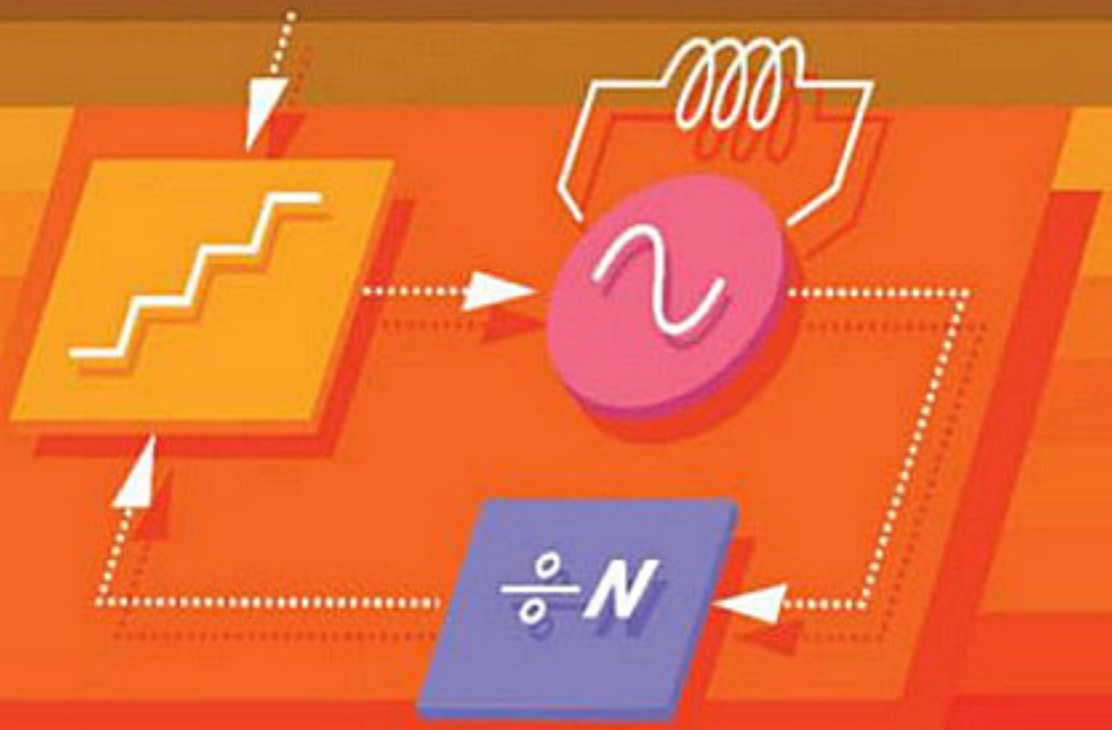
Behzad Razavi



DESIGN OF CMOS PHASE-LOCKED LOOPS

From Circuit Level to Architecture Level

Design of CMOS Phase-Locked Loops

Using a modern, pedagogical approach, this textbook gives students and engineers a comprehensive and rigorous knowledge of CMOS PLL design for a wide range of applications. It features intuitive presentation of theoretical concepts, built up gradually from their simplest form to more practical systems; broad coverage of key topics, including oscillators, phase noise, analog PLLs, digital PLLs, RF synthesizers, delay-locked loops, clock and data recovery circuits, and frequency dividers; tutorial chapters on high-performance oscillator design, covering fundamentals to advanced topologies; and extensive use of circuit simulations to teach design mentality, highlight design flaws, and connect theory with practice. Offering over 200 thought-provoking examples that demonstrate best practices and common pitfalls, 250 end-of-chapter homework problems to test and enhance the readers' understanding, and solutions and lecture slides for instructors, this is the perfect text for senior undergraduate and graduate-level students and professional engineers who want an in-depth understanding of PLL design.

Behzad Razavi is Professor of Electrical Engineering at The University of California, Los Angeles. He has received numerous teaching and education awards, and is a member of the US National Academy of Engineering and a Fellow of the IEEE. His previous textbooks include *Fundamentals of Microelectronics, RF Microelectronics* and *Design of Analog CMOS Integrated Circuits.*

Design of CMOS Phase-Locked Loops

From Circuit Level to Architecture Level

BEHZAD RAZAVI
University of California, Los Angeles

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre, New Delhi – 110025, India

79 Anson Road, #06–04/06, Singapore 079906

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning, and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9781108494540
DOI: 10.1017/9781108626200

© Cambridge University Press 2020

This publication is in copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published 2020

Printed in the United Kingdom by TJ International Ltd, Padstow Cornwall

A catalogue record for this publication is available from the British Library.

Library of Congress Cataloging-in-Publication Data

ISBN 978-1-108-49454-0 Hardback

Cambridge University Press has no responsibility for the persistence or accuracy of URLs for external or third-party internet websites referred to in this publication and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

To my brother Hossein,
who has always been there for me

Contents

Preface xiii

Acknowledgments xiv

About the Author xvi

1 Oscillator Fundamentals 1
1.1 Basic Concepts . . . 1
1.2 Oscillatory Feedback System . . . 2
1.3 A Deeper Understanding . . . 4
1.4 Basic Ring Oscillators . . . 9
1.4.1 Inverter-Based Rings . . . 12
1.5 Basic LC Oscillators . . . 13
1.5.1 LC Circuit Concepts . . . 13
1.5.2 LC Oscillators as Feedback Systems . . . 17
1.5.3 LC Oscillators as One-Port Systems . . . 23
1.6 Voltage-Controlled Oscillators . . . 27
1.7 Appendix I . . . 30

2 Introduction to Jitter and Phase Noise 34
2.1 Brief Review of Noise . . . 34
2.1.1 Noise in Time and Frequency Domains . . . 34
2.1.2 Device Noise . . . 35
2.1.3 Propagation of Noise . . . 37
2.1.4 Average Power of Noise . . . 38
2.1.5 Approximation of Noise Spectrum . . . 39
2.1.6 Accumulation of Noise with Time . . . 39
2.2 Basic Jitter and Phase Noise Concepts . . . 40
2.2.1 Jitter . . . 41
2.2.2 Phase Noise . . . 44
2.2.3 Limitations of Narrowband FM Approximation . . . 49
2.2.4 Relationship between Jitter and Phase Noise . . . 50
2.2.5 Types of Jitter . . . 52
2.3 Trade-Off Between Phase Noise and Power . . . 53
2.4 Basic Phase Noise Mechanisms . . . 55
2.4.1 Phase Noise versus Frequency Noise . . . 55
2.4.2 Ring Oscillators . . . 56
2.4.3 LC Oscillators . . . 58
2.5 Effect of Jitter on Performance . . . 58
2.6 Effect of Phase Noise on Performance . . . 59

3 Design of Inverter-Based Ring Oscillators 64
3.1 Phase Noise in Ring Oscillators . . . 64
3.1.1 General Equation . . . 64
3.2 Preliminary Design Ideas . . . 70
3.3 Obtaining the Desired Frequency . . . 72
3.3.1 Greater Node Capacitances . . . 73
3.3.2 Greater Number of Stages . . . 74
3.3.3 Greater Transistor Lengths . . . 74
3.3.4 Frequency Division . . . 75

3.4 Phase Noise Considerations . . . 75
3.4.1 Transistor Noise Simulations . . . 76
3.4.2 Reference Oscillator Phase Noise . . . 78
3.4.3 First 2-GHz Oscillator Phase Noise . . . 80
3.4.4 Second 2-GHz Oscillator Phase Noise . . . 80
3.4.5 Third 2-GHz Oscillator Phase Noise . . . 81
3.4.6 Fourth 2-GHz Oscillator Phase Noise . . . 82
3.5 Frequency Tuning . . . 83
3.5.1 Tuning Considerations . . . 83
3.5.2 Continuous and Discrete Tuning . . . 84
3.5.3 Tuning by Variable Resistance . . . 85
3.5.4 Tuning by Variable Capacitance . . . 90
3.6 Discrete Frequency Tuning . . . 93
3.7 Problem of Supply Noise . . . 96
3.7.1 Voltage Regulation . . . 96
3.7.2 Current Regulation . . . 97

4 Design of Differential and Multiphase Ring Oscillators 103
4.1 General Considerations . . . 103
4.2 Phase Noise Considerations . . . 105
4.3 Basic Differential Ring Design . . . 108
4.3.1 Initial Design . . . 108
4.3.2 Design Improvements . . . 111
4.4 Obtaining the Desired Frequency . . . 112
4.4.1 Method 1: Greater Node Capacitances . . . 113
4.4.2 Method 2: Larger Transistors . . . 114
4.4.3 Method 3: Greater Number of Stages . . . 115
4.5 Two-Stage Ring Oscillators . . . 116
4.5.1 Basic Idea . . . 116
4.5.2 Design Example . . . 120
4.6 Linear Scaling . . . 121
4.7 Tuning Techniques . . . 122
4.7.1 Resistive Tuning . . . 122
4.7.2 Varactor Tuning . . . 123
4.7.3 Tuning the Number of Stages . . . 124
4.8 Comparison of Inverter-Based and Differential Rings . . . 126
4.9 Inverter-Based Oscillators with Complementary or Quadrature Outputs . . . 126
4.9.1 Coupled Oscillators . . . 126
4.9.2 Phase Noise Considerations . . . 128
4.9.3 Direct Quadrature Generation . . . 129
4.9.4 Quadrature Generation by Interpolation . . . 130
4.10 Ring Oscillators with LC Loads . . . 134

5 LC Oscillator Design 138
5.1 Inductor Modeling . . . 138
5.2 Phase Noise Analysis . . . 143
5.2.1 A Simple Case . . . 144
5.2.2 Cyclostationary Noise . . . 146
5.2.3 Noise Injected by Cross-Coupled Pair . . . 149
5.2.4 Phase Noise Calculation . . . 153
5.3 Tail Noise . . . 155
5.3.1 Tail Thermal Noise . . . 155

5.3.2 Tail Flicker Noise . . . 158
5.4 Effect of Tail Capacitance . . . 160
5.5 Step-by-Step Design . . . 161
5.5.1 Preliminary Thoughts . . . 161
5.5.2 Design Example . . . 162
5.5.3 Frequency Tuning . . . 166
5.5.4 Summary of Oscillator Design Procedure . . . 173

6 Advanced Oscillator Concepts 178
6.1 Phase Noise Analysis by Impulse Response . . . 178
6.1.1 Phase Impulse Response . . . 178
6.1.2 Effect of Flicker Noise . . . 184
6.1.3 Cyclostationary Noise . . . 186
6.2 Current-Limited versus Voltage-Limited Phase Noise . . . 186
6.3 Oscillators with Complementary Cross-Coupled Pairs . . . 186
6.3.1 Design Issues . . . 188
6.3.2 Design Example . . . 190
6.4 Class-C Oscillators . . . 192
6.4.1 Design Example . . . 193
6.5 Phase Noise Reduction by Frequency Division . . . 193
6.6 Quadrature Generation Techniques . . . 195
6.6.1 Frequency Division . . . 195
6.6.2 Quadrature LC Oscillators . . . 196

7 Basic PLL Architectures 209
7.1 Phase Detectors . . . 209
7.2 Phase Control by Feedback . . . 211
7.3 Analysis of Simple PLL . . . 213
7.3.1 Static Behavior . . . 213
7.3.2 Frequency Multiplication . . . 215
7.3.3 Dynamic Behavior . . . 215
7.3.4 PLL Transfer Function . . . 219
7.3.5 Drawbacks of Simple PLL . . . 223
7.4 Phase/Frequency Detector . . . 224
7.5 Charge-Pump PLLs . . . 227
7.5.1 Charge Pumps . . . 227
7.5.2 PFD/CP/Capacitor Cascade . . . 228
7.5.3 Basic Charge-Pump PLL . . . 228
7.5.4 PFD/CP/Capacitor Transfer Function . . . 229
7.5.5 Phase Margin Calculation . . . 236
7.6 Higher-Order Loops . . . 238
7.7 Basic Charge Pump Topologies . . . 242
7.8 Settling Time . . . 244

8 PLL Design Considerations 246
8.1 More on PLL Transfer Functions . . . 246
8.1.1 Limitations of Continuous-Time Approximation . . . 250
8.2 PFD Issues . . . 251
8.3 Charge Pump Issues . . . 251
8.3.1 Up and Down Skew . . . 251
8.3.2 Voltage Compliance and Channel-Length Modulation . . . 252
8.3.3 Random Mismatches . . . 254

8.3.4 Clock Feedthrough and Charge Injection . . . 255
8.3.5 Other Charge Pump Nonidealities . . . 255
8.4 Improved Charge Pumps . . . 257
8.5 PLLs with Discrete VCO Tuning . . . 259
8.6 Ripple Reduction by Sampling Filter . . . 260
8.7 Loop Filter Leakage . . . 260
8.8 Filter Capacitor Reduction . . . 261
8.9 Trade-Off Between Bandwidth and Spur Level . . . 262
8.10 Phase Noise in PLLs . . . 263
8.10.1 Shaping of Input Phase Noise . . . 263
8.10.2 Shaping of VCO Phase Noise . . . 266
8.10.3 Charge Pump Noise . . . 269
8.10.4 Loop Filter Noise . . . 272
8.10.5 Supply Noise . . . 274

9 PLL Design Study **277**
9.1 Design Procedure . . . 277
9.2 PFD Design . . . 278
9.3 Charge Pump Design . . . 278
9.3.1 First CP Design . . . 278
9.3.2 Second CP Design . . . 282
9.3.3 Third CP Design . . . 282
9.3.4 Fourth CP Design . . . 283
9.3.5 PFD/CP Interface . . . 283
9.4 Behavioral Simulations of PLL . . . 283
9.4.1 Loop Simplification . . . 283
9.4.2 Loop Dynamics . . . 286
9.4.3 Effect of Ripple . . . 287
9.5 Simulation of the PLL Transfer Function . . . 289
9.5.1 One-Pole Approximation . . . 289
9.5.2 Use of Input FM Source . . . 289
9.5.3 Use of Random Phase Modulation . . . 291
9.6 Effect of VCO Phase Noise . . . 293
9.6.1 VCO Phase Noise Model . . . 293
9.6.2 VCO Phase Noise Suppression . . . 293
9.7 Loop Filter Noise . . . 295
9.8 Doubling the Reference Frequency . . . 296
9.8.1 Doubler Design . . . 296
9.8.2 Frequency Doubling Issues . . . 296
9.8.3 PLL Redesign with Doubled Reference . . . 298
9.8.4 PLL Simulations . . . 298
9.9 Feedback Divider Design . . . 299
9.9.1 Topology Selection . . . 299
9.9.2 Divider Circuit Design . . . 300
9.10 Use of Lock Detectors for Calibration . . . 301
9.11 Design Summary . . . 303

