that creates FEXT. For instance, a signal with a 1-ns rise time switching between 0.6V and 1.6V will induce the same

FEXT amplitude as a 1-ns rise time signal switching from 0V to 1V since they both have a 1V-per-nanosecond swing.

It is apparent from (10.1b) that the FEXT amplitude grows as *Kf* is made larger, as the length of the coupling region increases, as the voltage swing (*dv*) increases, and when rise time (*dt*) is small. Therefore, FEXT is combated by lowering the coupling between traces since that lowers *Kf*, keeping the coupled length short, reducing the voltage swing of the aggressor, and lengthening its rise time.

10.3.2 What Are Typical *Kf* Values?

To find the forward crosstalk for a set of traces, a 2D field solver is used to determine the mutual capacitance and inductance, and *Kf* and then FEXT are found with (10.1a) through (10.1c).

Figure 10.4 shows typical *Kf* values for a pair of 5-mil (0.13-mm)-wide, half-ounce 65Ω and 50Ω solder mask–covered microstrips on FR4 ($\varepsilon_r = 4.2$). Since forward crosstalk is not present on low loss stripline, the graph only shows microstrip.

The separation is given as multiples of the trace width. A separation of 3 represents 15 mils since the trace is 5 mils wide (or 0.39 mm for a 0.13-mm trace).

We find from Figure 10.4 that *Kf* is –160 pS/m for a 5-mil (0.13-mm)-wide, 50Ω microstrip separated by 5 mils ("separation" of 1) from its aggressor. Using this in (10.1b) and assuming that the launched pulse has a 2V swing (*dv* in the equation) and a 1-ns rise time (*dt*), the forward crosstalk pulse is found to be –320 mV if the traces are coupled for a length of 1 meter. The crosstalk would be –160 mV if the coupled region was a half-meter in length.

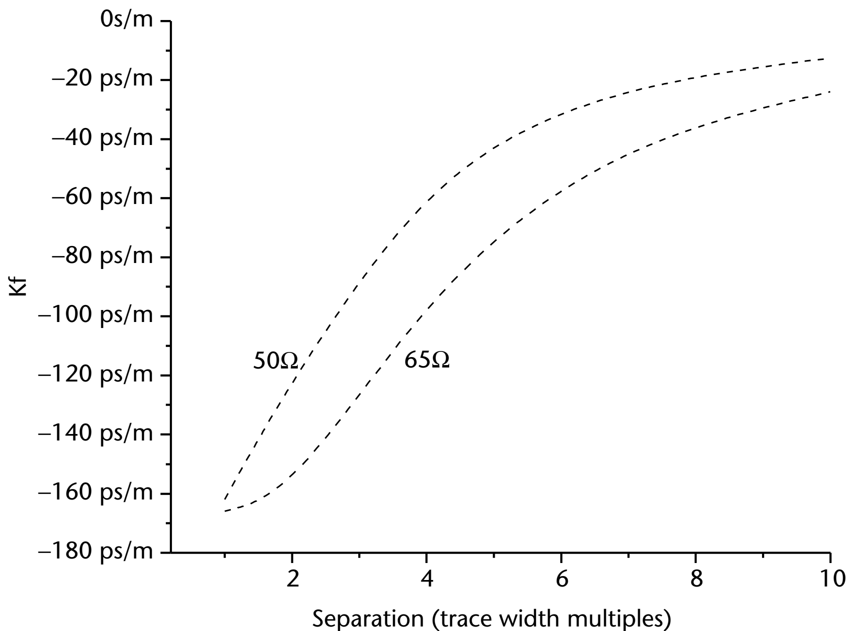


Figure 10.4 *Kf* for half-ounce, 5-mil (0.13-mm)-wide 50Ω and 65Ω solder mask–covered microstrips on FR4. (Raw data from the Polar Si9000 2D field solver [11].)

Because the 65Ω trace is further away from the ground plane than the 50Ω trace, more of the field lines can couple to the victim and less couple to the ground plane. This makes *Kf* larger (more negative) for the 65Ω trace.

Because for these two cases inductive coupling outweighs capacitive coupling, *Kf* is negative. It can become positive when the microstrip is covered with a thick dielectric such as a heavy application of solder mask or conformal coating (provided that it has a higher dielectric constant).

10.3.3 FEXT Summary

To summarize FEXT behavior:

- FEXT is 0 in striplines, provided that the transmission line is not very lossy. FEXT is always possible in microstrips.
- The FEXT pulse will have the opposite polarity of the aggressor pulse when the magnetic coupling is greater than the capacitive coupling. It will have the same polarity when the magnetic coupling is less than the capacitive coupling.
- The FEXT pulse will have a higher amplitude if the aggressor rise time is small or if the aggressor has a large change in voltage.
- The FEXT pulse amplitude increases as the coupled length increases.
- The FEXT pulse width is equal to the aggressor rise time.

10.4 Near-End Crosstalk

Referring to Figure 10.2, it is evident that in NEXT the inductively and capacitively coupled currents are flowing in the same direction and add when they reach the load at the near end. For this reason the NEXT pulse induced on the victim line always has the same polarity as the aggressor pulse.

Because the currents flow away from the aggressor pulse, the current is continually induced in the victim as long as the aggressor pulse travels down the line. At each instant as it moves, the aggressor pulse reaches a fresh mutual capacitor and inductor, allowing current to be continually injected into the victim trace.

As illustrated in Figure 10.5, if the coupled region is electrically long (that is, if the aggressor rise time is shorter than twice the delay of the coupled region), the net result is a NEXT pulse having a width equal to twice the electrical length of the coupled line, and an amplitude that saturates at a value determined by the aggressors change in voltage. However, if the coupling region is short, the NEXT amplitude cannot reach its full value. In this case the NEXT amplitude depends in part on the rise time (*dt*) of the aggressor pulse.

10.4.1 Calculating NEXT

The backwards crosstalk coupling factor (*Kb*) is presented in (10.2a). It is has no units and is used with (10.2b) when the line is short or (10.2c) when it is long to find the actual NEXT voltage [4, 5, 8–10].

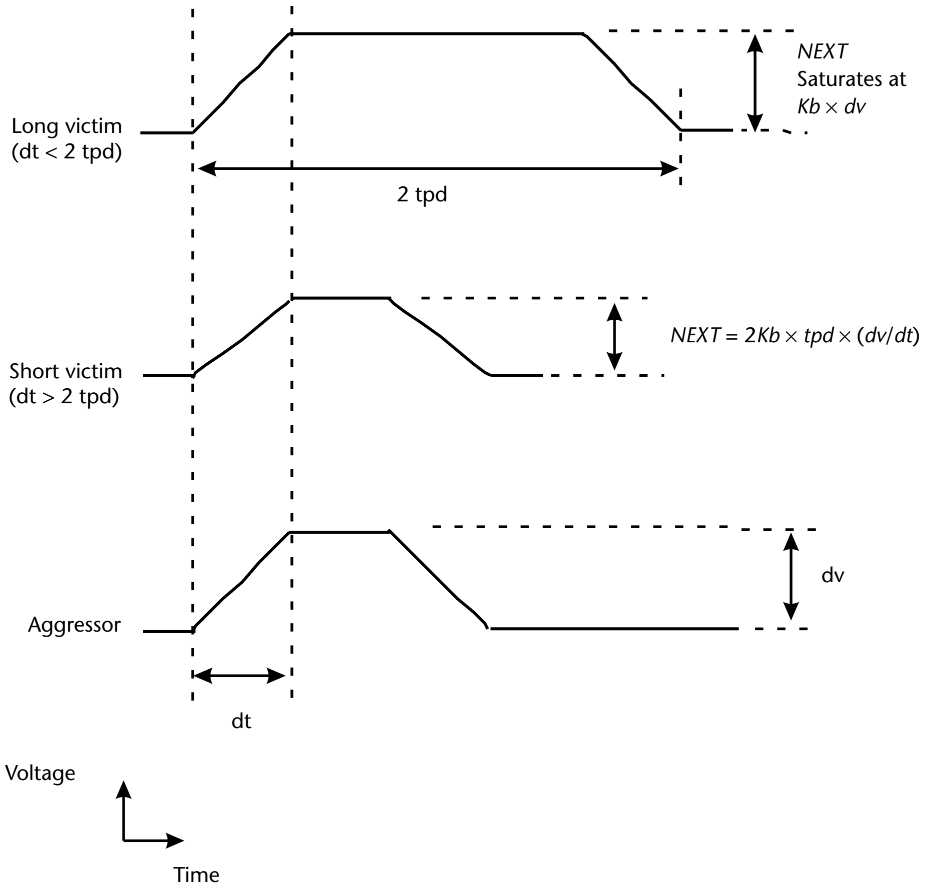


Figure 10.5 Width and amplitude of NEXT pulse appearing on victim depends on length of coupled region. NEXT on short lines grows in proportion to the length. It saturates at a maximum value when the line is long.

$$Kb = \frac{1}{4\sqrt{L_s \times C_s}} \times \left(\frac{L_m}{Z_o} + (Z_o \times C_m) \right) \tag{10.2a}$$

$$NEXT = 2 \times Kb \times tpd \times \frac{dV}{dt} \qquad (\text{when} \quad dt \geq 2tpd) \tag{10.2b}$$

$$NEXT = Kb \times dV \qquad (\text{when} \quad dt < 2tpd) \tag{10.2c}$$

$$tpd = length \times \sqrt{L_s \times C_s} \tag{10.2d}$$

It is evident from (10.2a) that the relationship between the trace impedance and the coupling inductance (L_m) and capacitance (C_m) determines the NEXT

amplitude. NEXT will always be present because, unlike with FEXT, the mutual inductance and capacitance terms cannot cancel each other.

As shown in the Problems, when more than one trace aggresses a victim, L_m and C_m represent the sum of the mutuals from all the aggressing traces.

Equations (10.2b) and (10.2d) show that, when the coupled length is small, NEXT grows by increasing the length of the coupled region (*tpd*), by increasing voltage swing of the aggressor (*dV*), or by reducing the aggressor rise time (*dt*).

It is apparent from (10.2c) that for long lines (defined as where the signal rise time is less than twice the coupled length) the NEXT amplitude is only affected by *Kb* and by the aggressor's voltage swing. Therefore, NEXT is combated by lowering the coupling between traces since that lowers *Kb*, keeping the coupled length short, reducing the voltage swing of the aggressor, and lengthening its rise time.

10.4.2 What Are Typical *Kb* Values?

Reverse crosstalk is shown in Figure 10.6 for a pair of 65Ω and 50Ω 5-mil (0.13-mm)-wide, half-ounce thick striplines and solder mask–covered microstrips on FR4 ($\varepsilon_r = 4.2$). The separation is given as multiples of the trace width: a separation of 3 represents 15 mils since the trace is 5 mils wide (or 0.39 mm for a 0.13-mm trace).

It is evident that *Kb* (and reverse crosstalk) is greater for higher-impedance traces, and reverse crosstalk may be larger in stripline than in microstrip.

To estimate the reverse crosstalk amplitude without using a field solver to determine L_m and C_m, *Kb* is found from the figure and is used in either (10.2b) or (10.2c), depending on the length of line. The Problems present this in more detail,

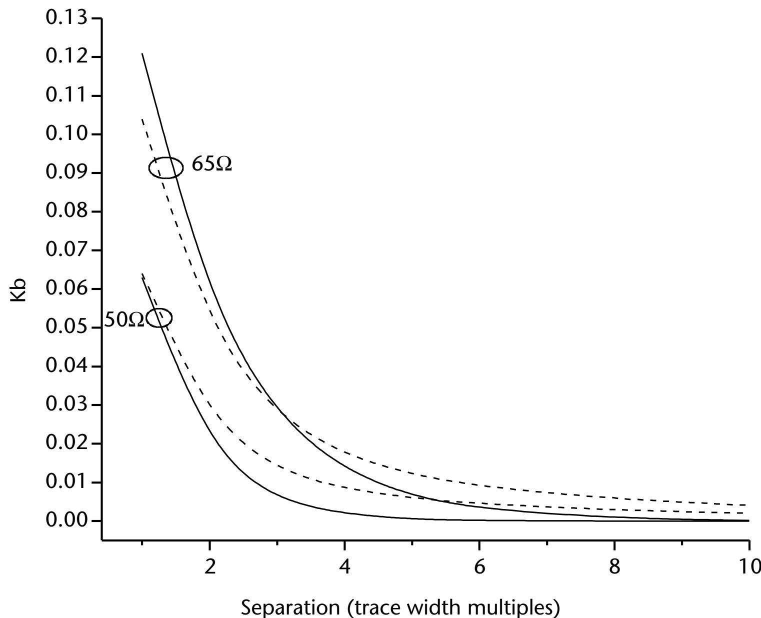


Figure 10.6 *Kb* to an adjacent trace for 5-mil (0.13-mm)-wide half-ounce stripline (solid lines) and solder mask–covered microstrip (dotted lines) on FR4. (Raw data from Linpar [12].)

but to illustrate, *Kb* for a 65Ω, 5-mil-wide stripline is found from Figure 10.6 to be 0.03 when the victim is a distance equal to three traces away. If the rise time of the aggressor pulse is small compared to the electrical length of the trace, the crosstalk voltage is found with (10.2c). For instance, NEXT will be 60 mV if the aggressor pulse is 2V.

10.4.3 NEXT Summary

To summarize NEXT behavior:

- The NEXT pulse has the same polarity as the aggressor pulse.
- NEXT is possible for either stripline or microstrip traces.
- Increasing the aggressor signal swing increases the NEXT pulse amplitude.
- If the coupled region is long (aggressor rise time is less than twice the coupled length, where $dt < 2\ tpd$), the NEXT amplitude saturates at a maximum value.
- If the coupling occurs over a short distance ($dt \geq 2\ tpd$), the NEXT pulse amplitude depends on the aggressor pulse rise time and will be smaller than the amplitude of the long line case.

10.5 How Closely Do Calculation and Simulation Agree?

Simulation results using the SmartSpice circuit simulator [13] lossy coupled transmission line appear in Figure 10.7 for a pair of 65Ω 4-mil (0.11-mm)-wide microstrips separated by 4 mils on FR4.

We see that the simulation predicts a NEXT amplitude of 191 mV, and as shown in the Problems, (10.2) calculates 192 mV. As expected, the NEXT pulse has the same polarity and is as wide as the aggressor pulse. In fact, it appears as a miniature version of the aggressor.

Two FEXT pulses are created because coupling occurs on each edge of the aggressor pulse. The FEXT pulses are opposite polarity because (as we find in the Problems) in this case *Kf* is negative. The simulation shows the FEXT pulses to be 385 mV, versus the 388 mV calculated.

The close agreement between the simulation and the equations only occurs when the transmission line has low loss and when both ends of all the transmission lines are terminated in resistors having a value equal to Z_o. Even with this excellent correlation, the simulation and calculations are still only estimates of the amount of crosstalk that will occur in actual hardware. This is because we usually do not know the true shape or the actual dimensions of the fabricated traces. Nor do we know the dielectric constant of the actual laminate or the actual solder mask thickness. These things mean that the topology we specify in a field solver only approximates the trace configuration that will actually be obtained in production.

10.5.1 How Well Do 2D Field Solvers Agree?

The results shown in Figure 10.7 show that simulation and calculation give the same results when using identical data. The correlation does not prove the field solver has produced accurate and consistent results even if the trace topology is known precisely. In fact, the calculation of the mutual capacitance and inductance requires the field solver to process very small numbers, especially when the spacing between traces is large. As shown in Table 10.1, numerical imprecision can result in discrepancies between field solvers. To illustrate this, Table 10.1 shows the worst-case range of *Kb* for a 5-mil (0.13-mm)-wide, 50Ω stripline across various spacing as reported by multiple 2D field solvers.

Table 10.1 shows that even if the trace topology is known perfectly, the simulation results obtained from different field solvers will not match exactly, especially when the spacing is large. For instance, the field solvers agree to within 0.5% or better for spacing of up to two times (10 mils), and the agreement is within 12% for spacing up to six times (30 mils). It becomes significantly worse for larger distances.

10.6 Guard Traces

Crosstalk can be reduced by placing a grounded trace between a victim and aggressor to create a guard trace [7, 14–16]. When done properly, the aggressor has higher coupling to the guard trace than to the victim, reducing the voltage induced on the victim. Guard traces are especially effective with stripline, but they also work well with microstrip.

One example showing the efficacy of a stripline guard trace is shown in Figure 10.8.

The top view of 4-mil (0.11-mm)-wide 65Ω stripline traces placed side by side with a separation of 4 and 12 mils (0.11 mm and 0.32 mm) is shown. The pulse

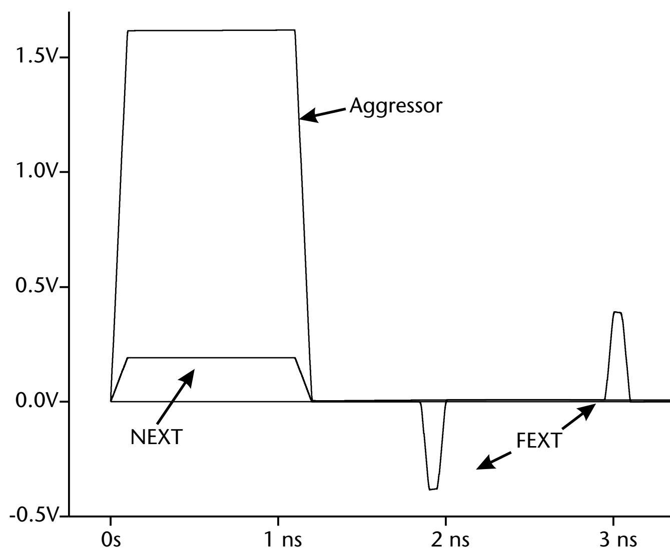


Figure 10.7 Simulated microstrip crosstalk results. NEXT is +191 mV (versus +192 calculated). FEXT is ±385mV (versus ±388 calculated).

Table 10.1 Differences in Kb Between Multiple 2D Field Solvers for a 5-Mil (0.13-mm)-Wide, Half-Ounce 50Ω Stripline on FR4

S	*% Difference*
1	0
2	0.5
4	3
6	12
8	45
10	170

Spacing (S) is in multiples of the trace width.

generator has an open circuit voltage of 3.3V with a 100-pS rise time. The resistors are all 65Ω, causing a 1.65-V, 100-pS signal to be launched down each aggressor.

As Figure 10.9 shows, with a 4-mil separation NEXT is 215 mV. It falls to 55 mV when the separation is 12 mils. Inserting a 4-mil-wide guard trace uses the same lateral space, but NEXT falls to 10 mV.

10.6.1 Connecting the Guard Trace

An improperly grounded guard trace can make crosstalk worse by coupling energy reflecting along its length back to the victim [7, 14]. In this situation the guard trace becomes an aggressor rather than a shield.

This can be avoided by placing a via to ground at both ends of the guard trace and including ground vias along the length of the guard trace. These vias must have a distance between them of less than a quarter wavelength of the highest frequency present [14]. Shielding can be improved if the vias are placed closer than this, and the vias should have a diameter roughly equal to the width of the trace [7].

As demonstrated in the Problems and shown in (10.3), this requirement can be converted into a rule of thumb for FR4 if the distance (d, measured in meters) between vias is no greater than one-tenth of the signal rise time (measured from the 10/90% points in nanoseconds) [7].

$$d = \frac{t_r}{10} \tag{10.3}$$

For instance, the vias should be separated by no more than 1/10 meter (about 4 inches) when the signal rise time is 1 ns ($t_r = 1$ in the equation).

10.7 Main Points

- Crosstalk is formed by inductive and capacitive coupling from one or more aggressors to a victim.

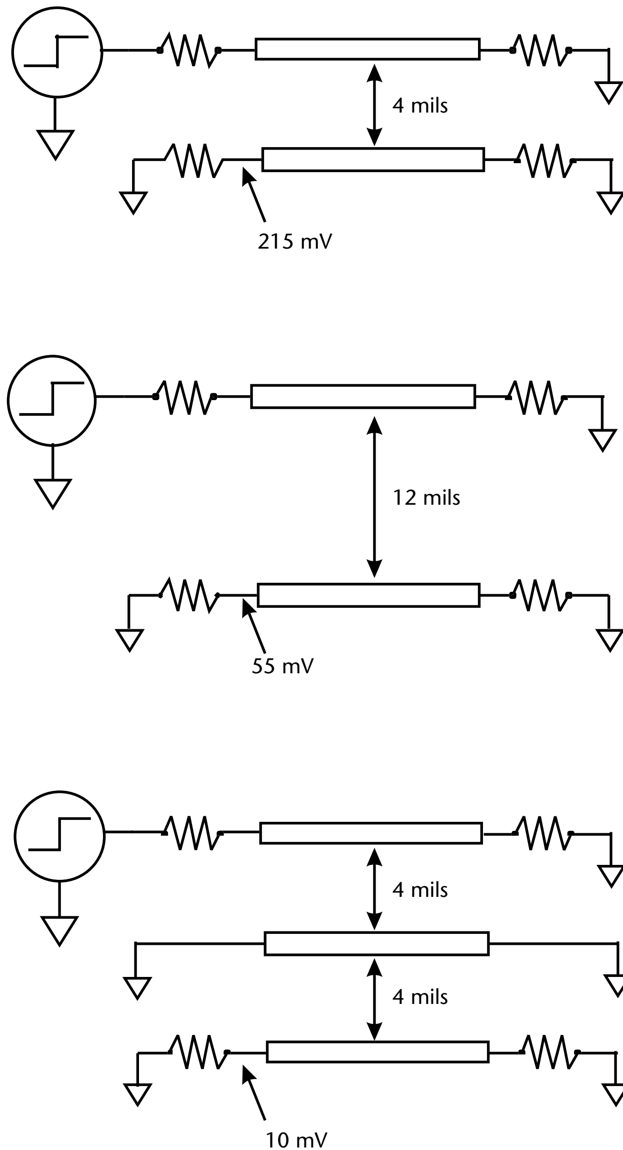


Figure 10.8 Top view of 4-mil-wide stripline traces. NEXT falls from 215 mV at minimum spacing to 55 mV when the spacing is increased to three times the trace width. With that same overall spacing, inserting a guard trace further reduces NEXT to 10 mV.

- Crosstalk appearing at the end of the trace closest to the aggressor transmitter is called near-end crosstalk (NEXT).
- Crosstalk appearing at the receiving end of the victim trace is called far-end crosstalk (FEXT).
- Microstrip can experience both NEXT and FEXT.
- Stripline can experience NEXT, but FEXT will be 0 if the line has low loss.
- The NEXT pulse has the same polarity as the aggressor pulse, and it grows with the amplitude of the aggressor. The value saturates at a maximum value when the line is long.

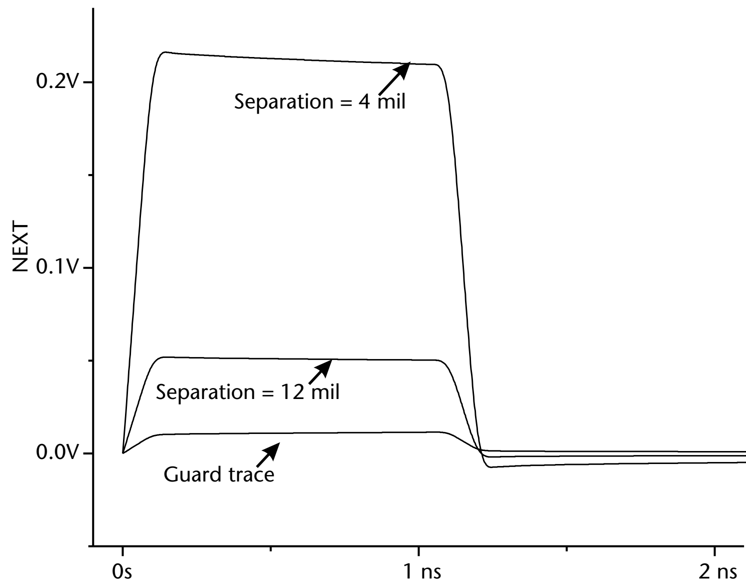


Figure 10.9 Simulation results for Figure 10.8. NEXT falls from 215 mV when the aggressor and victim separation is 4 mils (0.13 mm) to 55 mV when it is increased to 12 mils (0.32 mm). Inserting a guard trace reduces NEXT to 10 mV.

- The FEXT pulse has the same polarity as the aggressor when capacitive coupling dominates, but has opposite polarity when inductive coupling dominates.
- Guard traces can be effective in reducing NEXT and FEXT, but only when properly grounded.

Problems

Answers to these problems are available on the Artech House Web site at http://www.artechhouse.com/static/reslib/thierauf/thierauf1.html.

10.1 A field solver has produced the following data for a pair of 8-mil-wide, half-ounce striplines, separated by 10 mils:

L MATRIX [H/M]

4.447E-07	8.28E-08
8.28E-08	4.447E-09

C MATRIX [F/M]

1.088E-10	–2.025E-11
–2.025E-11	1.088E-10

Find the crosstalk coupling factors.

10.2 10.2 A field solver has produced the following data for a pair of 8-mil-wide, half-ounce microstrips, separated by 10 mils:

L MATRIX [H/M]

3.670E-07	5.28E-08
5.28E-08	3.670E-09

C MATRIX [F/M]

9.00E-11	–5.70E-12
–5.70E-12	9.00E-11

Find the crosstalk coupling factors.

10.3 Find FEXT for the pair of traces described in Problems 10.1 and 10.2 if the coupled region is 40 cm (15.75 inches) in length. Assume that in each case the signal has a rise time of 250 ps and swings from 0.8V to 1.8V.

10.4 Find NEXT for the pair of traces described in Problems 10.1 and 10.2 if the coupled region is 40 cm (15.75 inches) in length. Assume in each case that the signal has a rise time of 250 ps and swings from 0.8V to 1.8V.

10.5 Derive the (10.3) rule of thumb appearing in Section 10.6.1, which shows that for guard traces on FR4 a via should be separated by no more than one-tenth of a rise time.

10.6 The inductance and capacitance matrices for the pair of 4-mil-wide microstrips separated by 4 mils on FR4 used in the simulation appearing in Figure 10.7 is presented here. The coupled region is 0.3 meter (11.8 inches) in length, and the aggressor swings from 0 to 1.68V in 0.1 ns. Calculate NEXT and FEXT.

L MATRIX [H/M]

4.009E-07	9.662E-08
9.662E-08	4.009E-09

C MATRIX [F/M]

1.029E-10	–2.233E-11
–2.233E-11	1.029E-10

10.7 Use Figure 10.6 to find *Kb* for a pair of 5-mil-wide half-ounce 50Ω striplines on FR4 that are separated by 10 mils.

10.8 The inductance and capacitance portion of an RLGC SPICE model appears below. What NEXT will appear on the center trace if the two flanking traces are aggressors, when the aggressors each launch 1.65V pulses with a 0.1-ns rise time? Assume that the coupled length is 0.3 meter (11.8 inches) and that the model data is presented in meters.

```
+ Lo = 4.380E-07
+ 1.124E-07 4.327E-07
+ 3.405E-08 1.124E-07 4.380E-07
+ Co = 1.143E-10
+ –2.935E-11 1.233E-10
+ –1.356E-12 –2.935E-11 1.143E-10
```

References

[1] Young, B., *Digital Signal Integrity*, Upper Saddle River, NJ: Prentice-Hall, 2001.

[2] DeFalco, J. A., "Reflections and Crosstalk in Logic Interconnections," *IEEE Spectrum*, July 1970, pp. 44–50.

[3] Paul, C. R., *Analysis of Multiconductor Transmission Lines*, New York: John Wiley & Sons, 1994.

[4] Hart, B. L., *Digital Signal Transmission Line Circuit Technology*, Oxford, U.K.: Van Nostrand Reinhold, 1988.

[5] Rosenstark, S., *Transmission Lines in Computer Engineering*, New York: McGraw-Hill, 1994.

[6] Catt, I., "Crosstalk (Noise) in Digital Systems," *IEEE Transactions on Electronic Computers*, Vol. EC-16, No. 6, December 1967, pp. 743–763.

[7] Thierauf, S. C., *High-Speed Circuit Board Signal Integrity*, Norwood, MA: Artech House, 2004.

[8] Feller, A., et al., "Crosstalk and Reflections in High-Speed Digital Systems," *Proceedings of the Fall Joint Computer Conference*, Washington, D.C., December 1965, pp. 511–515.

[9] Sengupta, M., et al., "Crosstalk Driven Routing Advice," *Proceedings of the 44th Electronic Components and Technology Conference*, Washington, D.C., May 1994, pp. 687–694.

[10] Shi, W., "Evaluation of Closed-Form Crosstalk Models of Coupled Transmission Lines," *IEEE Transactions on Advanced Computing*, Vol. 22, No. 2, May 1999.

[11] Polar Si9000 Field Solver, version V9.01.00, Polar Instruments, Inc. Beaverton, OR.

[12] Djordjevic, A. R., et al., *LINPAR for Windows*, Norwood, MA: Artech House, 1999.

[13] SmartSpice, Simucad Design Automation, Version V3.10.8.R, Santa Clara, CA.

[14] Ladd, D. N., et al., "SPICE Simulation Used to Characterize the Cross-Talk Reduction Effect of Additional Tracks Grounded with Vias on Printed Circuit Boards," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, Vol. 39, No. 6, June 1992.

[15] Bakoglu, H. B., *Circuits, Interconnections, and Packaging for VLSI*, Reading, MA: Addison-Wesley, 1990.

[16] Suntives, A., et al., "Using Via Fences for Crosstalk Reduction in PCB Circuits," *2006 IEEE International Symposium on Electromagnetic Compatibility*, Portland, OR, Vol. 1, August 1–18, 2006, pp. 34–37.

CHAPTER 11

Understanding Signal Reflections

11.1 Introduction

The previous chapters have assumed that the transmission line has the same impedance everywhere and that the load impedance is the same as the impedance of the line. Although this is usually the goal, especially when signaling at high speeds, it is often not achieved in practice.

Reflected energy waves (signal reflections) created when the transmission line impedance does not equal the receiver impedance affect the voltage and shape of the waveform appearing at the receiver terminals. The waveform becomes even more distorted when the reflections created by a mismatch in the impedance between the driver and the transmission line arrive back at the load. The change in the waveform can be quite dramatic, as illustrated in Figure 11.1.

Figure 11.1 shows the voltage appearing at a receiver having a high input impedance when a 3.3-V square wave is launched down a transmission line whose impedance does not match either the driver or the receiver. The distorted wave shape is obvious, but the voltage levels that the pulse reaches (+5V and –2V) and the length of time that it is at those levels are also important. As discussed in Chapter 1, these voltages and times may be great enough to damage the ASIC I/O cell.

In this chapter we will examine the behavior when the source and load impedances do not equal the transmission line impedance, and see how it is possible for waveforms such as those appearing in Figure 11.1 to be created. Chapter 12 expands on this and presents strategies on how to correct these mismatches and how to restore the waveform and maintain proper voltage levels.

11.2 How Are Reflections Created?

A pulse travels down a lossless transmission line in the setup appearing in Figure 11.2.

The launched voltage is observed at V_{ne} and a current pulse of the same shape is simultaneously launched. Together these two pulses represent the wave of energy (the incident wave) traveling down the line from the generator resistance (R_g) to the

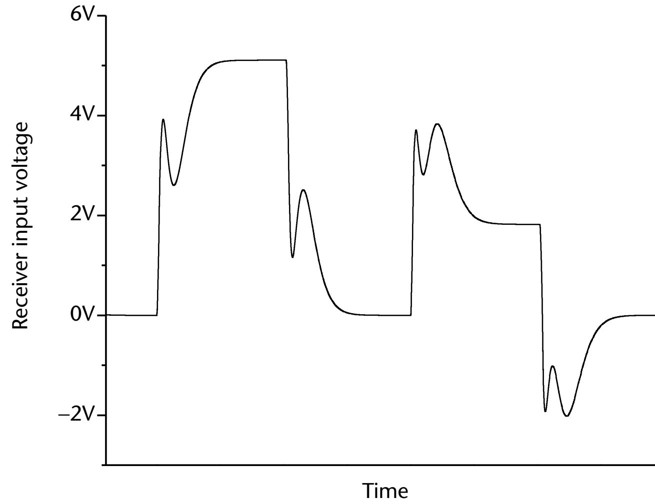


Figure 11.1 A launched 3.3-V square wave measured at the receiver when neither the transmitting nor the receiving impedances equal the transmission line impedance.

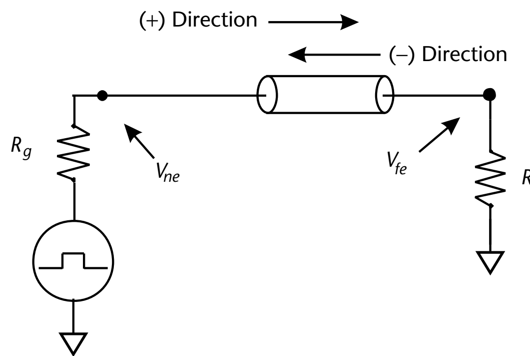


Figure 11.2 Transmission line connected to a pulse generator of resistance R_g and a load resistance R_l. Waves move in the positive direction when flowing left to right. The negative direction is right to left.

load resistance (R_l). Waves traveling from left to right are defined to be moving in the (+) direction, while right-to-left movement is in the (−) direction.

As first discussed in Chapter 6, and repeated in (11.1), the generator launches incident voltage (V_i) and incident current (I_i) waves in a ratio determined by the transmission lines characteristic impedance (Z_o).

$$Z_o = \frac{V_i}{I_i} \tag{11.1}$$

This ratio is maintained everywhere along the line unless there is a change in the impedance. However, difficulties arise when the load resistance does not match the line impedance, as demonstrated in Figure 11.3.

In Figure 11.3(a) an incident voltage pulse of 1V is launched down a 50Ω transmission line. From (11.1) an incident current pulse of 20 mA accompanies the

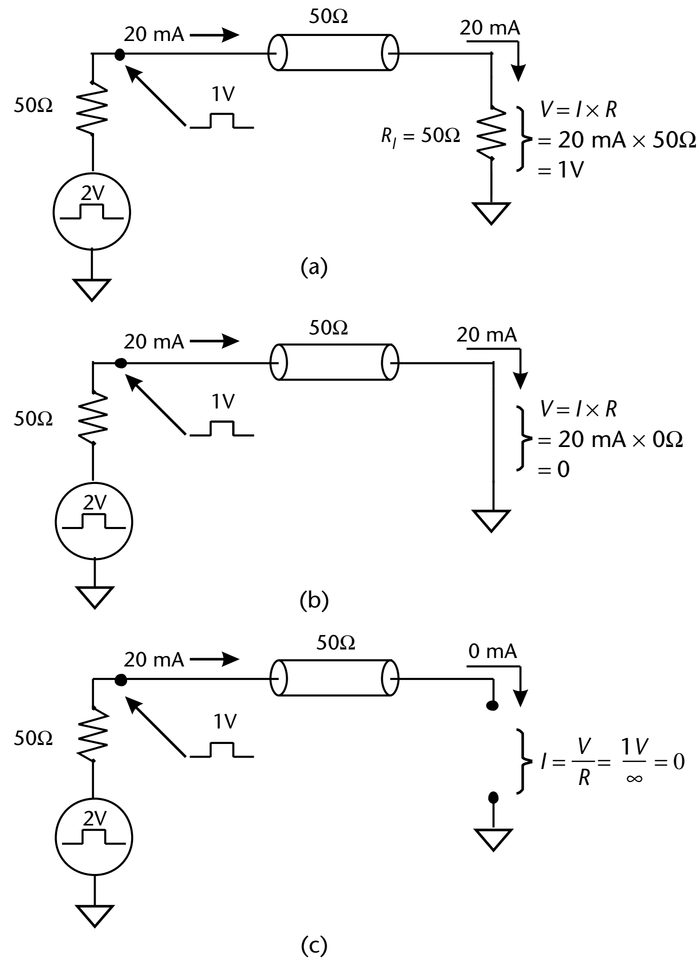


Figure 11.3 Launched voltage and current remains in the proper ratio (a) when the load resistance equals the transmission line resistance, but not when the load is (b) a short circuit or (c) an open circuit.

voltage pulse. Equation (11.1) is still satisfied when the pulses reach the 50Ω load resistor, but this is not the case when the load has any other value.

For instance, changing the load to a short circuit as is shown in Figure 11.3(b) requires the load voltage to be zero. However, the energy represented by the incident voltage and current pulses cannot simply vanish when the waves reach the short-circuited load. A similar problem shown in Figure 11.3(c) occurs when the load is infinite, representing an open circuit. There the current must become 0 at the load, even though 20 mA has been launched down the line.

The solution to these dilemmas centers on conservation of energy. The incident energy wave launched by the generator is perfectly absorbed by the load when R_l matches the transmission line impedance, but this is not so for any other value of load resistance. For instance, the launched energy cannot be absorbed by the open circuit, and because the line is lossless the energy is not dissipated there. Instead, when the energy wave encounters the infinite impedance of the open circuit, it rebounds and travels back up the line toward the generator. The rebound is a reflection, and wave theory is used to explain its behavior [1–4].

11.2.1 A Physical Analogy for Transmission Line Reflections

The reflection can be pictured by a physical analogy of a pulse traveling along a rope. For this analogy the transmission line in Figure 11.3(c) has been replaced with a taught rope. The far end is securely attached to a massive wall, representing the infinitely high impedance of an open circuit.

By holding and then abruptly shaking the near end, an incident pulse is launched that travels toward the far end. From elementary physics (for example, see [5]), when the incident pulse strikes the immovable wall, the pulse reflects, sending a smaller pulse of the same polarity back up the rope to the source.

In a similar way, the incident voltage and current pulses traveling in the (+) direction, down the transmission line, reflect from the open circuit. These reflected waves travel as freely along the line as do the incident waves and are called the reflected voltage (V_r) and the reflected current (I_r) waves. They travel at the same velocity as the incident waves, but in the (–) (that is, the opposite) direction.