10 Digital Phase-Locked Loops **305**
10.1 Basic Idea . . . 305
10.2 ADC Basics . . . 306
10.2.1 Quantization . . . 306
10.2.2 Flash ADC . . . 307

10.2.3 Interpolation . . . 308
10.3 Time-to-Digital Conversion . . . 309
10.3.1 Basic TDC Topology . . . 310
10.3.2 Effect of Quantization Noise . . . 312
10.3.3 TDC Dynamic Range . . . 314
10.3.4 TDC Imperfections . . . 315
10.4 Transistor-Level TDC Design . . . 316
10.5 Improved TDCs . . . 319
10.5.1 Vernier TDC . . . 319
10.5.2 Multi-Path TDCs . . . 321
10.6 TDC/Oscillator Combinations . . . 323
10.7 Digitally-Controlled Oscillators . . . 326
10.7.1 Problem of Discrete Frequencies . . . 326
10.7.2 DAC Principles . . . 328
10.7.3 Matrix Architecture . . . 331
10.7.4 Coarse/Fine DACs . . . 333
10.7.5 DCO Topologies . . . 333
10.8 Loop Dynamics . . . 341
10.8.1 Digital Filter Implementation . . . 341
10.8.2 Correspondence between Analog and Digital PLLs . . . 342

11 Delay-Locked Loops **346**
11.1 Basic Idea . . . 346
11.2 Loop Dynamics . . . 347
11.3 Choice of Number of Delay Stages . . . 349
11.4 Effect of Nonidealities . . . 349
11.4.1 PFD/CP Nonidealities . . . 349
11.4.2 Supply Noise . . . 350
11.4.3 Phase Noise . . . 351
11.5 Generation of Multiple Phases . . . 352
11.6 Frequency-Multiplying DLLs . . . 354
11.6.1 Basic Topologies . . . 354
11.6.2 Design Issues . . . 357
11.6.3 Use of Frequency Multiplication in False Lock Detection . . . 358
11.7 DLL/PLL Hybrids . . . 358
11.8 Phase Interpolation . . . 360
11.9 High-Speed PD Design . . . 363
11.10Duty Cycle Correction . . . 364

12 RF Synthesis **368**
12.1 RF Synthesis Requirements . . . 368
12.2 Integer-N Synthesizers . . . 370
12.3 Fractional-N Synthesizers . . . 371
12.3.1 The Need for Modulus Randomization . . . 373
12.3.2 Noise Shaping . . . 376
12.3.3 Discrete-Time Model . . . 379
12.3.4 $\Delta\Sigma$ Fractional-N Synthesizers . . . 382
12.3.5 Higher-Order $\Delta\Sigma$ Modulators . . . 384
12.4 Nonlinearities in Fractional-N Loops . . . 392
12.4.1 Charge Pump Nonlinearity . . . 393
12.4.2 Charge Pump Settling Behavior . . . 395
12.5 Reduction of Quantization Noise . . . 396

12.5.1 DAC Feedforward . . . 396
12.5.2 Noise Cancellation by DTC . . . 399
12.5.3 Reference Frequency Doubling . . . 400

13 Clock and Data Recovery Fundamentals 404
13.1 General Considerations . . . 404
13.2 Properties of Random Binary Data . . . 406
13.2.1 Spectrum of NRZ data . . . 406
13.3 Clock Recovery by Edge Detection . . . 408
13.4 Clock Recovery by Phase-Locking . . . 410
13.4.1 Bang-Bang Phase Detector . . . 411
13.4.2 Alexander Phase Detector . . . 416
13.4.3 Hogge Phase Detector . . . 423
13.5 Problem of Data Swings . . . 426

14 Advanced Clock and Data Recovery Principles 428
14.1 Half-Rate Phase Detectors . . . 428
14.1.1 Half-Rate Bang-Bang PDs . . . 428
14.1.2 Half-Rate Linear PDs . . . 431
14.2 Oscillatorless CDR Architectures . . . 434
14.2.1 DLL-Based CDR Circuits . . . 434
14.2.2 PI-Based CDR Circuits . . . 435
14.2.3 Digital CDR Circuits . . . 437
14.3 Frequency Acquisition . . . 443
14.4 Jitter Characteristics . . . 444
14.4.1 Jitter Generation . . . 445
14.4.2 Jitter Transfer . . . 445
14.4.3 Jitter Tolerance . . . 448

15 Frequency Dividers 453
15.1 General Considerations . . . 453
15.2 Latch Design Styles . . . 454
15.2.1 Static Latches . . . 455
15.2.2 Dynamic Latches . . . 459
15.3 Divide-by-2 Circuit Design . . . 464
15.4 Dual-Modulus Prescalers . . . 467
15.5 Divider Design for RF Synthesis . . . 472
15.5.1 Pulse Swallow Divider . . . 472
15.5.2 Vaucher Divider . . . 474
15.6 Miller Divider . . . 476
15.7 Injection-Locked Dividers . . . 478
15.8 Fractional Dividers . . . 480
15.9 Divider Delay and Phase Noise . . . 482

Index 486

Preface

A quick search on Google brings up nearly two dozen books on PLLs. So why another one? This book addresses the need for a text that methodically teaches modern CMOS PLLs for a wide range of applications. The objective is to teach the reader how to approach PLLs from transistor-level design to architecture development.

Based on 25 years of teaching courses on the subject and the latest trends in industry, this book deals with oscillators, phase noise, analog phase-locked loops, digital phase-locked loops, RF synthesizers, delay-locked loops, clock and data recovery circuits, and frequency dividers. The objective is to reach a broad spectrum of readers while maintaining a cohesive flow.

As with my past writings, I have implemented a multitude of pedagogical tools to help the reader learn efficiently—and experience the pleasure of learning. One principle that I uphold in writing is to start with the simplest possible arrangement, teach how it works and what shortcomings it has, and then add components to it to improve its performance. This approach allows the reader to see how a basic architecture evolves to a complex system. After laying the theoretical foundation for each topic, I present a step-by-step design flow and proceed to design the circuit.

And not all design efforts are successful. The reader can clearly see how certain decisions lead to a dead end and how we revise these decisions to reach a new, more practical solution. This exploratory mentality not only makes the process of learning more exciting but also helps the reader see why each component is necessary, what criteria govern its choice, and what not to do.

A unique aspect of this book is its extensive use of simulations to teach design and investigate agreement between theory and practice. For each design, I use the theoretical basis to choose certain parameters and predict the performance, and then I simulate the circuit. If the simulation results do not agree with the predictions, I delve into the details and determine why. Another unique aspect of this book is that it leverages concepts from one field (e.g., wireless technology) to another (e.g., wireline communications) by bringing the vast knowledge in these fields under one roof.

A website for the book provides additional resources for readers and instructors, including Powerpoint slides and a solutions manual.

Behzad Razavi
September 2019

Acknowledgments

It gives me great pleasure to express my gratitude to the many people who have contributed to this book. First, an army of reviewers:

Morteza Alavi, Delft University of Technology
Tejasvi Anand, Oregon State University
Pietro Andreani, Lund University
Baktash Behmanesh, Lund University
Amir Bozorg, University College Dublin
Farhad Bozorgi, University of Pavia
Francesco Brandonisio, Infineon
Brian Buell, IDT
Wei Sung Chang, Mediatek
Chienwen Chen, Realtek
Nick Young Chen, University of Macau
Peng Chen, University College Dublin
Yung-Tai Chen, Realtek
Dmytro Cherniak, Infineon
Murat Demirkan, Analog Devices
Jianglin Du, University College Dublin
Alper Eken, Analog Devices
Eythan Familier, Maxlinear
Zhong Gao, University College Dublin
Werner Grollitsch, Infineon
Jane Gu, UC Davis
Ehsan Hadizadeh, University of British Columbia
Zhiqiang Hang, Samsung
Shilei Hao, Qualcomm
Aravind Heragu, Semitech
Cheng Ru Ho, University of Southern California
Ali Homayoun, Movandi
Kenny Hsieh, TSMC
Chien Hsueh, UC San Diego
Yizhe Hu, University College Dublin
Cheng-Chang Huang, National Taiwan University
Rulin Huang, UCLA
Zue-Der Huang, Realtek
Jeongho Hwang, Seoul National University
Milad Kalantari, Hong Kong Univesity of Science and Technology
Nader Kalantari, ArianRF
Sivash Kananian, Stanford University
Alireza Karimi, UC Irvine
Ting-Kuei Kuan, TSMC
Tai-Cheng Lee, National Taiwan University
Guansheng Li, Broadcom
Bangan Liu, Tokyo Institute of Technology
Hanli Liu, Tokyo Institute of Technology
Lorenzo Lotti, UC Berkeley
Enrico Mammei, ST Microlectronics
Abishek Manian, Texas Instruments
Dennis Andrade Miceli, University College Dublin
Aravind Narayanan, Ericsson
Reza Nikandish, University College Dublin
Ali Nikoofard, UC San Diego
Minh Hieu Nguyen, University College Dublin
Viet Nguyen, University College Dublin
Fabio Padovan, Infineon
Kwanseo Park, Seoul National University
Fabio Quadrelli, Infineon
Ehsan Rahimi, University of Pavia
Wahid Rahman, Alphawave
Negar Reiskarimian, MIT
Sujiang Rong, TransaSemi
Jhoan Salinas, University of Toronto
Saurabh Saxena, IIT Madras
Mohamed Shehata, University College Dublin
Fei Song, Ubilinx
Mrunmay Talegaonkar, Inphi
Bortecene Terlemez, IDT
Howie Tu, Realtek
Eric Vandel, Semitch
Marco Vigilante, Qualcomm
Haisong Wang, Semitech
Zisong Wang, UC Irvine

Zhongzheng Wang, University College Dublin
Ying Wu, Delft University of Technology
Kai Xu, University College Dublin
Yu-Che Yang, Realtek
Sheng Ye, Maxlinear
Jun Yin, University of Macau
Hongyang Zhang, University of Pavia
Yan Zhang, UCLA

My special thanks go to Pietro Andreani, who heroically reviewed almost the entire book. Grazie, Piero! In addition, Pavan Hanumolu, Long Kong, and Bogdan Staszewski have kindly answered my many questions and helped enrich the book.

I also wish to thank the staff at Cambridge University Press for their support during the production of the book, particularly, Elizabeth Horne, Charles Howell, and Beverley Lawrence.

My wife, Angelina, typed the entire book. I am very grateful to her.

Behzad Razavi

About the Author

Behzad Razavi received the BSEE degree from Sharif University of Technology in 1985 and the MSEE and PhDEE degrees from Stanford University in 1988 and 1992, respectively. He was with AT&T Bell Laboratories and Hewlett-Packard Laboratories until 1996. Since 1996, he has been Associate Professor and subsequently Professor of electrical engineering at University of California, Los Angeles. His current research includes wireless and wireline transceivers and data converters.

Professor Razavi was an Adjunct Professor at Princeton University from 1992 to 1994, and at Stanford University in 1995. He served on the Technical Program Committees of the International Solid-State Circuits Conference (ISSCC) from 1993 to 2002 and VLSI Circuits Symposium from 1998 to 2002. He has also served as Guest Editor and Associate Editor of the *IEEE Journal of Solid-State Circuits, IEEE Transactions on Circuits and Systems,* and *International Journal of High Speed Electronics*. He presently serves as the Editor-in-Chief of the *IEEE Solid-State Circuits Letters*.

Professor Razavi received the Beatrice Winner Award for Editorial Excellence at the 1994 ISSCC, the best paper award at the 1994 European Solid-State Circuits Conference, the best panel award at the 1995 and 1997 ISSCC, the TRW Innovative Teaching Award in 1997, the best paper award at the IEEE Custom Integrated Circuits Conference in 1998, and the McGraw-Hill First Edition of the Year Award in 2001. He was the co-recipient of both the Jack Kilby Outstanding Student Paper Award and the Beatrice Winner Award for Editorial Excellence at the 2001 ISSCC. He received the Lockheed Martin Excellence in Teaching Award in 2006, the UCLA Faculty Senate Teaching Award in 2007, and the CICC Best Invited Paper Award in 2009 and in 2012. He was the co-recipient of the 2012 and the 2015 VLSI Circuits Symposium Best Student Paper Awards and the 2013 CICC Best Paper Award. He was also recognized as one of the top 10 authors in the 50-year history of ISSCC. He received the 2012 Donald Pederson Award in Solid-State Circuits. He was also the recipient of the American Society for Engineering Education PSW Teaching Award in 2014. Professor Razavi is a member of the US National Academy of Engineering. He received the 2017 IEEE CAS John Choma Education Award.

Professor Razavi has served as an IEEE Distinguished Lecturer and is a Fellow of IEEE. He is the author of *Principles of Data Conversion System Design* (IEEE Press, 1995), *RF Microelectronics* (Prentice Hall, 1998, 2012) (translated to Chinese, Japanese, and Korean), *Design of Analog CMOS Integrated Circuits* (McGraw-Hill, 2001, 2016) (translated to Chinese, Japanese, and Korean), *Design of Integrated Circuits for Optical Communications* (McGraw-Hill, 2003, Wiley, 2012), and *Fundamentals of Microelectronics* (Wiley, 2006) (translated to Korean, Portuguese, and Turkish), and the editor of *Monolithic Phase-Locked Loops and Clock Recovery Circuits* (IEEE Press, 1996), and *Phase-Locking in High-Performance Systems* (IEEE Press, 2003).

1

Oscillator Fundamentals

At the heart of every phase-locked loop lies an oscillator, playing a critical role in the performance that can be achieved. For this reason, we devote five chapters of this book to oscillator design. This chapter aims to build a solid foundation for general oscillator concepts before we delve into high-performance design in Chapters 3-6. We begin with basic concepts and discover how a negative-feedback system can oscillate. We then extend our view to ring and LC oscillators.

1.1 Basic Concepts

If we release a pendulum from an angle, it swings for a while and gradually comes to a stop. The "oscillation" begins because the original potential energy turns into kinetic energy as the pendulum reaches its vertical position (Fig. 1.1), allowing it to continue its trajectory to the other extreme angle (position 3), at which the

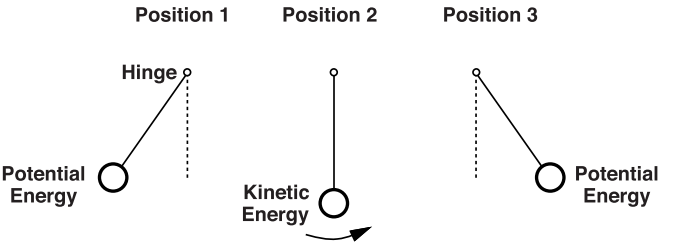


Figure 1.1 A pendulum acting as an oscillatory system.

energy is again in potential form. The oscillation stops because the friction at the hinge and the air resistance convert some of the pendulum's energy to heat in every oscillation period.

In order to sustain the oscillation, we can provide external energy to the pendulum so as to compensate for the loss caused by the hinge and the air. For example, if we give the pendulum a gentle push each time it returns to position 1, it will continue to swing. If the push is too weak, we undercompensate, allowing the oscillation to die; if the push is too strong, we overcompensate, forcing the swing amplitude to increase from one cycle to the next. We also note that the *period* of oscillation is independent of the amplitude.[1]

The above mechanical example points to several ingredients of an oscillatory system: (1) an initial "imbalance," i.e., an initial condition or packet of energy (provided by bringing the pendulum to position 1); (2) a tendency for one type of energy to turn into another and vice versa; and (3) a sustaining mechanism that replenishes the energy lost due to inevitable imperfections. Not all oscillator circuits contain all of these ingredients, but it is helpful to bear these concepts in mind.

[1] This is true only if the pendulum oscillates with a small amplitude.

Example 1.1

Repeat the foregoing experiments with a "lossless" pendulum.

Solution

If released from an angle, such a pendulum oscillates indefinitely. Now, if we give a push each time the pendulum reaches the left end, then the swing keeps increasing due to the additional energy that we inject into the system in each cycle. Note that this indefinite growth does not occur if we give a push at a frequency other than the pendulum's natural oscillation frequency.

The above example serves as a guide in our analysis: if a system has a tendency to oscillate at a frequency ω_0, then it creates a *growing* oscillatory output in response to an external injection at a frequency ω_0. From another perspective, such a system indefinitely amplifies a periodic input at this frequency.

1.2 Oscillatory Feedback System

We know from basic analog design that a negative-feedback system can become unstable. We exploit this property to construct oscillators.

Let us first study oscillation in the frequency domain. Consider the feedback system shown in Fig. 1.2(a), where the negative sign at one adder input signifies negative feedback at low frequencies. Depicted in Fig.

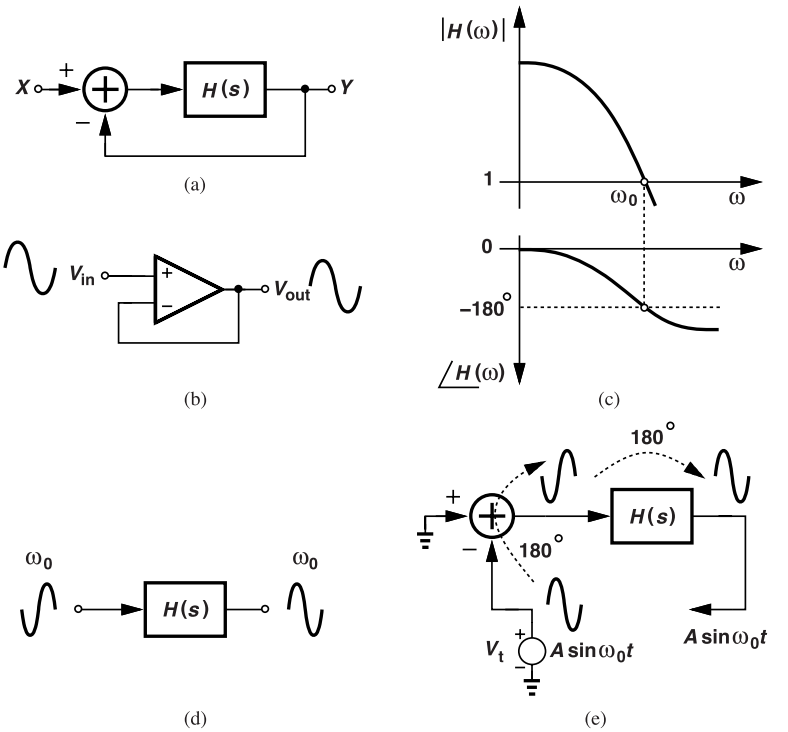


Figure 1.2 (a) Simple feedback system, (b) realization using an op amp, (c) open-loop frequency response showing zero phase margin, (d) signal inversion at ω_0, and (e) propagation of a sinusoid at ω_0 around the loop.

1.2(b) is an implementation example, which, in response to a low-frequency sinusoidal input, simply acts as a unity-gain buffer. Note that the op amp exhibits negligible phase shift at low frequencies.

How can the arrangements in Figs. 1.2(a) or (b) oscillate? Writing the closed-loop transfer function of the former as

$$\frac{Y(s)}{X(s)} = \frac{H(s)}{1+H(s)}, \tag{1.1}$$

we observe that the denominator falls to zero if $H(s) = -1$ for some value of s. If $X(s)$ is a sinusoid, then $s = j\omega_0$ and we must have $H(j\omega_0) = -1$. The open-loop frequency response thus exhibits a unity magnitude and a 180° phase shift at ω_0 [Fig. 1.2(c)]. We note that $(Y/X)(j\omega_0) \to \pm\infty$, concluding that the system provides an *infinite* gain for such a sinusoid. As surmised in the previous section, this scenario suggests an oscillatory loop.

Let us examine the condition $H(j\omega_0) = -1$ more closely: this equality means that $H(s)$ itself *inverts* the input at this frequency [Fig. 1.2(d)]. That is, $H(s)$ has so much phase shift (or delay) at ω_0 that the overall feedback becomes positive. This can be seen by setting the main input, X, to zero, breaking the loop, and applying a stimulus at this frequency [Fig. 1.2(e)], and following it around the loop. The returned signal is in phase with the test voltage, V_t. We say the loop contains a 180° phase shift due to the nominally negative feedback and another frequency-dependent 180° phase shift arising from $H(s)$. These two phase shifts must not be confused with each other.

The total phase shift of 360° at ω_0 implies that the signal returns to enhance itself as it circulates around the loop. This phenomenon results in amplitude growth because the returned signal is at least as large as the starting signal, i.e., because $|H(j\omega_0)| = 1$. We therefore summarize the conditions for oscillation as

$$|H(j\omega_0)| = 1 \tag{1.2}$$

$$\angle H(j\omega_0) = 180^\circ, \tag{1.3}$$

which are called "Barkhausen's" criteria. We also call $H(j\omega_0) = -1$ the "startup condition." Note that $H(j\omega)$ generally has a complex value at $\omega \neq \omega_0$ and becomes real only at ω_0.

The oscillation buildup can also be studied in the time domain. We begin with the arrangement shown in Fig. 1.3(a) and note that, with $H(j\omega_0) = -1$, the output is equal to the input but shifted by 180°. If the loop

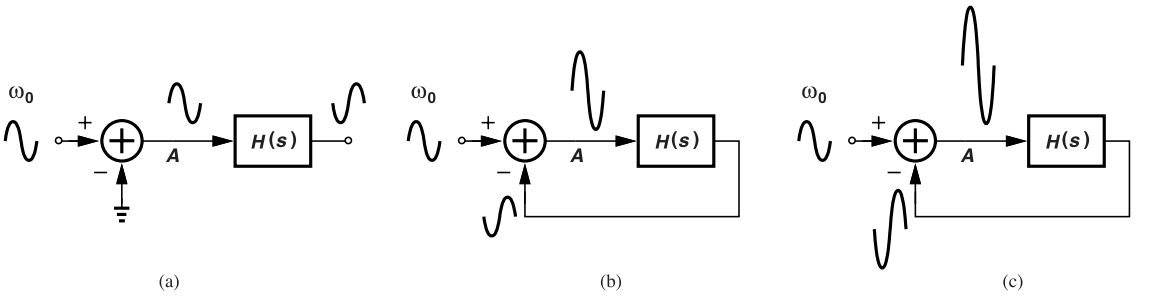


Figure 1.3 Growth of an input sinusoid around the loop with time.

is closed [Fig. 1.3(b)], the output is *subtracted* from the input, yielding a larger swing at A. This signal is again inverted and subtracted from the input, leading to indefinite grow of the amplitude [Fig. 1.3(c)].

In summary, a negative-feedback system can generate a growing periodic output in response to a sinusoidal input if its loop gain reaches -1 at a finite frequency, ω_0. But does such a system oscillate if we apply no input? Yes, the wideband noise of the devices within the loop exhibits a finite energy in the vicinity of ω_0, producing a small component that circulates around the loop and causes oscillation. For example, as shown in Fig. 1.4, a noise source, V_n, at the input of $H(s)$ yields an output given by

$$Y = V_n \frac{H(s)}{1+H(s)}, \tag{1.4}$$

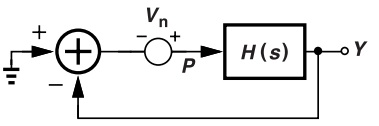


Figure 1.4 Effect of noise injected into a closed-loop system.

thereby experiencing infinite gain at $s = j\omega_0$. That is, even though V_n is infinitesimally small at ω_0, Y can assume a finite swing.

The foregoing analysis suggests that, to test for oscillation, we can inject a sinusoidal input at *any* point and observe the response at *any* point so long as both points are within the loop.[2] In Fig. 1.4, for example, V_n can be placed at the output of $H(s)$. Similarly, the point of observation can be P rather than Y. By the same token, the injection and observation points can be the same, pointing to another method of finding the oscillation conditions that is suited to some circuits. We inject a current at ω_0 into a node within the loop and examine the voltage at that node, i.e., we compute the impedance. If the voltage and hence the impedance go to infinity at ω_0, the circuit can oscillate. Figure 1.5 depicts the concept. We return to this point in Section 1.5.3.

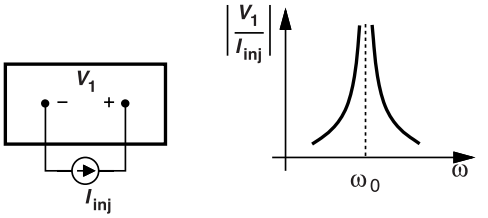


Figure 1.5 Infinite port impedance in an oscillatory circuit.

1.3 A Deeper Understanding

Analysis Methods In the analysis of oscillators, we first wish to determine how the devices and the bias conditions must be chosen so as to guarantee oscillation. Our previous studies point to three methods using the small-signal model of the circuit:

1. Open the loop and enforce the startup condition, $H(j\omega_0) = -1$, thus obtaining the circuit design requirements.
2. In analogy with the ideal pendulum example, release the closed-loop circuit with an initial condition and determine the design parameters for oscillation. The initial condition can be created by an impulse of current injected into a node within the loop or simply by assuming a finite voltage on a capacitor.
3. Inject a sinusoidal current into a node in the closed-loop circuit and compute the conditions necessary for the impedance seen at this node to go to infinity. This method is not universal but still proves helpful.

A given oscillator topology may lend itself more easily to one method than to another. The following examples illustrate these thoughts.

[2] An injection point is considered to be within the loop if the transfer function from that point to the output is not zero.

Example 1.2

A common-source (CS) stage is placed in a feedback loop as shown in Fig. 1.6(a). Can the circuit oscillate?

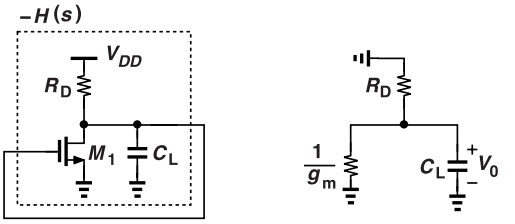


Figure 1.6 (a) Feedback around a CS stage, and (b) equivalent circuit.

Neglect other capacitances and channel-length modulation.

Solution

To retain consistency with the block diagram of Fig. 1.2(a) and noting that the CS stage inverts at low frequencies, we denote the circuit in the dashed box by $-H(s)$ (why?). Applying the first analysis method, we write

$$H(s) = g_m \left(R_D || \frac{1}{C_L s} \right) \tag{1.5}$$

$$= g_m \frac{R_D}{R_D C_L s + 1}. \tag{1.6}$$

Owing to its single pole, $H(s)$ can contribute a maximum phase of -90° (at infinite frequency), disallowing $H(j\omega_0) = -1$. Thus, the loop cannot oscillate.

Let us try the second method by applying an initial condition to C_L. Since M_1 operates as a diode-connected device, the small-signal model reduces to that shown in Fig. 1.2(b), revealing that C_L simply discharges through $R_D || g_m^{-1}$ and no oscillation occurs. Similarly, the third method gives an impedance of $R_D || g_m^{-1} || (C_L s)^{-1}$ seen at the output node, indicating that it cannot go to infinity at any $s = j\omega$.

Example 1.3

The common-source stage studied in the previous example does not exhibit enough phase shift to allow oscillation. It is possible to insert an additional delay in the loop in the form of a delay line as shown in Fig. 1.7(a). Here, a voltage change at the drain takes ΔT seconds to reach the gate. Determine the startup condition and the frequency of oscillation. Neglect all capacitances and channel-length modulation.

Solution

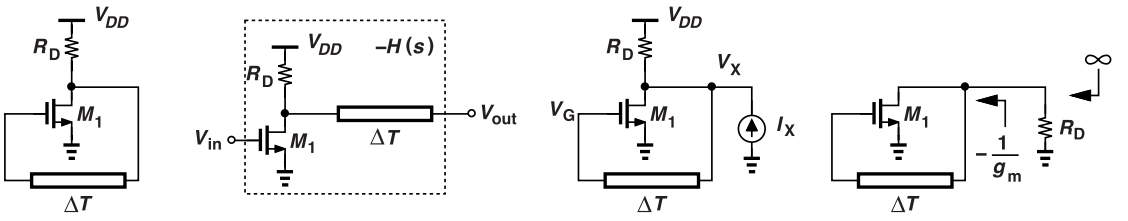
Using our first analysis method, we break the loop as illustrated in Fig. 1.7(b) and write

$$H(s) = -\frac{V_{out}}{V_{in}} \tag{1.7}$$

$$= g_m R_D e^{-s\Delta T}. \tag{1.8}$$

(The transfer function of the ideal delay line is equal to $e^{-s\Delta T}$.) We seek $H(j\omega_0) = -1$, i.e., $g_m R_D \exp(-j\omega_0 \Delta T) = -1$. It follows that, if $|H(j\omega_0)| = 1$ and $\angle H(j\omega_0) = 180^\circ$, then

$$g_m R_D = 1 \tag{1.9}$$



(a) (b) (c) (d)

Figure 1.7 (a) CS stage with a feedback delay line, (b) open-loop system, (c) computation of impedance at one node, and (d) illustration of infinite impedance at ω_0.

$$\omega_0 \Delta T = \pi. \tag{1.10}$$

The first equation is the startup condition, and the second can be expressed as

$$f_0 = \frac{1}{2\Delta T}, \tag{1.11}$$

where $f_0 = \omega_0/(2\pi)$. The circuit therefore oscillates with a period equal to $2\Delta T$. Note that the delay line introduces a phase shift of 180° at ω_0.