11.3 The Reflection Coefficient

The percentage of the incident voltage wave that is reflected from an impedance discontinuity is called the voltage reflection coefficient, ρ. The current reflection coefficient is ρ_i. These are used in (11.2), where V_i and I_i are the incident voltages and current waves, and V_r and I_r are the values of the reflected waves.

$$V_r = V_i \times \rho \tag{11.2a}$$

$$I_r = I_i \times \rho_i \tag{11.2b}$$

These relationships apply equally to the receiving and transmitting ends of the transmission line.

Although the voltages and currents have separate reflection coefficients, when signal integrity engineers refer to the reflection coefficient, they generally mean the voltage reflection coefficient. In this book whenever ρ is used it will represent the voltage reflection coefficient, and ρ_i will be used in those instances when it is necessary to indicate the current reflection coefficient.

11.3.1 What Determines the Reflection Coefficient Value?

The reflection coefficient for a transmission line with low loss is given in (11.3). Its derivation relates the ratio of the incident and reflected waves.

$$\rho = \frac{R - Z_o}{R + Z_o} \tag{11.3}$$

The transmission line impedance is Z_o, and R is either R_l (the resistance of the load resistor), or R_g (the generator resistance). The reflection coefficient at the load is ρ_l; it is called ρ_g at the generator.

For example, ρ_l is –0.25 when a 25Ω load is connected to a 50Ω transmission line, and it is 0 when the load equals 50Ω. The reflection coefficient at the load becomes +0.25 when the load resistance is 83Ω. The Problems have further examples, but these show that ρ is positive when the load resistance is greater than the transmission line impedance. When it is less, ρ is negative. It becomes zero when R is the same as the transmission line impedance. It can be shown [6, 7] that the current reflection coefficient has the same value but the opposite sign from the voltage reflection coefficient. These results are summarized in Table 11.1.

Equation (11.3) can still be used if the loads are complex (such as capacitors or inductors rather than resistors), or when the transmission line impedance is complex (as is the case when the line is lossy). In those cases ρ is replaced with Γ (uppercase gamma), and in Figure 11.3 R becomes the load and generator impedances Z_l or Z_g.

11.4 How Do the Reflected Waves Combine with the Incident Waves?

From wave theory [1, 3] we know that the total voltage at the load is the sum of the incident and reflected waves (the waves superimpose). As shown algebraically in (11.4), this combined wave superimposes with any portions of the wave still present.

$$V_t = V_i + V_r + V_p \tag{11.4a}$$

$$I_t = I_i + I_r + I_p \tag{11.4b}$$

The total voltage and currents are V_t and I_t, V_i and I_i are the incident pulses, and V_r and I_r are the reflected voltage and current pulses. V_p and I_p are the voltage and currents still present.

11.4.1 Difference Between the Response of a Pulse and a Step

The third term in (11.4) hints at the interaction between the pulse width and the electrical length of the line. If the pulse width is very wide, the reflected signal will arrive in time to combine with it. However, if the pulse is narrow, it will have changed state before the reflection arrives.

Table 11.1 Reflection Coefficient Values

R	ρ	ρ_i	*Notes*
0	−1	+1	R is much less than Z_o (short circuit)
Z_o	0	0	R equals transmission line Z_o (matched circuit)
∞	+1	−1	R is much larger than Z_o (open circuit)

By examining Figure 11.4(a), we see that when the pulse (V_p) is narrow enough to return to zero before reflected pulse arrives, the total near-end voltage (V_t) is the sum of the incident and reflected pulses.

However, as shown in Figure 11.4(b), if the pulse is wide enough to still be present when the reflection arrives, the incident and reflected pulses add during the period when the pulses overlap, combining to change the shape of the waveform. Although in Figure 11.4 the incident and reflected waves are positive going, they can be negative going.

11.5 What Is the Behavior When There Are Multiple Reflections?

Commonly in digital systems an impedance mismatch is present at both ends of the transmission line. This causes reflections to rebound from both ends, which can significantly alter the wave shape and signal amplitude at both ends of the line. A demonstration of this setup appears in Figure 11.5.

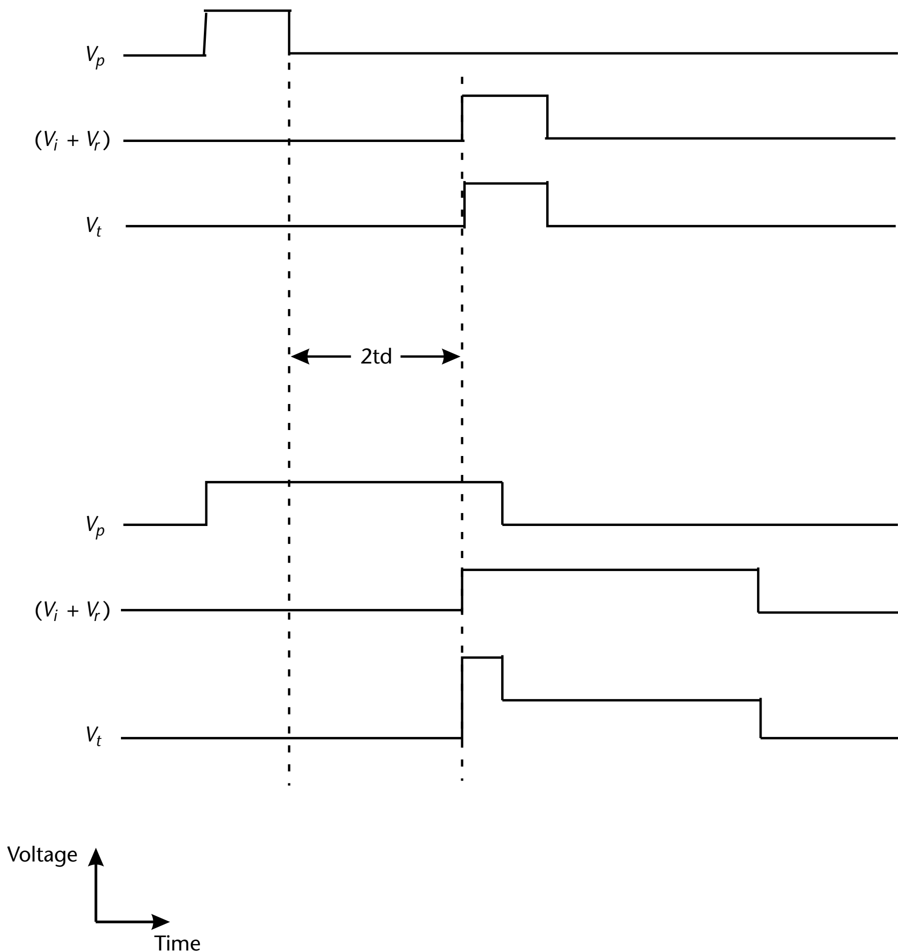


Figure 11.4 The total voltage (V_t) (a) when the launched (V_p) pulse is narrow and (b) when it is wide enough to combine with the reflection.

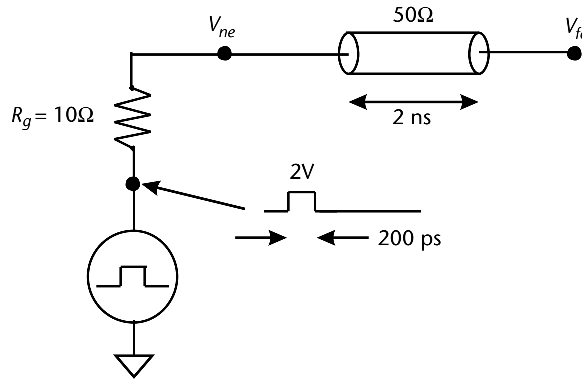


Figure 11.5 Circuit setup demonstrating reflections when the generator and load resistances do not match the transmission line impedance.

Figure 11.5 shows a 2-V, 10Ω driver connected to a 60Ω, 2-ns-long lossless transmission line. The far end of the line is open-circuited, representing the input of a high-impedance receiver. The driver launches a single pulse that has a width of 200 ps, which is shorter than the electrical length of the line, and very fast rise and fall times. A circuit simulation showing the behavior of this circuit appears in Figure 11.6.

Figure 11.6(a) shows the load voltage, V_{fe}. The voltage's near end, V_{ne}, is shown in Figure 11.6(b). Multiple reflections are present and it is apparent that the voltage at the near and far ends alternate between a (+) going and (–) going wave. The voltage at the load caused by the first reflection actually exceeds the driver's 2-V supply voltage.

Hand calculations can recreate the simulation results, but because there are multiple reflections, keeping track of the various sums becomes cumbersome. A

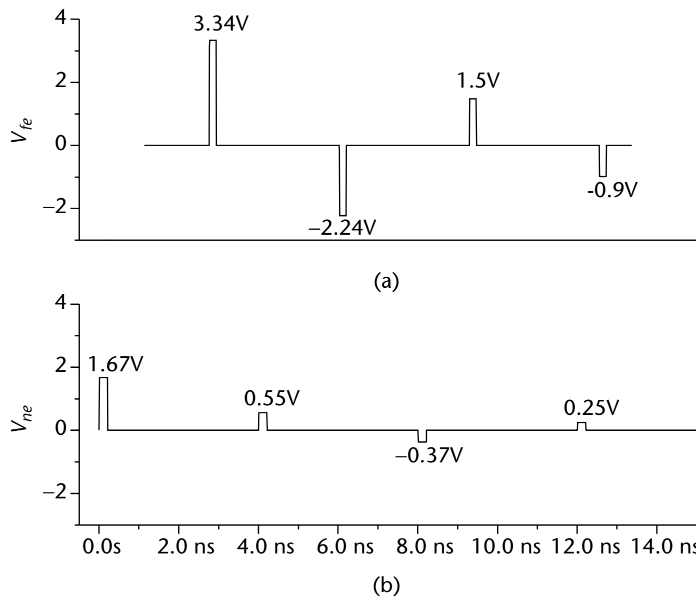


Figure 11.6 Simulation results for Figure 11.5. Voltages (a) at the load (V_{fe}) and (b) at the generator (V_{ne}).

reflection chart [1, 6, 7] will be used manage the sums and to show the timing and amplitude of the reflections.

We begin by using the voltage divider principle to determine that the driver launches 1.67V down the line. This is the incident voltage, V_i.

From (11.3) we find the reflection coefficient at the driver end of the line is $\rho_g = \frac{R_g - Z_o}{R_g + Z_o} = \frac{10 - 50}{10 + 50} = -0.67$, and from Table 11.1, ρ_l is +1.

We are now ready to construct the reflection chart shown in Figure 11.7.

The voltages present at the generator end of the line are recorded on the left side of the chart, and the load end voltages are recorded on the right side. Time increases vertically downward with the zigzagging lines representing the reflection bouncing between the two ends. Creating a reflection chart for the voltage wave is described here, but the method for the current chart is the same.

The chart is filled in as follows:

1. At $t = 0$ the 1.67V incident wave found with the voltage divider principle is launched.

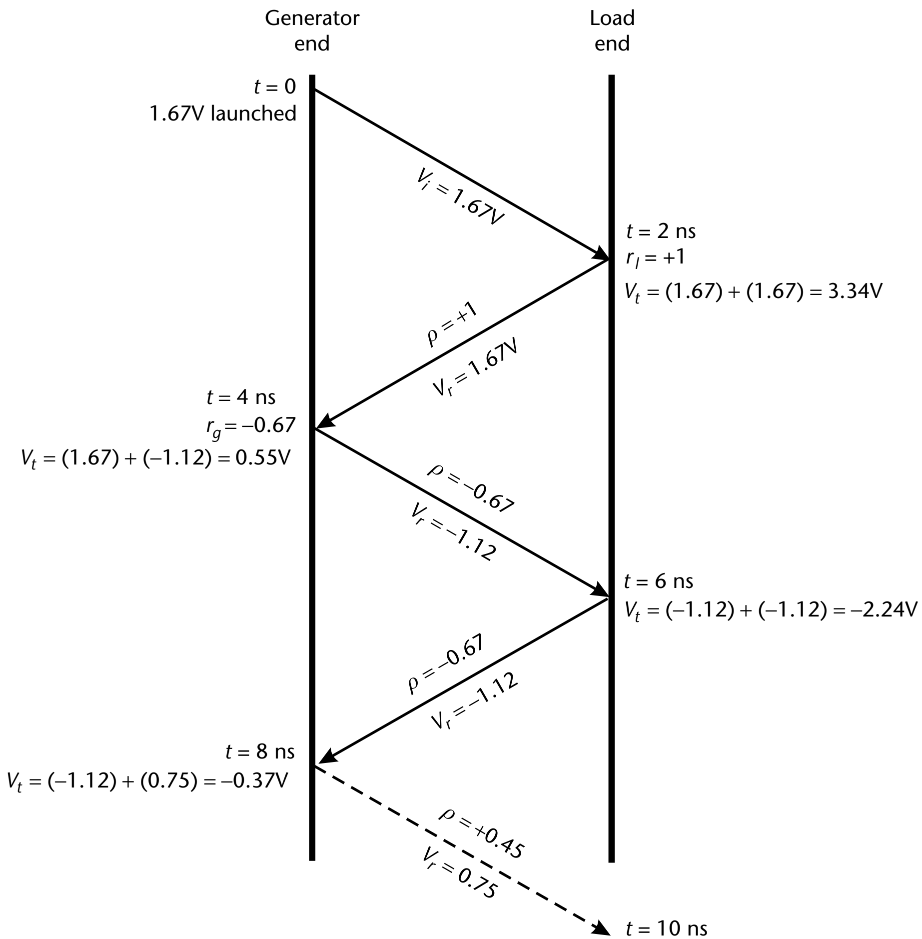


Figure 11.7 Reflection chart illustrating results from Figure 11.6. The load is separated from the generator by an electrical distance of 2 ns. Generator voltages appear on the left and load voltages appear on the right.

2. At $t = 200$ ps the output of the pulse generator returns to 0V.
3. The pulse arrives at the load at $t = 2$ ns. Since $\rho_l = +1$, a 200-pS-wide reflection of 1.67V is created. From (11.4) the load voltage is the sum of the incident and reflected voltages, noting that the load was previously 0V. This makes the load voltage 3.34V.
4. The +1.67V reflection (and not the 3.34V load voltage) travels back down the line and reaches the generator at $t = 4$ ns. Since $\rho_g = 0.67$, a 200-pS reflection of –1.12V is created from the +1.67V reflection impinging on the generator (the output of which is 0V). From (11.4) the generator voltage now becomes 0.55V.
5. The reflected –1.12V, 200-pS-wide pulse travels back up the line and at $t = 6$ ns arrives at the load where it sums with the –1.12V reflection created by the open circuit. Because the load has returned to zero by the time this reflection arrives, a total voltage of –2.24V is created.
6. The –1.12V, 200-pS-wide reflection travels down the line toward the generator, where it arrives at $t = 8$ ns.
7. When the –1.12V reflection reaches the generator, a +0.75V reflection is produced which reaches the load at $t = 10$ ns. As the chart shows, this is 45% of the initial launched voltage of 1.67V. In fact, this is the product of the percentages of all the proceeding reflections ($1 \times -0.67 \times 0.67$). This observation can be used to find the voltage after any number of reflections without resorting to constructing a reflection chart.
8. This process continues with the reflection becoming smaller and changing signs each time it rebounds from the generator impedance.

From the reflection chart it is apparent how it is the energy from the first reflection that makes its way up and down the transmission line. Because each subsequent reflection is multiplied by ρ (whose absolute value is never greater than 1), this means that, ignoring the effects of crosstalk, in a single pulse system the subsequent reflections will never have a magnitude worse than the first reflection. This knowledge is useful when setting up transmission line simulations and when making laboratory measurements.

The chart suggests a difference between pulse and step excitation. If the signal is a step longer than twice the round trip of the transmission line, the reflections would add to the launched voltage to create the composite voltage, V_{ne}. For instance, the near-end voltage at step 4 in the reflection chart (Figure 11.7) would have become 2.22V (the sum of the reflections plus the 1.67V still present from the generator) if the launched pulse was a step.

Reflection charts are no longer a practical design tool because even complex networks can easily be modeled with modern circuit simulators. However, because they show the behavior and characteristics of the reflected energy, they are a good learning tool that the signal integrity engineer should master.

11.5.1 What Is the Behavior When There Are Many Pulses?

The load voltage when a stream of pulses (rather than a single pulse) is launched down the line appears in Figure 11.8. The setup is identical to that used to create Figure 11.6, but in this case the 200-ps pulses repeat every 500 ps.

Assuming that the receiver switches at 1.65V (corresponding to $\frac{V_{dd}}{2}$), the receiver will properly detect the first eight pulses (V_{fe} = 3.34V), but the amplitude of the following eight pulses are too low to register as logic 1s. The next group will be properly detected, but the last group has marginal amplitude.

It is apparent from this how the data pattern influences the received signal, even when the transmission line has no losses. Proper termination can greatly mitigate these effects, but as part of the analysis process the signal integrity engineer should determine the set of data patterns that make the interconnect behave in the worst possible way. Once identified, any that can legitimately occur in the system should be included for use during design verification testing measurements on actual hardware.

11.6 Reactive Discontinuities

The loads connected to transmission lines in actual systems always include some amount of capacitance and inductance. For instance, the leads of an integrated circuit or discrete component are often inductive, and a via can appear capacitive. In Chapter 14 we will see how bends add excess capacitance and inductance to a trace. The way in which these and other discontinuities that can appear along a transmission line are sketched in Figure 11.9.

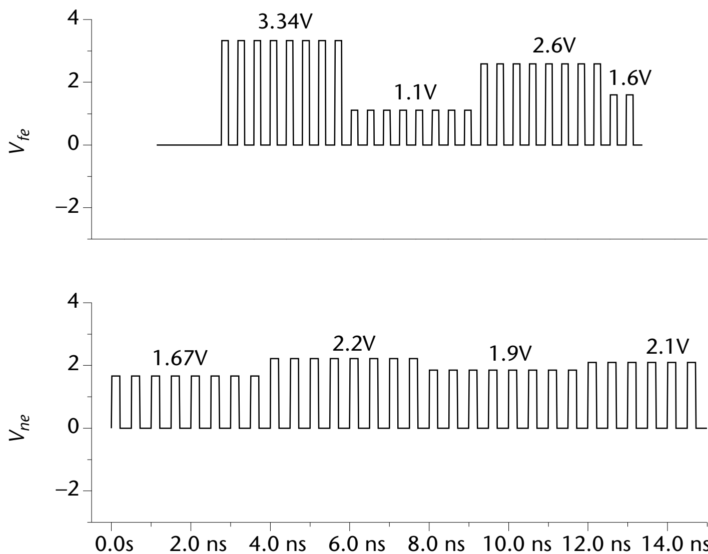


Figure 11.8 Effects of multiple pulses on the circuit of Figure 11.5. V_{fe} and V_{ne} are significantly different than the otherwise identical single pulse results illustrated in Figure 11.6.

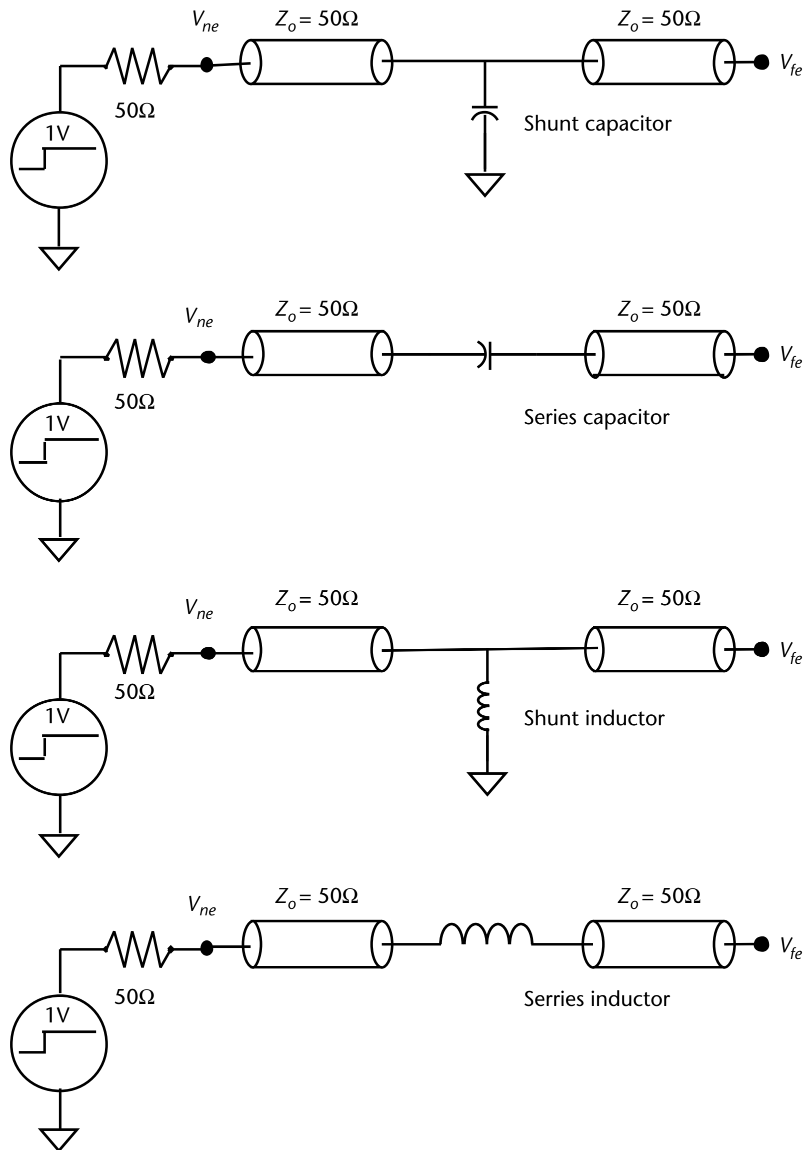


Figure 11.9 Types of reactive discontinuities.

From circuit theory we know that a capacitor initially acts as a short circuit, and how the voltage exponentially rises as it charges. Once fully charged, it appears as an open circuit. From an impedance standpoint the capacitor initially has a low impedance that exponentially increases. If given enough time, the impedance becomes infinite.

Similarly, an inductor initially acts as an open circuit and it charges exponentially to become a short circuit. Its impedance is initially infinite and falls exponentially to zero.

Table 11.1 showed that a short circuit creates a negative reflection, while an open circuit creates a positive one. The result is that a shunt capacitor appears as a sharp negative going reflection that exponentially returns to the base line voltage as the capacitor charges and the impedance increases. A series inductor initially

appears as a positive reflection that exponentially returns to the baseline as its impedance falls. These behaviors are evident in Figure 11.10(a, d).

The charging behavior of the series capacitor configuration is apparent in Figure 11.10(b), and the shunt inductor in Figure 11.10(c). Initially the capacitor is a low impedance and its impedance increases as it is charged by the launched current. The inductor is a high impedance, but as the inductor charges, its impedance decreases, shunting increasingly more of the signal energy to ground.

11.7 Main Points

- The energy transmitted down a transmission line consists of voltage and current waves.
- Conservation of energy requires that a reflection wave be created from the incident wave when a change in impedance is encountered.
- The reflection coefficient (ρ) indicates how much of the incident wave is reflected from a discontinuity.
- An open circuit creates a positive voltage reflection and a negative current reflection.

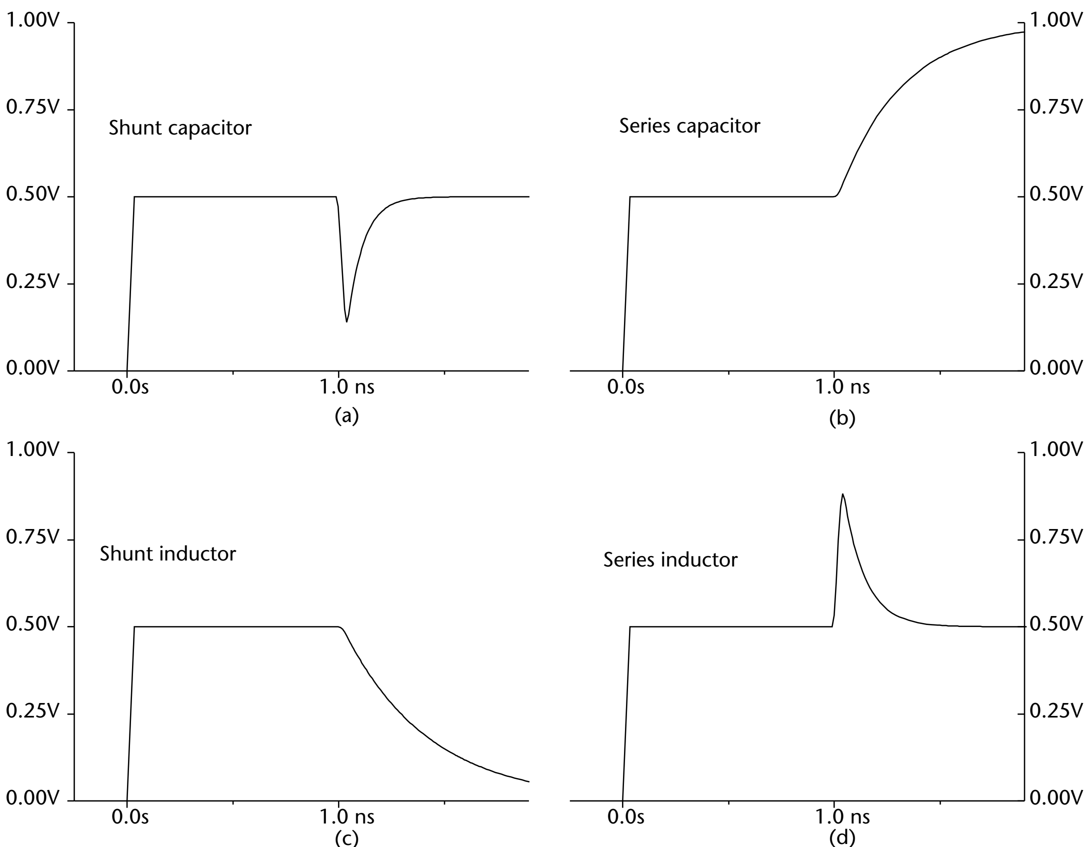


Figure 11.10 (a–d) Waveforms created by discontinuities as measured at V_{ne} for a step voltage.

- A short circuit creates a negative voltage reflection and a positive current reflection.
- At the discontinuity the incident and reflected waves add to create a composite wave.
- Multiple reflections can be simultaneously present on a line and can combine in complex ways.
- Reflections from earlier pulses can combine with later pulses to change their shape and amplitude.

Problems

Answers to these problems are available on the Artech House Web site at http://www.artechhouse.com/static/reslib/thierauf/thierauf1.html.

11.1 Find the voltage reflection coefficient at the load when $R_g = 10\Omega$, $R_l = 100\Omega$, and the transmission line has an impedance of 75Ω.

11.2 Using the same information from in Problem 11.1, find the voltage reflection coefficient at the generator.

11.3 Using the information from the previous problems, show how the incident wave does not satisfy Ohm's law when it reaches the load. Then show how the reflection determined by ρ causes the network to satisfy Ohm's law.

11.4 What value of voltage and current will be reflected back from the load in Figure 11.2 when the load is an open circuit? Assume that the pulse generator has an open circuit voltage of 2.2V and an impedance R_g of 25Ω and that the transmission line impedance is 50Ω.

11.5 Assuming in Figure 11.2 that V_{ne} is a 1-V pulse when the pulse generator is not connected to a load, determine the V_{fe} when $R_g = 50\Omega$, $Z_o = 50\Omega$, and $R_l = 100\Omega$.

11.6 Assuming in Figure 11.2 that the transmission line has an electrical length of 2 ns and that V_{ne} is a 200-ps-wide, 1-V pulse when the pulse generator is not connected to a load, create a reflection chart for V_{fe} when $R_g = 25\Omega$, $Z_o = 50\Omega$, $R_l = 100\Omega$.

11.7 Assume in Figure 11.2 that the transmission line is lossless and has an electrical length of 2 ns. If the pulse generator creates a 200-ps-wide, 1-V pulse when it is not connected to a load, describe the behavior at the near and far ends when $Z_o = 50\Omega$ and is $R_l = \infty$ (an open circuit). Assume that $R_g = 150\Omega$ when the pulse is first launched, but that R_g becomes an open circuit before the first reflection arrives back at the source.

References

[1] Brown, R. G., et al., *Lines, Waves and Antennas*, New York: John Wiley & Sons, 1973.

[2] Johnson, W. C., *Transmission Lines and Networks*, New York: McGraw-Hill, 1950.

[3] Skilling, H. H., *Transient Electric Currents*, New York: McGraw-Hill, 1952.

[4] Matick, R. E., *Transmission Lines for Digital and Communication Networks*, New York: IEEE Press, 1969.

[5] Serway, R. A., *Physics for Scientists and Engineers*, 4th ed., Philadelphia, PA: Sanders College Publishing, 1996.

[6] Rosenstark, S., *Transmission Lines in Computer Engineering*, New York: McGraw-Hill, 1994.

[7] Hart, B. L., *Digital Signal Transmission Line Circuit Technology*, Berkshire, U.K.: Van Nostrand Reinhold, 1988

CHAPTER 12

Termination Strategies

12.1 Introduction

In the previous chapters we implied that the load and transmission line impedances must match for proper electrical operation of an interconnect. However, the actual goal is to ensure that the received signal crosses the receiver threshold with adequate timing margin and, at least for clocks, that the signal transitions smoothly.

The signal integrity engineer accomplishes this by arraigning the wiring topology and selecting the termination scheme and I/O drivers so that reflections do not cause multiple pulses, glitches in the waveform, or unacceptable voltage levels. Terminations can be placed at the load end (far-end termination) or at the driving (source termination) end. The choice depends on available timing margins, the wiring topology, the number of loads on a net and their location compared to one another, and power dissipation.

12.2 Source Series Termination

Source series termination uses a resistance to connect the driver (the source) to the transmission line. This resistance raises the impedance of the transmitter (TX in Figure 12.1) and is either a discrete resistor placed on the circuit board (shown as R_s), or it is an integral part of the integrated circuit I/O cell. In some integrated circuits the built-in series impedance can be selected (programmed) from a range of values.

12.2.1 Selecting the Resistance Value

Until now, we have tacitly assumed that the driver impedance has the same value as the transmission line impedance. However, in some circumstances it can be advantageous for the driver impedance not to equal the transmission line impedance.

Three cases are shown in Figure 12.2 for the point-to-point configuration shown in Figure 12.1. The driver impedance is very low so that series resistor R_s sets the total impedance value.

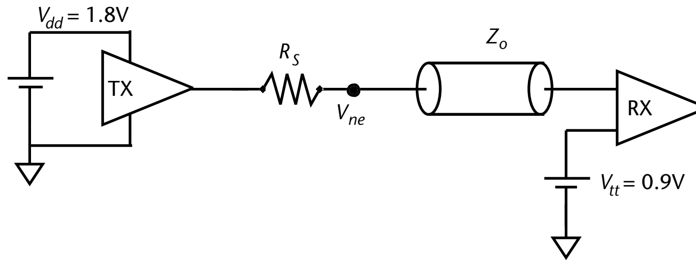


Figure 12.1 Source series termination. Resistor R_s may be incorporated in the driver or, as shown, may be a discrete resistor on the circuit board. V_{tt} sets the receiver switch point.

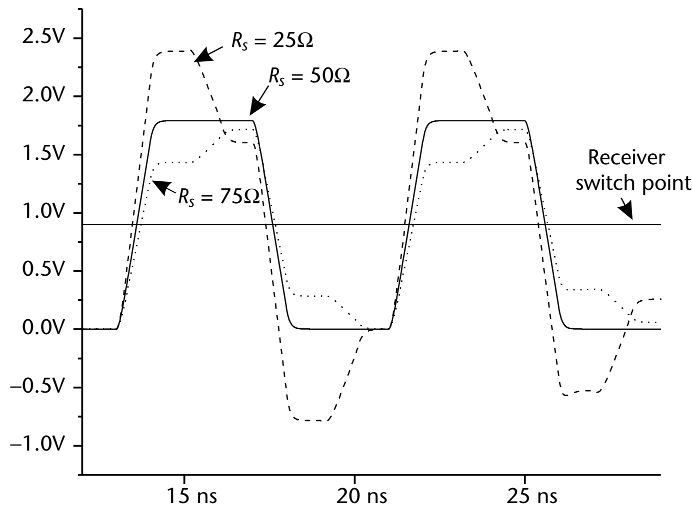


Figure 12.2 Response at the load when a 1.8-V driver is connected to a 50Ω transmission line with R_s equal to 25Ω (broken curve), 50Ω (solid), and 75Ω (dotted).

From Figure 12.2 we see that by making the series resistance (R_s) equal to the 50Ω impedance of the transmission line, the received signal switches cleanly from 0V to 1.8V (the value of V_{dd}, the voltage powering the driver).

By making R_s smaller than Z_o (in this case, 25Ω), the driver launches more current and a higher voltage down the transmission line, and the waveform crosses the receiver switch point sooner. However, the disadvantage is that any reflections arriving back at the driver from the load will be rereflected. In fact, the generator reflection coefficient (described in Chapter 11) will be negative, meaning that once the reflection arrives back at the load, it will subtract from the original wave. If large enough, this can cause the signal to fall (ring back) below the point where the receiver switches. In this case the switch point is set by V_{tt} to be half of V_{dd}, and since the signal only dips to 1.5V, the receiver does not falsely trigger. However, the signal exceeds V_{dd} (1.8V) and transitions significantly below ground. The signal integrity engineer would need to confirm that these voltages do not exceed safe operating limits for the integrated circuit technology.

Making R_s higher than Z_o (75Ω in Figure 12.2) causes less current and a lower voltage to be launched, and although the reflections do not exceed V_{dd} or go below ground, the lower current has caused the switch point to be crossed later.

If this timing degradation is acceptable, launching less current can be an advantage in low-power applications. However, the signal rise time will suffer if R_s is made too large, and care should be taken to ensure that the rise time is not degraded so much that the signal slowly transitions through the receiver switch point. Such slow transitions make the receiver susceptible to noise induced false triggering and can cause the receiver to switch multiple times before the signal is large enough to be properly detected. The resulting unpredictable timing and the creation of errant pulses (chattering) is never acceptable for clocks or strobe signals since it can lead to double clocking. Although the logic design may tolerate this type of behavior on signal lines, chattering is still undesirable as it can cause the receiver to draw excessive current and can generate power supply noise that may cause problems elsewhere in the system (including increased RF emissions).

12.2.2 Effect of Driver Resistance

The output transistors in practical CMOS I/O drivers always have some on resistance (R_{ds_on}) that adds to the resistor R_s placed on the circuit board. Because R_{ds_on} is nonlinear, its value depends on the current being launched by the driver. This is why it is poor practice to use a simple resistor to model I/O behavior rather than an actual circuit model of the driver.

Additionally, in CMOS I/O drivers R_{ds_on} generally increases as the power supply is reduced and as temperature increases. Circuit simulation shows these effects and should be used to determine the value of R_s for proper operation across the range of transmission line impedances, power supply voltages, and temperatures expected in the system. The tolerance and the temperature coefficient of the resistor type used for R_s should be included in the model.

12.2.3 Power Savings

If we assume that the connection from the resistor to the driver is electrically short, R_s simply adds to R_{ds_on}, limiting the amount of current that the driver can launch.

Since the current is less, the power dissipated by the driver is reduced. A second benefit of adding R_s is that the total power is shared between the driver and the resistor mounted on the circuit board. In effect, by adding R_s, the signal integrity engineer has moved some of the power dissipation from the ASIC to the circuit board. In some instances the increased number of circuit board components is worth the beneficial effect of helping to lower the ASIC junction temperature.

12.3 Parallel and Thevenin Termination

Parallel termination is placed at the load end of the transmission line. The two principal variants are shown in Figure 12.3.

The driver TX can be either a push-pull or an open drain type driver. The push-pull driver sources current when launching a logic high and sinks current to pull the line to a logic low. The open drain (or open collector when the I/O cell is made with bipolar logic) driver only sinks current to pull the line low. In this case the parallel termination supplies the current required to pull the line up to a logic high.

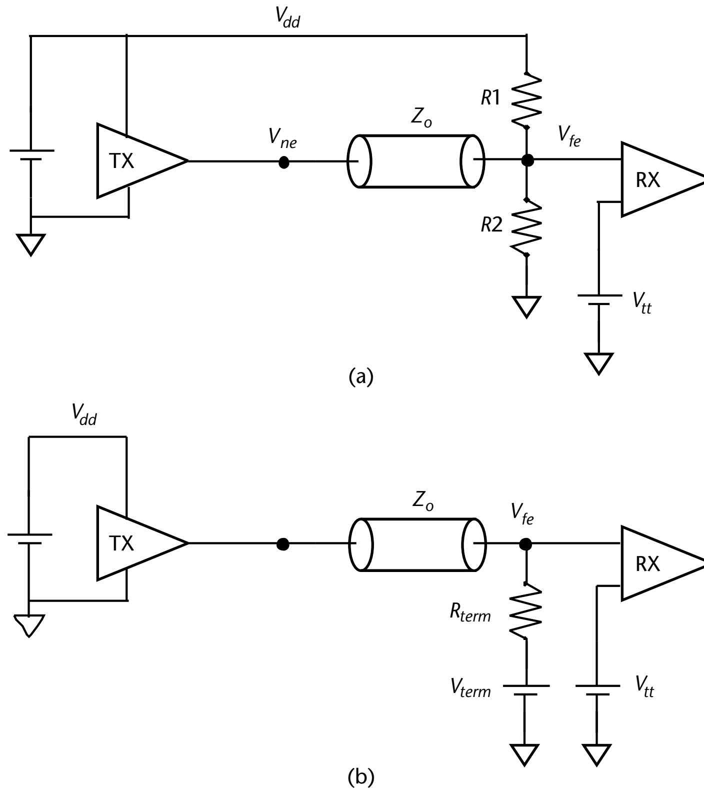


Figure 12.3 (a) Principal types of parallel termination. $R1$ and $R2$ form a voltage divider from V_{dd} with an impedance equal to Z_o. (b) The Thevenin equivalent.