We can apply the third method by computing the closed-loop impedance seen at, for example, the output node. From the arrangement depicted in Fig. 1.7(c), we have $V_G = V_X \exp(-s\Delta T)$ and hence a small-signal drain current of $g_m V_X \exp(-s\Delta T)$. A KCL at the output node gives

$$\frac{V_X}{R_D} + g_m V_X e^{-s\Delta T} = I_X \tag{1.12}$$

and hence

$$\frac{V_X}{I_X} = \frac{R_D}{1 + g_m R_D e^{-s\Delta T}}. \tag{1.13}$$

We observe that if the startup condition, $g_m R_D = 1$ is fulfilled, then the denominator goes to zero for $s = j\omega_0 = j2\pi/(2\Delta T)$. The reader is encouraged to apply the second analysis method as well.

To gain more insight, let us compute the output impedance of the circuit while excluding R_D. If $R_D = \infty$ in Eq. (1.13), we have

$$\frac{V_X}{I_X} = \frac{1}{g_m e^{-s\Delta T}}, \tag{1.14}$$

which reduces to $V_X/I_X = -1/g_m$ at $s = j\omega_0$. Interestingly, the loop comprising M_1 and ΔT presents a negative resistance, which cancels the "loss" due to R_D if $g_m R_D = 1$ [Fig. 1.7(d)].

Oscillation Growth It is important to distinguish between two cases when $H(j\omega_0) = -1$ and the loop is stimulated. In response to an initial condition (or an impulse), the circuit oscillates with a constant amplitude [Fig. 1.8(a)]—as did the ideal pendulum in the previous section. On the other hand, with a sinusoidal excitation at ω_0, the oscillation amplitude continues to grow [Fig. 1.8(b)] (unless some other mechanism, e.g., a nonlinearity, stops the growth).

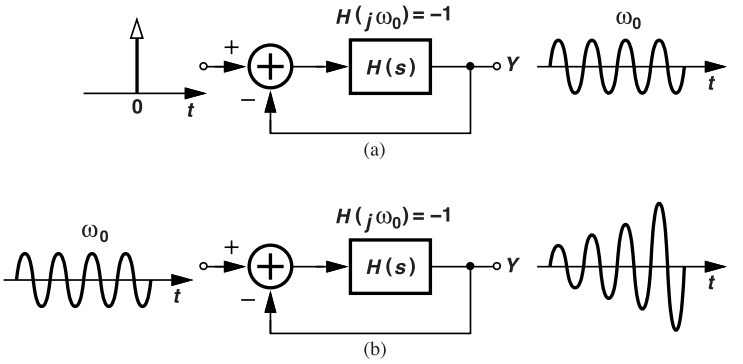


Figure 1.8 Response of an oscillatory system to (a) an impulse, and (b) a sinusoid at ω_0.

Startup Condition Revisited The condition $H(j\omega_0) = -1$ places the feedback loop at the edge of oscillation, failing to hold if process, voltage, and temperature (PVT) variations cause a slight drop in the loop gain. Moreover, this condition prohibits large-signal oscillations: if the oscillation amplitude grows to the extent that the circuit becomes nonlinear, the loop gain may drop below unity, violating the startup condition. The following example illustrates this point.

Example 1.4

Figure 1.9(a) shows a differential realization of the oscillator studied in Example 1.3. If $g_m R_D = 1$, explain

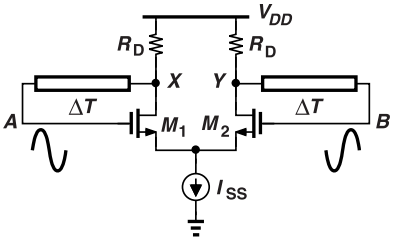


Figure 1.9 Differential pair with feedback delay lines.

why the oscillation amplitude remains small.

Solution

The circuit operates such that V_X and V_Y swing differentially and so do V_A and V_B. If V_A and V_B have small swings, the differential pair exhibits a unity voltage gain, sustaining the oscillation. The circuit cannot operate with large swings because the gain would then drop below unity at the peaks of V_A and V_B.

For the two reasons mentioned above, namely, PVT variations and gain drop due to nonlinearity, oscillators are typically designed with $|H(j\omega_0| > 1$ [and $\angle H(j\omega_0) = 180°$].

Example 1.5

The differential oscillator of Fig. 1.9 is redesigned for voltage swings that are large enough to ensure I_{SS} is entirely steered to the left or to the right. Determine the small-signal loop gain.

Solution

Suppose the four nodes carry a peak-to-peak swing of V_0. For M_1 or M_2 to carry all of I_{SS}, we have $|V_A - V_B|_{max} = V_0 = \sqrt{2}(V_{GS} - V_{TH})$ [1], where $V_{GS} - V_{TH}$ denotes the transistors' overdrive voltage in equilibrium (when $V_A = V_B$). With complete switching of the differential pair, we also have $|V_X - V_Y|_{max} = I_{SS}R_D = |V_A - V_B|_{max}$. It follows that

$$\sqrt{2}(V_{GS} - V_{TH}) = I_{SS}R_D \tag{1.15}$$

and hence

$$g_m R_D = \sqrt{2} \tag{1.16}$$

because $g_m = I_{SS}/(V_{GS} - V_{TH})$ in equilibrium (where $I_{D1} = I_{D2} = I_{SS}/2$).

With $|H(j\omega_0)| > 1$ and $\angle H(j\omega_0) = 180^\circ$ [Fig. 1.10(a)], we face an interesting puzzle. If, for example,

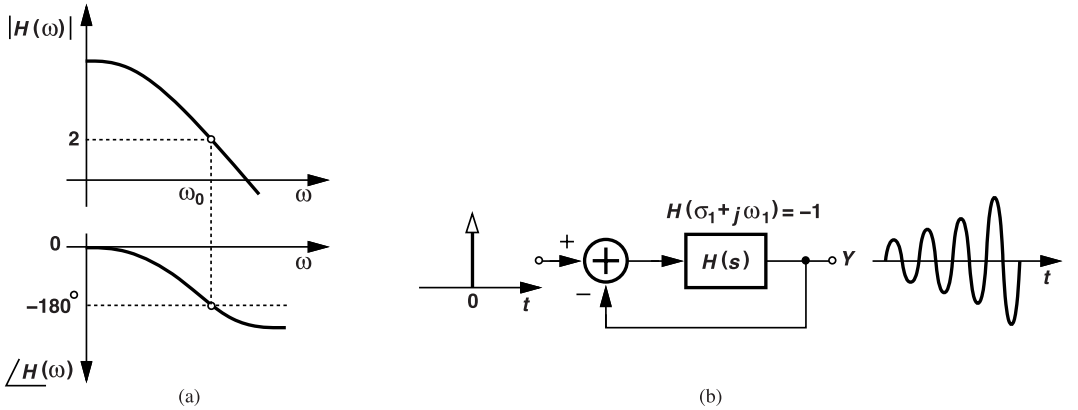


Figure 1.10 (a) Open-loop response with gain greater than unity at ω_0, and (b) impulse response of closed-loop system.

$H(j\omega_0) = -2$, then the closed-loop gain is equal to $H(j\omega_0)/[1 + H(j\omega_0)] = +2$ and not infinity! How then does the circuit oscillate? This result shows that the loop does not oscillate at ω_0. Rather, the circuit may find another value of s such that $H(s)/[1 + H(s)] \to \infty$, i.e., $H(s_1) = -1$, where s_1 has a *complex* value, $\sigma_1 + j\omega_1$, and $\sigma_1 > 0$. We study this case in Appendix I but should mention here that such a value of s leads to a growing sinusoid even with an impulse input [Fig. 1.10(b)], a point of contrast to the situation depicted in Fig. 1.8(a). [As explained in Appendix I, the condition $|H(j\omega_0)| > 1$ does not always guarantee oscillation, but suffices for typical oscillators.]

In the case of $|H(j\omega_0)| > 1$ and $\angle H(j\omega_0) = 180^\circ$, our node impedance test must also be revisited. For example, if $g_m R_D > 1$ in Eq. (1.13), then $|V_X/I_X|$ becomes *real* and *negative* at $s = j\omega_0$ (why?). In particular, the output resistance seen in Fig. 1.7(d) is now "stronger" than $-R_D$, leaving a residual negative component after canceling R_D. This component thus allows the oscillation amplitude to grow.

Positive Feedback at dc We have seen that positive feedback at a finite frequency, ω_0, can cause oscillation. But what happens if we have positive feedback at dc (low frequencies) as well? For example, if we cascade two CS stages as shown in Fig. 1.11 to obtain a greater phase shift, the feedback becomes positive at dc. If the low-frequency loop gain is greater than unity, this circuit latches up. This can be seen by assuming a small upward perturbation in V_X, which leads to a greater, downward change in V_Y. This change in turn causes even a larger upward change in V_X, etc. We say the circuit "regenerates" until V_X reaches V_{DD} and V_Y falls to a low value, turning M_1 off. This circuit is in fact used as a memory cell.

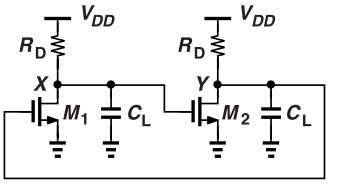


Figure 1.11 Two CS stages in a feedback loop.

To avoid latch-up, we design oscillators such that, at dc, the feedback is negative or, if it is positive, the loop gain is well below unity.

Oscillator Topologies Numerous oscillator topologies have been introduced over the years. Examples include "phase-shift," "Wien bridge," "relaxation," "multivibrator," "ring," and LC oscillators. In this book, we deal with primarily the last two as they are most commonly used in integrated circuit design.

1.4 Basic Ring Oscillators

Ring oscillators are popular in today's phase-locked system for their design flexibility and wide frequency tuning range. This section builds the foundation for these oscillators and Chapters 3 and 4 introduce advanced ring concepts.

We have seen that a single common-source stage does not provide sufficient phase shift to allow $\angle H(j\omega_0) = -180°$. Even a loop employing two CS stages fails to oscillate because $\angle H(j\omega)$ reaches $-180°$ only at $\omega = \infty$. We therefore surmise that a three-stage "ring" can satisfy both $\angle H(j\omega_0) = -180°$ and $|H(j\omega_0)| = 1$. Depicted in Fig. 1.12(a), this simple ring oscillator has negative feedback at low frequencies

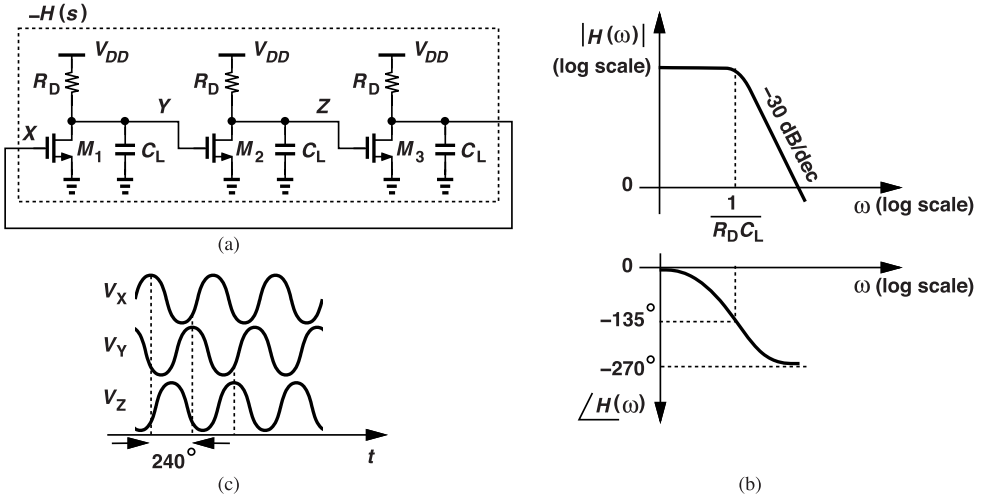


Figure 1.12 (a) Three CS stages in a feedback loop, (b) loop transmission, and (c) node waveforms.

and can be analyzed by assuming identical stages. The circuit in the dashed box is called $-H(s)$ to comply with the negative-feedback system shown in Fig. 1.2(a). We write

$$-H(s) = \left[-g_m \left(R_D || \frac{1}{C_L s}\right)\right]^3, \tag{1.17}$$

where channel-length modulation and other capacitances are neglected. It follows that

$$H(j\omega) = \frac{g_m^3 R_D^3}{(R_D C_L j\omega + 1)^3}. \tag{1.18}$$

Figure 1.12(b) sketches the magnitude and phase behavior. For $|H(j\omega_0)| = 1$, we have

$$\left(\frac{g_m R_D}{\sqrt{R_D^2 C_L^2 \omega_0^2 + 1}} \right)^3 = 1 \tag{1.19}$$

and hence

$$\omega_0 = \frac{\sqrt{g_m^2 R_D^2 - 1}}{R_D C_L}. \tag{1.20}$$

Also, $\angle H(j\omega_0) = -180^\circ$ yields

$$\tan^{-1}(R_D C_L \omega_0) = 60^\circ \tag{1.21}$$

and

$$\omega_0 = \frac{\sqrt{3}}{R_D C_L}. \tag{1.22}$$

Interestingly, (1.20) and (1.22) give

$$g_m R_D = 2. \tag{1.23}$$

In other words, each stage must provide a low-frequency voltage gain of 2 to guarantee oscillation.

A few properties of the above oscillator are worth noting. First, at ω_0, each stage exhibits a phase shift of 60° arising from its output pole plus 180° due to the low-frequency inversion of a CS amplifier. Thus, the waveforms at X, Y, and Z in Fig. 1.12(a) have a phase separation of 240° ($= -120^\circ$) [Fig. 1.12(c)]. Second, C_L can represent all of the transistor capacitances with reasonable accuracy. For example, at Y, C_L includes C_{GS2}, C_{DB1}, and the Miller effect of C_{GD2}. Since the waveforms in Fig. 1.12(c) suggest equal swings at the three nodes, we can assume a large-signal voltage gain of -1 from Y to Z and write the Miller capacitance as $C_{mill} = [1 - (-1)]C_{GD2} = 2C_{GD2}$.[3]

Example 1.6

A ring oscillator similar to Fig. 1.12(a) incorporates N identical CS stages, where N is an odd number. Determine the startup condition and the frequency of oscillation.

Solution

With an odd number of stages, the loop provides negative feedback at low frequencies, necessitating a frequency-dependent phase shift of $-180^\circ/N$ per stage for oscillation. That is,

$$-\frac{180^\circ}{N} = -\tan^{-1}(R_D C_L \omega_0) \tag{1.24}$$

and hence

$$\omega_0 = \frac{1}{R_D C_L} \tan \frac{180^\circ}{N}. \tag{1.25}$$

[3]This is an approximation because the phase shift from Y to Z is equal to -120° rather than -180°.

As N increases, ω_0 falls because each stage must exhibit less phase shift. To ensure a unity loop gain at this frequency, we must have

$$\left| g_m \left(R_D || \frac{1}{C_L j\omega_0} \right) \right| = 1, \tag{1.26}$$

obtaining

$$g_m R_D = \sqrt{\tan^2 \frac{180^\circ}{N} + 1}. \tag{1.27}$$

As N increases, the required low-frequency gain decreases.

Oscillation Amplitude As explained in the previous section, a loop gain of unity at ω_0 produces only a small oscillation amplitude. In practice, the three-stage ring of Fig. 1.12(a) must generate nearly rail-to-rail swings, requiring a small-signal, low-frequency gain higher than 2 per stage. The reader is encouraged to prove that the closed-loop poles of the circuit lie in the right half plane if $g_m R_D > 2$.

We can visualize the oscillation startup by assuming identical stages and hence equal initial voltages at all three nodes. As shown in Fig. 1.13, with a small perturbation (e.g., due to the noise of the devices), the circuit

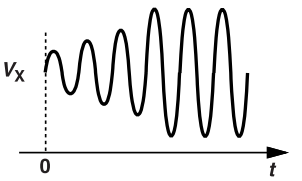


Figure 1.13 Initial growth of V_X.

begins to oscillate and the amplitude grows. The growth continues until the nonlinearity of each stage causes the gain to drop near and at the peaks. We say the "average" loop gain falls to unity due to the nonlinearity.

Example 1.7

Equation (1.20) implies that ω_0 is a function of g_m while Eq. (1.22) does not. Explain this discrepancy.

Solution

Since Eq. (1.22) derives from the phase shift criterion, it approximately applies even to large-signal operation. After all, the phase shift per stage must be 60° whether the swings are large or not. Equation (1.20), on the other hand, begins to lose its validity as the loop enters the large-signal regime.[4]

We recognize that the oscillation amplitude is determined by the circuit's nonlinearities if $|H(j\omega_0)| > 1$; otherwise, the amplitude would grow indefinitely. The calculation of the amplitude is difficult in the circuit of Fig. 1.12(a) because the additional 60° phase shift requires that the CS stage be treated as a nonlinear dynamic system. Fortunately, however, the oscillator topologies used in practice provide well-defined voltage swings, as seen in subsequent chapters.

[4] To improve the accuracy of Eq. (1.20), we can use an "average" transconductance. Since g_m varies periodically, it can be expressed as a Fourier series, with the first term yielding the average value.

Ideal Waveforms We have seen that ring oscillator waveforms are close to a sinusoid if $|H(j\omega_0)| = 1$ and begin to resemble a square wave if $|H(j\omega_0)|$ is well above unity. We can then ask what the desirable waveform of an oscillator should be. Most applications prefer sharp transitions to ensure rapid turn-on and turn-off of the devices that are driven by the oscillator. For example, a radio-frequency (RF) mixer sensing the output of an oscillator exhibits less noise with abrupt edges. Thus, if possible, we wish to design for square waveforms.

1.4.1 Inverter-Based Rings

A common implementation of ring oscillators employs CMOS inverters—rather than resistively-loaded CS amplifiers—as gain stages. Shown in Fig. 1.14(a) is an example. If the circuit begins with $V_X = V_Y = V_Z$,

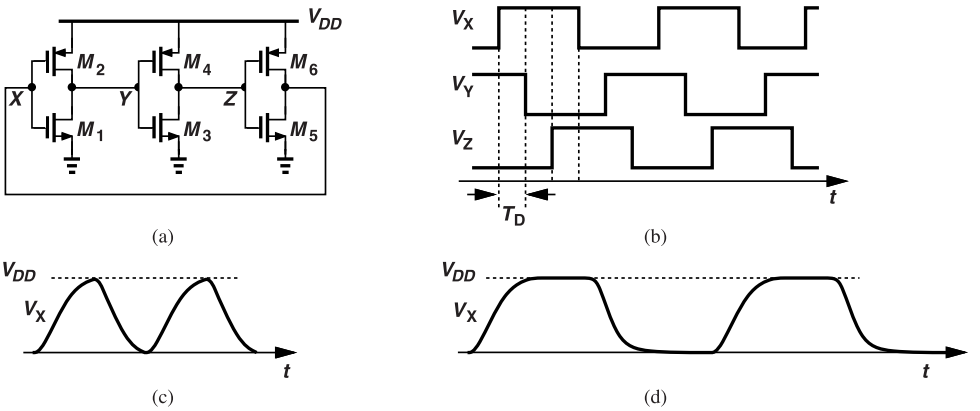


Figure 1.14 (a) Three-stage inverter-based ring oscillator, (b) node waveforms, (c) realistic signal shape, and (d) waveform for a five-stage ring.

each stage exhibits a small-signal voltage gain of $-(g_{mN} + g_{mP})(r_{ON}||r_{OP})$, which typically exceeds 2. The noise of the devices then causes regeneration until the oscillation reaches rail-to-rail voltage swings. Note that the negative feedback at dc prevents latch-up.

What if, upon power up, the ring begins with one node at zero? For example, if $V_X = 0$, then $V_Y = V_{DD}$ and $V_Z = 0$; the third inverter thus raises V_X toward V_{DD}. In other words, $V_X = 0$ cannot be a stable state. Figure 1.14(b) depicts the circuit's steady-state waveforms, revealing that a transition on one node causes another on the next node after one gate delay, T_D. The total oscillation period is therefore equal to $6T_D$.

The reader may wonder if the oscillation frequencies obtained from our small-signal and large-signal analyses are equal. For the three-stage ring of Fig. 1.14(a), the former analysis and Eq. (1.22) give $\omega_0 = 2\pi f_0 = \sqrt{3}/[(r_{ON}||r_{OP})C_L]$, where C_L denotes the total capacitance seen from each node to ground. The latter analysis, on the other hand, predicts $f_0 = 1/(6T_D)$. These two values are generally unequal, with $1/(6T_D)$ correctly predicting the frequency, but the small-signal result gives additional insight.

Some Properties The actual waveforms of the foregoing three-stage ring oscillator are less sharp than the square waves in Fig. 1.14(b) but not as gradual as a sinusoid. Illustrated as an example in Fig. 1.14(c), V_X changes its direction as soon as it reaches V_{DD} or ground because of the short delay around the loop. By comparison, a five-stage ring provides an oscillation period of $10T_D$, allowing some "relaxation" time for the high and low levels of V_X [Fig. 1.14(d)].

A few digital design principles reveal other properties of ring oscillators. First, since the delay of an inverter falls as V_{DD} increases, the oscillation frequency is inversely proportional to the supply voltage. As explained in Chapter 3, the supply sensitivity of ring oscillators proves problematic in most applications. Second, each inverter draws an average power of $f_0 C_L V_{DD}^2$, leading to a total power consumption of $nf_0 C_L V_{DD}^2$ for an n-stage ring.

1.5 Basic LC Oscillators

LC oscillators have some advantages over ring configurations, specifically, lower phase noise and the ability to operate at higher frequencies. We therefore employ LC topologies in many applications.

The analysis and design of LC oscillators heavily rely on the properties and modeling of LC tanks. It is thus necessary to begin our study with basic LC circuits.

1.5.1 LC Circuit Concepts

We know from basic circuit theory that the behavior of an ideal capacitor is expressed as $I = C_1 dV/dt$ in the time domain and as $V = I/(C_1 s)$ in the frequency domain. Similarly, an ideal inductor can be modeled by $V = L_1 dI/dt$ or $V = (L_1 s)I$. The parallel combination of these two components exhibits interesting properties. Illustrated in Fig. 1.15(a), such an impedance is dominated by L_1 at low frequencies and by C_1

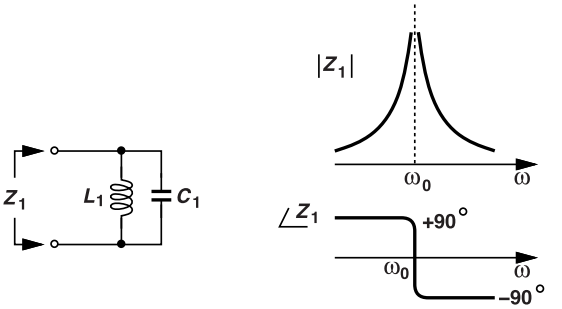


Figure 1.15 (a) Ideal parallel LC tank, and (b) its impedance characteristics.

at high frequencies (why?). We thus expect both $|Z_1|$ and $\angle Z_1$ to act accordingly. As shown in Fig. 1.15(b), Z_1 behaves inductively for $\omega < \omega_0 = 1/\sqrt{L_1 C_1}$ and capacitively for $\omega > \omega_0$. At the resonance frequency, ω_0, the two reactances cancel each other, producing $Z_1 = (L_1 j\omega_0)||[1/(C_1 j\omega_0)] = \pm j\infty$. The LC tank exemplifies a "resonator."

Let us now incorporate a more realistic inductor model. Implemented as a metal wire, L_1 suffers from a finite series resistance, R_S [Fig. 1.16(a)]. We say L_1 is "lossy" because the current flowing through it

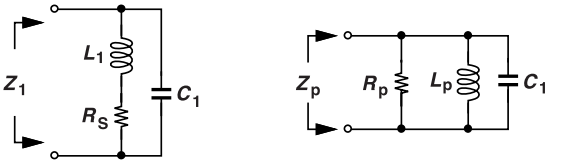


Figure 1.16 (a) Lossy LC tank, and (b) its equivalent model.

causes thermal dissipation in R_S. This physical model, however, gives little intuition, especially if we seek a resemblance between the behaviors of the ideal and the lossy tanks. As seen later, we prefer a model that represents the loss by a *parallel* resistance, R_p [Fig. 1.16(b)]. Of course, the tanks of Figs. 1.16(a) and 1.16(b) cannot be equivalent at all frequencies (why?) but one circuit can approximate the other for some frequency range.

To derive the values of L_p and R_p in terms of L_1 and R_S, we omit C_1 and equate Z_1 and Z_p:

$$L_1 s + R_S = (L_p s)||R_p \tag{1.28}$$

$$= \frac{L_p R_p s}{L_p s + R_p}. \tag{1.29}$$

For sinusoidal signals, $s = j\omega$, and

$$j\omega(R_S L_p + L_1 R_p - L_p R_p) + R_S R_p - L_1 L_p \omega^2 = 0. \tag{1.30}$$

This equation holds in general only if the imaginary and real parts are independently zero, yielding

$$R_S L_p + (L_1 - L_p) R_p = 0 \tag{1.31}$$

$$R_S R_p = L_1 L_p \omega^2. \tag{1.32}$$

Calculating R_p from the former and substituting in the latter, we have

$$L_p = \frac{R_S^2}{L_1 \omega^2} + L_1. \tag{1.33}$$

As explained later in this section, typically $R_S^2/(L_1\omega^2) \ll L_1$, and

$$L_p \approx L_1. \tag{1.34}$$

From Eq. (1.32), we obtain

$$R_p \approx \frac{L_1^2 \omega^2}{R_S}. \tag{1.35}$$

It is important to bear in mind that a *lower* R_S or a *higher* R_p translates to a more ideal tank.

Example 1.8

Explain which part of the above analysis fails if ω changes arbitrarily.

Solution

As ω falls, the approximation $R_S^2/(L_1\omega^2) \ll L_1$ eventually fails to hold. Without this approximation, we must substitute for L_p in (1.32) from (1.33), obtaining the exact value of R_p:

$$R_p = \frac{L_1^2 \omega^2}{R_S} + R_S. \tag{1.36}$$

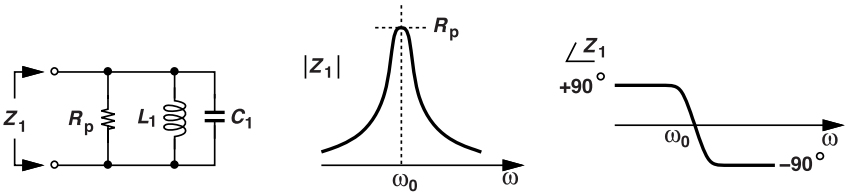
The parallel equivalent resistance given by (1.35) or (1.36) is a fictitious, mathematical quantity that simplifies our analyses. Interestingly, R_p is a frequency-dependent quantity, but we typically assume it has a constant value calculated at the tank resonance frequency. (As explained in Chapter 5, R_S itself varies with frequency due to the skin effect.)

Example 1.9

A parallel tank resonating at 5 GHz employs a 5-nH inductor with a series resistance of 20 Ω. Determine the value of R_p at 5 GHz and the error that it incurs if the tank resonance frequency is changed to 5.5 GHz. Assume R_S is constant.

Solution

From Eq. (1.36), we have $R_p \approx 1.25$ kΩ ($\gg R_S$). At 5.5 GHz, on the other hand, the actual R_p would be equal to 1.51 kΩ. Thus, a constant value of 1.25 kΩ incurs an error of about 18% at 5.5 GHz.



(a) (b) (c)

Figure 1.17 (a) Parallel lossy LC tank, (b) magnitude of impedance, and (c) phase of impedance .

Let us examine the impedance of an LC tank whose loss is modeled by a parallel resistor R_p [Fig. 1.17(a)]. The magnitude and phase follow those depicted in Fig. 1.15(b) at very low or very high frequencies (why?). Around the resonance frequency, however, $|Z_1|$ rises to only R_p while L_1 and C_1 cancel each other [Fig. 1.17(b)]. Similarly, the phase displays a gradual change [Fig. 1.17(c)].

An important conclusion here is that, as the tank becomes less lossy, both $|Z_1|$ and $\angle Z_1$ exhibit sharper changes.

Example 1.10

If R_p in Fig. 1.17(a) doubles while L_1 and C_1 remain constant, what happens to the phase plot in Fig. 1.17(c)?