In the general parallel termination scheme, two resistors form a voltage divider from a power supply (usually the supply powering the driver, V_{dd}, but this is not always the case). Ideally the supply has a low impedance and there is an AC short circuit from V_{dd} to ground. This places resistors $R1$ and $R2$ in parallel, but because in actual hardware the power supply has parasitic series inductance, it has a large impedance at high frequencies. To compensate for this, the AC short is created by adding decoupling capacitors between V_{dd} and ground close to $R1$ and $R2$. This makes the impedance of the termination (R_{term}) the parallel combination of $R1$ and $R2$, as shown in (12.1).

$$R_{term} = \frac{R1 \times R2}{R1 + R2} \tag{12.1}$$

The termination voltage (V_{term}) is determined by the voltage divider principle, shown in (12.2).

$$V_{term} = V_{dd} \times \frac{R2}{R1 + R2} \tag{12.2}$$

Resistors $R1$ and $R2$ can be found in terms of V_{dd} and V_{term} by simultaneously solving these two equations, producing (12.3a) and (12.3b).

$$R1 = \frac{V_{dd} \times R_{term}}{V_{term}} \tag{12.3a}$$

$$R2 = \frac{R1 \times V_{term}}{V_{dd} - V_{term}} \tag{12.3b}$$

The Problems show how to use these equations, but as an example, to obtain R_{term} of 50Ω and 1.65V for V_{term} when V_{dd} is 5V, we first use (12.3a) and find 152Ω for R1. This value is used in (12.3b) to find that $R2$ is 75Ω. Once these values are known, (12.1) and (12.2) can be used to determine how changes in resistance due to temperature or component tolerance affects R_{term} and how changes in V_{dd} affects V_{term}.

Resistors $R1$ and $R2$ can be separate, discrete resistors or a resistor network component. The matching between resistors is superior with the networks, but crosstalk and ground bounce can be worse than when using discrete devices. In general, surface mount devices have lower parasitic inductance and offer better high-frequency response than through-hole devices, with BGA devices giving the best response [1].

However, when using resistor networks, it is important to verify by simulation (and later during DVT) that crosstalk and simultaneous switching effects do not cause unacceptable amounts of noise for those terminators that share a common power and ground connection within the device.

12.3.1 Selecting V_{tt} and V_{term}

Generally in CMOS systems, the driver switches "rail-to-rail": the output swings between V_{dd} and ground. In these systems the receiver switch point is set by a termination voltage (V_{tt} in Figure 12.3) to be half the value of V_{dd} ($\frac{V_{dd}}{2}$), since this creates symmetrical low and high noise margins. For instance, 120Ω resistors would be used for $R1$ and $R2$ to terminate a 60Ω transmission line at $\frac{V_{dd}}{2}$. Assuming V_{dd} is 2.5V, this gives a 60Ω termination centered at $V_{tt} = 1.25$V.

However, depending on the switching technology, half of V_{dd} is not always the most appropriate value for V_{tt}. For instance, in TTL signaling (originally implemented with bipolar technology but now largely done with CMOS or BiCMOS), a logic low is sensed with a voltage close to ground and a high with a voltage nearly equal to half of V_{dd}. As explored in the Problems, in this situation the termination voltage is shifted to be midway between the two logic values.

The voltage for V_{term} must be chosen carefully when working with three-state buses (that is, a multidrop bus where the transmitters can be placed in a

high-impedance state). In some power-sensitive applications, all of the drivers are turned off when the bus is not actively driving data. As shown in the Problems, this greatly reduces the power dissipated by the termination compared to the situation where the bus is statically held in a state (often low) for extended periods. However, with all the transmitters turned off, resistors $R1$ and $R2$ force the bus voltage to become equal to V_{term}. If this is close to V_{tt}, the result may be receiver chattering and the creation of power supply or RF noise, along with an increase in receiver power dissipation. These effects are especially pronounced on bidirectional buses where each node on the bus has a three-state driver and receiver. In this situation the termination voltage should be selected that is guaranteed to exceed the receiver switch point under all environmental and worst-case tolerance conditions.

12.3.2 Thevenin Termination

From circuit theory the two resistors and the V_{dd} power supply in Figure 12.3 can be replaced with a Thevenin equivalent resistance in series with a Thevenin voltage supply [shown in Figure 12.3(b) as R_{term} and V_{term}]. This Thevenin termination is often used to terminate signals driven to memory devices, such as double data rate (DDR) memories.

For instance, a 60Ω Thevenin termination centered at 1.25V is formed with a single 60Ω resistor (R_{term}) connected to a 1.25-V power supply (V_{term}). This approach uses half the number of resistors and generally allows the single resistor to be placed closer to the point of termination than is possible with two resistors. When done properly, this improves signal quality and, when many lines are terminated, can be lower in cost. However, the creation of V_{term} must be done with care, and in small systems it may cost more to properly create V_{term} than is saved by eliminating half of the termination resistors.

It is generally inappropriate to use a simple three-terminal voltage regulator for the V_{term} supply. Those regulators usually only source current, but the termination requires that it both source and sinks current. Instead, regulators especially designed for termination applications should be used, and the manufacturer's guidelines concerning placement and value of decoupling capacitors must be closely followed.

High amounts of rapidly changing current flow when many signals are terminated, and to keep the termination voltage noise-free, the circuit board layout must provide a low-inductance path from the terminator resistor to V_{term}.

Narrow traces are too inductive to connect the regulator to the terminators. For this reason in large, high-performance systems, a voltage plane (or a portion of a plane) is sometimes dedicated to supplying V_{term}, much as a plane is used to supply V_{dd} or ground. Besides providing a low-inductance connection from R_{term} to V_{term}, a plane frees up routing channels that would otherwise be consumed by wide traces feeding V_{term} to various locations.

As explored in the Problems, Thevenin termination can dissipate less system power than parallel termination, especially when a three-state bus is intentionally placed in the high-impedance state to save power.

12.3.3 Connecting R_{term} Directly to V_{dd} or Ground

A further refinement on the Thevenin termination is possible by setting V_{term} to either 0V or equal to V_{dd}. As shown in Figure 12.4, this is equivalent to connecting the resistor directly to ground or to the power supply.

This scheme can be helpful in reducing power dissipation if the signal is predominately in one state. For example, by connecting the termination to ground, a driver sending a pulse that is most often at ground potential will only dissipate power during those infrequent times when the signal voltage pulses above ground.

12.4 Diode Termination

The scheme shown in Figure 12.5 shows that instead of preventing reflections by impedance matching the load to the line, reflections can be clipped by replacing resistors $R1$ and $R2$ with diodes. Signal integrity engineers occasionally refer to this as *diode termination*, but actually it is more properly called *diode clamping* because the diodes clamp the reflections; they do not prevent them.

Diode $D1$ conducts when the high-going reflection exceeds V_{dd} by a diode drop, and diode $D2$ conducts when the signal is more negative than ground by a diode drop. Together these diodes bound the signal to $(V_{dd} + V_d)$ and $(V_{ss} - V_d)$,

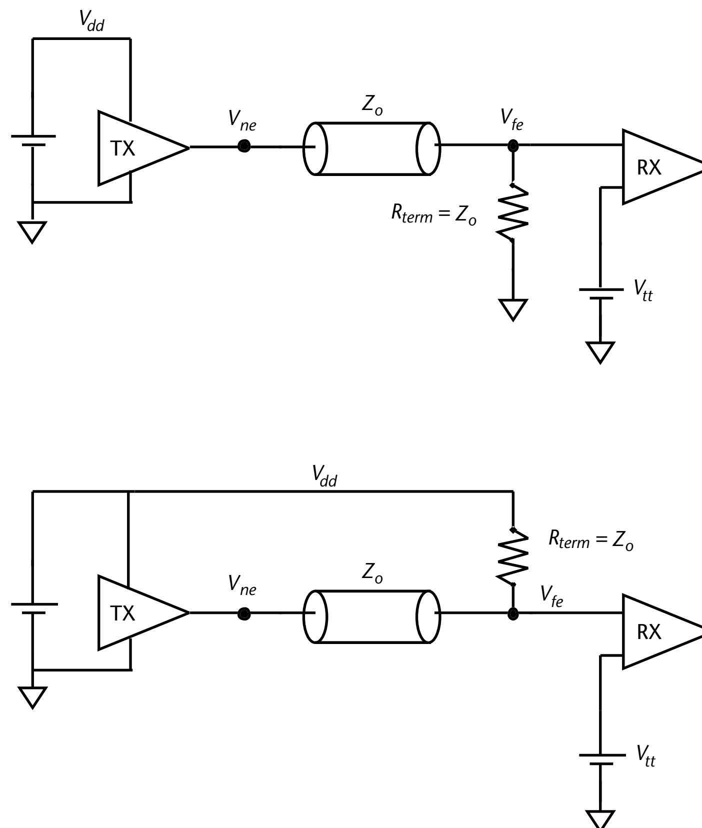


Figure 12.4 A proper parallel termination is formed with a single resistor equal to Z_o that is connected to ground or V_{dd}.

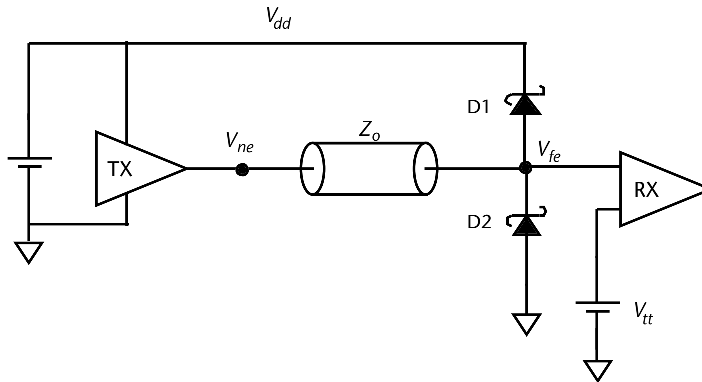


Figure 12.5 Clipping reflections with Schottky or low-capacitance, small-signal diodes.

where V_d is the drop across the diode when it is conducting the clamping current. Because practical diodes include parasitic series resistance that changes in value with the current passing through the diode, V_d is higher than the voltage needed to turn the diode on. For instance, at room temperature a small signal switching diode has a turn on voltage of about 0.6V, but the parasitic series resistance can cause the clamped voltage to be closer to a volt.

The diode clamping action means that in systems where the receiver operates from a power supply lower in voltage than the driver, $D1$ must be connected to the driver power supply and not to the receiver power supply. Otherwise, diode $D1$ will conduct whenever the driver transmits a logic high, even when reflections are not present. This static power dissipation can increase the junction temperature of the driver and wastes system power.

Because connecting $D1$ to the driver supply cause clamping to occur at a voltage higher than the receiver supply, in dual supply systems the clamping voltage may be not be low enough to prevent signal reflections from exceeding the maximum allowed for the receiver inputs. This is not a factor when the receiver supply is a higher voltage than the driver supply, but in that case clamping at the higher voltage will not be as effective. For instance, if the driver supply is 2.5V and the receiver is 3.3V, $D1$ will not clip reflections unless they reach roughly 4V. The 3.3V receiver is probably capable of withstanding this voltage, but the reflection sent back down the line may not be as easily tolerated by the driver and will encourage crosstalk to neighboring signal lines. For these reasons diode clamping is often not as effective as actual termination in situations where the driver and receiver operate at different voltages.

12.4.1 Diode Types

Diodes for signal clamping are either small signal switching diodes or, because of their lower turn-on voltage and faster switching response, Schottky barrier diodes [2, 3].

The diodes parasitic capacitance adds to the total receiver load capacitance. Diodes with low junction capacitance (under 2 pF) and small charge storage times (the diode transit time) are readily available, especially in surface mount packages, and are preferred to minimize this undesirable parasitic loading.

A load line analysis can be performed to determine how the diodes on resistance affects the clamped voltage, but the effects are seen best with a circuit simulator. The simulations should include diode models from different manufacturers because even for the same part number the circuit models often have different characteristics. Therefore, simulations must show that the clamping is acceptable with models from all the vendors supplying diodes for the build.

Although individual diodes can be used, dual diodes in three lead surface mount packages specifically made for signal clipping are available that include $D1$ and $D2$ with a common connection between them. The common lead connects to the signal trace and the other two leads connect to V_{dd} and ground. Some examples of small signal silicon dual diodes in surface mount packages include the MMBD1703, BAV99, MMBD4148SE, and the BAS40-04 (Schottky).

As part of the diode selection process, the circuit simulations should include the measurement of the current through diode. The signal integrity engineer must ensure that the clamping current is well within the diode current rating under all switching conditions and environmental extremes.

12.5 AC Termination

The scheme shown in Figure 12.6 places a capacitor in series with the termination resistor. This blocks the DC path so the network only dissipates AC power.

The acceptable value of blocking capacitor C_{term} falls within a wide range. The general requirement is for the capacitors reactance to be very much smaller than the resistance of R_{term}. In this way the capacitor acts as an AC short, connecting R_{term} directly to V_{term} (such as to ground, as is shown in the figure). However, if it is made too large, excessive time is required to establish the proper bias across the capacitor, adversely affecting the received signal amplitude.

This termination is best used with clocks or other repetitive signals, and is usually not suited for data buses or address lines where the frequency content of the data stream varies. Very low frequencies corresponding to the back-to-back transmission of many same sense bits require a large capacitor, but such a large value is not appropriate when the signal frequency is high (corresponding to an alternating 1/0 pattern, for instance) because of the prolonged time required for the capacitor

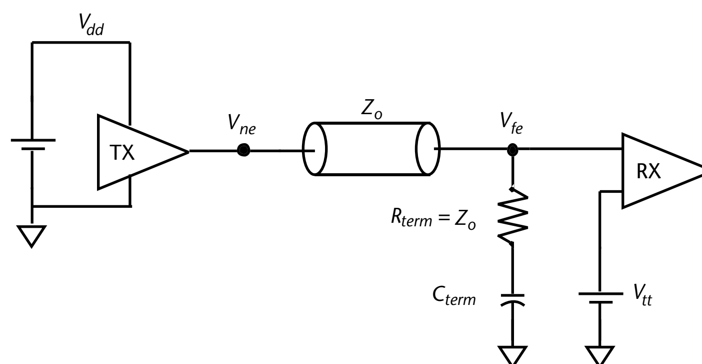


Figure 12.6 With AC termination TX only dissipates AC power. C_{term}'s reactance is much less than R_{term}.

to charge to the bias level. For this reason using AC termination on data buses results in significant data-dependent jitter [4] and is usually avoided.

Simple equations exist to calculate C_{term} from the signal rise time [5, 6], but these are misleading because they underestimate the capacitance needed for fast rise time, low-frequency signals. Equation (12.4) avoids this by providing an estimate based on the signal frequency.

$$C_{term} = \frac{3}{f \times Z_o} \tag{12.4}$$

The equation computes the capacitance with a reactance 1/20th of Z_o at the frequency f. The capacitor reaches its steady state bias level within 15 back-to-back transitions. By that time the circuit is properly terminating the signal, but pulses arriving before this time may not be.

For instance, C_{term} is 60 nF when Z_o is 50Ω and f is a 1-MHz clock. The capacitor has reached the steady state bias point by the 15th clock pulse (15 μs).

The calculated value should be used as a starting point when performing circuit simulations of the termination. The simulations ought to include the interconnect parasitics (such as the via inductance) that are in series with the RC network since these increase with frequency, diminishing the ability of C_{term} to act as an effective connection for R_{term}.

Discrete resistors and capacitors are usually used to create the AC terminator, but integrated RC networks are also available. As demonstrated in the Problems, the capacitance of these networks is small, which makes them most useful in terminating clocks with frequencies in the hundreds of megahertz range.

12.6 Topologies

In Figure 12.7 the wiring between surface mount integrated circuits U1 and U4 is a single load point to point connection, while the connection from U2 to U3, U5, and U6 is an example of point-to-point multiload topology. The connection between integrated circuits U7 through U10 is a multidrop topology.

These topologies have different electrical characteristics and, as we will see, different response to various types of termination.

12.6.1 Single-Load Point-to-Point Termination

The series and parallel termination of a point-to-point single-load connection is shown in Figure 12.8.

The series termination illustrated in Figure 12.8(a) shows a long transmission line $T2$ connecting a capacitive load C_L to a series terminating resistor R_s. Ideally, R_s is connected directly to the driver, but in general practice a length of trace (shown as transmission line $T1$) is used to make the connection. This pin escape trace is necessary because the physical size of R_s usually makes placing it immediately next to the driver pin impossible. This is especially true with large, fine-pitch multipin ASICs where the pins (or balls if in a BGA package) are arrayed in several

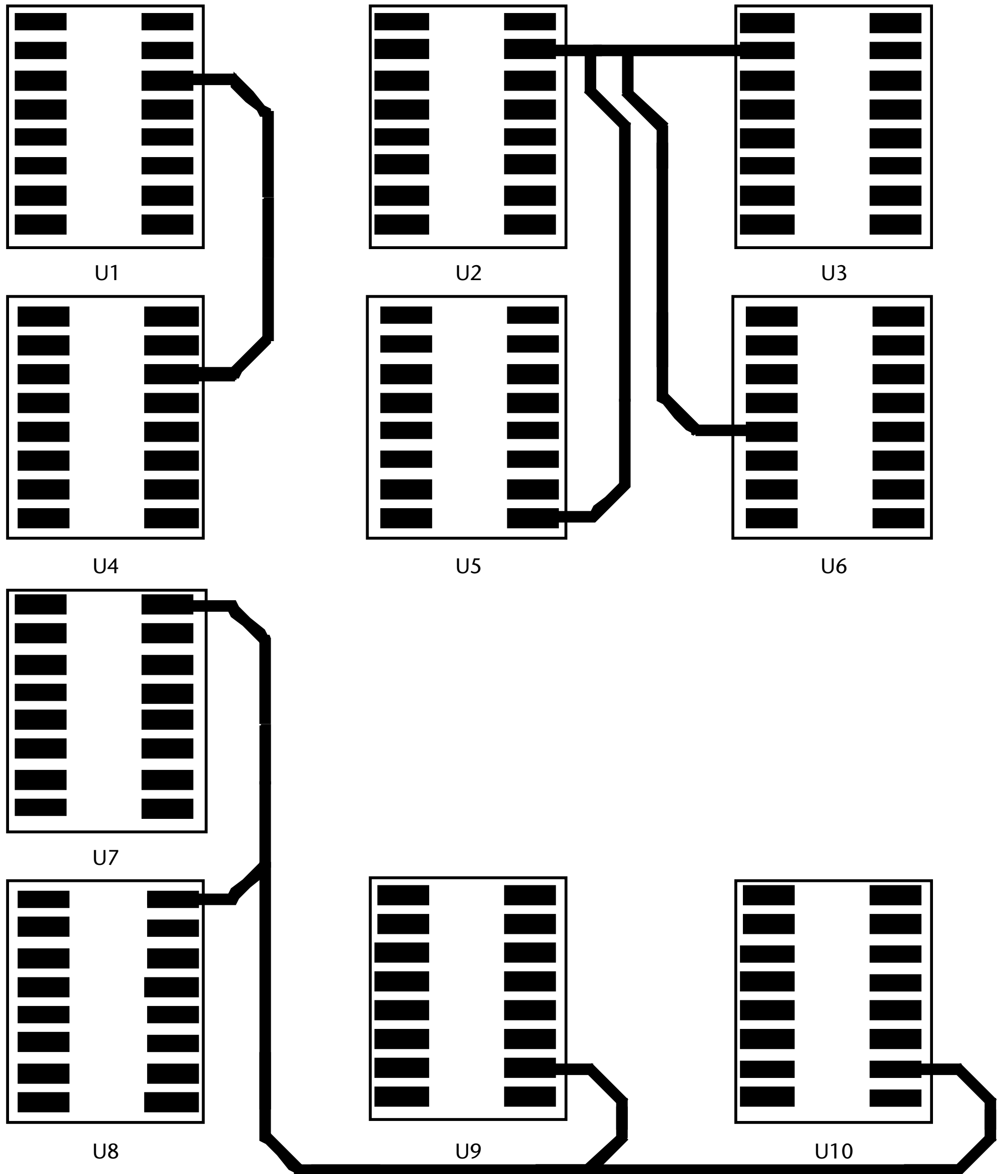


Figure 12.7 Circuit board artwork showing point-to-point topology from U1 to U4; single load point-to-point from U2 to U3 and U5 – U6; multidrop from U7 to U8 through U10.

rows. In that case $T1$ is long enough to escape the pin field and to place the signal in an open area where R_s can be mounted.

In lower-density boards the pin escape trace $T1$ is often routed as microstrip, but high-performance, high-density boards will frequently escape with stripline. In either case, the impedance of $T1$ (Z_{o1}) is often not the same as that of $T2$ (Z_{o2}). If $T1$ is too long, its impedance will determine the current launched by the driver and reflections will occur even if R_s properly matches $T2$. For this reason the length of $T1$ should always be short compared to the signal rise time, and it should always be less than the length of $T2$.

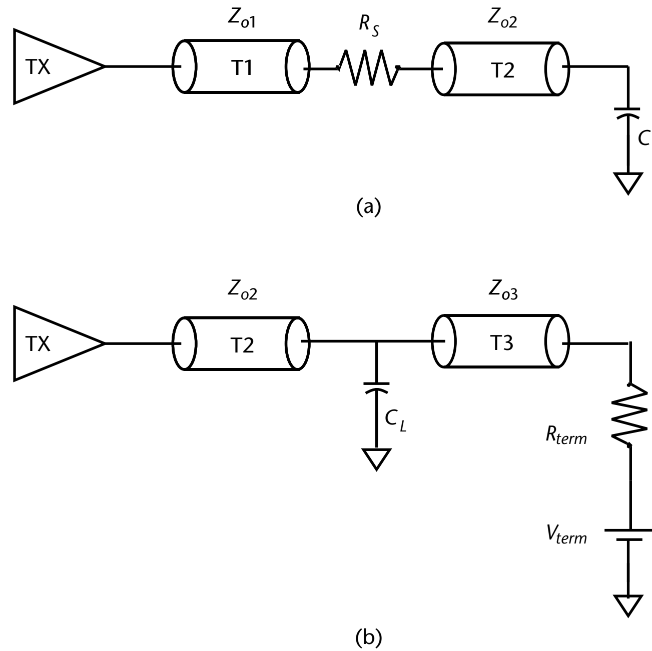


Figure 12.8 (a) Series and (b) parallel termination of a single point-to-point load. Capacitor C_L represents the receiver input capacitance.

The parallel termination illustrated in Figure 12.8(b) shows how the long connecting trace $T2$ directly connects to C_L and how the termination connects through transmission line $T3$. Although a Thevenin termination is pictured, it actually represents any of the parallel termination schemes we have already discussed.

In general, the termination is usually some distance from the load and trace $T3$ may be longer than $T2$. This fly-by termination is especially likely if the board is very densely populated because often space is not available to place the termination immediately at the pin of the receiving device. In this case $T3$ allows the signal to fly by the load and be terminated by the resistor, which is placed in an open area.

The length of $T3$ is not critical provided $T2$ and $T3$ have the same impedance. By making R_{term} equal in value to Z_{o3}, transmission line $T3$ is properly terminated and for any length it appears to C_L as a resistor of value R_{term}. This means $T2$ and $T3$ must be formed on the same routing layer.

Although $T3$ need not be short, very long lengths must be analyzed carefully, especially when signaling at high data rates. Very long lengths of trace will be lossy at high frequencies, which invalidates the idea that any length of $T3$ is the same as connecting R_{term} directly to the load. Also, crosstalk is more likely to occur on long lengths since there are more opportunities for $T3$ to be exposed to coupling from other signals. Regardless of $T3$'s length, it should be considered a critical net and carefully scrutinized for crosstalk.

12.6.2 Multiple Load Point-to-Point Termination

Multiload point-to-point termination schemes are shown in Figure 12.9.

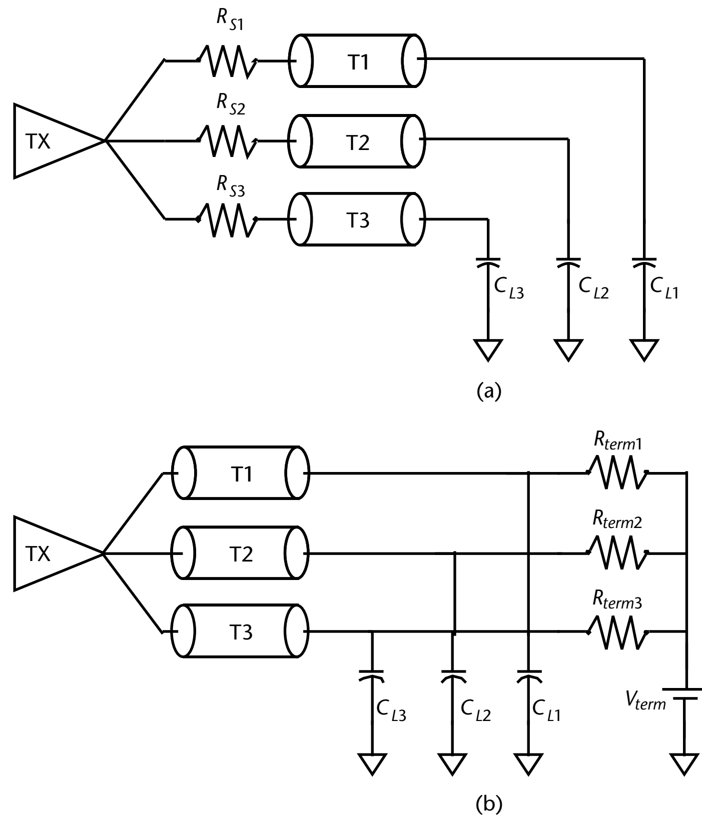


Figure 12.9 (a) Series. (b) Parallel multiload point-to-point termination schemes.

Series termination is shown in Figure 12.9(a), and parallel termination is shown in Figure 12.9(b), where we see that three loads are connected by three transmission lines to the driver, and in both cases the driver must simultaneously launch current down the three lines. This effectively places them in parallel. For instance, if the impedance of each line is 60Ω the driver experiences a load of 20Ω. Unless the driver has a very low output impedance, it may not be capable of launching a voltage high enough to reach the receiver switch point into such a low impedance.

With the series termination scheme, reflections from the impedance mismatch at the load will increase the received voltage, provided that the load impedance is very high. This means that the driver need not launch as high a voltage and can be weaker than in the parallel terminated case.

However, with this scheme the relative lengths of $T1$, $T2$, and $T3$ can become important. If the transmission lines are the same length, the reflections from all of the lines simultaneously arrive at the near end, and the waveform is the same at each load. However, if the lengths differ, the reflections arrive back at the driver at different times, which then combine in complex ways before they are relaunched to the loads. This alters the wave shape at each load. An example of this for the series termination scheme is shown in Figure 12.10.

In this example the driver has a nominal impedance of 12Ω and is connected to a 3.3-V source. The driver "fans out" a common clock to three loads that are located at different distances. Series resistors R_{S1} through R_{S3} are each 22Ω.

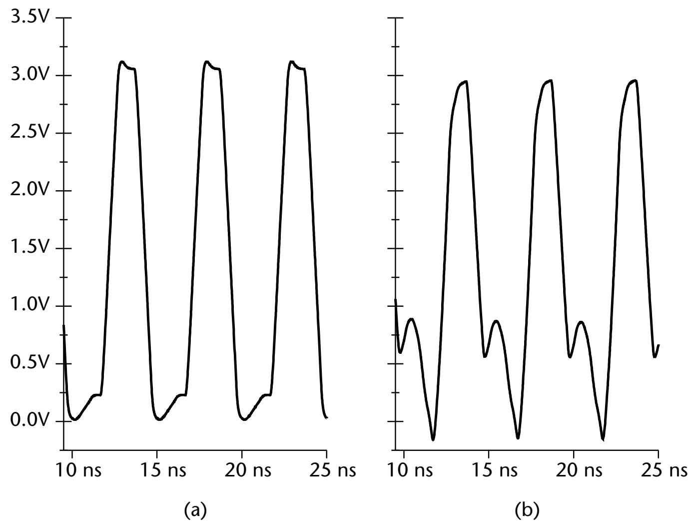


Figure 12.10 Series-terminated clock waveform at the end of the 3-inch (7.6-cm) trace (a) when all traces are the same length and (b) when the trace lengths differ (plotted with the same vertical scale).

Figure 12.10(a) shows the waveform at load capacitor C_{L3} when traces $T1$, $T2$, and $T3$ are all 3 inches (7.6 cm) in length. The rising and falling edges transition smoothly, without kinks (the signal is monotonic), and then high- and low-going overshoot is well controlled.

In contrast, Figure 12.10(b) shows the voltage at C_{L3} when transmission line $T1$ is 1 inch (2.5 cm), $T2$ is 2 inches (5.1 cm), and $T3$ is 3 inches (7.6 cm) in length.

The effect on the load of the unequal arrival time of the pulses at the near end is evident. This waveform is not acceptable for a clock. The falling edge glitches up to over 0.75V, and the negative-going overshoot is pronounced. To correct this waveform, the traces should all be length matched, as shown in Figure 12.10(a), or a clock fan-out device, a specialty integrated circuit made to distribute low skew copies of an input signal, could be used.

12.7 Multidrop Lines

As shown schematically in Figure 12.11, rather than using point to point connections, a multidrop connection can be used to send a signal from a driver to multiple loads.

The three load capacitors C_{L1} through C_{L3} are connected to the driver by tapping into a transmission line. Each of these lines may be very long compared to the signal rise time, or (more commonly) their electrical length may be comparable to the rise time. Without proper termination and attention to length matching, signals on multidrop lines are particularly likely to be nonmonotonic and to experience multiple reflections.

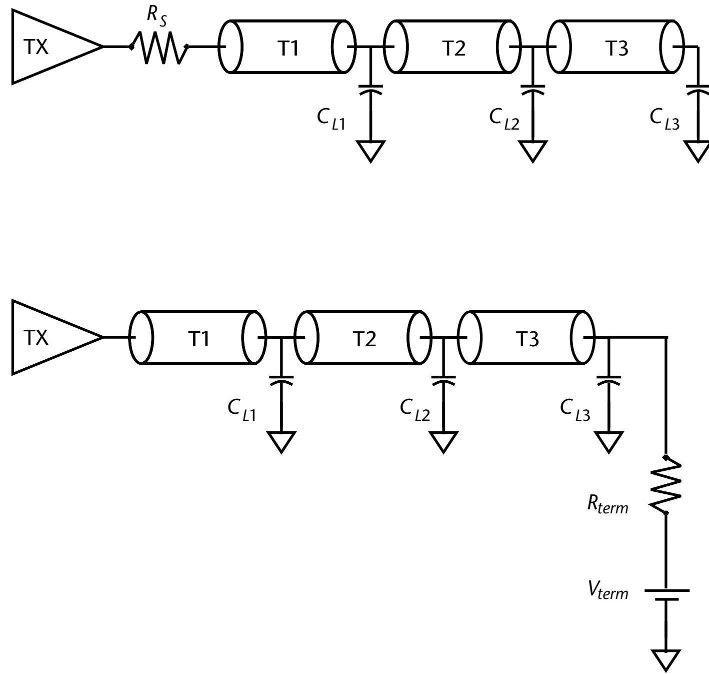


Figure 12.11 Multidrop connection. Either source series termination (using R_S) or one of the parallel schemes (Thevenin R_{term} and V_{term} as shown here) may be used.

12.7.1 Response When Signal Rise Time Is Very Short

The response of the circuit when the signal rise time is much smaller than the electrical length of the line is illustrated in Figure 12.12. Each transmission line is 50Ω and 6 inches (15.2 cm) in length. Load capacitors C_{L1} through C_{L3} are each 2 pF and represent the capacitance of a receiver.

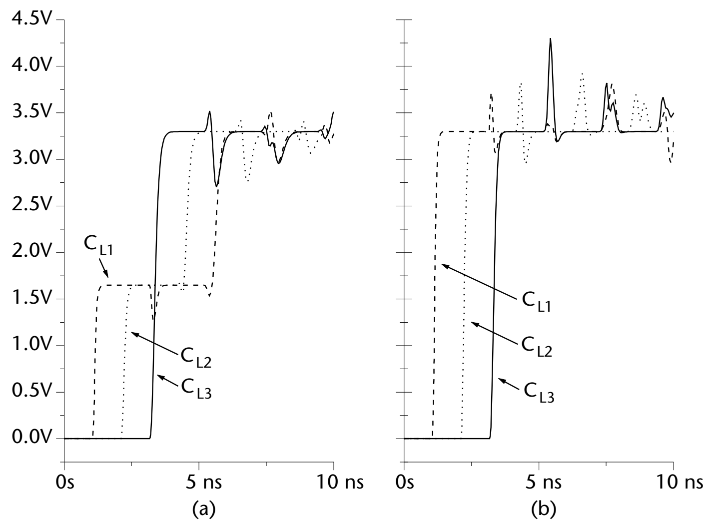


Figure 12.12 Response to the circuit shown in Figure 12.11 with (a) series termination and (b) parallel termination. Load C_{L1} is the last to fully switch when series terminated, but it is the first when parallel terminated.

It is evident from Figure 12.12(a) that when the line is series terminated, a plateau voltage is created at each load as the incident wave passes by. In this case, the plateau is not quite equal to half the value of V_{dd}, the power supply feeding the driver.

In general, if the plateau voltage is greater than the receiver switching threshold, the receivers will detect the voltage as a valid logic level and will switch on the incident wave. However, incident wave switching is not occurring in this particular case because the plateau voltage does not exceed the receiver switch point.

Instead, each of the loads must wait until the reflection created at the open circuit at the very end of the line arrives and increases the plateau to a higher voltage. Since the reflection is created at the furthest load, C_{L3} is the first to switch. We can see in Figure 12.12 how each of the remaining loads switch in reverse order as the reflection travels from C_{L3} toward C_{L1}.

In contrast, the response when parallel terminated [appearing in Figure 12.12(b)] shows that there is no plateau and reflections are not required for the signal to reach the switch point. In this case, the load closest to the driver (C_{L1}) switches before C_{L3}.

Although the parallel termination dissipates more power (including power in the driver), it gives the fastest overall response because the signal is high enough to exceed the receiver's switch point as it first passes by rather than having to wait for the reflection.

This timing difference is evident in Figure 12.12. For instance, if we use 2.4V as the receiver switch point, we can see that in this example load C_{L1} reaches that voltage at a time of about 1.2 ns when parallel-terminated and at roughly 5.7 ns when series-terminated. However, the load C_{L3} reaches 3.4 ns for both the series and parallel termination schemes.

12.7.2 Response When Signal Rise Time Is Comparable to the Transmission Line Delays

In Chapter 6 we saw that if the line was longer than half the signal rise time, the line components acted as distributed elements. We can look at this in the reverse and say that if the signal rise time is longer than twice the electrical length of the line than the elements appear as a single lump. For instance, a transmission line having a 0.5-ns electrical length would appear as a single RLC lump if the signal rise time was 1 ns, but it would appear more as a distributed circuit if the rise time was less (such as 750 ps).

We will use this insight to analyze the effects of loads evenly distributed along a transmission line when the rise time is comparable to the electrical length of the line.

A driver source series terminated to a 12 inch (30.5 cm) 65Ω transmission line is shown in Figure 12.13. Without loads the total one way delay of this line is 2.1 ns. A receiver (represented by load capacitors C_{L1} and C_{L2}) are placed midway and at the end of the line. A 65Ω resistor R_s is used to impedance match the driver to the transmission line.

A 3.3-V driver with an on resistance much lower than 65Ω launches a fast rise time signal down the transmission line. The dotted curve in Figure 12.14 shows that the voltage V_{ne} has an initial plateau of 1.65V. Since this is one half of the

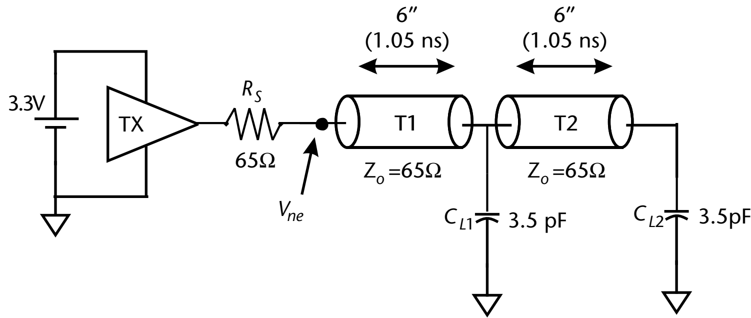


Figure 12.13 Capacitive loads C_{L1} and C_{L2} placed along and at the end of a series terminated transmission line. Signal rise time determines *T1*'s apparent impedance.

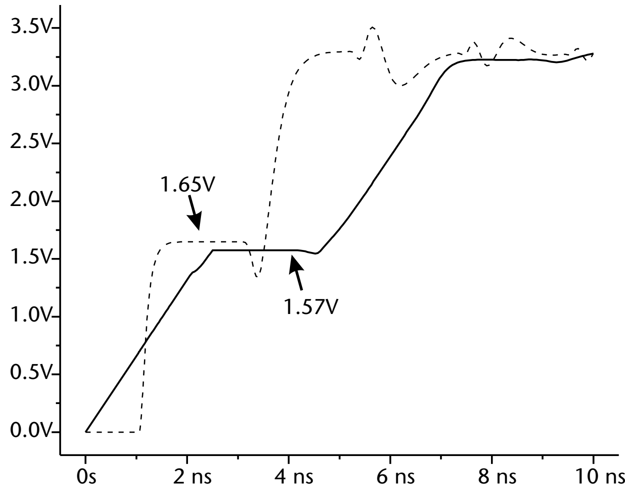


Figure 12.14 Voltage measured at V_{ne} when the signal has a very sharp rise time (dotted curve) and when it is longer (solid curve). The two voltages show that the impedance is not the same.

driver voltage this confirms the transmission line impedance is 65Ω. The negative going pulse caused by C_L is clearly visible.

The solid curve shows the response when the signal rise time is extended to 2.5 ns. The plateau is still present, but it has been shifted down to 1.57V. As we first saw in Chapter 6, the voltage divider principle can be used to determine that the transmission line impedance has fallen from 65Ω to 59Ω. The negative-going pulse caused by load capacitor C_{L1} that is so evident with the sharp-edged pulse is no longer visible.

Lengthening the signal rise time causes these changes because now the line's distributed capacitors and inductors become indistinct and blend together, appearing as a single inductor and capacitor. Capacitor C_L adds to the transmission line capacitance, increasing it, but it does not change the transmission inductance.

The transmission line inductance (L) and capacitance (C) affects the transmission line impedance (Z_o) as is shown in (12.5) (which was first presented in Chapter 6).

$$Z_o = \sqrt{\frac{L}{C}} \tag{12.5}$$

Since the capacitance has increased and the inductance has not changed, (12.5) shows how adding C_{L1} lowers the transmission line impedance. In fact, as is evident by the absence of the negative-going pulse at the end of the plateau region, C_{L1} no longer shows up as a separate element.

Usually the loads in CMOS system are mostly capacitive, making it possible to ignore the inductance. Doing this lets us calculate the loaded impedance, Z_L (12.6) [6, 7].

$$Z_L = \frac{Z_o}{\sqrt{1 + \frac{C_L}{C_o}}} \tag{12.6}$$

where L_s and C_s are the self-inductance and capacitance of the transmission line per unit length, and C_L is the load capacitance evenly distributed along the transmission line length.

Similarly, distributed loads increase the effective delay of a transmission line. The loaded delay time (tpd_L) is given in (12.7).

$$tpd_L = tpd\sqrt{1 + \frac{C_L}{C_o}} \tag{12.7}$$

Often for a circuit board trace we know Z_o and tpd and not C_o. Simultaneously solving the equations for delay and impedance presented in Chapter 6 lets us find C_o in terms of Z_o and tpd, presented here:

$$C_o = \frac{tpd^2}{\sqrt{tpd^2 \times Z_o^2}} \tag{12.8}$$

Using Figure 12.13 as an example, assume the 12-inch-long (30.5-cm), 65Ω line has an end-to-end delay of 2.1 ns without loads. Each 6-inch segment therefore has an unloaded delay of 1.05 ns. From (12.8), C_o for each segment is 16.2 pF. Using this and the value of C_L (3.5 pF) in (12.6), we find that Z_L is 59Ω, matching the results in Figure 12.14. By using (12.7), we find that the delay for each segment has increased from 1.05 ns/inch to 1.2 ns/inch (430 ps/cm to 480 ps/cm).

As the Problems illustrate, because the self-capacitance of low-impedance traces is so high, their impedance and delay are less affected by small value capacitive loads.

12.8 Stubs and Branches

In Figure 12.11 the loads are shown as being connected directly to the transmission line, but usually a short length of transmission line (a stub) is used to make the connection. This is shown schematically in Figure 12.15.

By using the definition developed in Chapter 6, we can say that the stub would be considered long if its electrical length is greater than half the signal rise time.

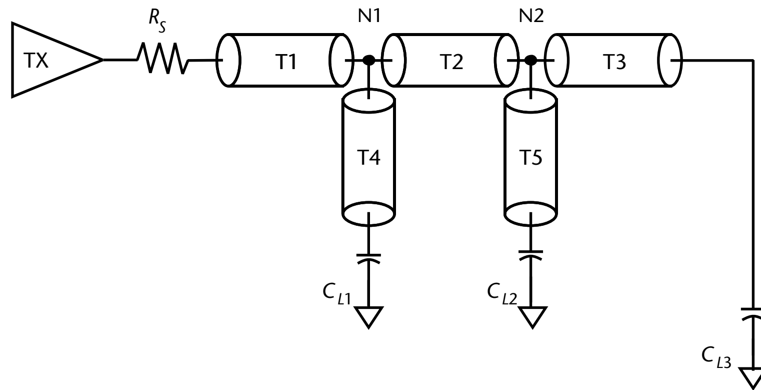


Figure 12.15 Stubs *T*4 and *T*5 connect the load capacitances to the main transmission line (*T*1 – *T*3).

For instance, if the signal rise time is 1 ns, transmission lines *T*4 and *T*5 would appear to be long if their electrical lengths were 0.5 ns or greater.

If the electrical length of *T*4 and *T*5 is small, then the capacitance of the stub is found with (12.8) and is added to the load capacitance as we did in the previous section. This increases the load capacitance and creates a larger impedance discontinuity at the point where the load connects to the line. However, because the stub is short, incident and reflected waves will not appear along its length.

12.8.1 Branches

When the trace connecting the load to the main transmission line is electrically long, it acts as a branch rather than a stub. The branch is then in parallel with the main transmission line, which lowers the impedance where the branch connects. Figure 12.15 is redrawn as Figure 12.16 to emphasize this parallel connection. The net can be series- or parallel-terminated.

It is evident from Figure 12.16 how an incident wave exiting *T*1 would simultaneously send energy down the parallel combination of *T*2 and *T*4. In this particular case *T*2 and *T*4 are each 50Ω, so the impedance experienced by the incident wave when it reaches *N*1 is 25Ω. Unless terminated at the load, the wave launched down *T*4 will reflect when it reaches load capacitor C_{L1}. That reflection will be rereflected when it reaches node *N*1, because from the point of view of this wave, the 50Ω transmission lines *T*1 and *T*2 are in parallel, and node *N*1 has an impedance of 25Ω. This same process occurs at node *N*2 with transmission lines *T*3 and *T*5. Depending on the relative lengths of *T*2 through *T*4, the reflected energy can create complex-looking waveforms at the loads, and the waveforms at each of the loads can be quite different.

We can see from this how ineffectual series termination will be at controlling these reflections. In fact, the best signal integrity will occur when each of the loads is individually terminated, but even so, the impedance mismatch between the 50Ω transmission line *T*1 and the apparent 25Ω impedance appearing at node *N*1 (and node *N*2) will still give rise to reflected energy.

This can be solved by lowering the impedance of *T*1 to 25Ω, matching *T*1 to the impedance at *N*1, and eliminating any reflections from the incident wave when

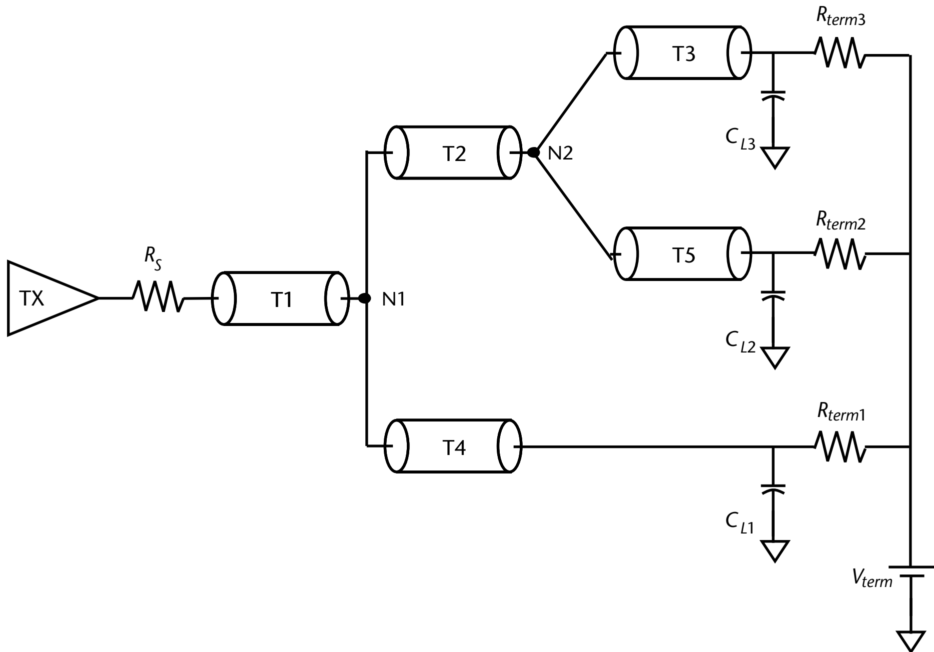


Figure 12.16 Electrically long transmission lines *T*4 and *T*5 are in parallel with transmission lines *T*2 and *T*3. This net may either be series-terminated with R_s or parallel-terminated using R_{term}.

it reaches *N*1. Provided that C_{L1} is properly terminated, reflected energy will not be present to travel through *T*4 back to *N*1, so there is no difficulty with the parallel impedance of *T*1 and *T*2 being so low (under 17Ω in this case).

These results are shown in Figure 12.17 for 50 transmission lines. Figure 12.17(a) shows the voltage across load capacitor C_{L1} and C_{L3} when a 15Ω driver is series-terminated with 50Ω (R_s = 50Ω). In Figure 12.17 parallel termination is not used. The poor signal quality is evident.

Figure 12.17(b) shows the load voltages with R_s removed and 50Ω, 1.65-V Thevenin terminators are connected to C_{L1}, C_{L2}, and C_{L3}. The signal quality is greatly improved, but the effects from the impedance mismatch at *N*1 and *N*2 are still apparent.

As is shown in Figure 12.17(c), further improvement can be obtained by lowering the impedance of T1 to 25Ω and retaining the 50Ω, 1.65V Thevenin terminators at each load. Lowering the impedance eliminates the mismatch at node *N*1, and the Thevenin termination reduces reflections from the loads.

Generally, this technique only works for unidirectional signals because it is usually not possible to properly impedance match all branches in all directions. For instance, replacing C_{L1} with a second transmitter would require that the impedance of *T*4 be lowered to the parallel combination of *T*1 and *T*2, but we have already seen that for proper operation when TX drives, transmission line *T*1's impedance has to be lowered to the parallel combination of *T*2 and *T*4.

Although Thevenin termination was used in Figure 12.17, any of the parallel termination techniques previously described can be used to terminate the branches. The fundamental point is that parallel termination prevents (or at least minimizes) reflections from being created by the loads, so the impedance mismatch created by

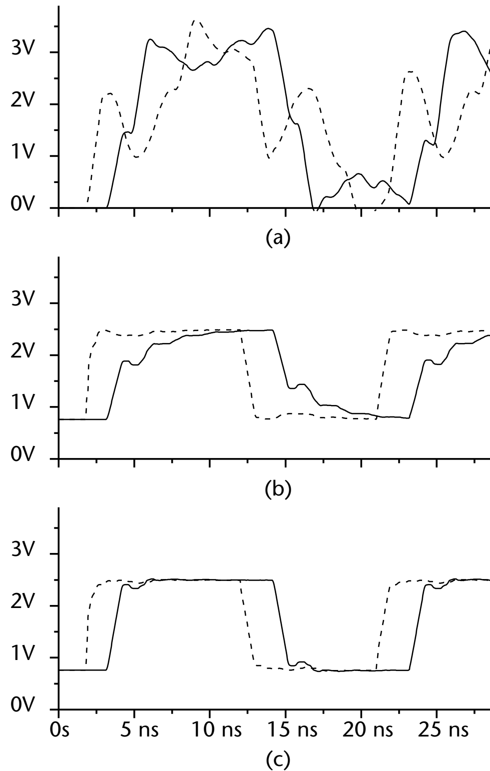


Figure 12.17 (a) Load capacitor C_{L1} voltage when $R_s = 25\Omega$, R_{term} not present and $T1 - T5 = 50\Omega$. (b) R_s removed, $R_{term} = 50\Omega$, $V_{term} = 1.65V$. (c) $R_{term} = 50\Omega$, $V_{term} = 1.65V$, $T1 = 25\Omega$.

the parallel combination of transmission lines at a branch is not as significant as it would be if the loads were not terminated. This means that because diode clamps clip the reflections and do not prevent them from being created, they usually are not a useful substitute in this situation for one of the other parallel terminations.

12.9 Main Points

- Transmission lines may be terminated at the source (source series termination) or at the load (parallel termination).
- Source series termination raises the impedance of the driver and can prevent reflected energy from rereflecting from the generator.
- Parallel termination impedance matches a high-impedance load to a lower-impedance transmission line. There are several variants, and circuit simulation should be used to determine which is best suited in a given application.
- Capacitive loads connected to a transmission line will either appear as a sharp discontinuity (if the signal rise time is very fast) or will combine with the transmission line capacitance to lower it.
- Connections made to a transmission line are either stubs (when short) or branches (when they are long). Their effects are very different.

Problems

Answers to these problems are available on the Artech House Web site at http://www.artechhouse.com/static/reslib/thierauf/thierauf1.html.

12.1 A parallel termination uses 120Ω, ± 5% resistors for $R1$ and $R2$ (shown in Figure 12.3). What is the likely impedance of the trace that this terminates?

12.2 Using the same setup described in Problem 12.1, determine if ± 5% tolerance resistors are adequate or if resistors with tighter tolerance are necessary.

12.3 Compare the performance of a 60Ω Thevenin termination shown in Figure 12.3 that uses a ± 5% resistor with the parallel termination illustrated in the previous problem.

12.4 Find the proper values for $R1$ and $R2$ (shown in Figure 12.3) to terminate a 75Ω coax cable to 1V when V_{dd} is 5V.

12.5 A system uses logic devices that have an input threshold of 2.4V. To guarantee that this exceeds the receiver threshold, the value of V_{term} in Figure 12.3 is set to be 3V. R_{term} must be no less than 130Ω to avoid exceeding the current handling ability of the driver. Find the value of $R1$ and $R2$ when V_{dd} is 5V.

12.6 Determine the difference in power dissipation in Figure 12.3 when the bus is held low if $R1$ and $R2$ are both 100Ω. Assume that V_{dd} is 3.3V and R_{ds_on} is 12Ω.

12.7 What is the minimum value capacitor that should be used to AC terminate a 65Ω trace that is carrying a 33-MHz clock?

12.8 An integrated RC network used for AC termination has a resistance R_{term} of 50Ω and a capacitance C_{term} of 470 pF. What is the lowest frequency that this network can reasonably be expected to terminate?

12.9 What is the proper value for R_{S1}, R_{S2}, and R_{S3} in Figure 12.9 when $T1$ has an impedance of 50Ω and the impedance of $T2$ and $T3$ is 65Ω? Assume that the signal rise time is 50 pS, the load capacitors are all 2.5 pF, and the transmission lines all have a delay of 2 ns.

12.10 Using the same information as in the previous problem, what should be the values of V_{term} and R_{term1} through R_{term3} when $V_{dd} = 3.3V$ and the receivers represented by C_{L1} through C_{L3} switch at $\frac{V_{dd}}{2}$?

12.11 Find the loaded impedance of the transmission line shown in Figure 12.13 when Z_o is lowered to 40Ω. How does this change compare to the 65Ω example given in Section 12.7.2?

12.12 What values should be chosen for transmission line $T1$ and resistor R_S and R_{term1} through R_{term3} in Figure 12.16 when $T2$, $T3$, and $T5$ are 50Ω and $T4$ is 65Ω?

12.13 Three state drivers are used to drive data between nodes with the topology shown in Figure 12.16. Assume that reducing power is important in this application, and to facilitate this, the bus will be made inactive (set to the high-impedance state) at frequent intervals. If the transmission lines are 50Ω and V_{dd} is 3.3V, what form of parallel termination (shown in Figure 12.3) is the best choice to reduce power?

References

[1] Seams, J., "A Comparison of Resistive Terminators for High Speed Digital Data Transmission," *High Frequency Electronics*, October 2005, pp. 18–26.

[2] Hewlett Packard Corp., "Applications Bulletin 14: Waveform Clipping with Schottky Diodes," undated.

[3] Hewlett Packard Corp., "Applications Bulletin 15: Waveform Clamping with Schottky Diodes," March 1989.

[4] Ethirajan, K., and J. Nemec, "Termination Techniques for High-Speed Buses," *EDN*, February 16, 1988, pp. 135–145.

[5] Vu, M., "Choose Termination and Topology to Maximize Signal Integrity and Timing," *EDN*, October 24, 1996, pp. 95–98.

[6] Bakoglu, H. B., *Circuits, Interconnections, and Packaging for VLSI*, Reading: Addison-Wesley, 1990.

[7] Fairchild, Inc., *ECL Data Book*, 1977.

CHAPTER 13

Differential Signaling

13.1 Introduction

The previous chapters discussed *single-ended* signaling, where a signal voltage is sent down a single transmission line and is compared with a reference voltage to determine its logical value. This scheme is shown in Figure 13.1(a). It is apparent how a voltage difference between the ground of the driver and the ground of the receiver (V_{noise}) can reduce the received noise margin, leading to false triggering. HSTL [1] and SSTL [2] are popular examples of single-ended signaling.

This contrasts with *differential signaling*, where a differential driver simultaneously sends a signal and its complement down two transmission lines (a differential transmission line or *diff-pair*). As shown in Figure 13.1(b), the differential driver switches one output from a logic low to high while simultaneously switching the complementary output from high to low. This differential signal is detected by a differential amplifier (a *diff-amp*), which uses the voltage difference between the two signals (and not the voltage measured to ground) to determine if it is a logic high or low. Common examples of differential signaling include LVDS [3], Infiniband [4], and RapidIo [5].

One of the diff-pairs is usually called the negative terminal even when the signaling is done with positive voltages. For instance, LVDS signals typically switch between +1.0V and +1.4V (and is centered around a common mode voltage of 1.2V [6]), but the driver outputs (*DO*+ and *DO*) and receiver inputs (*RI*+ and *RI*–) are referred as the "positive" and "negative" terminals.

In comparison to single-ended signaling, differential signaling:

- Requires more signal traces and I/O pins;
- Retains good noise immunity even when signaling with lower voltages;
- Requires different termination techniques;
- Is more immune to differences in voltages between the driver and receiver grounds;
- Is more immune to noise coupled on the signal traces (provided the noise couples equally to them both).

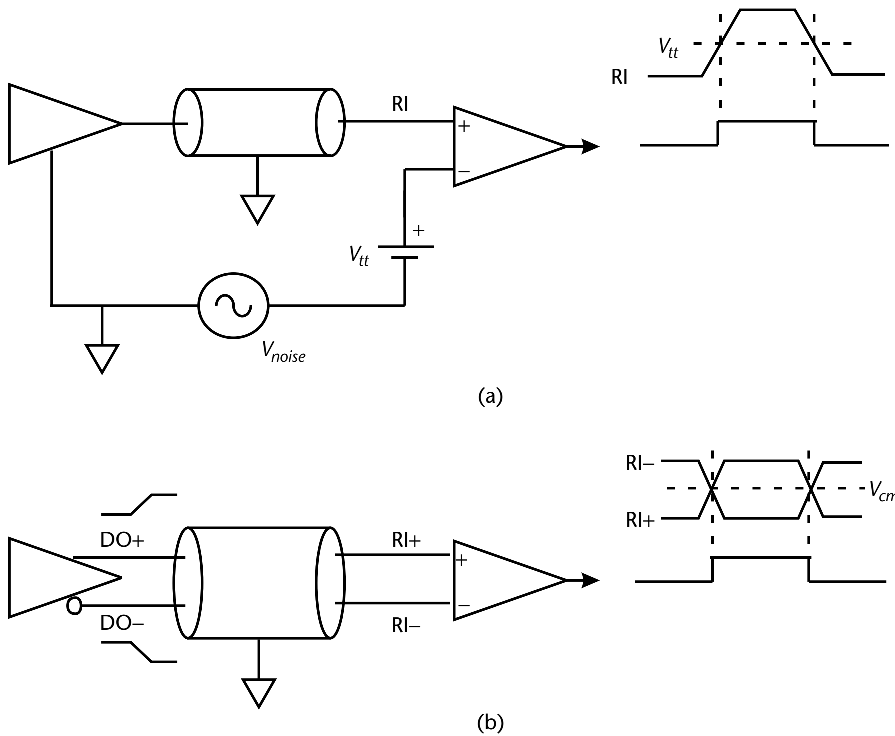


Figure 13.1 (a) In single-ended signaling the receiver switch point (V_{tt}) is with respect to the local ground. (b) With differential signaling the crossing of complementary signals determines the switch point.

13.2 What Are the Electrical Characteristics of Differential Signaling?

A diff-amp receiver amplifies the difference in voltage appearing on its inputs (*RI*+ and *RI*–) and rejects voltages that simultaneously appear on both inputs. This is why noise that couples equally to both inputs is ignored by the receiver. In fact, to get the most benefit from differential signaling, we will intentionally exploit this characteristic by deliberately routing the two traces forming the diff-pair so that noise is encouraged to couple equally to both traces.

Taking the receiver ground connection as a reference, the difference in the ground voltages between the transmitter and receiver appears to the receiver as a noise voltage that adds or subtracts from the incoming signal. However, because they do not use ground as a signal reference, the switch point of differential receivers is not affected by this type of *common mode* noise.

However, there are limits in the common mode voltage over which the circuit will properly operate. The receivers' *common mode range* specifies the maximum voltage that the signal may reach (measured with respect to the ground local to the receiver) before the transistors within the diff-amp become unbiased and it loses the ability to differentiate valid differential signals. This is a critical parameter to assess when performing signal integrity simulations and when making laboratory measurements.

13.2.1 How Is a Differential Pair Observed?

A probe with a ground clip is used to make single-ended oscilloscope measurements in the laboratory. The ground clip sets the reference for the measurement by specifying the location where zero volts is defined. This is also how SPICE simulations are customarily observed, since, unless told otherwise, SPICE calculates and displays its results with respect to node 0, the system ground.

In contrast, a differential probe is used to measure the differential signal. The displayed signal is truly differential: The difference between *RI*+ and *RI*– appearing in Figure 13.1 is measured rather than its value relative to ground. An example of a differential signal is shown in Figure 13.2.

We see that each of the signals swing between 0V and 1V with respect to ground, but the difference signal swings from +1V to –1V.

Because an oscilloscope only allows a few transitions at a time to be displayed, it is difficult to observe the interaction between data bits occurring at significantly different places in a long stream of data. Setting the oscilloscope time base to simultaneously display all the bits does not provide the resolution necessary to distinguish a single pulse. Rather than an ordinary oscilloscope, a device that captures and then overlays successive pulses (bits) is used. One example of a device that can do this is the digital signal analyzer such as the one shown in Figure 13.3.

13.2.2 What Is an Eye Diagram?

A digital signal analyzer creates an *eye diagram* [8–12] such as that shown in Figure 13.4 by capturing a data stream and overlying successive pulses. Ideally, all of the rising edges should exactly overlay, as should all of the falling edges. The space in between should be totally clear (void of any transitions). The *data mask* defines the minimum acceptable width and height of this opening and is a specification that the link must meet for proper operation.

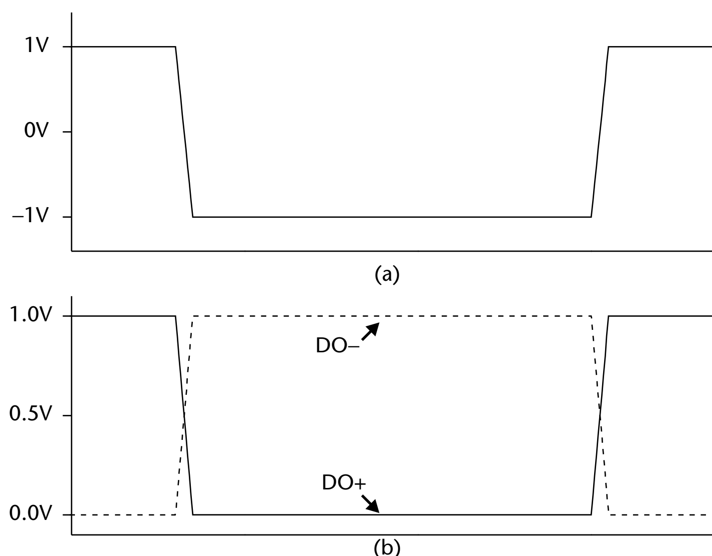


Figure 13.2 (a) The differential signal (*DO*+) – (*DO*–) swings between ± 1V even though (b) when measured respect to ground the signals switch from 0 to 1V.

Figure 13.3 A digital signal analyzer such as this Tektronix DSA 8200 [7] is used to display a differential data stream. (Courtesy of Tektronix, Inc. Used with permission.)

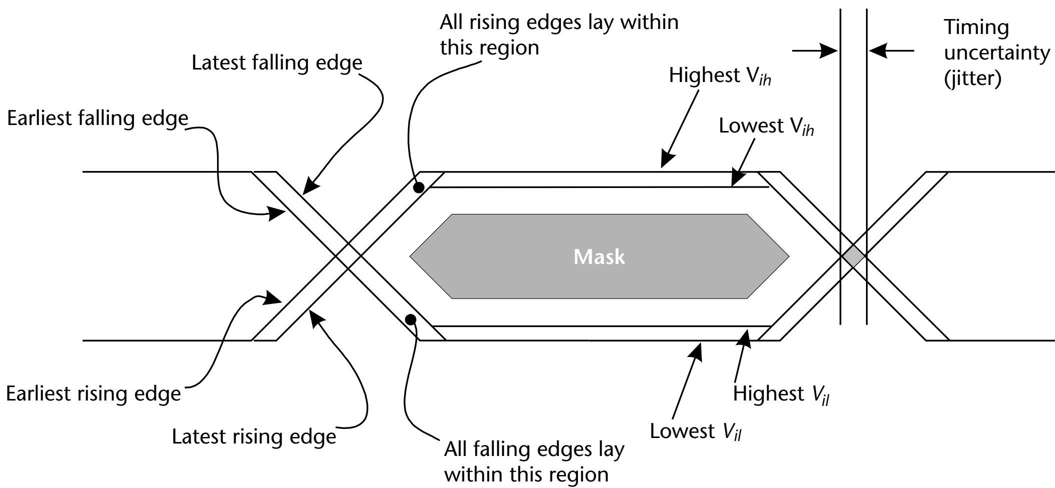


Figure 13.4 The earliest and latest appearances of the rising and falling edges and their amplitudes are simultaneously displayed in this eye diagram. (*From:* [13]. Used with permission.)

In reality jitter will prevent the pulses edges from perfectly lining up, and reflections and losses will prevent each pulse from having the same logic low (V_{il}) and high levels (V_{ih}). These things reduce the horizontal and vertical eye opening and, if severe enough, can prevent proper operation.

13.2.3 What Is Intersymbol Interference?

We saw in Figure 8.2 how transmission line losses cause a pulse to shrink in height (amplitude) and for its base to grow in width. This widening at the base is particularly serious because if it grows large enough, successive pulses (symbols) will blend together, making it difficult to distinguish between each one. Symbols can also

become distorted when they combine with residual energy remaining on the line caused by reflections from earlier symbols.

Any such interference between symbols [*intersymbol interference* (ISI)] alters the shape of the waveform appearing at the receiver [14–16] and is easily observed with an eye diagram. Fortunately, signal processing techniques have been developed that allow proper signal reception of very high-speed signals, even on lines where ISI and losses are high.

13.2.4 How Are the Effects of Loss Corrected?

We have seen in previous chapters how transmission line losses increase as frequency increases, and we also saw that a signal with fast rise times have a high-frequency content. Since high-data rate signaling requires sharp, fast rise time pulses, it is apparent how transmission line losses can become significant when signaling at high data rates. In fact, even in the absence of crosstalk or other kinds of noise, in high-speed signaling it is quite likely for frequency dependent losses to distort the pulses enough to prevent the link (the hardware and software connection between a transmitter and receiver) from operating properly.

It is very common to use an *equalizer* on high-speed links to improve the quality of the received signal [11, 12, 14, 15]. The equalizer can either be a passive circuit [13, 17] or a *precompensation* circuit built into the transmitter. With the precompensation technique [10, 12, 18–20], the transmitter monitors the data pattern that it is sending and intentionally predistorts the signal to compensate for the transmission line loss.

When done properly, the predistortion applied by the transmitter and the distortion caused by the transmission line interact in such a way as to produce in a much cleaner signal at the receiver. An example of the benefits of a properly equalized net is shown with the eye diagram appearing in Figure 13.5.

The eye in Figure 13.5(a) is difficult to find and is an example of unacceptable signal integrity. Activating equalization at the transmitter as shown in Figure 13.5(b) dramatically reduces jitter and improves ISI. The eye is clearly defined and is large enough to permit reliable link operation.

13.3 What Are the Electrical Characteristics of a Differential Transmission Line?

In Chapter 9 we saw how two traces that were carrying signals exactly out of phase with one another operated in the odd mode. Because this describes the way in which ideal differential signals switch, differential signaling operates in the odd mode. This suggests that to form a diff-pair the odd-mode impedance should be set to the desired impedance. However, as shown in (13.1), the *differential impedance* is actually twice the odd-mode impedance. This profoundly affects the way in which differential transmission lines are created and terminated.

$$Z_{diff} = 2Z_{oo} \tag{13.1}$$

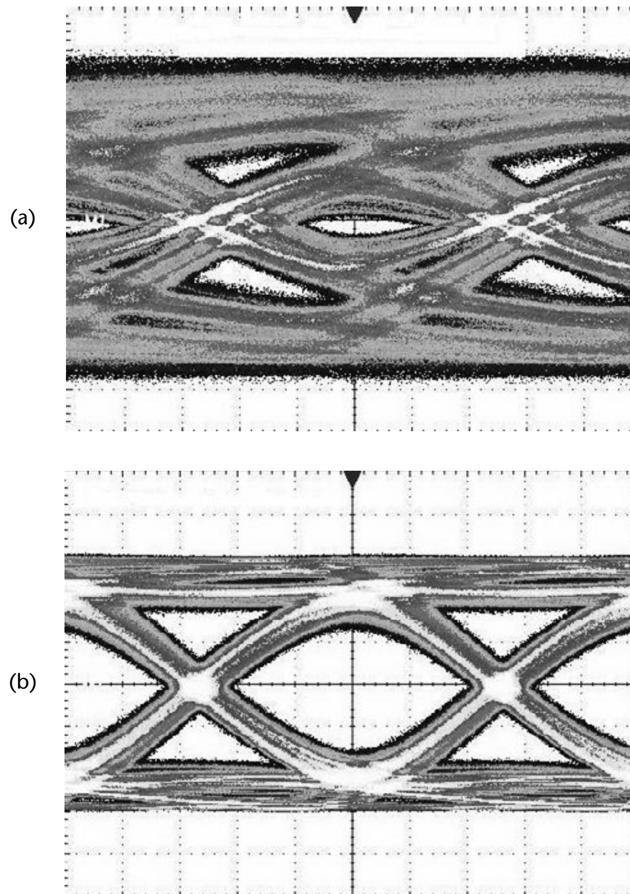


Figure 13.5 Measured eye diagram showing the link (a) with no equalization and (b) with equalization. (*From:* [13]. Used with permission.)

13.3.1 Why Is the Differential Impedance Twice the Odd-Mode Impedance?

Figure 13.6 shows a 50Ω differential driver connected to a pair of traces that have an odd-mode impedance of 50Ω. From (13.1) this setup has a differential impedance of 100Ω, and we will prove this by calculating the voltages and currents launched by the driver.

During the time t, the driver is transmitting a logic high value by sending a high voltage on the $DO+$ line and a low voltage on the $DO-$ line. To find the differential impedance, we must determine the values of these voltages and of the launched current. With this we will use Ohm's law to find the impedance.

The termination resistance R_{term} placed across the diff-pair converts the launched differential current into a differential voltage that is detected by the diff-amp receiver. The received voltage swings about a common-mode voltage of 0.5V. The diff-amp input terminals have very high impedance and they draw no current.

By using the voltage divider principle as described in Chapter 6, we might conclude that the 1-V, 50Ω driver would launch a 0.5-V signal down the 50Ω transmission line. In fact, 0.75V is launched. To understand why this is so, we redraw the circuit as shown in Figure 13.7, diagramming the common-mode voltage and the current flow. The diff-amp receiver inputs draw no current and so can be ignored.

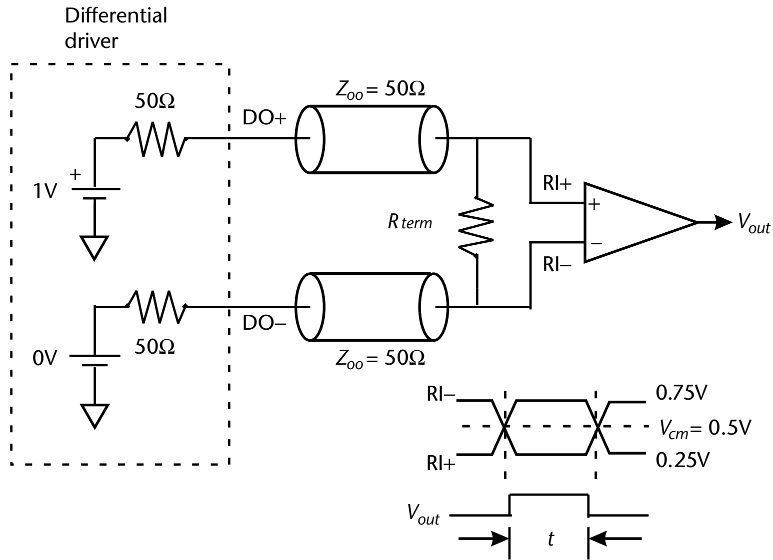


Figure 13.6 A 50Ω differential driver connected to a differential receiver. The signal appearing across R_{term} swings 0.25V above and below a 0.5-V common-mode voltage. The receiver output shown at the bottom of the figure swings rail to rail.

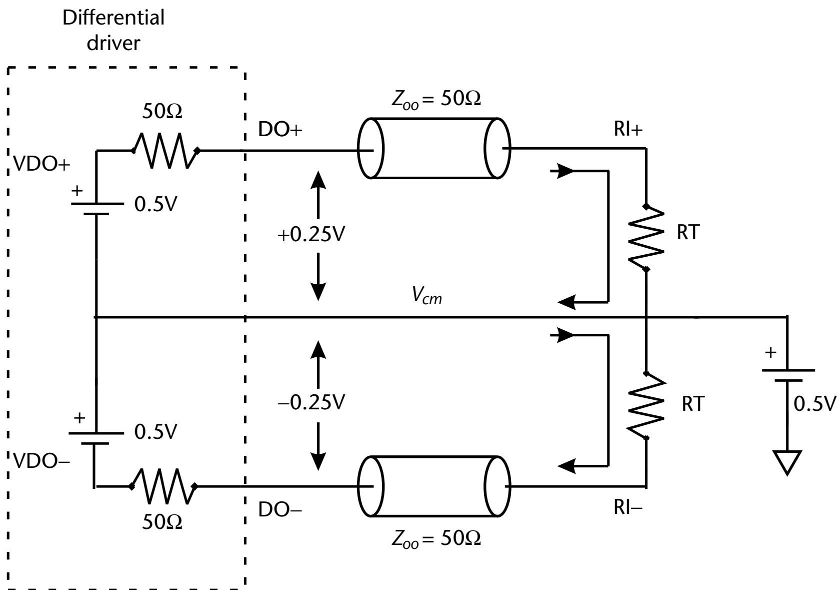


Figure 13.7 Differential driver connected to a differential pair. The outputs switch about the common-mode voltage *Vcm*. Voltages on *DO*+ and *DO*– are measured with respect to *Vcm*.

The circuit consists of a 0.5-V supply voltage around which the outputs switch (the common-mode voltage, V_{cm}), and two 0.5-V supplies and 50Ω resistances connected to the *DO*+ and *DO*– outputs. The termination resistor R_{term} shown in Figure 13.6 has been separated into two equal valued resistors called *RT*. They are connected together to the common-mode voltage supply.

When driving a logic high, the common-mode supply adds to the supply connected to the *DO*+ pin. When measured with respect to ground, *DO*+ is a 50Ω resistance in series with a 1-V source. Similarly, the supply connected to the *DO*– pin

subtracts from the common-mode voltage, making it a 50Ω resistance in series with a 0V source. This validates the model; the *DO*+ and *DO*– pins have the correct 50Ω impedance, and with respect to ground, they switch between the correct voltage levels (0 and 1V).

With this circuit arraignment the voltage divider principle can be applied to find the launched voltage. Since the driver and transmission line impedances are both 50Ω, the voltage on the *DO*+ and *DO*– pins will each be half the switched voltage. As shown, this is ±0.25V when measured with respect to the common-mode voltage.

Our analysis to this point is illustrated in Figure 13.8. For simplicity only the DO+ signal is shown. As expected the launched voltage is 0.25V higher than the common-mode voltage. Ohm's law can now be used as was done in Chapter 6 to find that 5 mA is sourced by *DO*+. The *DO*– analysis is the same, noting that the pin is 0.25V lower than the common-mode voltage. That driver sinks 5 mA from the common-mode supply.

The arrows in Figure 13.7 show that the currents from the *DO*+ and *DO*– outputs flow in the opposite directions. Because they are the same value, this makes the net current in the common-mode connection zero, and the connection can be removed without disturbing the circuit operation. This has been done in Figure 13.9, where we see the voltages of the two outputs measured with respect to ground rather than to the common-mode voltage. The *DO*+ output is 0.25V higher the 0.5-V common-mode voltage, making it 0.75V with respect to ground. The *DO*– output is 0.25V lower than the common-mode voltage, making it 0.25V.