Solution

We expect the phase transition around ω_0 to become sharper. We write

$$Z_1(s) = (L_1 s)||\left(\frac{1}{C_1 s}\right)||R_p \tag{1.37}$$

$$= \frac{R_p L_1 s}{L_1 C_1 R_p s^2 + L_1 s + R_p}. \tag{1.38}$$

For $s = j\omega$,

$$Z_1(j\omega) = \frac{jR_p L_1 \omega}{R_p(1 - L_1 C_1 \omega^2) + jL_1\omega}. \tag{1.39}$$

It follows that

$$\angle Z_1(j\omega) = \frac{\pi}{2} - \tan^{-1}\frac{L_1\omega}{R_p(1 - L_1 C_1 \omega^2)}. \tag{1.40}$$

To determine how the slope of $\angle Z$ depends on R_p, we differentiate (1.40) with respect to ω and replace $L_1 C_1 \omega^2$ with 1 in the result, arriving at

$$\frac{d}{d\omega}\angle Z_1(j\omega) \approx -2R_p C_1. \tag{1.41}$$

The slope around ω_0 is thus linearly proportional to R_p.

We should remark that the phase profile depicted in Fig. 1.17(c) is not exact at low frequencies if L_1 has a *series* resistance R_S. This is because the tank simply reduces to R_S near dc, displaying zero phase shift. This point is nonetheless unimportant in most of our studies.

Quality Factor To determine how "good" an inductor is, we can define a figure of merit for it, specifically, a "quality factor," Q. We expect the Q to be extremely high for an ideal inductor and lower as we include loss. For an inductor L_1 having a series R_S, we define the Q as

$$Q = \frac{L_1\omega}{R_S}. \tag{1.42}$$

Note that R_S is compared to the *impedance* of L_1 at the frequency of interest. In essence, the numerator represents the desired impedance and the denominator, the undesired impedance. The Q rises linearly with ω if R_S is assumed independent of the frequency.

We make three remarks. First, our previous approximation, $R_S^2/(L_1\omega^2) \ll L_1$, in fact translates to $Q^2 \gg 1$, which holds in most practical cases. Second, substituting $R_S = L_1\omega/Q$ in (1.35) gives

$$Q = \frac{R_p}{L_1\omega}, \tag{1.43}$$

revealing that we wish R_p to be as large as possible. Indeed, $R_p = \infty$ leads to an ideal parallel tank. This expression may suggest that Q *falls* as ω rises, but we must recall that R_p itself is frequency-dependent. Third, substituting $R_p = QL_1\omega$ in Eq. (1.41), we have an alternative definition for Q at the resonance frequency:

$$Q = \frac{\omega_0}{2}\left|\frac{d}{d\omega}\angle Z_1(j\omega)\right|. \tag{1.44}$$

That is, the Q can be viewed as proportional to the slope of the phase response.

Tuned Amplification The parallel tank exemplifies a "tuned load," a two-terminal circuit whose impedance reaches a peak only at a certain frequency. As with resistors and current sources, a tuned circuit can serve as the load in an amplifier, providing narrowband amplification. Shown in Fig. 1.18(a), such an arrangement has

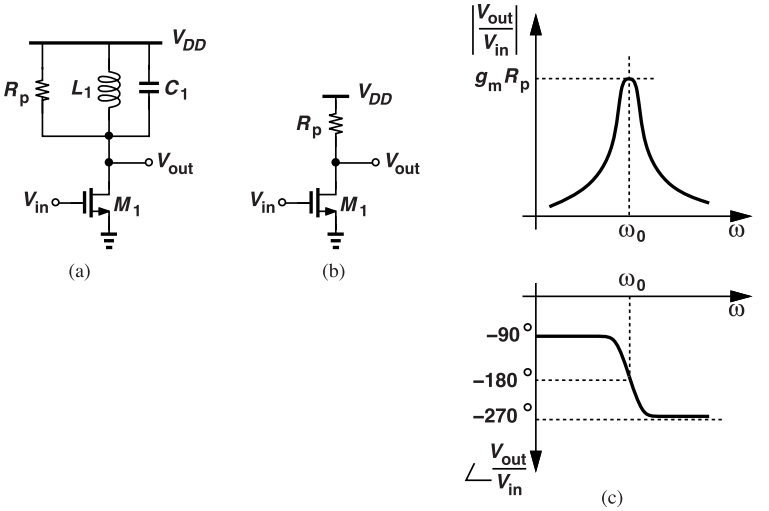


Figure 1.18 (a) CS stage using a tuned load, (b) equivalent circuit at resonance, and (c) frequency response.

a transfer function given by

$$\frac{V_{out}}{V_{in}}(s) = -g_m\left[(L_1 s)||R_p||\frac{1}{C_1 s}\right], \tag{1.45}$$

if channel-length modulation is neglected. We now appreciate the simplicity afforded by the parallel-resistance model: at the resonance frequency, ω_0, the circuit reduces to that in Fig. 1.18(b), yielding a voltage gain of $-g_m R_p$ and a phase shift of 180°. From the plots in Fig. 1.17, we readily arrive at the gain and phase profiles depicted in Fig. 1.18(c). The amplifier also acts as a band-pass filter (BPF).

Example 1.11

The tuned amplifier of Fig. 1.18(a) senses a sinusoidal input at ω_0 with a peak amplitude of 20 mV and a dc level of 500 mV. If $g_m R_p = 5$, plot the output waveform. Assume $V_{DD} = 1$ V.

Solution

The circuit inverts the signal and amplifies it by a factor of 5. But, what is the dc level at the output? If the series resistance of L_1 is neglected, the dc drop from V_{DD} to V_{out} must be zero. (An ideal inductor cannot sustain a finite dc voltage.) Thus, V_{out} has an average value equal to V_{DD}, exhibiting voltage swings *above* V_{DD} (Fig. 1.19). This property of parallel LC tanks proves useful in low-voltage design.

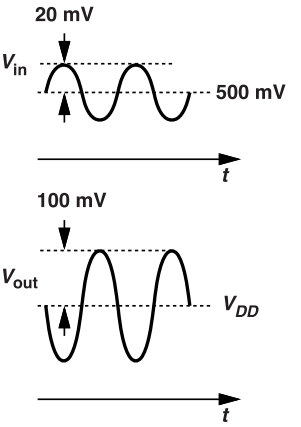


Figure 1.19 Tuned amplifier input and output waveforms.

1.5.2 LC Oscillators as Feedback Systems

As with the resistively-loaded CS stage in Section 1.4, the tuned amplifier of Fig. 1.18(a) can be placed in a feedback loop to obtain oscillation. To this end, we must establish a total phase shift of 360° around the loop at ω_0, which the circuit of Fig. 1.18(a) cannot. We therefore cascade two identical stages as shown in Fig. 1.20(a), noting that the open-loop transfer function exhibits a magnitude of $(g_m R_p)^2$ and a phase of -360° at $\omega_0 = 1/\sqrt{L_1 C_1}$. The closed-loop circuit thus oscillates at ω_0 if $(g_m R_p)^2 \geq 1$. We often redraw the oscillator as shown in Fig. 1.20(b). Symmetry implies that V_X and V_Y vary differentially with time.

Example 1.12

Can the above 1C oscillator latch up?

Solution

To determine the possibility of latch-up, we examine the circuit near dc, recognizing that the parallel resistance model no longer holds. If each inductor exhibits a low-frequency series resistance of R_S, the oscillator reduces to the configuration depicted in Fig. 1.21, where the loop gain is given by $(g_m R_S)^2$. To avoid latch–up, we must have $(g_m R_S)^2 < 1$, a condition readily met in typical designs.

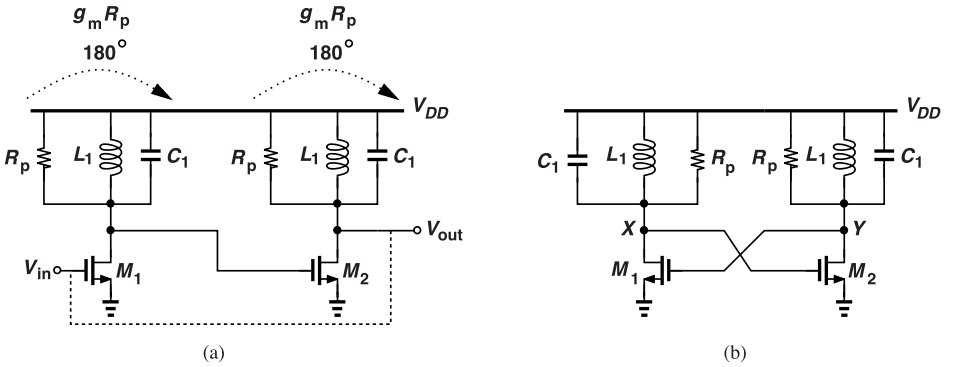


Figure 1.20 (a) Two cascaded tuned CS stages in a loop, and (b) the circuit redrawn as cross-coupled amplifiers.

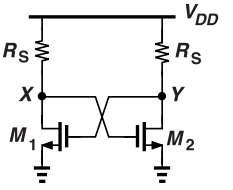


Figure 1.21 LC oscillator equivalent circuit at low frequencies.

It is instructive to sketch the waveforms of the oscillator of Fig. 1.20(b). As explained in Section 1.3, we generally choose the loop gain greater than unity so as to ensure large swings. In Fig. 1.20(b), M_1 and M_2 experience complete switching if $(g_m R_p)^2$ is sufficiently large; thus, as, for example, V_X falls and V_Y rises, M_2 begins to turn off and M_1 turns on more. The drain currents therefore swing between 0 and a maximum value (Fig. 1.22). Flowing through their respective tanks, these currents generate V_X and V_Y. We

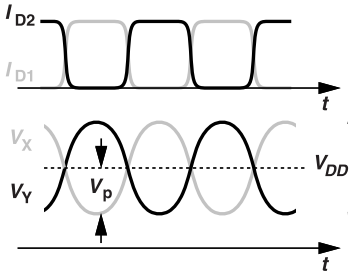


Figure 1.22 Cross-coupled oscillator waveforms.

observe that (1) the output common-mode level is approximately equal to V_{DD} (Example 1.11), and (2) the voltage waveforms resemble sinusoids whereas the current waveforms do not. This is because the parallel tank significantly filters components at $2\omega_0$, $3\omega_0$, etc., as it converts the current to voltage.

Example 1.13

What is the maximum drain-source voltage experienced by M_1 and M_2 in Fig. 1.20(b)?

Solution

For a single-ended peak swing of V_p, the drain voltage in Fig. 1.22 reaches approximately $V_{DD} + V_p$. This value of V_{DS} may not be allowed by the technology because it can "stress" the transistor, shortening its lifetime.

Example 1.14

One oscillation test that we devised in Section 1.3 is to measure the impedance between two nodes of the circuit and determine whether it goes to infinity at some frequency. Apply this test to the circuit of Fig. 1.20(b). Neglect channel-length modulation.

Solution

Let us examine the impedance between X and Y. Drawing the small-signal model as shown in Fig. 1.23(a), we observe that the net current injected into ground by the two tanks is zero because V_X and V_Y are differential. We therefore place the two in series and merge them as shown in Fig. 1.23(b). This allows us to compute the overall impedance as a parallel combination of two impedances, namely, the tank in the dashed box and the resistance produced between X and Y by M_1 and M_2. For the latter, we construct the model depicted in Fig. 1.23(c), note that $g_m V_X + g_m V_Y = 0$ (why?), and write a KVL, $V_Y + V_t = V_X$. That is, $V_t = 2V_X = -2V_Y$. Also, $I_t = -g_m V_X = -g_m V_t/2$. In other words,

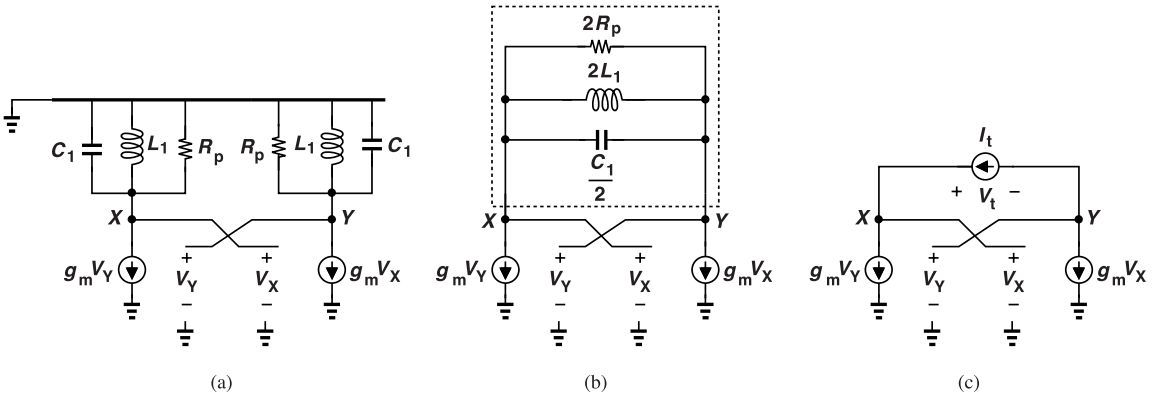


Figure 1.23 (a) Small-signal model of oscillator, (b) model with merged tanks, and (c) equivalent circuit for calculating output resistance of cross-coupled pair.

$$\frac{V_t}{I_t} = -\frac{2}{g_m}. \tag{1.46}$$

For oscillation to occur, the parallel combination of this (negative) resistance and the tank in Fig. 1.23(b) must go to infinity:

$$\left[(2L_1 s)||\left(\frac{2}{C_1 s}\right)\right] || \left[(2R_p)||\frac{-2}{g_m}\right] = \infty. \tag{1.47}$$

We intuitively predict that both the first parallel combination and the second must individually go to infinity.

To be more rigorous, however, let us return to Example 1.10 and apply Eq. (1.39) to the case at hand:

$$\frac{V_t}{I_t}(j\omega) = \frac{j(2L_1\omega)}{1 - L_1C_1\omega^2 + \dfrac{j(2L_1\omega)}{(2R_p)||\dfrac{-2}{g_m}}}. \tag{1.48}$$

This function goes to infinity if $\omega = 1/\sqrt{L_1C_1}$ and $2R_p||(-2/g_m) = \pm\infty$, i.e., $g_m = 1/R_p$. As expected, this result agrees with the startup condition $(g_mR_p)^2 = 1$ derived previously.

An interesting and useful observation here is that the cross-coupled transistors M_1 and M_2 in Fig. 1.20(b) produce a *negative* (small-signal) resistance between their drains. We return to this point in Section 1.5.3.

Example 1.15

In analogy with the three-stage ring oscillator of Fig. 1.12(a), a student constructs the multi-phase circuit shown in Fig. 1.24(a). Determine the frequency of oscillation and the startup condition.