The voltage difference between the two outputs is the *differential output voltage* and is given in:

$$V_{diff} = V_{DO+} - V_{DO-} = 0.75 - 0.25 = 0.5\text{V} \tag{13.2}$$

All the background information is now available to find the differential impedance; the differential voltage is 0.5V and the current is 5 mA. Ohm's law determines the relationship between the transmission line impedance, voltage, and current. In this case the differential impedance is 100Ω, as given in:

$$Z_{diff} = \frac{V_{diff}}{I_{diff}} = \frac{0.5\text{V}}{5\text{ mA}} = 100\Omega \tag{13.3}$$

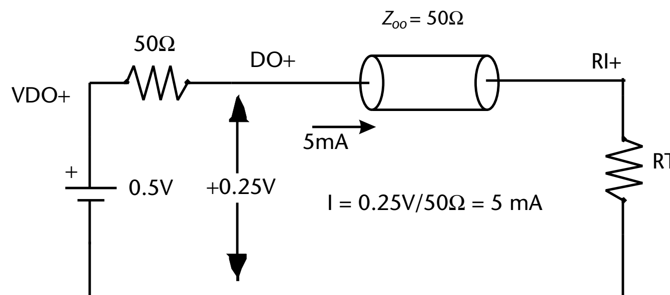


Figure 13.8 A 0.25-V, 5-mA signal measured with respect to V_{cm} is launched down a 50Ω transmission line.

From this we see that, although the traces were routed with a 50Ω odd-mode impedance, the differential impedance is actually twice that (100Ω).

13.4 How Are Differential Transmission Lines Terminated?

From Figure 13.8 we see that a proper termination is formed by connecting a resistor equal to the transmission line impedance between each traced and a common-mode voltage. For instance, for a diff-pair having a 50Ω odd-mode impedance, attaching to each of the traces 50Ω resistors connected to a supply voltage (such as the 0.5V common mode from the previous section) would properly terminate the line.

A second approach appears in Figure 13.9, which shows that a valid termination can be formed by placing resistance directly across the traces forming the diff-pair. It is evident the value for a single resistor would be twice the value of *RT*, and we have just seen that for proper termination *RT* should be made equal to the odd-mode impedance. For instance, attaching a 100Ω resistor across the diff-pairs will properly terminate them when they have an odd-mode impedance of 50Ω. This makes intuitive sense: Because the differential impedance is equal to twice the odd-mode impedance, a single resistor equal to the differential impedance should be a proper differential termination. These two ideas are illustrated in Figure 13.10.

Figure 13.10(a) shows the common-mode termination, while Figure 13.10(b) shows the differential termination technique. We can see in Figure 13.10(a) that if the common-mode voltage is made zero, the two resistors are effectively connected to ground. In fact, diff-pairs are sometimes terminated with a resistor equal to the odd-mode impedance (Z_{oo}) connected between each of the traces and ground. This scheme is attractive for those drivers that require a connection to ground to establish a DC bias for proper operation, but such a termination is not truly differential. Any difference in the voltage between the ground at the transmitter and the local ground at the receiver where the termination resistors are connected will appear as a voltage offset to the receiver, which can erode receiver noise margins.

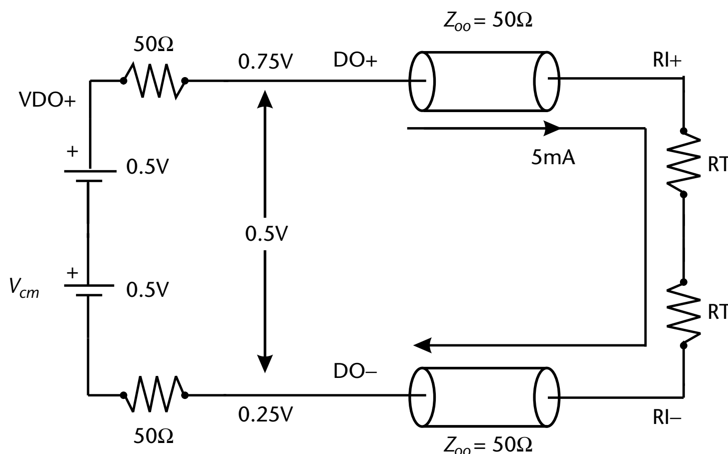


Figure 13.9 Output voltages measured with respect to ground.

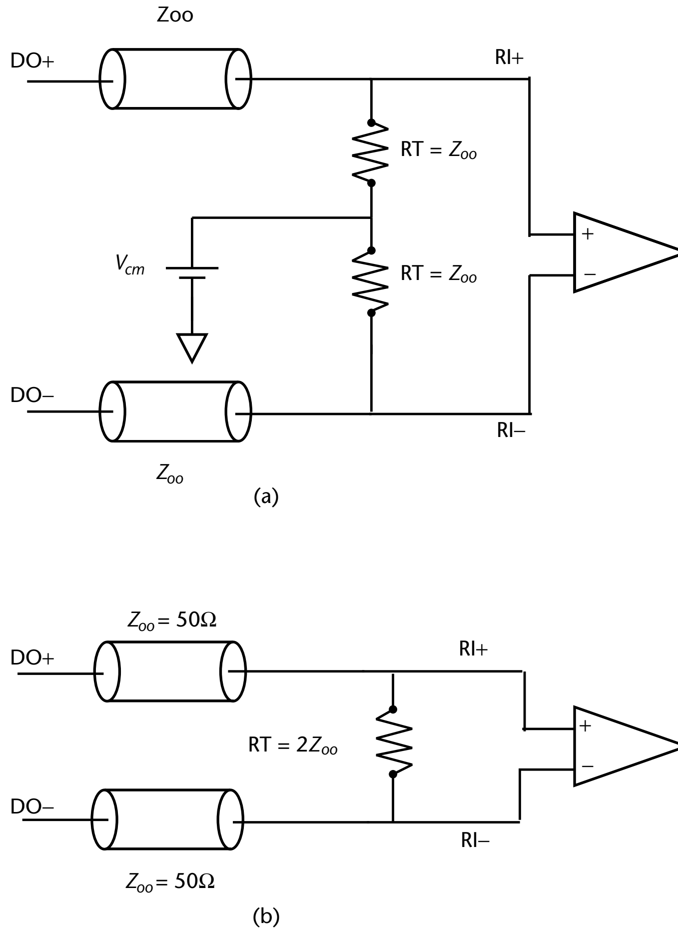


Figure 13.10 Diff-pair termination either (a) with a resistor to the common mode voltage or (b) between traces.

13.4.1 How Should a Diff-Pair Be Terminated When the Propagation Is Not Fully Odd Mode?

Assuming that signals propagate in the odd mode with no even-mode component is an important caveat that is usually not valid in practice, especially when working with high-speed signals. We now turn our attention to describing a few of the physical effects that can cause a differential signal to experience mode conversion (changing from fully odd mode to having some even-mode component) [21, 22].

In the previous section we only had a single propagation mode (odd mode), so a single resistor across the diff-pairs was sufficient to properly terminate the traces. However, when both the even and odd modes are present, it is necessary to properly terminate both modes. Figure 13.11(a) shows how to do this.

Three resistors are necessary: one on each trace to a voltage source to terminate the even mode ($R1$ in the figure) and one connecting the traces for the odd- mode termination ($R2$). The voltage source can be a termination voltage or it can be set to zero, which is the same as connecting the resistors to ground, as is shown in Figure 13.11(a).

The values for the resistors [23] are given in:

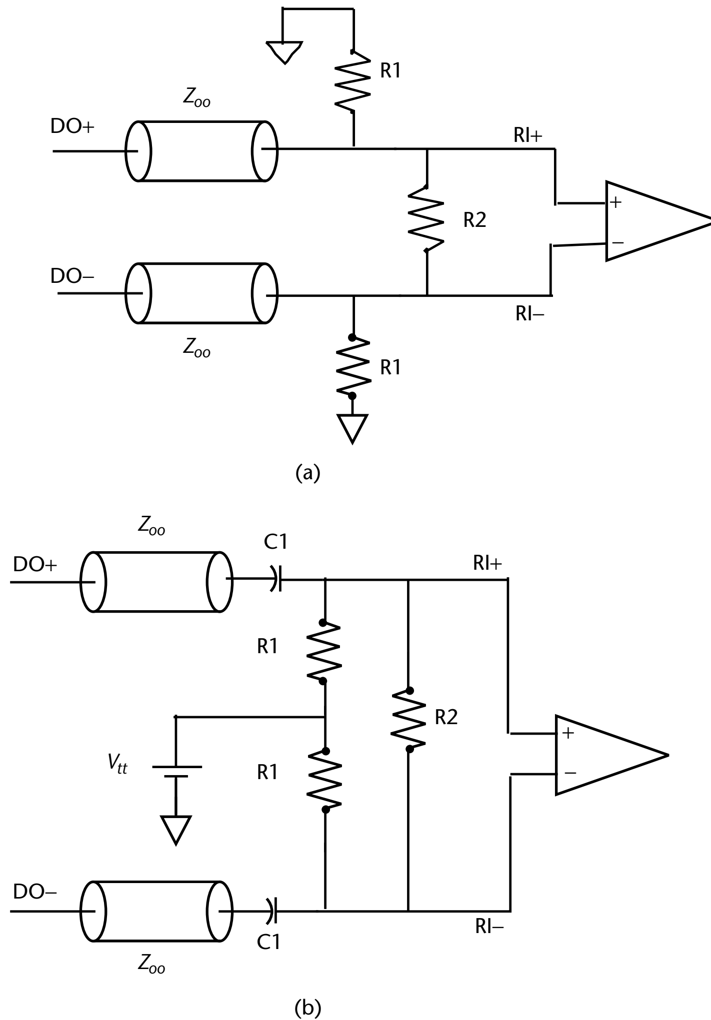


Figure 13.11 (a) Termination scheme when the even and odd modes cannot be ignored. (b) By connecting $R1$ to V_{tt} the line is simultaneously rebiased and terminated. Capacitor $C1$ allows the transmitter and receiver to operate at different bias voltages.

$$R_1 = Z_{oe} \tag{13.4}$$

$$R_2 = \frac{2Z_{oe}Z_{oo}}{Z_{oe} - Z_{oo}} \tag{13.5}$$

Figure 13.11(b) shows how capacitors and a power supply can be used to simultaneously rebias and terminate a diff-pair. The ability to both *level shift* and terminate is useful when the receiver and transmitter common-mode voltages are different. For instance, this technique is used when it is necessary to send signals between logic families operating at different supply voltages. In this case V_{tt} is set to the common-mode voltage required by the receiver and so puts the diff-amp at the optimum bias point for its operation. The transmitter can be signaling at a voltage significantly different from V_{tt}, but the capacitors only allow the signal amplitude

to pass and do not allow the receiver to be influenced by the transmitter's different common-mode voltage. This idea is explored in the Problems.

From the Problems we find that (13.4) and (13.5) can be checked by noting that when the traces are very loosely coupled, $R2$ becomes infinite and $R1$ becomes equal to Z_o. This makes intuitive sense: Very loosely coupled lines can be treated as isolated transmission lines and are properly terminated with a resistor equal to Z_o.

Some field solvers directly report the even- and odd-mode impedances, which makes using (13.4) and (13.5) easy. Alternatively, the even and odd impedances can be calculated by using (9.4) and (9.6) from Chapter 9, as we do in this chapter's Problems when finding $R1$ and $R2$.

13.4.2 What Prevents Creating Odd-Mode Signals?

As we have seen, the diff-pair signals must be exactly out of phase for odd-mode signaling. In fact, anything that causes the two signals to lose this phase relationship (even for a short distance) creates an even-mode voltage that, if large enough, requires the three-resistor scheme shown in Figure 13.11 for proper termination. Besides being a signal integrity problem, even-mode currents are a serious *electromagnetic interference* (EMI) concern that can hinder a product from passing radiated emission type testing [24].

Even mode signaling can arise when:

- The diff-pair trace lengths are not exactly the same length.
- The diff-pairs are routed in different dielectrics (microstrip versus stripline) or on different routing layers.
- The transmitter output drivers do not switch at precisely the same time.
- The output drivers have different rise and fall times.

Layout techniques (presented in the next section) can help minimize length matching problems, and proper design of the transmitter output driver (along with the logic that drives it) can minimize the remaining two items.

Circuit board signal integrity engineers perform simulations or electrical measurements to determine how closely the driver output impedance and rise times are matched, and the results from field solvers are used to determine if a significant even-mode component is present in the diff-pairs. If so, it is necessary to use the three-resistor scheme shown in Figure 13.11. By building detailed circuit models, the effects that the imperfect driver and unequal trace lengths have on the overall circuit operation can be determined.

13.5 How Are Differential Transmission Lines Created?

The two traces forming a diff-pair can either be routed side by side on the same layer (*edge-coupled*) or on top of each other on separate layers (*broadside-coupled*). Cross-sections of these two configurations (*topologies*) are shown in Figure 13.12. They may be formed as tightly or loosely coupled pairs, but, as we will see, these topologies have different electrical characteristics.

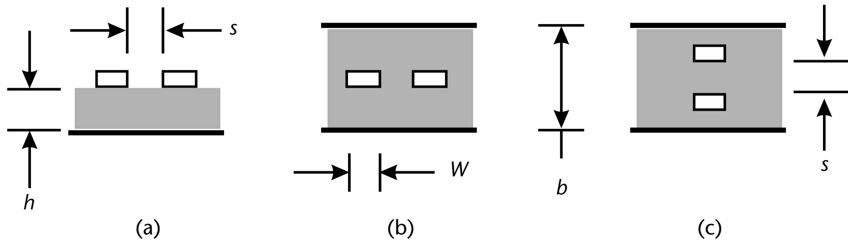


Figure 13.12 (a, b) Edge-coupled and (c) broadside-coupled diff-pairs in cross-section.

13.5.1 Manufacturing Trade-Offs Between Loosely and Tightly Coupled Pairs

Loosely coupled pairs are attractive to fabricate because the alignment and spacing between traces and to the return planes is not as critical as with tightly coupled traces. Also, the differential impedance of loosely coupled traces is not affected as much by over etching. This means that it is easier for the manufacturer to obtain the specified differential impedance even if the traces are not rectangular (see Chapter 8). Both of these factors reduce scrap and so makes manufacturing less costly.

Another consideration is that by making edge coupled traces loosely coupled the individual routing layers are thinner than they would be if the traces were tightly coupled. As illustrated in Table 13.1, this makes it possible to include more routing layers for a given overall board thickness.

The microstrip traces are covered with 1 mil of solder mask having the same dielectric constant as the laminate.

In Table 13.1 the tightly edge-coupled microstrip and stripline traces are defined as having a trace separation of (s) of equal to the trace width, while s is three times the trace width for loosely coupled traces. This designation is arbitrary, but it allows us to see how choosing the amount of coupling can affect the routing layer thickness.

On FR4 separating the 4-mil-wide broadside traces by less than about 5 mils requires impractical board dimensions, so for the broadside case tight coupling is defined in Table 13.1 as 7 mils. This requires in a plane separation (b) of 31 mils to obtain a 100Ω differential impedance. Loosely coupling the traces by increasing the trace separation to 20 mils results in a slight reduction in the overall layer thickness.

Table 13.1 Approximate Values for h and b (in Mils) for Tightly and Loosely Coupled 4-Mil-Wide, Half-Ounce 100Ω Diff-Pairs on FR4 ($\varepsilon_r = 4.2$)

Topology	*Tight*		*Loose*	
	h or b	*s*	*h or b*	*s*
Microstrip	4	4	3	12
Stripline	42	4	11	12
Broadside	31	7	29	20

The separation between traces is s (in mils). Figure 13.12 defines h, b, and s.

13.5.2 Electrical Trade-Offs Between Loosely and Tightly Coupled Pairs

There are three electrical characteristics to consider when choosing between loosely and tightly coupled traces: signal loss, EMI, and noise.

In general, tightly coupled diff-pairs have lower signal losses than loosely coupled pairs when fabricated on the same laminate and with the same trance width. Losses at 1 GHz are shown in Table 13.2 when the dielectric thicknesses are as illustrated in Table 13.1.

RF emissions will be reduced with tightly coupled traces [27, 28], so loosely coupled diff-pairs should be avoided (especially for microstrip diff-pairs) in those circumstances where EMI is a concern.

13.6 What Are Some Suggested Diff-Pair Layout and Routing Rules?

The electrical requirements for proper differential signaling are:

- The differential impedance must be consistent and of the correct value.
- The traces forming the diff-pairs must be the same length.
- Noise must be equally coupled to both traces.

These three electrical rules are satisfied with the following rules:

1. Use the trace width and spacing specified by the circuit board manufacturer to obtain the proper differential impedance.
2. Do not allow a signal trace to come between the two traces forming a diff-pair.
3. Keep signal traces (and vias) much further from the diff-pair than the spacing between the pairs themselves.
4. Avoid routing over splits in the return path.
5. Route the diff-pair traces on the same layer.
6. Minimize layer hopping.
7. Use differential vias when layers must be changed.
8. Match the overall trace lengths.
9. Match the length on each layer.
10. Route traces passing through pin fields and near vias so that noise is equally coupled on both pairs.

Table 13.2 Loss (in Decibels Per Meter) Simulated [25, 26] at 1 GHz for Tightly and Loosely Coupled 4-Mil-Wide 100Ω Diff-Pairs on FR4 ($\varepsilon_r = 4.2$)

Topology	*Tight*	*Loose*
Microstrip	7.9	8.5
Stripline	8.3	8.4
Broadside	8.2	8.9

An explanation and example violations for each rule appear in the following sections.

13.6.1 Obtaining the Proper Differential Impedance

Differential pairs are so common in modern designs that nearly all circuit board fabrication shops now have experience in creating them. This is especially true for edge-coupled 100Ω diff-pairs that are wider than the minimum allowed trace width. For instance, most shops can reliably create an edge coupled 100Ω diff-pair using 8-mil-wide (or wider) traces. High-end shops can reliably form 100Ω diff-pairs with traces less than 5 mils wide.

It is important not to overspecify when discussing the requirements with the circuit board shop. For instance, it is proper to specify that the traces are to be a 100Ω differential pair using traces 8 mils in width and that the traces are to be spaced as close together as possible. This gives the shop freedom to adjust the trace width (bias the trace), spacing, and laminate thickness based on their experience manufacturing similar boards. However, it is undesirable to tell the shop that the traces are to be 8 mils wide and have a 50Ω odd-mode impedance when spaced 8 mils apart. Although technically correct, this terminology not widely used by circuit board manufacturers and is so restrictive that it might increase the board manufacturing cost.

The second aspect of maintaining proper differential impedance involves trace routing. Once the manufacturer has specified the trace separation necessary to obtain the desired differential impedance, make sure the layout designer adheres to this everywhere.

One example of how this rule is violated is shown in Figure 13.13.

In this example a diff-pair jogs around a via, which for a short distance increase the trace-to-trace spacing. This encourages the traces to couple to the via rather than to each other, altering the impedance.

The third aspect of maintaining proper differential impedance involves trace width. It is sometimes necessary to intentionally make a trace narrow so it can be routed through a connector or BGA pin field. This is especially likely with long traces because they are often made very wide to reduce signal loss. Such wide

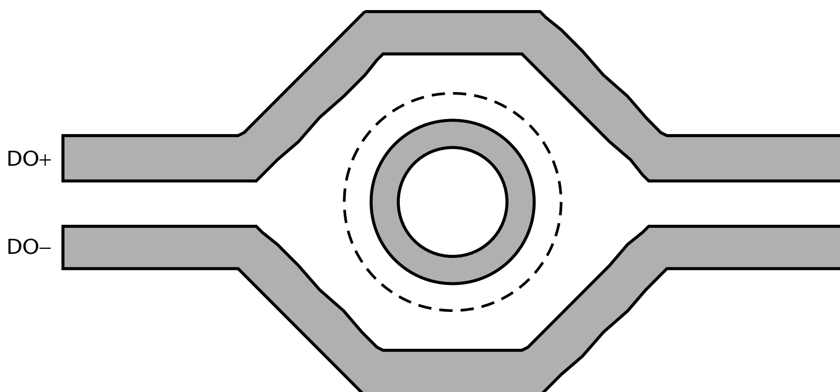


Figure 13.13 Top view of diff-pair traces jogging around a via and its antipad (shown as a broken circle). This routing increases the differential impedance.

traces must be reduced in width (necked down) to fit through the pin field, which increases the trace impedance through that region.

In practice the communication link may be robust enough to operate satisfactorily when the signal is exposed to any of these types of discontinuities. This is particularly true if the data rate is low enough to make the period of the discontinuity small compared to the signal bit time. However, even a low data rate link can fail if many of these discontinuities are simultaneously present.

13.6.2 Signal Trace Between Diff-Pairs

When routing edge-coupled diff-pairs, insure that there are no signal traces routed between the traces forming the diff-pair. It is not usually possible to violate this rule when the traces are tightly coupled, but violations can arise when the traces are widely separated (loosely coupled).

Violation of this rule usually comes about when space between the diff-pairs is large enough for the autorouter to place a signal trace in the area between the diff-pairs. Autorouting software is a tool often used by layout designers because it speeds the layout process by working from a list of nets to automatically place and connect traces to components. If there is enough space, the autorouter may route signals between the traces forming the loosely coupled diff-pair unless it is purposely restricted from doing so. This type of error can be hard to find on dense boards because the errant trace (or traces) may only run for a short distance.

13.6.3 Keeping Signal Traces Far Away from Diff-Pairs

The top view of either a stripline or microstrip edge coupled diff-pair is shown in Figure 13.14. Traces *A*1 and *A*2 are aggressors routed alongside the diff-pair (*DO*+ and *DO*−). It is apparent that trace *A*1 will preferentially couple onto *DO*+ and will couple much less to *DO*−. This asymmetric coupling changes the impedance of *DO*+ and makes the noise common mode, which will not be rejected by the receiver.

For these reasons the separation from the aggressor to the diff-pair (*sn*) must always be much greater than the separation between the traces forming the diff-pair (*sd*). This rule applies to either broadside or edge-coupled pairs.

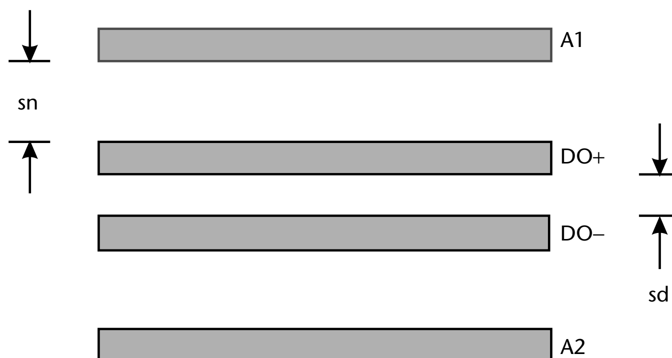


Figure 13.14 Top view showing four traces. Prevent common mode noise by making the spacing between the diff-pair (*sd*) smaller than the spacing to other traces (*sn*).

13.6.4 Avoid Return Path Splits

The most obvious application of this rule is when a diff-pair crosses over a split in a power or ground plane.

Splits are cuts in the plane that create isolated regions ("power islands"). They are often used to divide a power plane so that multiple voltages can be located on the same plane. For instance, the ground plane is sometimes split to isolate an analog ground from a digital ground.

High-speed, loosely coupled differential signals are more likely to have difficulty crossing a split. In contrast, slower-speed differential signals, especially when they are tightly coupled and are high-impedance, are better able to tolerate splits. For these types of traces, the return current is mostly in the diff-pair traces and not in the ground plane, so a void in the ground plane may not significantly alter the transmission line characteristics.

A more subtle version of the same effect is illustrated in Figure 13.15.

Figure 13.15 shows a trace crossing a void in the ground plane caused by a large antipad. Antipads (first introduced in Chapter 5) are holes in the power or ground planes that allow a via to pass through without making connection. Antipads are sometimes made with large diameters to reduce the coupling capacitance from the via to the plane. This is especially common on boards having many layers (including back planes) because reducing the parasitic capacitance of each layer by even a small amount can significantly reduce the total parasitic capacitance.

The difficulty illustrated in Figure 13.15 is that the local ground plane has been removed in the region where the trace crosses the antipad. This increases the trace inductance and lowers its capacitance, raising the trace impedance in that region. A trace routed through a deep pin field may cross over many of these discontinuities, which can result in severe signal degradation.

13.6.5 Route on Same Layers

The two traces forming an edge coupled diff-pair should always be routed on the same layer. Otherwise, the delay, impedance, and noise coupling are very unlikely to be identical.

This rule is sometimes violated when the two capacitors connecting a serial transmitter or receiver—generally, a *serializer/deserializer* (SERDES)—to a diff-pair are placed on opposite sides of a circuit board. Often this technique allows the capacitors to be placed closer to the SERDES, but doing so makes the noise and impedance on the two traces unequal. It also usually results in different via lengths, making the overall length of the two traces unequal.

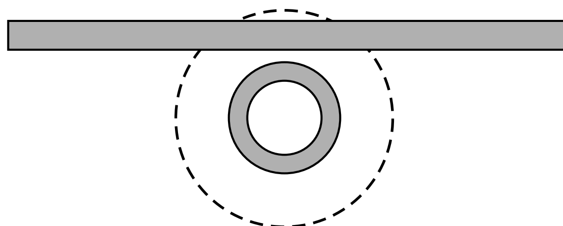


Figure 13.15 A trace crossing over a large antipad (dotted). The trace impedance changes in the region where it crosses the split.

13.6.6 Minimize Layer Hopping

Vias represent an impedance discontinuity unless their impedance is intentionally designed to match the trace impedance (described in Chapter 14). This gives rise to reflections and at high speeds can significantly impair the ability of a serial interconnect from operating properly.

13.6.7 Use Differential Vias

A pair of vias can be designed to have differential impedance equal to the differential impedance of the diff-pair traces. As described in Chapter 14, designing these structures requires a 3D field solver to model the effects of via size, antipad diameter, pad diameter, and laminate thickness on the differential impedance [29]. Such via structures greatly reduce (and may eliminate) the impedance discontinuity presented by vias not as carefully designed, thereby reducing signal reflections and allowing the signal to operate at very high speeds.

13.6.8 Match Overall Trace Lengths

Violations of this rule can be subtle and detecting them often requires a detailed understanding of the routed path.

For instance, right-angle multirow connectors have different length pins in the upper and lower rows. If these two rows are used to form a diff-pair, the length skew occurs on each of the connectors forming a mated pair, doubling the difference. For this reason routing diff-pairs through these connectors should be done with pins on the same row. This length mismatch can be present when routing either edge-coupled or broadside traces.

This rule can be unintentionally broken when routing broadside diff-pairs. Because they are routed on different layers, vias or interlayer connections made by connector pins moving signals to the board's outer surface naturally will be of different lengths [13]. When two connectors are used (as in a midplane or daughter card situation) this inevitable length skew can sometimes be corrected by swapping the layers on the two cards. If the two cards have the same stackup thickness the via (or pin) lengths become matched, eliminating the skew.

A third subtle way in which this rule is violated is shown in Figure 13.16 where a trace includes one or more bends as it changes direction.

Figure 13.16(a) shows how the bends cause a difference in trace length. In high-performance signaling back-to-back bends can be used to readjust the overall length.

13.6.9 Match Lengths on Each Layer

Besides ensuring that the total length of trace from the transmitter to receiver is the same for the two traces forming a diff-pair, it is important that the lengths are matched on each of the layers on which they are routed. For instance, a diff-pair may exit a BGA pin field on the surface, using microstrip, and then transition with a pair of vias to a stripline layer. The lengths of the two microstrip traces should be

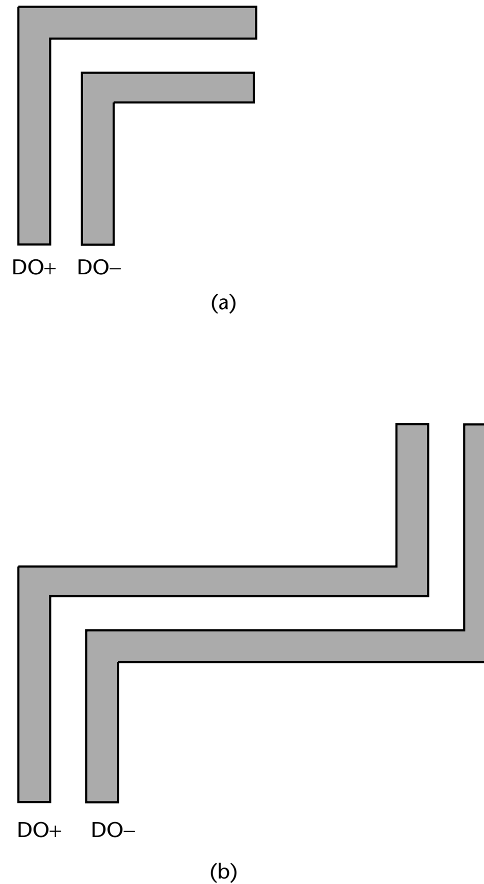


Figure 13.16 (a) Bends in diff-pairs cause *DO+* to be longer than *DO–*. (b) Back-to-back bends can readjust this skew.

matched, and the lengths of the two stripline traces should be matched. This ensures that the delay and impedance of each segment is the same.

13.6.10 Ensure That Noise Is Equally Coupled to Both Traces

Subtle mechanisms can cause unequal noise coupling on the diff-pair traces. An example showing the effects on a broadside-coupled pair placed between two return planes appears in Figure 13.17.

Noise present on plane $L1$ will preferentially couple onto the $L2$ trace, and noise on $L4$ will couple onto $L3$. The traces forming the diff-pair are subjected to different noise unless the planes have a low impedance between them and have the same noise voltages. If both $L1$ and $L4$ are ground planes, placing numerous vias between the planes (plane stitching) can lower the impedance between the planes. Alternatively, if one is a power plane, decoupling capacitors can be used to provide an AC connection between the planes. In this way the same noise appears on both planes, making the noise coupled from $L1$ to $L2$ identical to the noise coupled from $L4$ to $L3$.

Another effect that can occur with edge coupled diff-pairs (either stripline or microstrip, but not with broadside pairs) is shown in Figure 13.18.

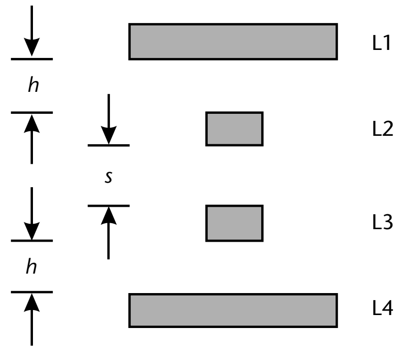


Figure 13.17 Side view of broadside coupled diff-pair formed on routing layers *L*2 and *L*3. Noise on the *L*1 plane will preferentially couple onto *L*2, and noise on plane *L*4 will more readily couple onto *L*3.

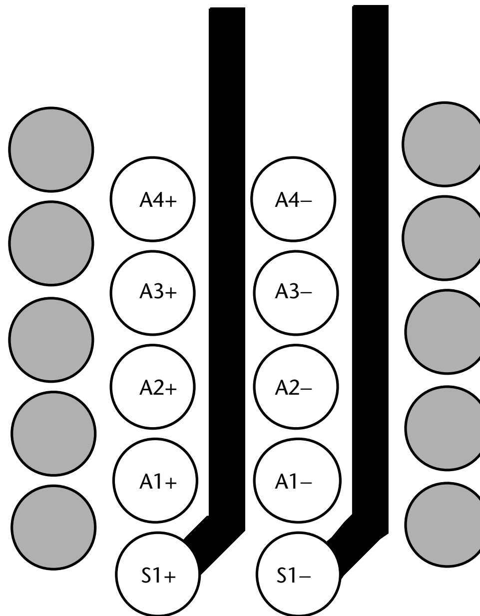


Figure 13.18 Top view of a diff-pair routing through a pin field. The noise environment for *S*1+ is not the same as for *S*1–. (*From:* [13]. Used with permission.)

In Figure 13.18 an edge-coupled diff-pair is routed through a pin field, such as a dense connector field or a BGA. The circles represent vias to which the connector is inserted, although this also occurs with vias connecting to surface mount lands. Ground connections are represented by dark circles, and the remaining circles are differential signals.

The signal connected to the *S*1+ pin experiences a very different noise environment than its complement (*S*1–). Signal *S*1+ is routed between rows of differentially switching signals, making the net transfer of noise zero. However, the *S*1– signal is routed between ground and the switching signals. In this case the switching noise does not cancel and noise from the *A*1– through *A*4– pins accumulates on the *S*1– signal. Since they are on top of one another, broadside pairs are immune from

this as they will experience the same pin-to-trace coupling no matter how the are routed.

13.7 Main Points

- Differential signaling uses a differential amplifier to reject noise present on both its inputs, making it important to couple noise equally to both inputs.
- The differential impedance of a differential transmission line is twice the odd-mode impedance.
- A single resistor placed across the diff-pair is a proper termination for diff-pairs with negligible even-mode energy.
- A three-resistor termination scheme is used when the even-mode component is too large to be ignored.
- Even-mode behavior is caused by poor routing practices or poor differential driver characteristics.
- Diff-pairs can be formed as edge or broadside coupled, and as either loosely or tightly coupled.
- Loosely coupled diff-pairs are easy to manufacture but tightly coupled traces may be electrically necessary.
- Special test equipment is necessary to analyze differential signals in detail.
- Eye diagrams clearly show the effects of noise and interference present on the channel.

Problems

Answers to these problems are available on the Artech House Web site at http://www.artechhouse.com/static/reslib/thierauf/thierauf1.html.

13.1 A field solver produced the following output file for a diff-pair.

```
L MATRIX [H/INCH]
1.052E-8 2.040E-9
2.040E-9 1.052E-8
C MATRIX [F/INCH]
2.835E-12 –5.490E-13
–5.490E-13 2.835E-12
```

Find the values for termination resistors *R*1 and *R*2 in Figure 13.11.

13.2 A diff-pair is described by the following SPICE transmission line model:

```
+ Lo = 8.92E-09
+ 1.32E-09 8.92E-09
+ Co = 2.50E-12
+ -2.34E-13 2.50E-12
```

What is the differential impedance?

13.3 The traces described in Problem 13.2 show a difference in the odd- and even-mode impedances of more than 25%. Is it necessary to use the two resistor termination shown in Figure 13.11, or will a single resistor across the diff-pair [as shown in Figure 13.10(b)] be sufficient?

13.4 It has been decided that a single termination resistor, placed across the diff-pair (as shown in Figure 13.10) will be used to terminate the diff-pair described in Problem 13.2. What should its value be?

13.5 It has been decided that resistors to ground from each trace in the diff-pair will be used to terminate the traces described in Problem 13.2. What values should these resistors have?

13.6 A loosely coupled diff-pair is described by the following SPICE transmission line model:

```
+ Lo = 8.47E-09
+ 2.30E-10 8.47E-09
+ Co = 3.08E-12
+ -2.43E-14 3.08E-12
```

What are the proper termination values if the termination appearing in Figure 13.11 is used?

13.7 A clock oscillator has a differential output that swings from 0.8V to 1.8V peak-to-peak (p/p). It is necessary to connect this to an ASIC differential receiver that requires the signal to have a 1V p/p swing centered on 1.65V. How should the diff-pairs connecting the oscillator to the ASIC be terminated?

References

[1] "High-Speed Transceiver Logic (HSTL) A 1.5V Output Buffer Supply Voltage Based Interface Standard for Digital Integrated Circuits," EIA/JESD8-6, Electronics Industries Association, August 1995.

[2] "JDEC STANDARD: Stub Series Terminated Logic for 1.8 V (SSTL_18)," JESD8-15A, JEDEC Solid State Technology Association, September 2003.

[3] National Semiconductor, *LVDS Owner's Manual,* 4th ed., Santa Clara, CA: National Semiconductor Corporation, 2008.

[4] Infiniband™ Trade Association, "Infiniband ™ Architecture Specification, Vol. 2, Release 1.1," November 2002.

[5] Fuller, S., *RapidIO: The Embedded System Interconnect*, New York: John Wiley & Sons, 2005.

[6] Hug, S., and J. Goldie, "An Overview of LVDS Technology," National Semiconductor Applications Note 971, National Semiconductor Corporation, 1998.

[7] Tektronix Incorporated, Beverton, OR.

[8] Lauterbach, M., "Getting More Out of Eye Diagrams," *IEEE Spectrum*, Vol. 34, No. 3, March 1997.

[9] Bateman, A., *Digital Communications*, Reading, MA: Addison-Wesley Longman, 1998.

[10] Dally, W. J., and J. W. Poulton, *Digital Systems Engineering*, Cambridge, U.K.: Cambridge University Press, 1998.

[11] Couch, L. W., II, *Digital and Analog Communication Systems*, 5th ed., Upper Saddle River, NJ: Prentice-Hall, 1997.

[12] Bissell, C. C., and D. A. Chapman, *Digital Signal Transmission*, Cambridge, U.K.: Cambridge University Press, 1996.

[13] Thierauf, S. C., *High-Speed Circuit Board Signal Integrity*, Norwood, MA: Artech House, 2004.

[14] Horowitz, M., et al., "High-Speed Electrical Signaling: Overview and Limitations," *IEEE Micro*, Vol. 18, No. 1, January/February 1998, pp. 12–24.

[15] Proakis, J. G., *Digital Communications*, New York: McGraw-Hill, 2000.

[17] Terman, F. E., *Radio Engineers Handbook*, New York: McGraw-Hill, 1943.

[16] Haykin, S., and B. Van Veen, *Signals and Systems*, 2nd ed., New York: John Wiley & Sons, 2002.

[18] Dally, W. J., Poulton, J., "Transmitter Equalization for 4-Gbps Signaling," *IEEE Micro*, Vol. 17, No. 1, January/February 1997.

[19] Turudic, A., et al., "Pre-Emphasis and Equalization Parameter Optimization with Fast, Worst-Case/Multibillion-Bit Verification," *Design Con 2007*, Santa Clara, CA, January 29–February 1, 2007.

[20] Chandrakasan, A., et al., *Design of High-Performance Microprocessor Circuits*, New York: IEEE Press, 2001.

[21] Young, B., *Digital Signal Integrity*, Upper Saddle River, NJ: Prentice-Hall, 2001.

[22] Simbeor Applications Note #2009_1, "Practical Notes on Mixed-Mode Transformations in Differential Interconnects," Simberian Corporation, Seattle, WA, January 2009.

[23] Marx, K. D., "Propagation Modes, Equivalent Circuits, and Characteristic Terminations for Multiconductor Transmission Lines with Inhomogeneous Dielectrics," *IEEE Transactions on Microwave Theory and Techniques*, Vol. MTT-21, No. 7, July 1973, pp. 450–457.

[24] Paul, C. R., *Introduction to Electromagnetic Compatibility*, New York: John Wiley & Sons, 1992.

[25] Polar Si9000 Field Solver, Polar Instruments, Inc. Beaverton, OR.

[26] Djordjevic, A. R., et al., *LINPAR for Windows*, Norwood, MA: Artech House, 1999.

[27] Ott, H. W., *Noise Reduction Techniques in Electronic Systems*, 2nd ed., New York: John Wiley & Sons, 1988.

[28] Rainal, A. J., "Transmission Properties of Balanced Interconnections," *IEEE Transactions on Components, Hybrids and Manufacturing Technology*, Vol. 16, No. 1, February 1993, pp. 137–145.

[29] Simbeor Application Note #2007_07, "Modeling Differential Via-Holes Without Stitching Vias," Simberian Corporation, Seattle, WA, October 2007.

CHAPTER 14

Trace and Via Artwork Considerations for Signal Integrity

14.1 Introduction

Practical considerations can cause the actual circuit board artwork to differ from the design that was simulated. Some of these things (such as routing close to the board edge or the way in which signals are length matched with serpentine traces) are visible in the artwork, but other choices (such as the sizes of the vias and anti-pads) are not as immediately obvious. The signal integrity engineer who is aware of these practical concerns can deliberately look for these structures as part of the layout artwork checking process and update the simulation model as necessary.

14.2 Mitered Corners

Nearly all traces will have one or more bends as they change direction or jog around obstacles. As shown in Figure 14.1, a sharp right angle bend increases the trace area in the vicinity of the bend. This adds a small amount of excess capacitance [1, 2] and excess inductance [3, 4] to the trace.

The excess inductance and capacitance are very small, but alter the impedance at the bend, which can cause reflections at high frequencies. Usually the excess capacitance is the dominate effect and is typically in the hundreds of femto farad range [5].

Mitering (or chamfering) is commonly used to reduce the excess capacitance and is shown in Figure 14.2. Mitering is not necessary until the signals have rise times in the subnanosecond range [5], but it can reduce reflections for those traces carrying frequencies (including harmonics) greater than several gigahertz [6]. Notches [6, 7] are sometimes used at microwave frequencies, but these exotic approaches are generally not seen on digital circuit boards.

The theoretical optimum miter occurs when the excess capacitance has been reduced to where the capacitive and inductive effects are equal [8], but in practice the actual angle and length of the miter (m) are not usually critical. As shown in Figure 14.2, an angle of 45° is commonly used, and the miter length can range from

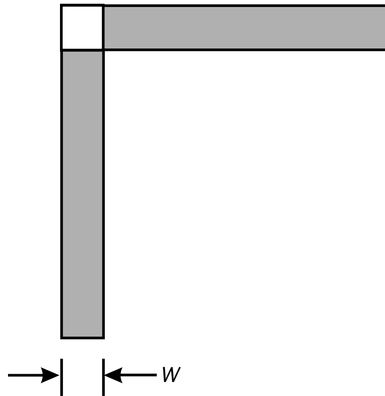
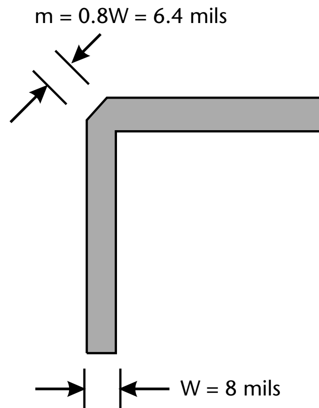


Figure 14.1 A trace of width *W* making a right-angled turn. The extra trace area (shown in white) adds capacitance and inductance, which can cause reflections at high frequency.



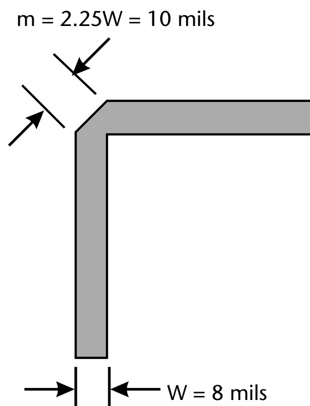


Figure 14.2 Mitering the trace outside edge at 45° to reduce excess capacitance. Optimal miter length (*m*) is between 0.8*W* and 1.25*W*. Dimensions are shown for an 8-mil (0.21-mm)-wide trace.

0.8*W* [9] to 1.25*W* [6]. For instance, *m* for an 8-mil (0.21-mm)-wide trace would be between 6.4 and 10 mils (0.17 mm and 0.25 mm) at an angle of 45°. In practice

the layout designer often selects these values such that the manufacturer's minimum trace width rules are not violated, especially on narrow traces.

14.3 Routing Near the Board Edge

Some of the electric and magnetic field lines escape traces which are routed close to the circuit board edge, increasing the trace inductance and decreasing its capacitance.

These changes reduce the delay but increase the trace impedance and conductor losses. The effects are most pronounced with wide, high-impedance traces because they are further from the return planes than narrow, low-impedance traces. Losses increase because the return current cannot spread as far, increasing return path resistance.

A top view of two 10-mil-wide, 65Ω stripline traces running parallel to the edge of the board is shown in Figure 14.3(a). One trace is 5 mils (0.13 mm) from

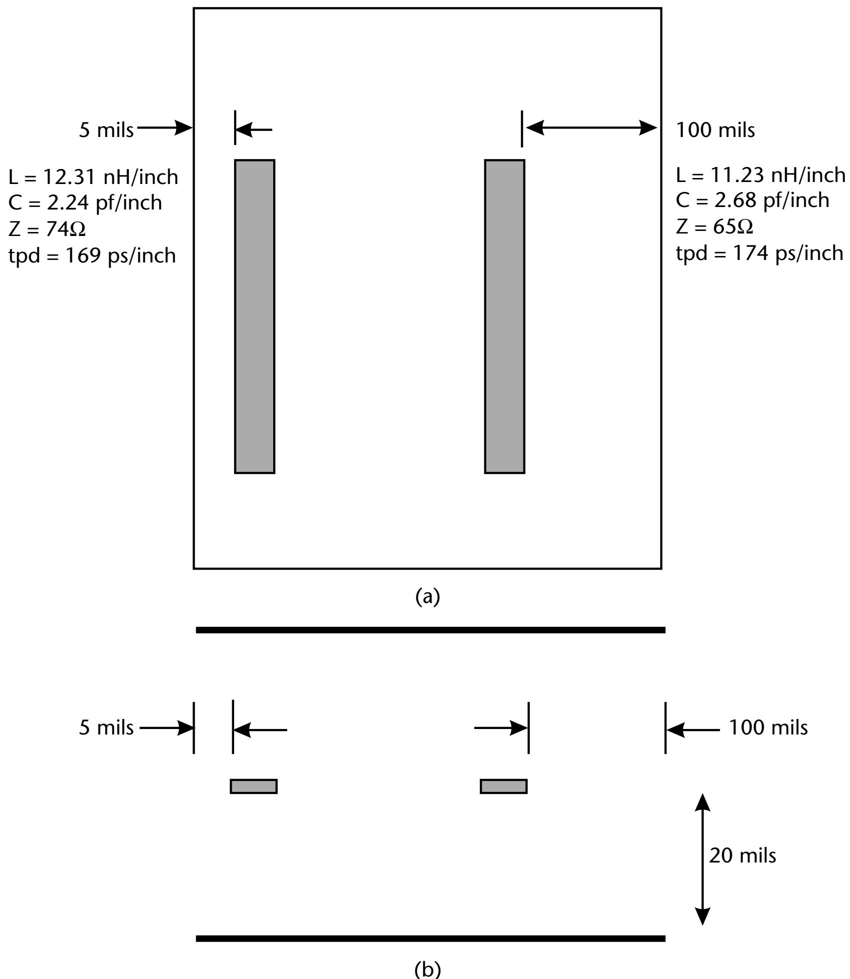


Figure 14.3 10-mil (0.25-mm)-wide stripline routed 5 mils (0.13 mm) from the board edge has different electrical properties than identical stripline routed much further in (a). (b) Side view shows the traces are 20 mils (0.5 mm) from either plane.

the edge, while the other, identical trace is 100 mils (2.5 mm) from the other edge. A side view appears in Figure 14.3(b).

The inductance of the stripline routed close to the board edge is about 10% higher and the capacitance is over 15% lower than the other stripline. This causes the impedance to grow by nearly 14% and the delay to fall by nearly 3%. Although not shown, the impedance and delay of a 4-mil-wide, 50Ω stripline (which is significantly closer to the return plane then the traces in Figure 14.3) only change by about 2% under these same conditions.

Microstrip behaves in the same general way as stripline, with the biggest effects seen on wide, high-impedance traces. Often 2D field solvers tacitly assume the return plane is very much wider than the trace. Figure 14.3 shows that circuit models created with this assumption will not represent the characteristics of the traces routed close to the edge. This can be a significant problem when signaling at high speeds, particularly if the signaling is occurring with differential pairs.

14.4 Serpentine Traces

Often high-performance integrated circuits include delay locked loops (DLLs) or similar circuits to digitally adjust the timing between signals [11, 12]. When these circuits are not present or when it is necessary match the arrival of individual signals within a bus, a long trace can be used to add a known delay based on the transmission lines time of flight (*tpd*). Because tpd is typically hundreds of picoseconds per inch, adjusting timing usually requires several inches of trace.

These delay lines are created by folding the trace back and forth to create a serpentine trace. They are also referred to as meander or trombone traces. A five-segment serpentine is diagrammed in Figure 14.4.

Each of the five segments has a delay equal to *tpd*, making the expected total delay 5 *tpd* plus the delays of the short vertical connecting straps. However, the actual delay will be less than this unless care is taken with the layout.

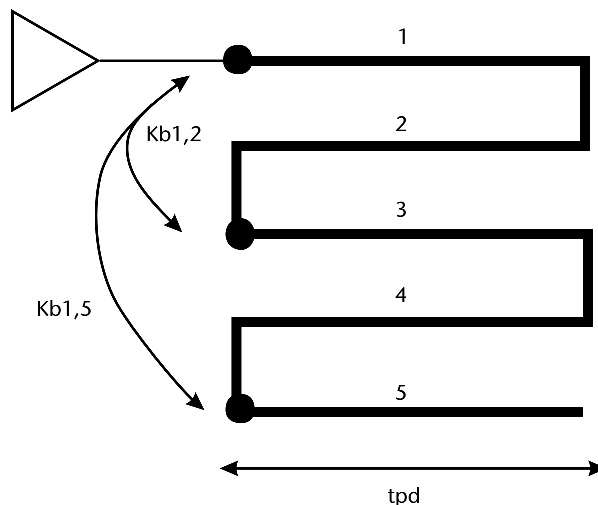


Figure 14.4 Five-segment serpentine delay line. Each long segment has a delay of *tpd*.

Near-end crosstalk (discussed in Chapter 10) between traces 1 and 5 (*Kb*1,5) causes trace 1 to induce a small voltage on trace 5 and a larger voltage on trace 2 from (*Kb*1,2). This coupling is shown as occurring on trace 3 because here the vertical connecting strap is assumed to be short enough to be ignored. The *Kb*1,5 signal arrives at the serpentine output *tpd* seconds after traveling the length of trace 5, and the stronger *Kb*1,2 signal arrives 2 *tpd* seconds later, after traveling the entire length of traces 3 and 4. This voltage adds to the crosstalk already present, causing the output to increase. The output reaches the full value at 5 *tpd* with the arrival of the launched signal. A ladder-like waveform [13] has been created at the serpentine output, as illustrated in Figure 14.5(a). The expected waveform is shown in Figure 14.5(b).

Laddering can occur with either stripline or microstrip because each are susceptible to near end crosstalk.

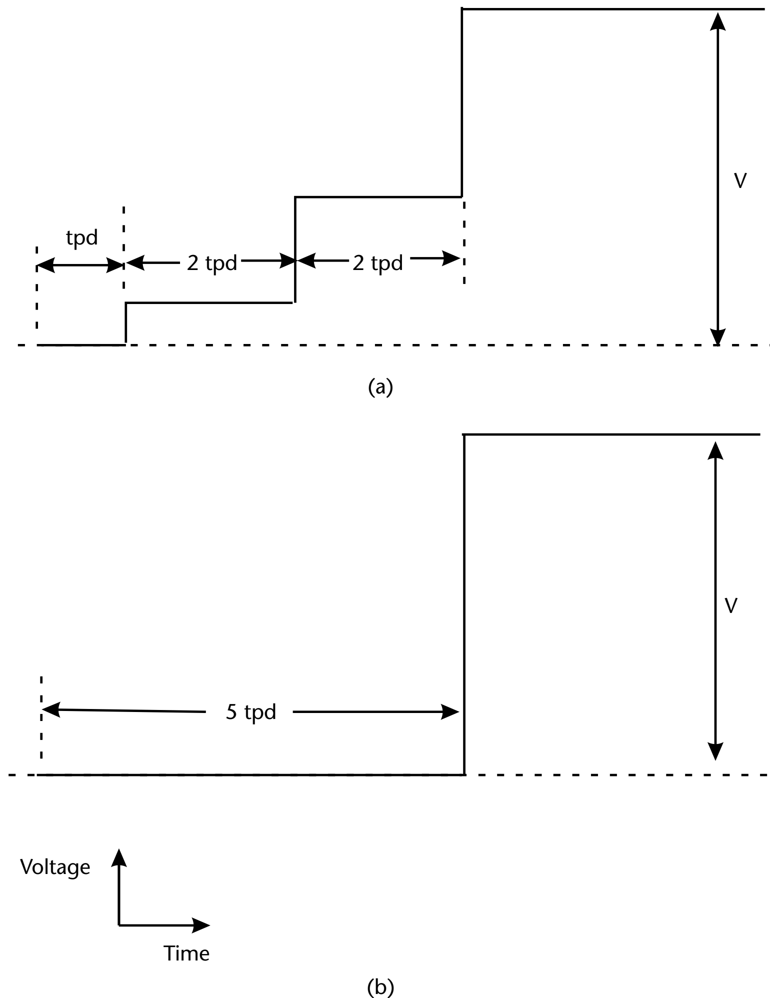


Figure 14.5 (a) Ladder-like waveform created by crosstalk appearing at the output of a five-segment serpentine delay line contrasts with (b) the expected waveform.

14.4.1 How Are Serpentine Delay Lines Modeled?

Multiconductor lossy circuit models of the serpentine can be created with a 2D field solver. However, this technique is only an approximation because it does not account for the coupling occurring from a horizontal segment to a vertical strap. A 3D solver that includes the entire serpentine structure is necessary to account for these effects.

Once created, the models are used in a circuit simulator with models of the actual driver and receiver to determine if a given spacing or number of segments will create laddering severe enough to falsely trigger the receiver. A multiconductor transmission line model represents the long horizontal segments, and the short vertical straps are modeled with direct connections. If they are long enough, transmission lines can be used for these straps. This technique is illustrated in Figure 14.6.

14.4.2 Effects of Corners

The serpentine corners (bends) have less delay than an equivalent amount of straight trace and so are said to speed up the signal. Mitering the corners can help correct this timing uncertainty [14].

Excess capacitance causes the impedance to be lower in the corners than the segments. This can be significant if there are many corners or when signaling at high speeds.

Corner effects and the coupling from the vertical segments to the horizontal straps are not properly accounted for in circuit models created by a 2D field solver. However, a 3D solver will capture them [15].

14.4.3 How Long Should the Segments Be?

Because the maximum coupled voltage grows with the number of segments in the serpentine [16], when laying out a serpentine, it is best to use a fewer number of long segments instead of a greater number of short ones [15]. Fewer segments also mean fewer corners and less uncertainty in the timing and impedance.

For these reasons the segments should be long (typically greater than the signal rise time) and few in number. Also, because crosstalk increases as the traces

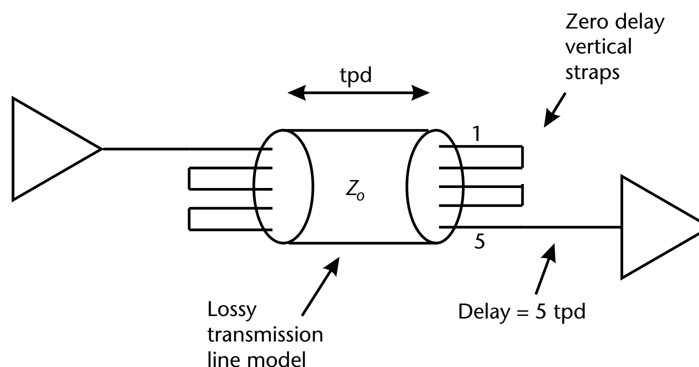


Figure 14.6 A multiconductor transmission line modeling a five-segment serpentine delay line. The vertical straps can be wires as shown or transmission lines.

are tightly packed together, laddering can be reduced by increasing the separation between segments.

14.4.4 Guard Traces

Placing a properly grounded guard trace between segments in the serpentine can reduce or eliminate laddering since this reduces crosstalk between the segments. Chapter 10 describes how to properly ground a guard trace. Adding them slightly reduces the serpentine delay [13, 17], but makes the delay more uniform across a broad range of frequencies [17].

14.5 Vias

A 3D rendering of a through hole via extending through a plane in a circuit board is shown in Figure 14.7.

On each end the via, a circular ring of metal ("the pad") connects to the signal trace, and a hole (the antipad) in the plane permits the conducting barrel to pass through without making a connection.

14.5.1 Via Circuit Model

Figure 14.8 shows how capacitance is present from the via barrel and the pads to the plane. All of these capacitors are summed into a single capacitance, *Ct*. An inductor (*Lt*) models the inductance of the barrel and its return path.

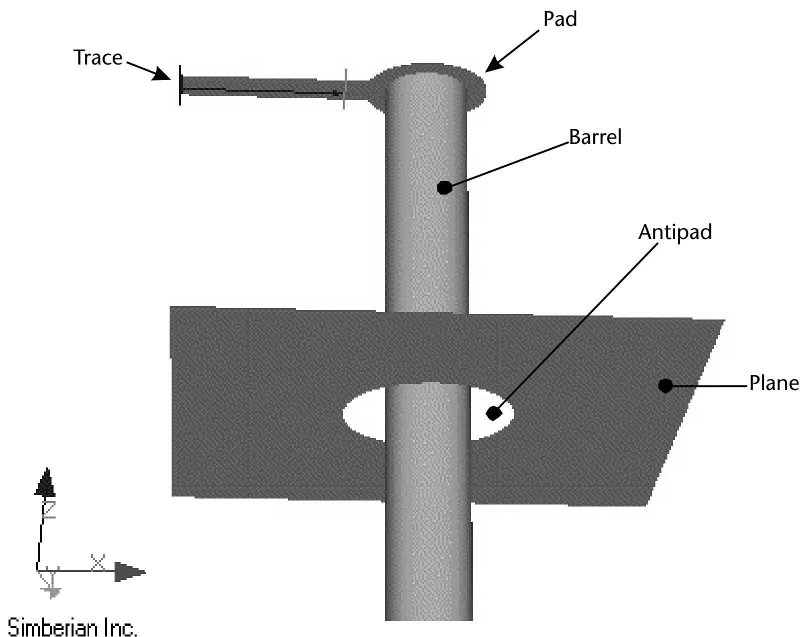


Figure 14.7 A through hole via passing through a plane. The antipad prevents the barrel from connecting the plane. A pad connects the signal trace to the barrel. (Figure generated by the Simbeor field solver [18].)

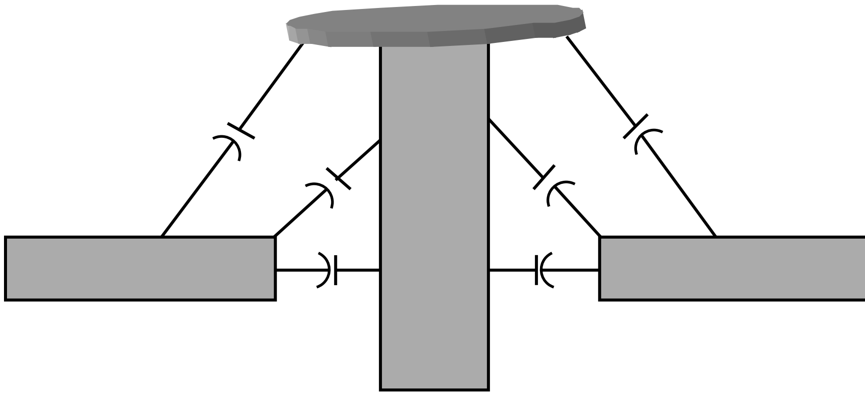


Figure 14.8 Side view of a through hole via passing through a solid plane showing the parasitic capacitance. Not shown is the parasitic inductance, which is also present.

A simple PI circuit, shown in Figure 14.9, can be used to model the impedance and delay through the via when the frequency content is not too high.

The PI is formed by using the total inductance (*Lt*) and applying half of the total capacitance (*Ct*) to each leg. This model is only valid when the voltage drop across *Lt* and the current through the capacitors is small compared to the current flowing through the via [19]. Typically this means the model is most accurate when the signals have rise times of about 100 pS or longer [20]. At higher frequencies a chain of these models, or a transmission line, may be used. For best results at still higher frequencies, an S-parameter-type model can be used.

14.5.2 What Factors Determine Via Capacitance?

The total capacitance is determined by the diameter of the barrel and pads, by the via length, and by the antipad diameter (and thickness and number of planes that the barrel passes through) [21, 22].

The pad capacitance increases with the area of the pad, while the via barrel capacitance increases as the square root of the length of the via [21]. These things lower the via's impedance. However, increasing the antipad size increases the distance from the barrel to the ground plane, which is equivalent to increasing the plate separation in a parallel plate capacitor. This reduces the capacitance of the barrel and tends to raise the via impedance.

Depending on the board thickness and the number of planes, a 10-mil (0.25-mm) diameter through-hole via with a 30 mil (0.75 mm) antipad and 22-mil

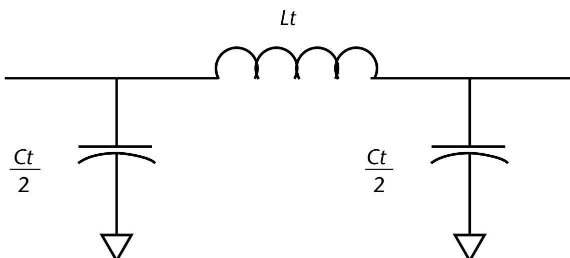


Figure 14.9 This PI circuit is a good low-frequency circuit model of a via.

(0.55-mm) pad has a capacitance from between roughly 500 fF when the via is 60 mils (1.5 mm) long passing through five planes to 1.1 pF when it is 180 mils (4.5 mm) long [23].

14.5.3 What Factors Determine Via Inductance?

The return current will always find a path having the lowest impedance, and this path determines the via's total loop inductance (*Lt*).

The return current will preferentially flow through a nearby return via since the small loop area has a low inductance. However, in the absence of a nearby return via, the return current will seek alternate paths to lower the impedance. This extended path is not accounted for in simple hand formulas that calculate via inductance.

Additionally, as the impedance of the via path grows, to minimize the overall return path impedance increasingly more of the return current flows through the capacitance formed between the power and ground planes. This analysis is best done by using a 3D field solver that includes the layer to layer capacitance and the actual location of any return path vias or decoupling capacitors.

A 10-mil (0.25-mm) diameter through-hole via, with a 30-mil (0.75-mm) antipad and 22-mil (0.55-mm) pad and no local vias to act as return paths, has an inductance varying from roughly 800 pH when it passes through five planes and is 60-mils (1.5 mm) long to 3 nH when it is 180-mil (4.5 mm) long [23].

14.5.4 What Are the Effects of Different Antipad Diameters?

A capacitor is formed each time the via passes through a plane. This is why a long via with small antipads passing through many plane layers has a higher capacitance and lower impedance than a short via with large antipads passing through a few planes.

The effect of antipad size is illustrated in the time domain reflectometry (TDR) plot appearing in Figure 14.10. The TDR shows the impedance in the vertical axis against the distance in the horizontal direction. The circuit board used for this measurement has 20-mil (0.5-mm) diameter vias with 40-mil (1-mm) pads. The board is 18 layers, 10 of which are power or ground planes.

From Figure 14.10 we see that in this one example that increasing the size of the antipad from 45 mils (1.13 mm) to 60 mils (1.5 mm) increases the vias' impedance from 32Ω to 43Ω.

14.5.5 What Are Nonfunctional Pads?

Especially in thick boards, manufacturers generally place pads on a via at each routing layer, even if a trace is not connecting to the via on that layer.

These nonfunctional pads are used to help relive thermal stresses placed on the via during assembly operations [25, 26]. The stresses can cause via cracking when the board is heated during soldering. Vias that are four or more times longer than their diameter are most susceptible to cracking, especially when lead free soldering is employed because of the high temperatures necessary for that process [25].

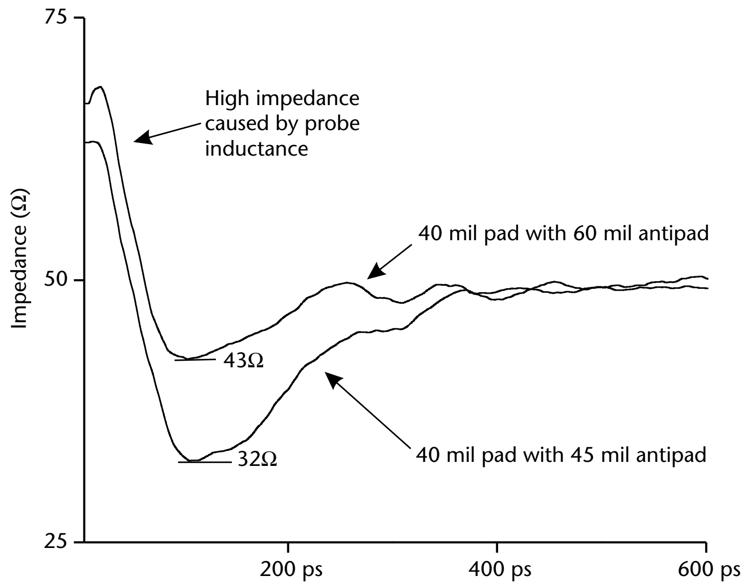


Figure 14.10 TDR showing that a via with large antipad has higher impedance than one with a small antipad. (*From:* [24]. Used with permission.)

The added capacitance of these pads lower the via impedance, which can cause significant reflections when signaling at high speeds. The circuit board manufacturer can be requested to remove nonfunctional pads from those nets carrying high data rate signals, but they should confirm that the board reliability will not be unduly impacted. Removal is done by the circuit board manufacturer and not in the artwork by the layout designer.

14.5.6 Can the Via Impedance Be Adjusted?

The impedance of a signal carrying via can be controlled by the careful placement of return path vias near the signal carrying via [27, 28]. This technique is of most use when signaling at high data rates and requires the use of a 3D field solver to adjust the placement of the vias and to select the pad and antipad sizes.

14.5.7 What Are Differential Vias?

A typical via presents a problem when differential signals change layers because the via pair will not have the same differential impedance as the traces.

To correct this, a pair of vias can be designed for a specific differential impedance [29, 30] by adjusting their spacing and dimensions and occasionally even the antipad shape. This is best done with a 3D field solver.

An example of a differential via appearing in a 3D field solver is shown in Figure 14.11.

Coupling between the vias occurs in the region between the planes, so no other structures (such as vias, microvias, traces, or nonplated holes) can be placed between the via pair. It is also important that other vias (including nonplated holes

Figure 14.11 A pair of vias forming a differential pair passing through two planes. Ground stitching vias are visible on either side. The rectangular tabs coming from the vias tops are microstrip traces. (Figure generated by the Simbeor 3D solver [18].)

used by metal screws to secure the circuit board) not be placed close to either of the vias forming the pair.

14.6 Main Points

- Mitered corners can be used on traces to remove the excess capacitance and inductance caused by bends.
- The degree of chamfering usually is not critical.
- Chamfering is important when signaling at high-speeds.
- Routing traces near the board's edge changes its electrical characteristics and can increase EMI.
- Crosstalk must be minimized when using serpentine traces to adjust timing.
- It is better to use a few long serpentine elements than many short ones.
- The via and antipad sizes strongly determine the vias electrical characteristics.
- Nonfunctional pads help secure the vias in a stackup, but are not visible in the artwork and can be removed on high-speed nets.

References

[1] Silvester, P., and P. Benedek, "Microstrip Discontinuity Capacitances for Right-Angle Bends, T Junctions, and Crossings," *IEEE Transactions on Microwave Theory and Techniques*, Vol. MTT-21, No. 5, May 1973, pp. 341–346.

[2] Silvester, P., and P. Benedek, "Corrections to Microstrip Discontinuity Capacitances for Right-Angle Bends, T Junctions, and Crossings," *IEEE Transactions on Microwave Theory and Techniques*, Vol. MTT-23, No. 5, May 1975, p. 456.

[3] Thompson, A. F., and A. Gopinath, "Calculation of Microstrip Discontinuity Inductances," *IEEE Transactions on Microwave Theory and Techniques*, Vol. MTT-23, No. 8, May 1975, pp. 648–655.

[4] Harms, P. H., and R. Mittra, "Equivalent Circuits for Multiconductor Microstrip Bend Discontinuities," *IEEE Transactions on Microwave Theory and Techniques*, Vol. MTT-41, No. 1, January 1993, pp. 62–69.

[5] Rainal, A. J., "Reflections from Bends in a Printed Conductor," *IEEE Transactions on Components, Hybrids and Manufacturing Technology*, Vol. 13, No. 2, June 1990, pp. 407–413.

[6] Simbeor Application Note #2008_05, "Minimal-Reflection Bends in Micro-Strip Lines," Simberian Corporation, Seattle, WA, September 2008.

[7] Widell, B. C., *Transmission Line Design Handbook*, Norwood, MA: Artech House, 1991.

[8] Douville, R. J. P., and D. S. James, "Experimental Study of Symmetric Microstrip Bends and Their Compensation," *IEEE Transactions on Microwave Theory and Techniques*, Vol. MTT-26, No. 3, March 1978, pp. 175–182.

[9] Chadha, R., and K. C. Gupta, "Compensation of Discontinuities in Planar Transmission Lines," *IEEE Transactions on Microwave Theory and Techniques*, Vol. MTT-30, No. 12, December 1982, pp. 2151–2156.

[10] Archambeault, B., "Effects of Routing High-Speed Traces Close to the PCB Edge," *Printed Circuit Design & Manufacture*, January 2008.

[11] Dally, W. J., and J. W. Poulton, *Digital Systems Engineering*, Cambridge, U.K.: Cambridge University Press, 1998.

[12] Chandrakasan, A., et al., *Design of High-Performance Microprocessor Circuits*, New York: IEEE Press, 2001.

[13] Wu, R., and F. Chao, "Laddering Wave in Serpentine Delay Line," *IEEE Transactions on Components, Packaging, and Manufacturing Technology*, Part B, Vol. 18, No. 4, November 1995, pp. 644–650.

[14] Rubin, B. J., and B. Singh, "Study of Meander Line Delay in Circuit Boards," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 48, No. 9, September 2000, pp. 1452–1460.

[15] Orhanovic, N., et al., "Characterization of Microstrip Meanders in PCB Interconnects," *Proceedings 50th IEEE Electronic Components and Technology Conference*, Las Vegas, NV, May 21–24, 2000, pp. 508–512.

[16] Shiue, G., et al., "Improvements of Time-Domain Transmission Waveform in Serpentine Delay Line with Guard Traces," *IEEE International Symposium on Electromagnetic Compatibility, EMC 2007*, Honolulu, HI, July 9–13, 2007, pp. 1–5.

[17] Nara, S., and K. Koshiji, "Study on Delay Time Characteristics of Multilayered Hyper-Shielded Meander Line," *IEEE International Symposium on Electromagnetic Compatibility, EMC 2006*, Vol. 3, Portland, OR, August 14–18, 2006, pp. 760–763.

[18] Simbeor Electromagnetic Simulation Environment, Simberian Corporation, Seattle, WA.

[19] Wang., T., et al., "Quasi-Static Analysis of a Microstrip Via Through a Hole in a Ground Plane," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 36, No. 6, June 1988, pp. 1008–1013.

[20] LaMeres, B., and T. S. Kalkur, "Time Domain Analysis of a Printed Circuit Board Via," *Microwave Journal*, November 1, 2000.

[21] Kok, P., and D. D. Zutter, "Capacitance If a Circular Symmetric Model of a Via Hole Including Finite Ground Plane Thickness," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 39, No. 7, July 1991, pp. 1229–1234.

[22] Johnson, H. H., and M. Graham, *High-Speed Digital Design*, Upper Saddle River, NJ: Prentice-Hall, 1993.

[23] Buck, T. J., "PWB Design and Manufacturing Considerations for High Speed Digital Interconnection, "Applications Material, Dynamic Details Incorporated, undated.

[24] Thierauf, S. C., *High-Speed Circuit Board Signal Integrity*, Norwood, MA: Artech House, 2004.

[25] Engelmaier, W., "Non-Functional Lands: Keep Them Vs. Remove Them," Global SMT & Packaging, June/July 2006, pp. 48–50.

[26] Birch, B., "Discussion on Non Functional Pad Removal/Backdrilling and PCB Reliability," PWB Interconnect Solutions Inc., Applications note, undated.

[27] Neu, T., "Designing Controlled Impedance Vias," *EDN*, October 2, 2003, pp. 67–72.

[28] Simbeor Application Note #2009_05, "Designing Localizable Minimal-Reflection Via-Holes for Multi-Gigabit Interconnects," Simberian Corporation, Seattle, WA, April, 2009.

[29] Simbeor Application Note #2007_08, "Design of Optimal Differential Via-Holes For 6-Plane Board," Simberian Corporation, Seattle, WA, October 2007.

[30] Wang, Y., and W. Hong, "Design of Differential Vias in High Speed Digital Circuit," *Microwave Conference Proceedings, APMC 2005 Asia-Pacific Conference Proceedings*, Vol. 5, December 4–7, 2005, p. 4.

CHAPTER 15

Identifying Common Signal Integrity Problems

15.1 Introduction

Previous chapters have covered the circuit and field theory necessary for understanding the ideas fundamental to signal integrity. This chapter and Chapters 16 and 17 focus on identifying, solving, and calculating results for many common signal integrity problems.

A select group of questions that a project manager or group leader should consider while reviewing a signal integrity analysis is presented in this chapter. The material can also be used when developing interview questions for signal integrity engineer candidates. The previous chapters can be referred to for detailed explanations.

The novice signal integrity engineers will find this chapter helpful as they review their own work or as they help prepare a project plan.

This chapter cannot cover all situations or all possible solutions. Instead, it is a tool to help think about the kinds of problems and hidden assumptions that can become buried in a complex signal integrity task. The term ASIC is used here to mean any integrated circuit that is connected to a microstrip or a stripline trace.

15.2 Questions to Ask When Reviewing PCB Stackup

- Is the board thickness impacting the via size?
    - The minimum allowable via diameter is determined by the overall board thickness and the ability of individual fabrication shops.
    - Vias that are too large cannot be placed in the pin field of small geometry packages (such as dense BGA pin fields). This can require the use of higher-end processing such as blind, buried, or micro-vias.
- Does the design require thin dielectrics?
    - If thin core dielectrics are used to increase the capacitance between power and ground planes, has a careful decoupling capacitor study been per-

formed showing this technique is cost effective compared to discrete capacitors?
    - Are the cores so thin that the number of shops capable of fabricating the board is limited and thus jeopardizing second source opportunities?
- Is an advanced laminate system being specified?
    - If so, do the signal integrity simulations show an improvement in overall board thickness, signal loss, crosstalk, or signal routing density significant enough to warrant the added cost? If not, why isn't FR4 (or an enhanced FR4 system) used?
    - Can the chosen laminate be used by multiple fabrication shops to manufacture the circuit board? If not, do the electrical benefits outweigh the risk of limited second sourcing?
- Does the stackup include many power planes?
    - If so, can power planes be removed from the stackup (reducing cost and overall thickness) and can components be rearranged to allow one or more planes to be split and shared between more than one voltage?
- Are the most critical signals segregated to a specific signaling layer (or layers)?
    - Often routing density and stackup improvements can be had by routing the majority of critical (often impedance controlled) signals on specially designed layers. This also allows the signal return paths to be carefully designed.
- Are dual stripline layers used to maximum benefit?
    - This can be an effective way of adding routing layers with only slight increase to the stackup thickness.
    - Coupling between layers is minimized when signals on alternate layers are run at right angles to one another (for instance, north-south on one layer and east-west on the adjacent layer).
    - Careful signal integrity analysis is required to determine the amount of crosstalk a bus will experience.
- Are the proper voltage planes used to form stripline and microstrip?
    - A voltage plane unrelated to the signaling voltage must be carefully analyzed and modeled. Otherwise, the actual circuit will experience noise not predicted by simulation.
- Are mounting holes located in a critical routing area?
    - If so, verify that signaling has not been impaired by routing too close to the hole or to a metallic screw.
- Are sharp rise time or critical signals routed close the board edge?
    - If so, these signals may experience increased jitter and poor impedance control, and the board may have increased EMI.
- What is the via/antipad strategy?
    - Has the layout designer been allowed to select the via and antipad sizes, or have these been designed? Often the layout designer selects the via size based on layout and routing considerations. These may be optimum

from a manufacturing standpoint, but may not provide the best electrical behavior.
  - Large antipads can remove extensive metal from the power or ground plane, especially in dense pin fields such as under BGAs. This can result in routing and (with high-power devices) power distribution problems.