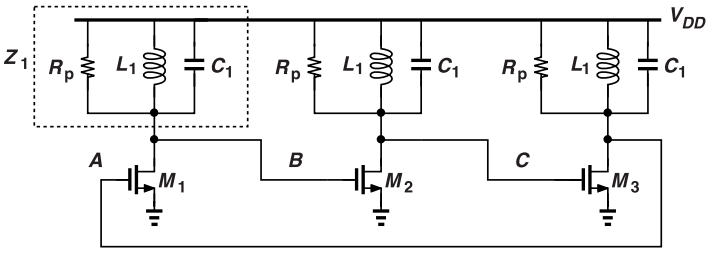


Figure 1.24 Three-stage ring using tuned amplifiers.

Solution

For the total phase shift around the loop to be equal to 360°, each stage must contribute -120°. Since $V_B/V_A = -g_mZ_1$, we have $\angle(-g_mZ_1) = -120^\circ$. From Fig. 1.18(c), we observe that this occurs at a frequency below the resonance frequency, ω_0. Returning to Eq. (1.40) and subtracting π from it to account for the inversion produced by a CS stage, we have

$$-\frac{\pi}{2} - \tan^{-1}\frac{L_1\omega_1}{R_p(1 - L_1C_1\omega_1^2)} = -\frac{2\pi}{3}, \tag{1.49}$$

where ω_1 denotes the oscillation frequency (not to be confused with the tanks' resonance frequency). In this case, the phase difference between each tank's current and voltage is equal to 60° rather than zero. Thus,

$$\frac{L_1\omega_1}{R_p(1 - L_1C_1\omega_1^2)} = \frac{1}{\sqrt{3}} \tag{1.50}$$

and hence

$$\omega_1 = \frac{-\sqrt{3}L_1 + \sqrt{3L_1^2 + 4L_1C_1R_p^2}}{2L_1C_1R_p}. \tag{1.51}$$

This result can be simplified if we express the tank Q at the resonance frequency as $Q = R_p/(L_1\omega_0)$, note that $\omega_0 = 1/\sqrt{L_1C_1}$, and write

$$\omega_1 = \frac{-\sqrt{3} + \sqrt{3 + 4Q^2}}{2Q}\omega_0. \tag{1.52}$$

The reader can prove that $\omega_1 < \omega_0$. For example, if $Q = 5$, then $\omega_1 = 0.84\omega_0$. To arrive at a simpler expression, we factor $4Q^2$ from $3+4Q^2$ in the numerator and write $\sqrt{1 + 3/(4Q^2)} \approx 1+3/(8Q^2)$, obtaining

$$\omega_1 \approx \left(1 - \frac{\sqrt{3}}{2Q}\right)\omega_0 \tag{1.53}$$

if $8Q^2 \gg 3$. As expected, the higher the Q, the less ω_1 departs from ω_0; this occurs because the higher Q translates to a greater slope for $\angle Z$ in Fig. 1.17(c) and hence requires only a small deviation from ω_0 to establish a phase shift of 60° in the tank.

As with the resistively-loaded ring oscillator, each stage in this circuit displays a *complex* transfer function at ω_1 because the tanks operate away from resonance. The cascade of the three stages, however, has a real transfer function.

For the startup condition, the magnitude squared of the transfer function of one stage must be at least unity at $\omega = \omega_1$. Multiplying the magnitude of Z_1 in Eq. (1.39) by g_m and squaring the result, we have

$$\frac{g_m^2R_p^2L_1^2\omega_1^2}{R_p^2(1 - L_1^2C_1^2\omega_1^2)^2 + L_1^2\omega_1^2} = 1. \tag{1.54}$$

Let us write $\omega_1 = \alpha\omega_0$ from Eq. (1.53) and

$$g_m^2R_p^2L_1^2\alpha^2\omega_0^2 = R_p^2(1 - \alpha^2)^2 + L_1^2\alpha^2\omega_0^2. \tag{1.55}$$

Thus,

$$g_m^2R_p^2 - 1 = \frac{R_p^2}{L_1^2\omega_0^2}\frac{(1 - \alpha^2)^2}{\alpha^2}. \tag{1.56}$$

That is,

$$g_mR_p \geq \sqrt{1 + \left(\frac{1}{\alpha} - \alpha\right)^2 Q^2}. \tag{1.57}$$

Note that α itself is a function of Q, and both Q and R_p are computed at the resonance frequency, ω_0. For example, if $Q = 5$, then $\alpha = 0.84$ and g_mR_p must be at least 2.02. We observe that this oscillator demands a greater g_mR_p than does the two-stage counterpart.

The oscillator of Fig. 1.20(b) suffers from a poorly-defined bias current: when the gate voltage reaches $V_{DD} + V_p$, the drain current strongly depends on the transistor characteristics and on V_{DD}. To resolve this issue, we modify the circuit as described below.

Differential Feedback Oscillator The foregoing oscillator developments have focused on single-ended tuned CS stages. We can alternatively consider a differential pair with tuned loads and explore whether it can form an oscillator. Shown in Fig. 1.25(a), such an arrangement exhibits a small-signal differential transfer function similar to that of the CS stage of Fig. 1.18(a), providing a peak gain of $-g_mR_p$ at ω_0.

It is instructive to first study the large-signal behavior of the circuit with a differential sinusoidal input at ω_0. As illustrated in Fig. 1.25(b), M_1 and M_2 experience complete switching if the input amplitude is large

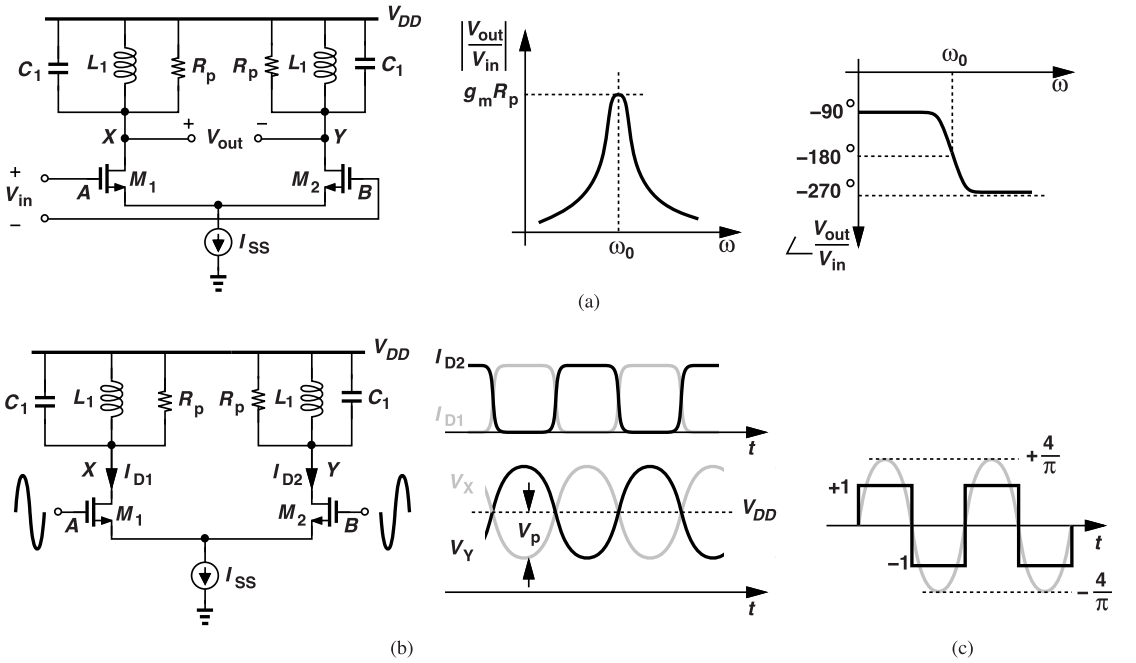


Figure 1.25 (a) Differential pair with LC loads, (b) large-signal input and output waveforms, (c) relation between square wave and its first harmonic.

enough, with I_{D1} and I_{D2} swinging between 0 and I_{SS}. Approximating each drain current waveform by a square wave, we observe that the fundamental sees a load impedance of R_p while the higher harmonics see a low impedance. Thus, V_X and V_Y resemble sinusoids.

We can calculate the peak output voltage swing in Fig. 1.25(b) as follows. A square wave toggling between $+1$ and -1 contains a fundamental with a peak value of $4/\pi$ [Fig. 1.25(c)]. (This is a useful property to remember. It is also helpful to remember that the fundamental has a *larger* amplitude than the square wave.) The peak fundamental amplitude of I_{D1} and I_{D2} is therefore equal to $(4/\pi)(I_{SS}/2)$ (why?). Since this component has a frequency of ω_0, it flows only through R_p on each side, generating a sinusoidal voltage with a peak single-ended amplitude of

$$V_p = \frac{2}{\pi} I_{SS} R_p. \tag{1.58}$$

The peak-to-peak differential output swing is equal to $4V_p = (8/\pi)I_{SS}R_p$. The reader is cautioned not to confuse the various factors of 2 that arise in these amplitude calculations.

Let us now place the tuned differential pair of Fig. 1.25(a) in a feedback loop. In our first attempt, we tie X to A and Y to B [Fig. 1.26(a)]. Why does this circuit not oscillate? The overall loop must provide a total phase shift of $360°$, but we know from Fig. 1.25(a) that this is not possible. We then swap the feedback connections as shown in Fig. 1.26(b), adding another $180°$ to the phase profile of Fig. 1.25(a). Now, at $\omega = \omega_0$, the phase shift around the loop is equal to $0°$ (or $360°$). The circuit thus oscillates if $(g_m R_p)^2 \geq 1$. Note that g_m denotes the small-signal transconductance of M_1 or M_2 when each carries a current of $I_{SS}/2$.

The oscillator derived above resembles that in Fig. 1.20(b) except for the tail current source, which defines the bias current much more accurately. From the analysis leading to Eq. (1.58), we have learned that the new topology provides a differential peak-to-peak swing of $(8/\pi)I_{SS}R_p$ (if M_1 and M_2 experience complete switching) with a CM level equal to V_{DD}.

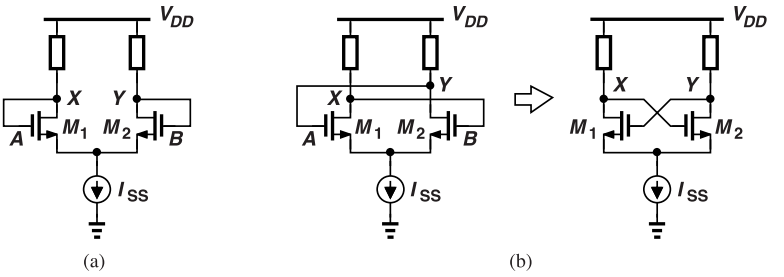


Figure 1.26 Differential pair with (a) negative, and (b) positive feedback.

Called the "cross-coupled oscillator," the circuit of Fig. 1.26(b) is the most widely-used choice today. Its robust operation and the modest startup condition, $g_m R_p = 1$, have made this topology a relatively fool-proof solution.

Example 1.16

Determine the necessary loop gain in the above circuit for M_1 and M_2 to switch completely. Assume square-law devices and neglect channel-length modulation.

Solution

For the transistors to steer I_{SS} completely, the differential voltage applied to their gates must reach $\sqrt{2}(V_{GS} - V_{TH})_{eq}$, where $(V_{GS} - V_{TH})_{eq}$ denotes the equilibrium overdrive, i.e., when I_{SS} is split equally. The equilibrium transconductance is $g_m = 2I_D/(V_{GS} - V_{TH})_{eq} = I_{SS}/(V_{GS} - V_{TH})_{eq}$. The peak differential voltage driving the transistors is equal to $2V_p$ in Fig. 1.25(b) and $V_p = (2/\pi)I_{SS}R_p$. We must have $2V_p = \sqrt{2}(V_{GS} - V_{TH})_{eq}$ and hence

$$\frac{4}{\pi} I_{SS} R_p = \sqrt{2}(V_{GS} - V_{TH})_{eq}. \tag{1.59}$$

That is,

$$g_m R_p = \frac{\sqrt{2}\pi}{4} \tag{1.60}$$

$$\approx 1.11. \tag{1.61}$$

This is equivalent to saying that the loop gain, $(g_m R_p)^2$, must exceed $\pi^2/8 \approx 1.23$.

The foregoing calculation is optimistic in that it assumes the drain currents are square waves. Since each transistor barely turns off, these currents carry a first harmonic amplitude smaller than $(2/\pi)I_{SS}$, requiring that $g_m R_p$ be substantially greater than $\sqrt{2}\pi/4$.

1.5.3 LC Oscillators as One-Port Systems

Our development of LC oscillators in the previous section has drawn upon instability in a feedback system, i.e., one whose output port is connected to its input port. An alternative perspective views oscillators as one-port networks and confers additional insights.

At the beginning of this chapter, we saw that an ideal pendulum with an initial condition oscillates indefinitely. We also observed that tapping a lossy pendulum in each period with our finger can compensate for the loss. These concepts readily apply to LC tanks as well. Let us release the ideal tank shown in Fig. 1.27(a) with an initial condition of V_0 across C_1. The capacitor begins to discharge through L_1, i.e., the

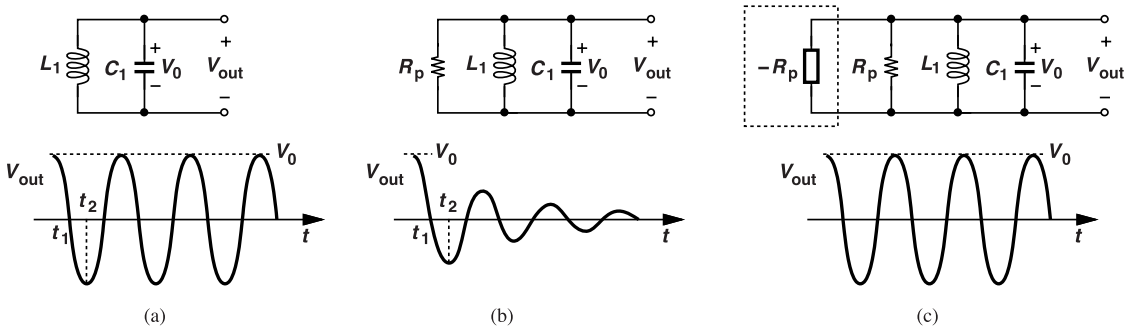


Figure 1.27 (a) Response of a lossless LC tank to an initial condition, (b) response of a lossy LC tank, and (c) use of a negative resistance to sustain oscillation.

electric energy, $(1/2)C_1V^2$, gradually converts to magnetic energy, $(1/2)L_1I^2$. At $t = t_1$, $V_{out} = 0$ and the initial energy given to the system now resides entirely in the inductor's magnetic field. This point corresponds to when the pendulum reaches the vertical position, carrying the energy in kinetic form. We can write $(1/2)C_1V_0^2 = (1/2)L_1I^2(t_1)$, obtaining $I(t_1) = \sqrt{C_1/L_1}V_0$. This is the current flowing through L_1 (and C_1) at this moment.