- How was the trace impedance value selected for each layer?
  - Has an analysis been done trading off ASIC driver strength and termination strategies and trace impedance?
  - Lower-impedance traces result in a thinner overall stackup and make the signals more immune to crosstalk, but increase driver current.
  - Higher-impedance traces require less drive current and are more immune to conductor loss but more sensitive to dielectric losses.
- How was the trace width arrived at for each signal or signal class?
  - Often the minimum trace width allowed by the manufacturer is chosen first, before any signal analysis is performed, but in some circumstances a wider width is preferable. For instance, widening the trace reduces conductor loss, and the crossover frequency where the dielectric loss becomes larger than the conductor loss will change.
  - Wider traces are easier to manufacture and generally result in less scrap (lowering costs).
  - Wide traces have lower inductance and larger capacitance than narrow ones. This is advantageous when routing power to terminators, for example.
  - A trace that is too narrow will limit the number of second sources capable of manufacturing the board in volume.
  - Narrow traces often have more variability in shape (and therefore impedance and resistance) than do wide traces.
- How was the minimum trace spacing determined?
  - Have crosstalk simulations (see Section 15.3) been performed that demonstrate proper operation with the selected spacing?
  - Spacing that is too small will limit the number of second sources capable of manufacturing the board in volume.

15.3 Questions to Ask When Reviewing Crosstalk Simulations

- Does the I/O model produce the fastest signal rise time?
  - The fast (sometimes called “best”) process corner produces the fastest rise time.
  - CMOS drivers produce the fastest rise time when the lowest permitted temperature and highest permitted voltage are used with the fast transistor models.
- Does the circuit board trace model include both conductor and dielectric losses?
- Has the termination resistance and voltage been set to the worst case?

  - Using nominal values will produce the least crosstalk.
- Have the effects of more than one neighboring trace been included in the simulation?
  - Unless the neighbors are widely spaced, the influence of multiple neighbors is required to obtain the worst-case crosstalk.
- Was it necessary to include the contributions of signals from other buses?
  - If the timing is just so, a signal from a bus unrelated to the bus being simulated can add to the total crosstalk if they are routed close to one another.
- Is near-end or far-end crosstalk the bigger concern in this design?
  - By mistakenly focusing on far-end crosstalk, problems at the near end can be overlooked.
- Do the receiver circuit models properly include parasitic capacitances on the inputs and outputs?
  - Not including input capacitance affects the receiver response time and the way in which reflections are formed, which will change the crosstalk magnitude and timing.
  - Not including output capacitance affects the driver rise time.
- Do the receiver models include the effects of clamping circuits?
  - Unrealistically high voltages can be reported if these are not included.
  - The effects of multiple reflections will not be properly reported if the clamps are missing.
- Does the induced voltage exceed the safe values allowed for the process technology?
  - To avoid long-term or catastrophic damage to the I/O circuits, the input voltage must not exceed the maximum allowed.
- Is the induced voltage large enough to cause the receiver output to change state when the effects of power and ground noise caused by a nearby driver switching are present?
  - If so, is the system timing such that this can be ignored?
- Is it possible for the bus to be put in a high-impedance (nondriven) state?
  - If unterminated, this case will probably yield the highest crosstalk.
  - Do the induced voltages remain within the safe range for the process technology?

15.4 Questions to Ask When Reviewing Signal Quality Simulations

The task here is to insure that signal quality has been tested under all possible valid worst-case conditions. Poor signal quality may be caused by a signal that is driven too weakly (underdriven) or by one that is driven too hard (overdriven). The signal integrity simulations must test for both cases.

Underdriven signals only gradually reach the receiver switch point and may not provide adequate noise margin. In high-performance systems underdriven signals are usually an unintentional consequence of the design and are seen as errors to

be corrected. However, sometimes signals are intentionally underdriven to remove high-frequency harmonics as a way to reduce EMI and to save power in low-power applications.

Overdriven signals are sharp edged and overshoot past the static high value. The overshooting may be severe enough for the signal to momentarily fall (to ring back) below the receiver switch point before it settles at the steady state value. For large overdrive conditions the signal may oscillate multiple times between successively smaller overshoot and ringback values before it finally comes to rest at a steady state value.

Therefore, the following questions should be asked when reviewing signal integrity simulations:

- What analysis was performed to guarantee that both the worst-case underdrive and worst-case overdrive situations have been identified?
  - In some systems the difference in signal quality between these two will be small, but usually the difference is demonstrable.
  - In CMOS systems underdrive is usually the worst when the slowest process corner is used at the highest temperature and lowest power supply voltage and the highest series resistance (if the signal is series terminated).
  - In CMOS systems worst-case overdrive usually occurs when the fastest process corner is used at the highest voltage and lowest temperature. If series terminated, the minimum resistance value is used.
- If the signal experiences ringback, has its impact on system timing been thoroughly analyzed?
  - For example, a worst case may be for a data strobe or clock to overshoot (and reach the switch point sooner than expected) while the data signal its clocking is underdriven and so only gradually reaches the switch point.
- If the signal overshoots, does the peak value remain within the manufacturer's safe values specification?
- If the signal is underdriven, how has the effect of power supply noise or noise on adjacent signals been tested?
  - Since underdriven signals only gradually reach the receiver switch point, they are especially vulnerable to false triggering caused by crosstalk or by noise coupled to the signal from the power supply.
- Have all the sources of signal coupling been identified and included in the simulation?
  - See the earlier discussion on crosstalk.
- Have the worst-case data patterns been identified?
  - The worst-case switching pattern depends on the exact circumstances of the design, but in general, a simple 1/0 "square wave" pattern is usually not the worst case. Often the worst pattern involves timing the transmission of a bit so that it combines with energy still present on the line, but the "lone one" [1, 2] can also be significant, especially in situations where the line is lossy.

- Bidirectional buses are especially challenging to simulate and the interaction when different drivers (presumably connected with unequal lengths of trace) are asserting the bus must be tested under all combinations of timing and impedances variations. Monte Carlo [3, 4] or some other statistical method is valuable in reducing the number of simulations, but some SI CAD can be configured to run all possible permutations. Software (either acquired as part of the SI CAD suite or a simple script written by the SI engineer) is critical in sifting through the data and selecting the worst cases.
- Often bidirectional buses have multiple worst-case patterns that depend on the direction of data flow and trace lengths. The SI engineer must be persistent to find all the worst-case patterns and not to stop looking once one or two have been identified.

- Has SSN been properly accounted for?
  - A common mistake when checking the switching behavior of an ASIC is to assume that the worst-case "signal noise" occurs when all the I/O pins drive simultaneously, often in the same direction. While this usually creates the most crosstalk and noise coupling, it does not necessarily create the worst reflections on the circuit board. In fact, the current demand when many drivers simultaneously switch may cause the power supply inside the ASIC feeding the I/O drivers to momentarily drop. In CMOS I/O drivers this current starvation retards the driven rise time. The net effect is that the rise time may be fastest when only one (or a few) drivers switch and may be the slowest when many simultaneously switch.

15.5 Questions to Ask When Reviewing Prelayout Simulations

- What tests have been done to validate each component model?
  - A common mistake is to accept the vendor's circuit model for each component without independently testing it for compliance.
  - At a minimum the model should be run in a simple test circuit to verify that it is operating as expected.
  - In critical designs the model should be validated with actual laboratory measurements before beginning detailed signal integrity simulations.
- What has been done to demonstrate that models from various vendors will function in the same simulation?
  - Circuit-level models sometimes require unique settings and parameters to run properly. The simulation will not operate correctly if the same parameters must be set differently for various models coexisting in the simulation.
- What is the strategy for determining the worst-case simulation conditions?
  - Realistic estimates of the junction temperatures must be made. If the die temperature of interconnected ASICs is not the same, a strategy addressing the behavioral difference these causes must be developed.
  - Have the effects of power supply noise been accounted for?
- Have trace lengths been estimated properly, and what assumptions have been made regarding component placement?

    - Proper estimates of component placement are critical for accurate simulation results because the amplitude and timing of reflections depend on the trace lengths.
- What assumptions have been made regarding termination placement and the inductance connecting them?
    - Component models that include inductance are necessary to obtain the correct high-frequency behavior.
    - Models of vias and any trace connecting the terminators to a power or ground system must also be included.
- Have the worst-case patterns been identified?
    - Were a series of manual simulations run, or was a statistical approach (such as Monte Carlo) used? What has been done to demonstrate that the worst-case patterns have actually been found?
    - Do the simulations include signal loss, or have lossless lines been assumed?
    - Do the simulations include crosstalk effects, or have their effects been accounted for by including a noise "adder"?
- Have models been created for the vias?
    - If so, what stackup assumptions have been made?
    - What process is in place to ensure that the vias in the actual product match the mechanical details used in the electrical model?

15.6 Questions to Ask When Reviewing Postlayout Simulations

- Has the layout artwork been fully checked for compliance with the layout rules?
- Have timing and signal quality simulations been performed with data from the fully routed circuit board?
    - Has the timing on all critical nets been verified?
- Is power supply decoupling adequate and is the capacitor placement and layout optimal?
    - Verify that ASIC manufacturers' guidelines have been followed for the placement and values of decoupling capacitors.
- Are previously unsimulated stubs present on any net, especially high-speed nets?
    - If long enough, stubs will create reflections that can interfere with proper signaling.
- Do the postroute SI simulations include the actual number of layer changes?
    - Vias increase signal loss and can cause reflections. Ensure that the actual layout matches the simulation model.
- Is the location of series or parallel termination resistors acceptable?
    - Does it match the simulations?
    - Do the simulations include the range of trace lengths and the impedance of the traces connecting the terminators?

- If a fly-by trace is used to connect a parallel terminator to a load, has that net been analyzed for coupling and crosstalk?
- Are any vias shared?
  - Vias connecting decoupling capacitors should never be shared: each decoupling capacitor should be connected to power and ground through individual vias.

15.7 Questions to Ask When Reviewing a Termination Strategy

- Has termination located on the ASIC die (ODT) been investigated?
  - This is generally superior to termination placed on the circuit board.
  - It may not allow as much flexibility in make adjustments.
- If a parallel termination strategy is used, in this application is it more cost-effective to use a resistor pair, or should a Thevenin resistor/supply be used?
  - The two-resistor scheme is simple and does not require creating and routing a separate supply voltage.
  - Often the single resistor Thevenin termination can be placed closer to the load and is especially attractive in large, densely packed designs.
- Does the simulation model account for the proper location of the series termination resistors?
  - The model should include the actual distance from the I/O pin to the termination resistor, and the impedance of that trace (which is might be higher than the remainder of the signal trace) should be included.

15.8 Questions to Ask When Reviewing PCB Layout Design Rules

- Have the recommendations of the ASIC manufacturer been incorporated into the design rules?
  - The quality and practicality of ASIC vendor-supplied signal and power integrity rules vary greatly. Nevertheless, the ASIC manufacturer's goals should be observed when developing signal integrity rules for an end product.
- Has an ASIC manufacturer's reference design been studied?
  - Valuable insight regarding the unique layout requirements of an ASIC can be obtained by reviewing the reference design supplied by the manufacturer. Especially valuable are stackup and decoupling (or other power supply filtering) information, but termination and trace impedance factors can also be discovered.

15.9 Questions to Ask When Reviewing ASIC Driver Selection Choices

- Have drivers with various drive strengths and rise times been tested?

    - Has the weakest driver with the least aggressive rise time that still meets the timing and receiver noise budgets been selected?
- Has the effect of many drivers simultaneously switching been simulated?
    - Have the circuit board traces' odd/even modes been properly simulated?
    - Does the simulation include parasitic inductances, capacitance, and resistance for the ASIC package and socket?
    - Does the simulation include some portion of the power distribution network on the ASIC and on the circuit board?
- Has the case of only one driver switching been simulated?
    - Does the signal rise time become so fast that transmission line effects are unacceptable?
- Have the simulations been performed across the full range of voltage, process, and junction temperatures?
    - Does the simulation include the effects of manufacturing tolerance and temperature changes on the board level termination resistors and termination supply voltages?

References

[1] Dally, W. J., and J. W. Poulton, *Digital Systems Engineering*, Cambridge, U.K.: Cambridge University Press, 1998.

[2] Thierauf, S. C., *High-Speed Circuit Board Signal Integrity*, Norwood, MA: Artech House, 2004.

[3] Hall, S. H., and H. L. Heck, *Advanced Signal Integrity for High-Speed Digital Designs*, New York: John Wiley & Sons, 2009.

[4] Li, P. M., *Jitter, Noise, and Signal Integrity at High Speed*, Upper Saddle River, NJ: Prentice-Hall, 2007.

CHAPTER 16

Solving Common Signal Integrity Problems

16.1 Introduction

This chapter outlines strategies for resolving several common signal integrity problems.

It is best to solve these problems during the simulation analysis phase, before the circuit board is manufactured. Often signal integrity errors are systematic in nature. All signals in a net class are usually misrouted in the same way and so they all have the same weaknesses. However, it is not unusual for the "noise problem" to initially appear to be on only a few signals. Once the problem on these signals is corrected, the weaknesses in the remaining signals become apparent, and they too must be fixed.

The general rule is that once a layout problem has been identified for a specific signal, that fix must be applied to all of the signals in that same net class. If the problem has been found on actual hardware (rather than in simulation, before the board has been manufactured), this often means that the circuit board must be relayed out.

The suggestions listed here are intended to identify the problem's underlying cause and are not an exhaustive list of remedies. They can be used when performing signal integrity simulations and during laboratory debug. They are also useful for discussion points when interviewing signal integrity candidates.

This chapter broadly summarizes the material presented in earlier chapters.

16.2 Reducing Crosstalk

Crosstalk can be minimized by reducing the parasitic coupling between traces or by reducing the signal strength of the aggressors. Often a combination of these is used in practice.

In particular, to reduce crosstalk:

- Increase the spacing between traces.
- Move critical signals to a separate routing layer.
- Ensure that the signal reference planes also connect to the I/O drivers (do not use an unrelated voltage plane as a reference).
- Add a properly grounded guard trace.
- Minimize the coupled length.
- Make the victim trace as low impedance as is practical.
- Lengthen the signal rise time (make it slower).
- Reduce the aggressor signal voltage swing.
- Properly terminate the aggressor and victim traces.
- Adjust timing so that the crosstalk event does not cause false triggering (this reduces the effect but does not eliminate the actual coupled voltage).
- Adjust timing so that multiple aggressors do not simultaneously couple onto a victim (this reduces the peak voltage coupled onto the victim).

16.3 Reducing Reflections

Reflections occur when the load impedance does not match the transmission line impedance.

Termination is required if the line is electrically long. Do the following when excessive ringing, overshooting, and ringback are present on a net that has been terminated:

- Verify that the manufactured trace impedance is the expected value and that it matches the termination.
- Perform an odd/even mode analysis to determine if the difference between the odd/even impedances is high. If so, consider increasing the spacing between traces or adding a guard trace to control the impedance under switching conditions.
- On multidrop nets perform a simulation analysis to determine the best way to connect the loads and to find the optimum termination type and value.
- Reduce the I/O cell drive strength to decrease the launched rise time.
- Shorten the trace length.
- If properly matching the load impedance to the transmission line impedance is not possible (for example, a large lumped capacitive load), eliminate multiple reflections by ensuring that the driver impedance closely matches the line impedance.
- Make termination more effective by breaking up a multiload net into several identical copies, each of which is driven from separate pins on the ASIC.
- Ensure that series termination resistors are placed close to the driving pin.
- Ensure that parallel terminations are properly connected to the load.

- Ensure that the signal reference planes also connect to the I/O drivers (do not use an unrelated voltage plane as a reference).
- Ensure that adequate decoupling is placed between power and ground planes.
- If a Thevenin termination is used, ensure that the Thevenin voltage is stable. It must not oscillate and should have little droop even when many signals connected to it simultaneously switch.

16.4 Reducing and Eliminating Simultaneous Switching Noise (SSN)

Eliminating SSN from a poorly designed circuit board can be difficult and, in some cases, impossible. It is best to analyze SSN effects before prototype hardware is fabricated, especially that of the micropackage and any connectors (including sockets).

The following circuit board fixes can be helpful when an integrated circuit experiences high SSN in those cases where little or no SSN analysis has been done. This list is also useful during debug to help identify SSN.

- Use power and ground planes to connect the ASIC to the power supply.
- Provide adequate decoupling capacitance adjacent to the IC pins. This is extremely critical.
- Reduce the number of pins that simultaneously switch.
- Arrange for some signals to always switch in the opposite direction (odd-mode parity, for instance) or stagger the timing so that some signals switch later than others.
- Use differential signaling.
- Use a less powerful (lower current drive) I/O cell.
- In CMOS systems decrease the power supply voltage and increase the junction temperature. This will reduce SSN but will also adversely affect timing.
- Increase the transmission line impedance, since this reduces the launched current.
- Use series termination on the circuit board since this reduces the launched current.

16.5 Improving Inadequate Timing Margins

Assuming that the root cause for the poor timing margin is signal integrity related and that it does not involve poor logic design, one or more of the following circuit board fixes can be helpful:

- Guarantee that ringback and reflections are not eroding setup or hold times.
- Use a higher current drive I/O cell, remove series termination, and provide parallel termination at the load (but these may increase crosstalk and simultaneous switching noise).

- If receiver setup time is inadequate, reduce the signal delay by increasing the driver I/O power supply voltage and decreasing the driving ASIC's temperature. Shorten the trace length and use microstrip rather than stripline.
- If receiver hold time is inadequate, increase the signal delay by lowering the driver I/O power supply voltage and/or increasing the driving ASIC's temperature. Lengthen the trace.
- Use serpentine traces to adjust the timing of a strobe relative to its data.
- Small adjustments in timing can be made by switching to a laminate system with a different dielectric constant (but this will change the trace impedance unless the stackup is also adjusted).

16.6 Correcting Intersymbol Interference (ISI)

ISI is caused by a combination of dispersion, crosstalk, signal reflections, and jitter. Ways to reduce crosstalk and reflections are covered earlier in this chapter. In addition to those general guidelines, some specific steps to take include:

- Reduce the frequency content of the signal by degrading the signal rise time and slowing the data rate to increase the bit time.
- Decrease microstrip and stripline conductor losses by using wide, high-impedance traces (but this will increase the stackup height and via length).
- Reduce dielectric losses in microstrip (but not stripline) by using wide, high-impedance traces (but this will increase the stackup height and via length).
- Use a low loss tangent laminate system (but this may not be cost-effective for price-sensitive products).
- Use a low dielectric constant laminate system.
- Adjust I/O driver precompensation and impedance control circuits. These circuits automatically correct for frequency-dependent losses and they better match the I/O cell impedance to the line impedance.
- When working with cables or very long traces, consider adding passive equalization [1–5].
- Minimize the number of vias in the signal chain since vias can create reflections and can act as lowpass filters that attenuate high-frequency harmonics.
- Use impedance matched vias.
- Ensure that the differential impedance is within specification and that the proper termination is used.
- Ensure that nets where ISI is of concern are routed well away from other traces, away from cutouts in the planes, are not near the board edge, and are away from unrelated vias.

16.7 Steps to Take When Circuit Board Traces Are Too Lossy

At high frequency the total interconnect losses are improved by reducing dielectric losses and conductor losses (losses caused by the resistance of the traces).

- Reduce the trace length, since loss increases exponentially with length.
- Use a 2D field solver to determine if the major contributor to the total loss is due to conductor loss or dielectric loss. Do not use a high-cost low dielectric loss laminate system if the loss is predominately due to the conductor.
- If conductor losses dominate, consider using low surface roughness copper.
- If dielectric losses dominate, consider using a low loss tangent laminate system (but unless very high-performance laminate systems are used, this only slightly reduces dielectric loss and so may not be acceptable in cost-sensitive applications. However, this may be the only viable solution when signaling at high speeds).
- Increase the trace width, but note that thicker traces will not improve high-frequency losses.
- Use high-impedance traces. This is especially effective in lowering dielectric and conductor losses in wide microstrips, but conductor losses in striplines will also improve.
- Consider routing critical signals as microstrip. This can be especially effective when signaling at high-data rates since the dielectric loss usually dominates at high-frequency, and dielectric losses are smaller for microstrip than stripline. However, because manufacturing tolerances are generally looser on the outer layers, impedance control may not be as good and reflections may increase.
- If the traces are microstrip, consider removing the solder mask and treating the exposed traces with a surface finish instead. This creates an exposed microstrip rather than an embedded microstrip and reduces dielectric loss. However, this can cause assembly problems. Be sure to check with manufacturing to verify that this option is acceptable.
- If only a few critical signals are experiencing unacceptable loss, consider using a mixed laminate system where a high-performance laminate is used on one layer in the stackup, while the remaining layers are traditional FR4 material. By segregating the critical stripline routes to the high-performance layer, the loss is reduced for those signals and the remaining lower-cost layers are used to route the majority of the signals.
- Verify that the signals are referenced to the proper power plane.

16.8 Options to Reduce Circuit Board Thickness

Using more routing layers than necessary is the most typical cause of thicker than desired circuit boards. To reduce board thickness:

- Use dual stripline (but this can cause crosstalk unless signals are routed orthogonally).
- Use thin core laminate technology between the power and ground planes (but this may not be cost-effective in price-sensitive products).
- Improve routing density to reduce the number of routing layers by using narrow traces (but this increases signal loss on long traces and may reduce the number of vendors capable of second sourcing the board).
- Use low impedance traces on inner layers (but this increases the current driven by I/O devices and may create SSN problems).
- Use a low dielectric constant laminate system (this works best when the board has many layers).
- Consider using half-ounce or quarter-ounce copper for traces and planes (this only provides a slight reduction; the loss and current handling effects must be analyzed).
- Consider removing some power planes by splitting them for use by more than one supply voltage. This may require components to be rearranged on the board.
- Consider replacing power planes with power islands located on the board surface.
- Consider routing slow, noncritical stripline signals on a layer where one of the planes is not the I/O voltage. For acceptable operation this requires careful design and analysis.
- If the board is too thick to accommodate a particular connector at the edge of a board, the circuit board fabricator can reduce the thickness in selective regions. This technique is not always cost-effective, but it allows the stackup to accommodate many planes and routing layers while simultaneously being thinner at the board edge.

16.9 Steps to Take When There Are Not Enough Routing Layers

- Use dual stripline to increase the number of routing layers without adding too much thickness.
- Improve routing density to reduce the number of routing layers by using narrow traces (but this increases signal loss on long traces and may reduce the number of vendors capable of second sourcing the board).
- Use low-impedance traces on inner layers and decrease the separation between traces (but this increases crosstalk and the current driven by I/O devices and may require that the nets be terminated).
- Consider removing some power planes by splitting them for use by more than one supply voltage, or using power islands, and using the free layer for routing.

- Consider using serial signaling technology to eliminate wide, synchronous buses.
- Consider using a cable or flexible circuit board assembly to jumper signals from one end of a large circuit board to another. This can remove the need to route critical, highly sensitive signals on the board, or it can be used to route many low-performance signals.

16.10 Steps to Take When a Circuit Board Must Be Cost Reduced

- Send the board stackup out to several vendors to determine if it can be simplified.
- Determine if a slight reduction in the board dimensions will significantly reduce cost. Laminate panels come in standard sizes and cost will be lower if the board is small enough to allow several copies to fit within one of the standard sizes with little waste. A slightly smaller board that allows several copies to be placed on a larger panel will cost less than a larger board that can only fit one copy per panel.
- Understand what characteristics are increasing manufacturing costs of the raw board. For instance, the need for advanced via technology or very fine lines can drive up cost.
- Study the decoupling very carefully. Often decoupling is not rigorously analyzed during design, and more capacitors than necessary are used. Use simulation to determine the worst-case conditions for power supply noise and replicate that on prototype hardware. Then selectively remove decoupling on the prototype to determine the minimum necessary for proper operation.
- Verify that all of the terminated nets actually require termination. Removing unnecessary termination components frees up board space (possibly allowing the board to be smaller or to use fewer layers) and, when using parallel termination, power can be reduced, and the costs related to parts and assembly are also reduced.
- If controlled impedance traces are used, verify that they are necessary. Requiring controlled impedance traces increase the bare board manufacturing cost.
- Verify that trace impedance specifications are not overly restrictive and are easy to manufacture. For example, many manufacturers are accustomed to producing 50Ω single ended and 100Ω differential pairs. Specifying different values than this will usually increase the number of boards scrapped and increase cost.
- Verify that spacing between traces is not wider than actually necessary to reduce crosstalk. Spacing that is too great reduces the number of available wiring channels and may increase the number of required routing layers.
- Investigate the use of dual stripline. This can reduce the layer count and board thickness, which can have a cascading effect on reducing board cost.

16.11 Steps to Take When Data-Dependent Errors Are Causing System Failures

- Rule out crosstalk (refer to Section 16.2).
- Rule out SSN (refer to Section 16.4).
- Verify by laboratory measurement or simulation that pulse smearing caused by dispersion (especially on long traces) is not preferentially distorting some bits. Dispersion can occur on parallel buses and is not only an artifact of serial signaling.
- Verify that the signal is properly terminated. Improper termination can result in extreme voltages being present at receiver inputs, which can require successive bits excessive time to assert to valid logic levels. Additionally, signals transitioning well below ground or well above the power rail can require a long recovery time from the receiver before proper timing is restored.
- Verify that data-dependent signal loss is not causing pulse distortion. If the signal return path is inadequate, the apparent AC resistance of a trace will increase (causing the signal loss to increase) as the return current from many signals commingle. Review the circuit board artwork to see if many signals are returned by a narrow strip of ground plane, or if multiple signals are forced into a small, isolated channel. If a power plane is one of the return paths, verify that it is the I/O voltage and is not unrelated.
- Verify that connectors have adequate numbers of ground pins (ideally one ground per signal, but fewer in practice with the actual number required being determined by simulation or laboratory measurement). If the connectors are used to signal between boards, verify that the connector ground pins are connected by low-impedance paths to the signal return planes on their respective boards.

References

[1] Couch, L. W., *Digital and Analog Communication Systems*, 5th ed., Upper Saddle River, NJ: Prentice-Hall, 1997

[2] Bennett, W. R., and J. R. Davey, *Data Transmission*, New York: McGraw-Hill, 1965.

[3] Thierauf, S. C., *High-Speed Circuit Board Signal Integrity*, Norwood, MA: Artech House, 2004.

[4] Proakis, J. G., *Digital Communications*, New York: McGraw-Hill, 2000.

[5] Federal Telephone and Radio Corp, *Reference Data for Radio Engineers*, 3rd ed., New York: Knickerbocker Printing Corp., 1949.

CHAPTER 17

Calculating Trace and Plane Electrical Values

17.1 Introduction

With the formulas presented in this chapter, the signal integrity engineer can estimate the impedance, time of flight, loss, and circuit values (R, L, C) for traces and the circuit values for planes. Ways to calculate conductance (G) appear in Chapter 7 and are not repeated here.

These are very helpful in estimating the parasitic values of traces without resorting to a field solver. In fact, the results from these equations can be used as a first-order check of the results from field solvers when validating circuit models.

Some of the equations (especially for impedance and loss) are too complicated to be performed reliably by hand on a calculator. However, they are easily solved with numerical software such as MATLAB [1] or MathCAD [2]. Spreadsheets can also be used, or the equations can be coded in a simple programming language (such as C or even a scripting language such as PERL, particularly for the simpler equations).

The dimensions for stripline and microstrip traces used in the formulas are shown in Figure 17.1.

17.2 How to Convert from Decibel Loss to Signal Swing

Chapter 8 showed how to calculate loss in decibels (dB). That equation is repeated here, in (17.1):

$$\text{Gain in dB} = 20 \times \log_{10}\left(\frac{V_{received}}{V_{transmitted}}\right) \tag{17.1}$$

For example, the gain is –6 dB when $V_{received}$ is 0.5V and $V_{transmitted}$ is 1V. Since the sign is (–), this value represents a loss.

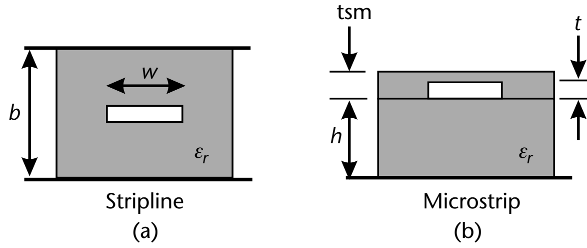


Figure 17.1 (a) Stripline trace is centered between two return planes, separated by distance *b*. (b) Microstrip is raised *h* above the return plane and the solder mask thickness is *tsm*. The traces have a width of *w* and a thickness of *t*, and the dielectric constant is ε_r.

Some signaling specifications give the total loss budget for the link in decibels. In those situations converting from loss in decibels to a voltage ratio is useful for understanding the required launched and received voltages.

Solving (17.1) for the ratio of $V_{received}$ to $V_{transmitted}$ produces:

$$\left(\frac{V_{received}}{V_{transmitted}}\right) = 10^{\frac{\text{dB}}{20}} \tag{17.2}$$

For example, the $V_{received}$ to $V_{transmitted}$ ratio is 0.71 when the loss is 3 dB. Here again, a (–) sign is used to show loss. If the minimum valid receiver voltage is 0.8V, with a 3-dB transmission line loss, the transmitter must launch 1.1V for the link to properly operate.

17.2.1 Nepers

Rather than using common logarithms to express loss in decibels, natural logarithms can be used to express the loss in nepers (Np) [3, 4]. Nepers are not as frequent as decibels in signal integrity work, but the conversion shown in (17.3) is useful when it is necessary to convert between the two.

$$\text{dB} = Np \times 6.868 \tag{17.3}$$

For instance, a loss of 27.5 dB can also be represented as a 4-Np loss.

17.3 Estimating DC Resistance

Equation (17.4) shows how the DC resistance of a trace or plane depends on the metals resistivity (ρ, see Chapter 5), cross-sectional area (A), and length.

$$R_{dc} = \frac{\rho}{A} \times Length \tag{17.4}$$

The trace dimensions are shown in Figure 17.2.

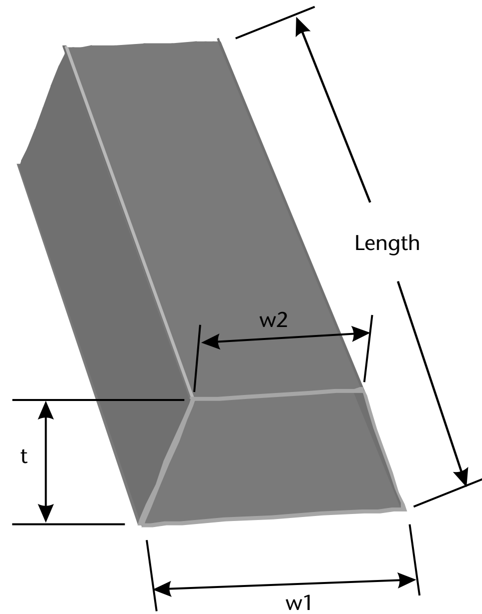


Figure 17.2 Trace dimensions for calculating DC resistance. Set $w_1 = w_2$ for rectangular traces.

The resistance of rectangular or trapezoidal copper traces in ohms per inch or per meter can be found by using:

$$R_{dc} = \frac{K_{rdc}}{t(w_1 + w_2)} \tag{17.5}$$

Set K_{rdc} to 1.41 to find the resistance in ohms per inch length when t, w_1, and w_2 are in mils. To find the resistance per meter length, use 352 for K_{rdc} and set t, w_1, and w_2 to micrometers.

For a perfectly rectangular trace, make w_1 equal to w_2. For example, to determine the resistance of a 5-inch-long perfectly rectangular 5-mil-wide, 0.65-mil-thick trace, K_{rdc} is 1.41, t is 0.65, and w_1 and w_2 are each 5. The equation yields 0.216 per inch, making the 5-inch-long trace resistance 1.1Ω.

17.4 Finding Inductance and Capacitance When the Physical Dimensions Are Not Known

The impedance (Z_o) and delay (tpd) equations (presented in Chapter 6) can be simultaneously solved to find the trace inductance (17.6) and capacitance (17.7).

$$L = Z_o \times tpd \tag{17.6}$$

$$C = \frac{tpd}{Z_o} \tag{17.7}$$

For example, the inductance L of an 86Ω transmission line having a delay of 5.6 ns/meter is found from (17.6) to be 482 nH/meter, and from (17.7) its capacitance is 65 pF/meter.

17.5 Finding Stripline Inductance and Capacitance When Impedance Is Known

Because the dielectric is homogeneous in stripline (see Chapter 7), the inductance and capacitance equations can be determined for that specific case when only the impedance is known.

The inductance of a stripline is found with (17.8). Set $K1$ to 0.085 to obtain the inductance in nH/inch. Use 0.033 to obtain the inductance in nH/cm.

$$Lsl = Z_o \times K1\sqrt{\varepsilon_r} \tag{17.8}$$

The capacitance of a stripline is calculated with (17.9). Set $K2$ to 85 to obtain the inductance in pF/inch. Use 33 to obtain the inductance in pF/cm.

$$Csl = \frac{K2\sqrt{\varepsilon_r}}{Z_o} \tag{17.9}$$

For example, the capacitance of a 50Ω stripline of any width or thickness is 3.4 pF/inch length when the dielectric constant (ε_r) is 4.

17.6 Finding Stripline Inductance and Capacitance When Trace Geometry Is Known

Inductance of a stripline trace can be estimated with (17.10). This equation has been simplified and curve-fitted [3] from [5]. It assumes that the trace is centered between the two planes. Since it does not account for trace thickness, it is only accurate to within about 20% of the actual value.

$$L = K3 \times \frac{b}{w+b} \tag{17.10}$$

where $K3$ is 16 for results in nH/inch, or 6.3 for nH/cm.

The reciprocity principle [3] can be used to find the capacitance of stripline (but not microstrip) from the inductance calculated in (17.10). Doing so yields (17.11), and since it is derived from (17.10), it has the same level of accuracy.

$$C = K4 \times \varepsilon_r \times \left(\frac{w+b}{b}\right) \tag{17.11}$$

where $K4$ is 0.45 for results in pF/inch or 0.18 for results in pF/cm.

Because in both of these equations the units cancel, $K3$ and $K4$ can have different units than h and w.

For example, the capacitance of a 5-mil-wide stripline centered between two planes spaced 66 mils apart when $\varepsilon_r = 4.0$ is 0.77 pF/cm ($K4 = 0.18$, $w = 5$, $h = 66$). This is 15% higher than the actual value of 0.67 pF/cm.

17.7 Finding Microstrip Inductance and Capacitance When Trace Geometry Is Known

Often the width, thickness, and impedance of a microstrip are known, but the time of flight is not. In that case the (17.12), (17.13) and (17.14) can be used to estimate the inductance and capacitance.

17.7.1 Microstrip Inductance

Inductance of a microstrip trace not covered in solder mask can be estimated (without accounting for trace thickness) with (17.12). This has been simplified and curve-fitted [3] from [5].

$$L = K5 \times \ln\left(6.3 \times \frac{h}{w}\right) \tag{17.12}$$

Set $K5$ to 5.1 to obtain the inductance in nH/inch or to 2 for nH/cm. Because the units cancel, $K5$ can have different units than h and w.

For example, the inductance of a 5-mil-wide microstrip spaced 11.5 mils above the ground plane is 5.35 nH/cm ($K5 = 2$, $w = 5$, $h = 11.5$). This is 7% lower than the actual value of 5.46 nH/cm.

17.7.2 Microstrip Capacitance

Because the dielectrics are not homogeneous, it is not proper to use the reciprocity principle to find the capacitance when the inductance is known, as was done with stripline.

Equation (17.13) is simplified from [6] and used to find the capacitance once the effective dielectric constant has been found with (17.14) [7] (repeated here from Chapter 7).

$$\mathrm{C} = \frac{K6 \times \varepsilon_{r_eff}}{\ln\left(\frac{8h}{w} + \frac{w}{4h}\right)} \tag{17.13}$$

$$\varepsilon_{r_eff} = \frac{\varepsilon_r + 1}{2} + \left(\frac{\varepsilon_r - 1}{2} \times \frac{1}{\sqrt{1 + \frac{10h}{w}}}\right) \tag{17.14}$$

Because the units cancel, these two equations may be used for traces dimensioned in either metric- or inch-based units. Set $K6$ to 1.414 to calculate the capacitance in pF/inch. Use 0.556 for $K6$ to determine capacitance in pF/cm.

For example, the capacitance of a 5-mil-wide microstrip-spaced 11.5 mils above the ground plane when ε_r is 4.0 is 0.535 pF/cm [$K6 = 0.556$, $w = 5$, $h = 11.5$, and from (17.14) $\varepsilon_{r_eff} = 2.806$]. This is 3% lower than the actual value of 0.55 pF/cm.

17.8 Estimating Inductance and Capacitance of a Plane

The following parallel plate capacitance and inductance equations can be used to estimate the inductance and capacitance of power and ground planes (see Figure 17.3). Because these equations do not take fringing into account, they are most accurate when w is extremely greater than h. This makes them unsuitable for use with traces.

17.8.1 Inductance

Parallel plate inductance is found with:

$$L = K7 \times \frac{Length}{w} \times h \qquad (17.15)$$

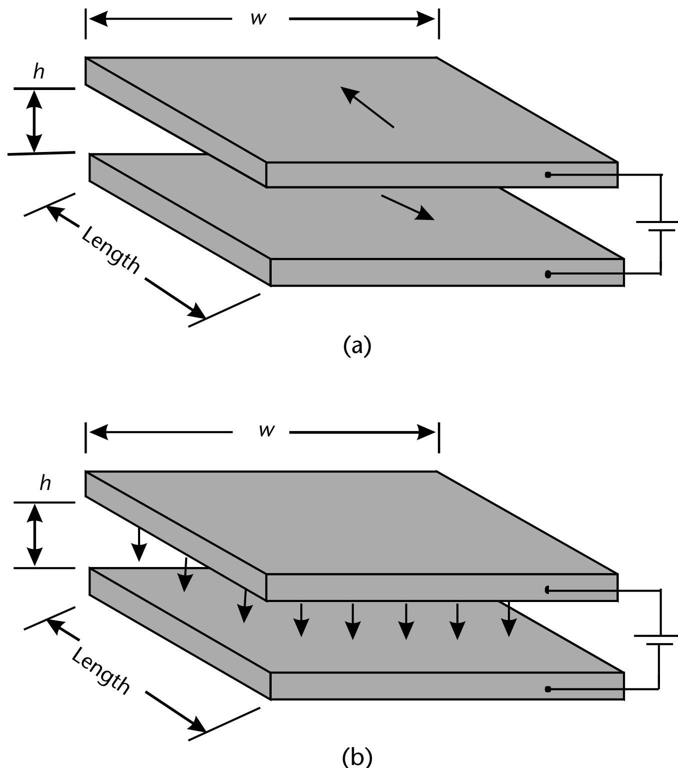


Figure 17.3 Parallel plate (a) inductance and (b) capacitance present between parallel plates. The selection of width and length matters in determining the inductance, but the capacitance is determined by the area.

To obtain the inductance in nH/inch, set *K*7 to 32; set it to 12.6 for nH/cm.

The equation is written this way to show the importance of the length-to-width ratio in determining the inductance. As with the capacitance, the *h* and *w* units cancel, so *L* can be found in terms of either inches or meters for *h* and *w* in any units, provided that the units for *K*7 and length match.

For example, the inductance of a 10-cm-long, 50-mil-wide plane placed on a 1-mil-thick block of laminate is 2.5 nH ($K7 = 12.6$, $Length = 10$, $w = 50$, $h = 1$) when the current is injected along the plane's wide edge. This is 0.25 nH/cm. Because the length is in centimeters, 12.6 was chosen for *K*7 rather than 32.

The inductance of those same planes when current is injected along the narrow edge is found to increase to 63 nH ($K7 = 12.6$, $Length = 50$, $w = 10$, $h = 1$). This is 1.3 nH/cm.

This example shows that simply by swapping *Length* and *w*, the inductance of the setup has increased fivefold. This illustrates the advantage in using wide conductors rather than narrow ones when it is important to reduce inductance.

17.8.2 Capacitance

The capacitance is found with:

$$C = K8 \times Length \times \frac{w}{h} \times \varepsilon_r \tag{17.16}$$

where *K*8 is 0.225 to calculate the capacitance in pF/inch length or use 0.885 for pF/cm length. Because the units for *w* and *h* cancel, they can have inch or metric units when using either value for *K*8.

For example, the capacitance of a 10-cm-long, 50-mil-wide plane placed on a 1-mil-thick block of laminate with a dielectric constant of 4 is 177 pF ($K8 = 0.885$, $Length = 10$, $w = 50$, $h = 1$, $\varepsilon_r = 4$). Because the length is in centimeters, 0.885 was chosen for *K*8 rather than 0.225. Notice that since the capacitance is determined by the area, it does not change if *Length* and *w* are swapped (but be sure to keep track of the units so the correct value for *K*8 is used).

17.9 Calculating Trace Loss from a Circuit Model

Lossy transmission line models in SPICE use the AC resistance (R_{ac}) and conductance (G) to model conductor and dielectric loss (see Chapter 8) in transmission lines. These lines may be circuit board traces or cables. Equation (17.17) shows how to calculate the total loss (α_t) in decibels when these values are known:

$$\alpha_t = 4.3\left(\frac{R_{ac}}{Z_o} + \left(G \times Z_o\right)\right) \tag{17.17}$$

Since R_{ac} and G are given for a specific frequency and length of line, α_t is only valid at that one frequency and length.

For example, α_t is 0.04 dB if R_{ac} is 400 mΩ, and G is 40 μS at 100 MHz for a 60Ω trace.

17.10 Calculating Stripline Loss from Trace Dimensions

Conductor loss in dB/inch is given in (17.18) for the frequency (*f*) in gigahertz. This equation assumes perfectly smooth copper and is a simplified and converted version [3] of the conductor loss equation originally given in [8].

To calculate the loss in dB/inch, set *K*8 to 0.002 and use mils for *w*, *b*, and *t*. To find the loss in dB/cm, set *K*8 to 0.02 and specify *w*, *b*, and *t* in micrometers.

$$\alpha_c = \frac{K8 \times \varepsilon_r Z_o \sqrt{f}}{b}\left[K9 + \frac{\frac{2w}{b}}{\left(1-\frac{t}{b}\right)^2} + 0.318 \times \frac{1+\frac{t}{b}}{\left(1-\frac{t}{b}\right)^2}\ln\left(\frac{K9+1}{K9-1}\right)\right] \tag{17.18}$$

The constant *K9* is given in (17.19):

$$K9 = -\frac{b}{t-b} \tag{17.19}$$

For example, on FR4 at 1 GHz, the conductor loss of a 60Ω, 5-mil-wide, 0.65-mil-thick stripline when the plane separation is 18 mils ($K8$ = 0.002, $K9$ = 1.037, w = 5, t = 0.65, b = 18, ε_r = 4, Z_o = 60, f = 1) is 0.083 dB/inch. This is 2.5% higher than the actual value of 0.081 dB/inch.

The error increases for larger values of *K9*. For instance, decreasing *b* to 12 mils in the above example causes *K9* to increase to 1.057 and gives an impedance of 50Ω. In this case the equation calculates α_c as 0.11 dB/inch, which is nearly 7% higher than the actual value of 0.104 dB/inch.

17.10.1 Dielectric Loss

The dielectric loss for stripline is given in Chapter 8. For convenience it is repeated here as (17.20), slightly altered from its original form in [8]:

$$\alpha_d = K10 \times f\sqrt{\varepsilon_r} \times LT \tag{17.20}$$

where *f* is the frequency in gigahertz and *LT* is the loss tangent (see Chapters 1 and 7) at that same frequency. Use 2.32 for *K*10 to get the dielectric loss in dB/inch. Set *K*10 to 0.91 for losses in dB/cm.

For example, by setting *K*10 to 2.32, at 1 GHz (f = 1), the dielectric loss on FR4 (ε_r = 4.0, LT = 0.02) is found to be 0.093 dB/inch. Setting *K*10 to 0.91 gives a dielectric loss of 0.036 dB/cm. These values are for striplines of any width, thickness, or impedance.

17.11 Calculating Microstrip Loss from Trace Dimensions

Conductor loss in dB/inch is given in (17.21) for the frequency (f) in gigahertz. This is a simplified and converted version [3] of the conductor loss equation originally given in [9, 10] and is most accurate for microstrips not covered with solder mask of 50Ω and higher [3].

$$\alpha_c = \frac{K11\sqrt{f}}{h \times Z_o}\left(\left(1-\left(\frac{w_p}{4h}\right)^2\right)\left(1+\frac{h}{w_p}+\frac{h}{\pi \times w_p}\left(\ln\left(\frac{2h}{t}\right)-\frac{t}{h}\right)\right)\right) \tag{17.21}$$

To calculate the loss in dB/inch, set $K11$ to 11.4 when the distance from the plane to the microstrip (h) is in mils. To find the loss in dB/cm, set $K11$ to 114 and specify h in micrometers.

The width correction factor w_p is given by (17.22):

$$w_p = w + \left(\frac{1}{\pi}\left(\ln\left(\frac{2h}{t}\right)+1\right)\right) \tag{17.22}$$

For example, for a 5-mil-wide, 0.65-mil-thick microstrip spaced 11.5 mils above the ground plan, w_p is 5.74 mils ($w = 5$, $t = 0.65$, $h = 11.5$). Using that value for w_p in (17.21) when $Z_o = 100\Omega$, and setting $K11$ to 11.4 and $\varepsilon_r = 4.0$, the conductor loss $\alpha_c = 0.051$ dB/inch. This is nearly 20% lower than the actual value of 0.063 dB/inch.

Lowering the impedance to 50Ω by reducing h to 2.6 mils cuts this error in half: w_p becomes 5.455, and α_c becomes 0.122 dB/inch (9% lower than the actual value of 0.123 dB/inch)

17.11.1 Dielectric Loss

The equation for dielectric loss in stripline can be used to estimate the losses in microstrip by replacing the dielectric constant ε_r with the effective dielectric constant (ε_{r_eff}) found with (17.14). This is shown in:

$$\alpha_d = K10 \times f \times \sqrt{\varepsilon_{r_eff}} \times LT \tag{17.23}$$

where f is the frequency in gigahertz and LT is the loss tangent (see Chapters 1 and 7) at that same frequency. Use 2.32 for $K10$ to get the dielectric loss in dB/inch. Set $K10$ to 0.91 for losses in dB/cm.

For example, for a 5-mil-wide trace ($w = 5$) spaced 11.5 mils ($h = 11.5$) above a ground plane, ε_{r_eff} is found to be 2.806 from (17.14) when ε_r is 4.0. Using this value in (17.23) and setting $K10$ to 2.32, at 1 GHz ($f = 1$) the dielectric loss on FR4 ($LT = 0.02$) is found to be 0.078 dB/inch.

This is nearly 20% higher than the actual value of 0.065 dB/inch. The accuracy improves to better than 10% for wider traces.

17.12 Finding Stripline Impedance

The impedance of a stripline trace can be found very accurately even for narrow traces with (17.24) [11, 12]. It accounts for effects of trace thickness and fringing and is accurate to better than 2% for wide traces under about 65Ω, provided that b is 2.6 mils (66 μm) or more. The trace must be wide enough so that $\frac{w}{b-t} \geq 0.35$, and accuracy improves for larger ratios. This is usually the case for 65Ω or lower traces on FR4. Other equations are presented in [3, 11–13] for higher impedance traces.

Equation (17.24) has been recast in a simpler form similar to that presented in [13]:

$$Z_o = \frac{94.15}{\sqrt{\varepsilon_r}\left(\frac{w}{b}K9 + \frac{C_f}{8.854\varepsilon_r}\right)} \tag{17.24}$$

The constant $K9$ is found with (17.19), and the fringing capacitance C_f is found with (17.25):

$$C_f = 5.656 \times \varepsilon_r \times \left[K9 \times \ln(K9+1) - (K9-1)\ln\left(K9^2 - 1\right)\right] \tag{17.25}$$

Although the fringing capacitance is in pF/m, the impedance equation (17.24) has been scaled so that w, b, and t can be in either mils or centimeters.

For example, for a 5-mil-wide, 0.65-mil-thick stripline on FR4 ($\varepsilon_r = 4.0$) located midway between ground planes spaced 66 mils apart, $K9$ is found from (17.19) to be 1.01. From (17.25) the fringing capacitance, C_f is found to be 17.34 pF/meter ($w = 5$, $t = 0.65$, $b = 66$).

By using that value for C_f in (17.24), Z_o is found to be 88Ω. This is 12% lower than the actual value of 100Ω. This inaccuracy is expected because this trace fails the $\frac{w}{b-t} \geq 0.35$ test.

Lowering the impedance by reducing the plane spacing to 12 mils passes the test and makes $K9 = 1.057$ and $C_f = 18.85$. Using this in (17.24) produces an impedance of 48.8Ω, about 1.5% lower than the actual value.

17.13 Finding Exposed Microstrip Impedance

The impedance of a microstrip not covered in solder mask can be found with (17.26) [11, 12]. It works best for traces under 60Ω on FR4 that are at least twice as wide as the distance from the return plane (h in Figure 17.1). The accuracy is better than 5% [3] in those cases and improves for wide, lower impedance traces.

$$Z_o = \frac{377}{\sqrt{\varepsilon_{r_eff}}\left[\frac{w}{h} + 1.393 + 0.667\ln\left(\frac{w}{h} + 1.444\right)\right]} \quad (17.26)$$

Although one of several equations can be used to determine ε_{r_eff}, this equation works best when (17.14) is used [3].

Because the units cancel, these two equations may be used for traces dimensioned in either metric- or inch-based units.

For example, for a 5-mil-wide, 0.65-mil-thick stripline on FR4 ($\varepsilon_r = 4.0$) placed 2.75 mils above the ground plane, ε_{r_eff} is found to be 3.09 from (17.14) ($w = 5$, $t = 0.65$, $h = 2.75$), and the impedance is found from (17.26) to be 54Ω. This is 4% higher than the actual value of 57Ω.

Adding a 1-mil-thick solder mask reduces the impedance to 47Ω. Other equations are presented in [3, 14, 15] for higher impedance traces.

17.14 Finding Solder Mask Covered Microstrip Impedance

The impedance of a microstrip covered in solder mask can be found with (17.27) and (17.28), which for clarity are slightly modified from [14]. Equation (17.27) reduces the impedance of the exposed microstrip trace given in (12.26) by the ratio of the effective dielectric constant with and without solder mask.

$$Z_{o_sm} = Z_o\sqrt{\frac{\varepsilon_{r_eff}}{\varepsilon_{r_eff_sm}}} \quad (17.27)$$

In the equation Z_o is the impedance of the exposed microstrip given with (12.26) and ε_{r_eff} is the effective dielectric constant of the exposed microstrip given with (17.14). The effective dielectric constant when the solder mask is applied is given by (17.28), where h is the height above the return plane and tsm is the thickness of the solder mask, measured from the base of the trace (see Figure 17.1). The units cancel, which means if tsm and h use the same units, they may be measured in either inch-based or metric-based units.

The equations assume that the dielectric constant of the solder mask and the laminate have the same values.

$$\varepsilon_{r_eff_sm} = \varepsilon_{r_eff} \times e^{-\frac{2\times tsm}{h}} + \varepsilon_r \times \left(1 - e^{-\frac{2\times tsm}{h}}\right) \quad (17.28)$$

For instance, the impedance on FR4 ($\varepsilon_r = 4$) of a solder mask covered 5-mil-wide, half-ounce (0.65-mil-thick) trace 3 mils above the return plane covered with a 1-mil-thick coating of solder mask on top of the trace is found as follows.

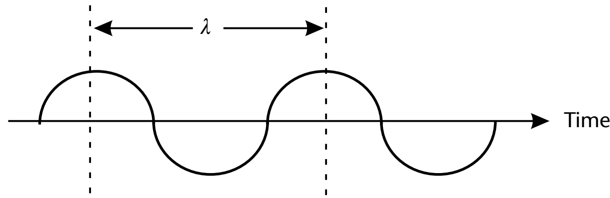


Figure 17.4 Wavelength of a sine wave is the distance between identical successive points.

1. Equation (17.14) is used to find that the effective dielectric constant of the exposed microstrip $\varepsilon_{r_eff} = 3.067$ ($h = 3$, $w = 5$, $\varepsilon_r = 4$).
2. The impedance of the exposed microstrip is found from (17.26) to be $Z_o = 56.4\Omega$.
3. The effective dielectric constant of the solder mask–covered microstrip is found from (17.28) to be $\varepsilon_{r_eff_sm} = 3.689$ (from step 1 $\varepsilon_{r_eff} = 3.067$, and *tsm* is the total thickness of the solder mask, which is the thickness of the trace plus the coating thickness, or 1.65 mils in this case).
4. The impedance is then found from (17.27) to be:

$$Z_{o_sm} = Z_o \sqrt{\frac{\varepsilon_{r_eff}}{\varepsilon_{r_eff_sm}}} = 56.4\Omega \sqrt{\frac{3.067}{3.689}} = 51.4\Omega$$

This is about 2.5% higher than the actual value of 50.2Ω.

17.15 Wavelength

As shown in Figure 17.4, the distance separating identical points on a sine wave (such as points of maximum amplitude) is called the sine waves wavelength.

Because the sine wave's frequency determines how frequently the wave repeats, the wavelength (λ) is shorter for high frequencies. The dielectric constant of the media through which the wave is passing also alters the sine wave's wavelength. It is longest when traveling through free space and becomes shorter when traveling through a dielectric such as FR4. For instance, a 300-MHz sine wave has a "free space wavelength" of 1 meter, but it is less than half that when it travels through FR4.

Wavelength in meters is calculated with (17.29) when the frequency f is in megahertz.

$$\lambda = \frac{300}{f \times \sqrt{\varepsilon_r}} \tag{17.29}$$

For stripline the dielectric constant of the laminate is used for ε_r, but ε_{r_eff} can be used for microstrip. For example, ε_r is 4.2 for a 5-mil-wide half-ounce 65Ω stripline on FR4, but ε_{r_eff} is only 3.6 for microstrip. At 300 MHz this makes the

wavelength on stripline 0.49 meter ($f = 300$, $\varepsilon_r = 4.2$), and 0.53 meter on microstrip ($f = 300$, and replacing ε_r in the equation with $\varepsilon_{r_eff} = 3.6$).

References

[1] The MathWorks, Inc., Natick, MA, http://www.mathworks.com.

[2] Parametric Technology Corp., Needham, MA, http://www.ptc.com.

[3] Thierauf, S. C., *High-Speed Circuit Board Signal Integrity*, Norwood: Artech House, 2004.

[4] Sinnema, W., *Electronic Transmission Line Technology*, Upper Saddle River, NJ: Prentice-Hall, 1988.

[5] Walker, C. S., *Capacitance, Inductance and Crosstalk Analysis*, Norwood, MA: Artech House, 1990.

[6] Baloglu, H. B., *Circuits, Interconnections, and Packaging for VLSI*, Reading, MA: Addison-Wesley, 1990.

[7] Schneider, M. V., "Microstrip Lines for Microwave Integrated Circuits," *Bell System Technical Journal*, May-June 1969, Vol. 48, No. 5, pp. 1421–1443.

[8] Cohn, S. B., "Problems in Strip Transmission Lines," *IEEE Microwave Theory and Techniques*, Vol. MTT 3, No. 2, March 1955, pp. 119–126.

[9] Pucel, R. A., D. J. Masse, and C. P. Hartwig, "Losses in Microstrip," *IEEE Transactions on Microwave Theory and Techniques*, Vol. MTT-16, No. 6, June 1968, pp. 324–350.

[10] Pucel, R. A., D. J. Masse, and C. P. Hartwig, "Correction to Losses in Microstrip," *IEEE Transactions on Microwave Theory and Techniques*, Vol. MTT-16, No. 12, December 1968, p. 1064.

[11] Cohn, S. B., "Characteristic Impedance of the Shielded-Strip Transmission Line," *IEEE Microwave Theory and Techniques*, Vol. MTT 2, No. 2, July 1954, pp. 52–57.

[12] Cohn, S. B., "Problems in Strip Transmission Lines," *IEEE Microwave Theory and Techniques*, Vol. MTT 3, No. 2, March 1955, pp. 119–126.

[13] Liao, S. Y., *Engineering Applications of Electromagnetic Theory*, New York: West Publishing Company, 1988.

[14] Kaupp, H. R., "Characteristics of Microstrip Transmission Lines," *IEEE Transactions on Electronic Computers*, Vol. EC-16, No. 2, April 1967, pp. 185–193.

[15] Bogatin, E., "Design Rules for Microstrip Capacitance," *IEEE Transactions on Components, Hybrids, and Manufacturing Technology*, Vol. 11, No. 3, September 1988.

About the Author

Stephen C. Thierauf is a signal integrity and design engineer in private practice (http://www.thieraufdesign.com), specializing in circuit and circuit board design, signal integrity consulting, and training. As the senior consulting hardware engineer at Infiniswitch/Fabric Networks, he was responsible for the simulation, design, and laboratory measurement of high-performance backplanes and high-speed serial interconnect. While working at Compaq Computer and Digital Equipment Corporation, he was a senior member of the technical staff responsible for the design and signal integrity analysis of high-speed I/O circuitry and interconnect on the Alpha and VAX series of microprocessors. Formerly a visiting scholar at Northeastern University, he holds a B.S. in electrical engineering technology from Wentworth College of Technology, Boston, Massachusetts. He coauthored five papers regarding the VAX and Alpha microprocessors, contributed to the book *Design of High Performance Microprocessor Circuits* (IEEE Press, 2001), and authored the book *High-Speed Circuit Board Signal Integrity* (Artech House, 2004).

Index

A

Aggressor, 95–98, 176
  see also Crosstalk
Antipad, 28–39
  3-D model example, 191
  diff-pair routing, 175, 178
  impedance effect on, 193–194
  poor return path, 177
  via capacitance effect on, 192
  via inductance effect on, 193
Attenuation, 79–81
  surface roughness, 86

B

Bandwidth, 6–8
  3 dB, 7
  requirements, 6
Bounce diagram, *see* Reflection chart

C

CAD, 13–17, 21–28
Capacitance
  2-D and 3-D differences, 27–28
  conductance and, 67
  coupling effects on, *see* Crosstalk
  decoupling, 57, 211
  dielectric constant, 9–10, 40–41, 65–66
  excess, *see* Chamfering
  impedance and, *see* Transmission line circuit model
  interlayer, 57–58
  matrix, 94
  microstrip, *see* Microstrip capacitance
  microstrip and stripline differences between, 66
  mutual, *see* Mutual capacitance
  parallel plate, 50, 57, 64–65, 222–223
  signal trace, 9
  SSN with regard to, 8–9
  stripline, *see* Stripline capacitance
  trace shape effects, 89
  transmission line, *see* Transmission line circuit model
Chamfering, 185
Characteristic impedance, 51, 54, 100, 124
Circuit simulator, *see* SPICE circuit simulator
Clock
  monotonic, 150
  termination, *see* Termination, clock
  quality, 139
Common mode
  noise, 162, 176
  range,162
  termination, 169
  voltage, 161–162, 166–172
Conductance, 23, 27, 51, 67
Conductivity, 64
  losses and, 86, 223
  see also Resistivity
Conductor loss, 83–86
  board edge effects, 187
  calculating, 224–225
  DC resistivity and, 63
  modeling of, 23, 27
  remedies, 213
  surface roughness effects on, 27, 45–47, 84–85
  trace shape effects on, 89
Connectors length matching and, 178–179
Constant-k filter, 51
Contract manufacturers, 35
Copper foils
  electrodeposited, 45–47
  rolled, 45
Copper weight, 43–46
  thickness of, 43

Core construction, 37–38, 224
Cores, 37–38
Corner effects, 83
Corners
mitering, *see* Mitering
serpentine and, 190
worst case circuit models, *see* Worst case analysis
Crosstalk, 107–117
calculating, 109–110, 113
circuit model, 108
comparison with simulation, 115–116
diode clamping effects on, 144
far end, 109–112
fly-by termination, 148
near end, 108, 112–115
questions to ask when reviewing, 201–202
reduced with guard trace, 116–117
reducing, 209–210
reviewing simulations, 201
serpentine traces effects on, 189
SPICE T element and, 23
SPICE W element and, 23–24

D

Data dependent errors, 216
Data mask, 163–164
see also Eye diagram
Data pattern, 132
worst-case, 203–204
Decibel, 80–83
calculating, 217–218
Nepers conversion, 218
reflection loss and, 26
Decoupling capacitors, *see* Capacitance, decoupling
Delay lines, *see* Serpentine traces
Design validation and test, 14–16
Dielectric constant, 9–10
effective, *see* Effective dielectric constant
environmental effects on, 87
relative, 9, 40
resin and glass weave effects, 89
signal loss affected by, *see* Dielectric losses
typical values of, 40–41, 65
using to calculate delay, 73–75
Dielectric loss, 86–88
calculating, 223–225
conductance and, 67
electrical model of, 23, 67, 83
reducing, 212–13
stripline vs. microstrip, 87–88
Differential
amplifier, 161–162, 166, 171
impedance, 165–166, 168–169, 173, 175–178
signaling, 161–182
termination, *see* Termination, differential
transmission line, 161, 165–181
vias, *see* Vias, differential
voltage, 166 –169
Differential pairs, 161–182
broadside coupled, 43, 172–174, 176, 178–180
edge coupled, 172–173, 175–180
odd mode and, 99, 170–172
simulating, 23–24
termination, 169–172
Digital signal analyzer, 163
Dispersion, *see* Pulse, dispersion
Dissipation factor, *see* Loss tangent
Distortion, *see* Attenuation
DK, *see* Dielectric constant
Double treat foils, 45
see also Copper foils
Drill ratio, 39
Dual stripline, 32, 42–43, 200, 214–215

E

E glass, 39–40
Effective dielectric constant, 65–66
calculating, 221, 227–228
delay and, 73–74
impedance and, 71–72
signal loss and, 87
Electrical overstress, 13, 19, 107
Electrodeposited foil, *see* Copper foils
Electromagnetic interference, 14
odd mode and, 172
Electrostatic discharge (ESD), 19
EMI, 14–15
odd mode and, 172
rise time effects and, 200, 203
tightly coupled advantages of, 174

Equalization
 IBIS models and, 22
 effectiveness of, 165–166
 see also Intersymbol interference
Equalizer, *see* Equalization
Evaluation boards, 31, 33–34
Even mode, 99–103
 delay, 100–101
 impedance, 100
 termination, 170–172
Excess capacitance, 185
Excess inductance, 185
Exposed microstrip, *see* Microstrip
External inductance, *see* Inductance, external
Eye diagram, 163–166

F

Far end crosstalk, *see* Crosstalk, far end
FEXT, *see* Crosstalk, far end
Field solvers, 21, 26–28
 2-D, 14, 26–7, 116–117, 188
 3-D, 14, 26, 27–28
 agreement between, 116–117
 conductivity use of, 64
 DC resistance, 63
 negative capacitance and, 95
 return plane assumptions when using, 188
 serpentine corners modeling of, 190
 via inductance modeling of, 193–194
 workflow used in, 14, 172
 see also Capacitance, matrix
 see also Inductance, matrix
Fly by termination, *see* Termination, fly-by
Foil construction, 37–38
Fourier analysis, 4, 33
Fourier decoupling capacitor efficacy, 33
FR4, 37, 39–40
 alternatives to, 82, 86
 see also Laminate
Frequency domain, 4–7
 3dB bandwidth, 7–8, 26
 line spectrums, 4
Fundamental frequency, 4

G

Glass to resin ratio, 40–1, 65, 87
Glass weave, 40, 89
Glass-to-resin, *see* Resin, glass-to-resin ratio
Guard trace, 116–119
 connecting to, 116–117
 reflections, used to reduce, 210
 serpentine, used with, 191

H

Harmonics, 4–7
 loss effects, 79, 82
Hot Carrier Effects (HCE), 19
 see also Electrical overstress

I

I/O
 drivers, 6
 electrical overstress in, *see* Electrical overstress
 models, 17, 21–22
 Rds_on effects of, 139
 return path current, 56–57
 SSN and, 8–9
 transmission line effects on, 103–104
 worst case conditions, *see* Worst case analysis
 see also IBIS models
IBIS models, 21–22
Impedance, 50–53, 59
 antipad size effects, 193–194
 calculating, 71–72, 219–220, 226–228
 changes due to skin effect, 71
 conductance and, 67
 differential, *see* Differential impedance
 influence on crosstalk, *see* Crosstalk
 launched voltage and, 54–55
 loaded, 154
 loss effects, 83–84
 mismatch, *see* Reflections
 modal, 100, 102–103
 Thevenin, 142
 trace shape effects, 89
 see also Characteristic impedance
 see also Reflection coefficient
Incident
 current, 54–55, 125–126
 voltage, 54, 124–126, 130–131
 wave, 123, 126–128, 152, 155
 wave switching, 152

Inductance
- excess, *see* Chamfering
- external, 71
- impedance relationship of, *see* Transmission line circuit model
- matrix, 94
- microstrip, *see* Microstrip inductance
- microstrip and stripline differences, 68
- mutual, *see* Mutual inductance
- parallel plate, 68, 222
- resistance and, 71
- return path relationship of, 50, 52, 55
- stripline, *see* Stripline inductance
- trace shape effects, 89
- transmission line, *see* Transmission line circuit model

Inner layer traces, *see* Stripline
Insertion loss, 24–25
Interlayer capacitance, 57–58
Internal inductance, 68, 71
Intersymbol interference, 165, 212

J

Junction temperature, 17, 139, 144, 204, 207, 211

L

Laddering, 189–191
Laminate, 37–42
- alternate types characteristics of, 82, 86
- choosing advanced types, 200, 212–214
- dielectric constant values, 65
- glass weave effects, 89
- humidity and temperature effects, 87

Launched voltage, 54–55, 103–104, 123–124
Layout designer
- as part of deign flow, 15–16
- roll, 34–35

Level shift, 171
Line spectrum, 4–5
Loop resistance, 69–70, 75, 83–86
Loosely coupled
- crosstalk assumptions, 108
- differential pairs, 172–177
- effect on impedance, 102

Loss 50–51, 63–65, 79–90
- board edge increase, 187
- calculating, 223–225
- conductor, *see* Conductor loss
- correcting, 165
- dielectric, *see* Dielectric loss
- insertion, *see* Insertion loss
- reflection, *see* Reflection loss
- transmission line models, 23–24, 67

Loss Tangent, 9–10
- frequency effects on, 67
- loss effects, *see* Loss
- typical values, 40, 65

Low pass filter, 5, 7, 51, 212
LT, *see* Loss tangent

M

Meander trace, *see* Serpentine traces
Microstrip, 42–43, 63–75
- capacitance, 220, 221–222
- crosstalk, 109–119
- delay, 73–74, 100
- differential, 194–195
- effects of modes, 101–102
- embedded, 42–43, 101
- exposed impedance, 226–127
- inductance, 220, 221–122
- losses, 80, 82–88, 174, 225
- plating, 43–44
- solder mask covered impedance, 227–228

Mils, 43
- trace thickness relationship, *see* Copper weight

Mitering, 185–187
Models
- IBIS models, 21–22
- LTRA model, 27
- O element, 23
- RLGC, 23, 27–28
- T element, 23
- transistor level, 14, 21–22
- via, 191–93
- W element, 23–24
- worst case, 13, 16–18
- *see also* CAD
- *see also* Field solvers

Moisture uptake, 10, 41, 67, 87
Multidrop bus, 146–47, 150–154
- selecting Vterm for, 141–142

Mutual capacitance, 93–97
see also Crosstalk
see also Even modes
see also Odd modes
Mutual inductance, 94–95, 97
see also Crosstalk
see also Even mode
see also Odd modes

N

Near end crosstalk, *see* Crosstalk, near end
Negative capacitance, 95
Nepers, 218
NEXT, *see* Crosstalk, near end
Nonfunctional pad, *see* Pads, nonfunctional

O

O element, 23
see also Models
Odd mode, 99–103
delay, 100–101
impedance, 100
termination, *see* Differential pairs, termination
Offset stripline, 43, 56
Outer layer traces, *see* Microstrip
Overshoot, 3, 13, 17–18, 54–55, 150, 203, 210
upper bandwidth, 7

P

Pads, 38-39
capacitance of, 192
nonfunctional, 193
ee also Antipad
Parallel plate capacitor
transmission line circuit model, 50, 64–65
see also Antipad
Pin escape, 146–147
Plane capacitance
calculating, 222–223
Plane inductance
calculating, 222
Plane stitching, 179
Plated through holes, 37
Power delivery network, 8
see also Simultaneous switching noise
Power integrity, *see* Simultaneous switching noise
Power islands, 177, 214
Precompensation, 22, 165, 212
see also Equalization
Prepreg, 37–38, 41, 46
Propagation time, 51
coupling effects on, 98–101
dielectric constant effects on, 73–74
Proximity effect, 69–70
PTH, *see* Plated through holes
Pulse, 3
dispersion, 4, 80–81, 212, 216
frequency domain representation, *see* Frequency domain
overshoot, *see* Overshoot
reconstruction of, 6
reflection effects on, 127–128, 132
ringback *see* Ringback
transmission line effects and, 58

R

Ra, 46, 85–86
see also Surface roughness
Reactive discontinuities, 132–134
Reflection chart, 130–131
Reflection coefficient, 126–127
reflection chart and, 130
Reflection loss, 24, 26
S-Parameters and, 24
Reflections, 54, 58, 123–135
overstress voltage and, 19
S-parameters and, 24, 26
Relative dielectric constant, *see* Dielectric constant
Resin
Alternatives to FR4, 40
circuit board construction, 37
content, 40–41
delay effects on, 89
FR4, 39–41
glass-to-resin ratio, 40–41, 65, 87
see also Laminate
Resistance, *see* Trace, nominal DC resistance of
Resistivity, 45, 63–64, 218
see also Conductivity
Resistor network, 141

Return path, 49–50
avoiding splits in, 177
current, 55–58
loss relationship to, *see* Skin effect
see also Proximity effect
Reverse crosstalk coupling factor, *see* Crosstalk
Reverse treat foils, 45
see also Copper foils
Ringback 3, 17, 203, 210–211
Rise time, 3–4
bandwidth relationship, *see* Bandwidth
effects on crosstalk, *see* Crosstalk
harmonics effect on, *see* Harmonics
loss effects on, *see* Loss
lumped loads and, 152
mitered corners and, 185
power supply noise and, 17–18
SSN and, 8–9
series resistance effects on, 139
stubs effects on, 154
transmission line effects on, 58–59
worst case settings, 18
RLGC, *see* Models
Rolled copper, *see* Copper foils
Roughness, *see* Conductor loss, surface roughness

S

S glass, 40
see also Laminate
see also Resin
Scattering parameters, 24–26
Touchstone files and, 26
Schottky clamp, 144–45
Self capacitance, 93–7
see also Capacitance, matrix
see also Even mode
see also Odd mode
Self inductance, 93–95, 97
see also Inductance, matrix
see also Even mode
see also Odd mode
SERDES, 177
Serializer/deserializer, 177
Serpentine traces, 188–191, 212
corners and, 190
guarding of,191
laddering effects in,189
modeling of,190
segment length selection,190–191
Signal integrity, 1
workflow, 14–16
worst case analysis, importance of, 16–18
Signal return, *see* Return path
Simultaneous switching noise, 8, 55, 211
Single ended signaling, 161
Skin effect
inductance effect on, 68, 71
proximity effect relationship to, 69–70
resistance effect on, 69
surface roughness relationship to, 84–85
Solder mask, 41, 63
calculating impedance with, 227–228
effective dielectric constant, 65–66
effects on delay, 74–75
uncertainty when simulating, 115
see also Microstrip
S-parameters, *see* Scattering parameters
SPICE circuit simulator
agreement with calculation, 115
capacitance matrix and, 95
transmission line models, 4, 23, 26–27, 67
Square wave, 4–7
data pattern, 203
SSN, 8–9
board fixes, 211
data pattern, 204
Stripline, 38, 42–43, 56–58, 63–75
board edge effects on, 187–188
capacitance, 220
crosstalk, 109–110, 112, 114–118
delay,73, 75, 98–102
dual, 32, 200, 214–15
modal effect on, 98–101
impedance of, 71–72
impedance calculation of, 226–227
inductance of, 71, 220
losses, 67, 82–85, 87–88, 174, 212, 224
proper reference planes, 200
S-Parameter values for, 24–25
wavelength, 228–229
Surface roughness, 27, 45–46
influence on loss, 84–86, 213
photo illustrating, 47, 85

Symbols,164–165
    see also Intersymbol interference

T

TDR, *see* Time Domain Reflectometry
Termination, 32–33, 137–157
    AC,145–146
    clock, 145–146, 149–150
    differential,166–167, 169–172
    diode, 143–145
    fly-by, 148, 206
    in workflow, 15–17
    multidrop, 150–154
    multiple load point-to-point, 148–150
    parallel,139–141
    Schottky clamp, 144–145
    single-load point-to-point, 146–148
    source series, 137–139
    strategy, 206
    stubs and branches, 154–157
    Thevenin, 142–143, 148,151, 156, 211
    see also Reflections
Test patterns, 16, 32
    worst case, 18, 203–205
Through hole via, *see* Plated through holes
Time Dependent Dielectric Breakdown
        (TDDB), 19
    see also Electrical overstress
Time domain, 4–7
Time Domain Reflectometry, 35, 193–194
Tooth profile, 46
    see also Surface roughness
Touchstone files, 26
Trace, 42–43
    aggressor, *see* Aggressor
    buried, *see* Stripline
    calculating resistance, 219
    capacitance of microstrip, *see* Microstrip
        capacitance
    capacitance of stripline, *see* Stripline capacitance
    circuit model, 50–52
    guard, *see* Guard trace
    inductance of microstrip, *see* Microstrip
        inductance
    inductance of stripline, *see* Stripline
        inductance
    nominal DC resistance of, 63–64
    on surface, *see* Microstrip
    shape of, 46–47
    thickness of, *see* Copper weight
Transatlantic telegraph cable, 1–3
Transmission line
    calculating loss from, 223
    circuit model, 50–51, 83, 94
    importance of return path, 49–51
    modeling, 23
    SPICE matrix, 95
    see also Models
Trombone lines, *see* Serpentine traces

V

Via, 38
    aspect ratio, 39
    blind, 38–39
    buried, 38
    capacitance of, 192–193
    differential, 178
    inductance, 193
    model of, 191–192
    plating to form, 38
Voltage divider
    launched voltage and, 54, 103, 130, 153,
        166, 168
    termination, 140

W

Wavelength, 228–229
Workflow, *see* Signal integrity, workflow
Worst case analysis, 13, 16–18
    DVT, 19
    Patterns for creating, 203–205
    Termination and, 142, 201
    Trace environmental effects and, 63, 103
    Silicon conditions and, 17, 21