After t_1, the inductor current continues to flow through the capacitor, charging it in the opposite direction and forcing V_1 to become negative. At $t = t_2$, all of the energy is returned to C_1 but with an opposite voltage polarity. The oscillation thus goes on.

We now extend this time-domain view to a lossy tank. As illustrated in Fig. 1.27(b), so long as $V_{out} \neq 0$, R_p carries a finite current and dissipates energy. Consequently, the inductor current at $t = t_1$ is less than $\sqrt{C_1/L_1}V_0$ and the absolute peak voltage at $t = t_2$ less than V_0. The oscillation therefore dies away.

How do we "tap" the lossy tank in Fig. 1.27(b) and replenish the energy dissipated by R_p? If we add a negative resistance equal to $-R_p$ in parallel with the tank [Fig. 1.27(c)], then $|R_p||(-R_p)| = \infty$, as if the tank were ideal. Consequently, the circuit oscillates in response to an initial condition. We call this topology a "one-port oscillator" because it results from attaching the one-port tank to a one-port network, $-R_p$. Of course, the idea is applicable to any lossy resonator.

We have already encountered a one-port oscillator example. As studied in Example 1.14, a cross-coupled transistor pair can act as a negative resistance, thus compensating for the loss of an LC tank. Note that this is true whether the circuit contains a tail current source or not.

Example 1.17

The tank in Fig. 1.27(c) has an R_p equal to 1 kΩ. Explain what happens if the negative resistance is equal to (a) -1.2 kΩ, or (b) -800 Ω.

Solution

With $-1.2\,\text{k}\Omega$, the net tank resistance is equal to $+6$ kΩ. Since the negative resistance is not "strong" enough, the circuit fails to oscillate. With $-800\ \Omega$, on the other hand, the LC combination sees a parallel resistance equal to -4 kΩ, i.e., the negative resistance is too strong. In this case, the oscillation amplitude grows until the negative resistance saturates, yielding an "average" value of $-1\,\text{k}\Omega$.

Negative-Resistance Circuits In order to realize the one-port oscillator of Fig. 1.27(c), we must design a circuit that exhibits a negative resistance between two nodes. We should remark that a small-signal negative resistance is not against the laws of physics; it simply refers to a port whose small-signal current *decreases* as the voltage applied to it increases. The cross-coupled pair is one such example.

Another topology that yields a negative resistance is shown in Fig. 1.28. Note that the source of M_1 need

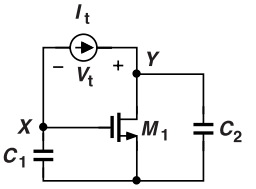


Figure 1.28 Circuit exhibiting negative resistance.

not be grounded. Neglecting the transistor capacitances, we write $V_{GS} = -I_t(C_1 s)^{-1}$ and hence $I_{D1} = g_m(-I_t)(C_1 s)^{-1}$. Since C_2 carries a current equal to $I_t - I_{D1} = I_t[1 + g_m(C_1 s)^{-1}]$, it sustains a voltage given by $I_t[1 + g_m(C_1 s)^{-1}](C_2 s)^{-1}$. A KVL gives

$$-\frac{I_t}{C_1 s} + V_t = I_t(1 + \frac{g_m}{C_1 s})\frac{1}{C_2 s}. \tag{1.62}$$

That is,

$$\frac{V_t}{I_t}(s) = \frac{1}{C_1 s} + \frac{1}{C_2 s} + \frac{g_m}{C_1 C_2 s^2}. \tag{1.63}$$

Equation (1.63) reveals that the small-signal impedance seen between X and Y consists of the series combination of C_1, C_2, and a third impedance equal to $g_m/(C_1 C_2 s^2)$. For $s = j\omega$, this term assumes a negative, *real* value equal to $-g_m/(C_1 C_2 \omega^2)$ and can be used to cancel the loss of a resonator. In contrast to the resistance produced by a cross-coupled pair, this quantity is a function of frequency but still a resistance.

The two capacitive terms in Eq. (1.63) can contribute to the capacitance needed in the tank. In fact, we can simply attach an inductor between X and Y and obtain a circuit called the "three-point oscillator." As illustrated in Fig. 1.29, such a topology reduces to the series combination of the inductor, C_1, C_2, and the

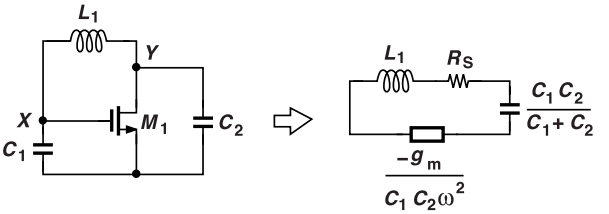


Figure 1.29 Basic three-point oscillator.

negative resistance. In this case, we prefer to model the loss of L_1 by a series resistance, R_S, and express the startup condition as

$$R_S = \frac{g_m}{C_1 C_2 \omega_0^2}. \tag{1.64}$$

If this condition is satisfied, the tank oscillates at a frequency equal to

$$\omega_0 = \frac{1}{\sqrt{L_1 \dfrac{C_1 C_2}{C_1 + C_2}}}. \tag{1.65}$$

Noting that the parallel equivalent resistance of L_1 is given by $R_p = L_1^2\omega_0^2/R_S$, we rewrite (1.64) as

$$\frac{L_1^2\omega_0^2}{R_p} = \frac{g_m}{C_1C_2\omega_0^2} \tag{1.66}$$

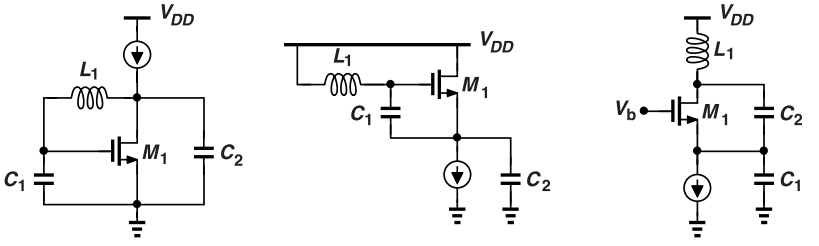
and hence

$$g_mR_p = L_1^2\omega_0^4C_1C_2 \tag{1.67}$$

$$= \frac{(C_1+C_2)^2}{C_1C_2}. \tag{1.68}$$

The minimum required g_mR_p occurs if $C_1 = C_2$ and is equal to 4. That is, for a given inductor Q, this circuit demands a greater transistor transconductance than does the cross-coupled oscillator.

The arrangement of Fig. 1.29 leads to three oscillators if one of the three nodes is tied to ac ground. Depicted in Fig. 1.30 along with bias components, these topologies follow Eqs. (1.64) and (1.65) but can also



(a) (b) (c)

Figure 1.30 Oscillators derived from the three-point topology.

be viewed as single-stage amplifiers with feedback. The following example illustrates this point.

Example 1.18

Considering the oscillator of Fig. 1.30(a) as a common-source stage with feedback, determine the oscillation frequency and the startup condition.

Solution

We draw the open-loop circuit as shown in Fig. 1.31, modeling the loss of L_1 by R_S. For oscillation to

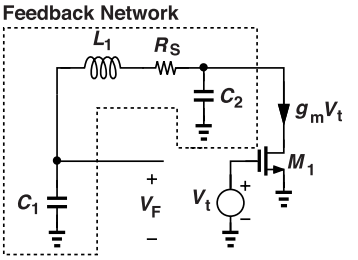


Figure 1.31 Open-loop circuit for analyzing the three-point oscillator.

occur, the feedback network must provide a phase shift of $180°$. We compute the loop transmission, $H(s)$, as $-V_F/V_t$. The drain current of M_1, g_mV_t, splits between C_2 and the L_1-R_S-C_1 combination. The current

flowing through the latter generates V_F across C_1:

$$-g_m V_t \frac{\dfrac{1}{C_2 s}}{\dfrac{1}{C_2 s} + L_1 s + R_S + \dfrac{1}{C_1 s}} \frac{1}{C_1 s} = V_F. \tag{1.69}$$

It follows that

$$H(s) = -\frac{V_F}{V_t} \tag{1.70}$$

$$= \frac{\dfrac{g_m}{s}}{L_1 C_1 C_2 s^2 + C_1 + C_2 + R_S C_1 C_2 s}. \tag{1.71}$$

Setting $H(s)$ to -1 and s to $j\omega$, we have

$$-L_1 C_1 C_2 \omega^2 + C_1 + C_2 + j\omega R_S C_1 C_2 + \frac{g_m}{j\omega} = 0. \tag{1.72}$$

The real and imaginary parts must independently fall to zero, yielding

$$\omega_1 = \frac{1}{\sqrt{L_1 \dfrac{C_1 C_2}{C_1 + C_2}}} \tag{1.73}$$

$$R_S = \frac{g_m}{C_1 C_2 \omega^2}, \tag{1.74}$$

as obtained previously.

The oscillators in Figs. 1.30(b) and 1.30(c) can be respectively considered as a source follower and a common-gate stage with feedback. The latter is called the "Colpitts oscillator." In practice, the cross-coupled topology of Fig. 1.26(b) proves more versatile and more robust than these three oscillators.

An interesting property of the oscillator shown in Fig. 1.30(b) is that it can drive a load resistance through the drain of M_1 without any additional current and without loading the tank. In fact, we can create a differential version of the circuit as depicted in Fig. 1.32 and directly drive a load, for example, an antenna. In such a case, the large bias current necessary for delivering power to the antenna also leads to lower phase noise (Chapter 2), but the power efficiency of this topology is lower than simple common-source stages acting as drivers. In Problem 1.20, we explore the role of R_b and why the two halves of this circuit oscillate differentially.

1.6 Voltage-Controlled Oscillators

In most applications, an oscillator's frequency must be variable because (a) it may be a function of the process, supply voltage, or temperature (PVT) and must therefore be restored to the desired value by "tuning," and (b) the system may require different oscillation frequencies at different times. For example, a microprocessor runs with a high-speed clock for computationally-intensive tasks but reverts to a low-speed clock during less demanding operations. The frequency can be tuned by varying electronically a parameter within the oscillator, e.g., a resistance, a capacitance, or an inductance. If the control is a voltage quantity, we call the circuit a voltage-controlled oscillator (VCO). As explained in Chapter 3, we can also construct current-controlled oscillators.

The tuning characteristic of a VCO is ideally a straight line given by $\omega_{out} = K_{VCO} V_{cont} + \omega_0$ (Fig. 1.33), where K_{VCO} is called the "gain" of the VCO and is expressed in rad/s/V. We also sometimes express

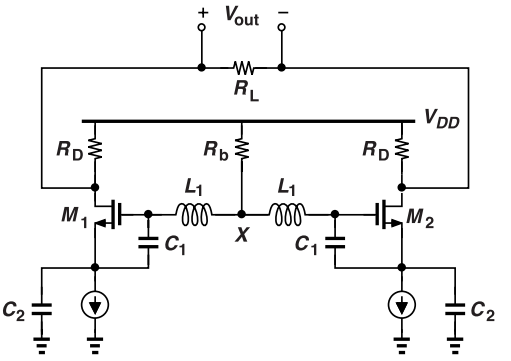


Figure 1.32 Differential oscillator derived from the topology in Fig. 1.30(b).

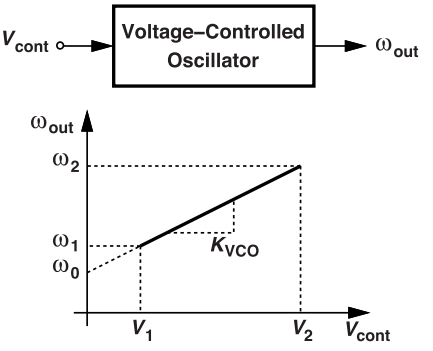


Figure 1.33 VCO tuning characteristic.

K_{VCO} in Hz/V, but it is important to remember that such a value must eventually be multiplied by 2π for phase-locked loop (PLL) analysis. The characteristic $\omega_{out} = K_{VCO}V_{cont} + \omega_0$ represents a static system; i.e., the output frequency changes instantaneously if V_{cont} jumps abruptly. While not exact, this model suffices in most cases.

We are also interested in the time-domain waveform produced by a VCO. For a general periodic (sinusoidal) signal, we would write $V_{out} = V_0 \cos \omega_1 t$. We call the argument of the cosine the "phase" of the signal. In this simple case, the phase, $\omega_1 t$, rises linearly with time at a slope equal to ω_1. As shown in Fig. 1.34, the

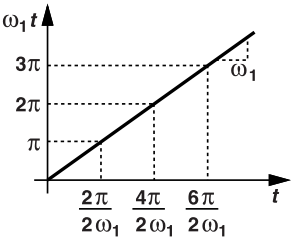


Figure 1.34 Growth of phase with time.

phase crosses multiples of π at multiples of the half period, $T/2 = 2\pi/(2\omega_1)$. Similarly, if the control voltage

of a VCO is constant, we have $V_{out} = V_0 \cos(\omega_0 t + K_{VCO} V_{cont} t)$ and hence the slope of the phase is given by $\omega_0 + K_{VCO} V_{cont}$. But, what if V_{cont} changes with time? In this case, we view $\omega_{out} = \omega_0 + K_{VCO} V_{cont}$ as the "instantaneous" output frequency, and write the total output phase, $\phi(t)$, as the time integral of ω_{out}:

$$V_{out}(t) = V_0 \cos \phi(t) \tag{1.75}$$

$$= V_0 \cos \left(\int \omega_{out} dt \right) \tag{1.76}$$

$$= V_0 \cos \left[\omega_0 t + K_{VCO} \int V_{cont}(t) dt \right], \tag{1.77}$$

where it is assumed K_{VCO} remains constant with time. As such, a VCO acts as a frequency modulator, i.e., a signal applied to its control modulates its output frequency. We also define the "excess phase," ϕ_{ex}, as

$$\phi_{ex} = K_{VCO} \int V_{cont}(t) dt. \tag{1.78}$$

In PLLs, we are primarily interested in ϕ_{ex} as the VCO output. Note that this relation corresponds to a transfer function $\phi_{ex}/V_{cont} = K_{VCO}/s$.

Example 1.19

A VCO is driven by a square wave as shown in Fig. 1.35(a). Plot the output waveform, the total output phase, and the excess output phase as a function of time.

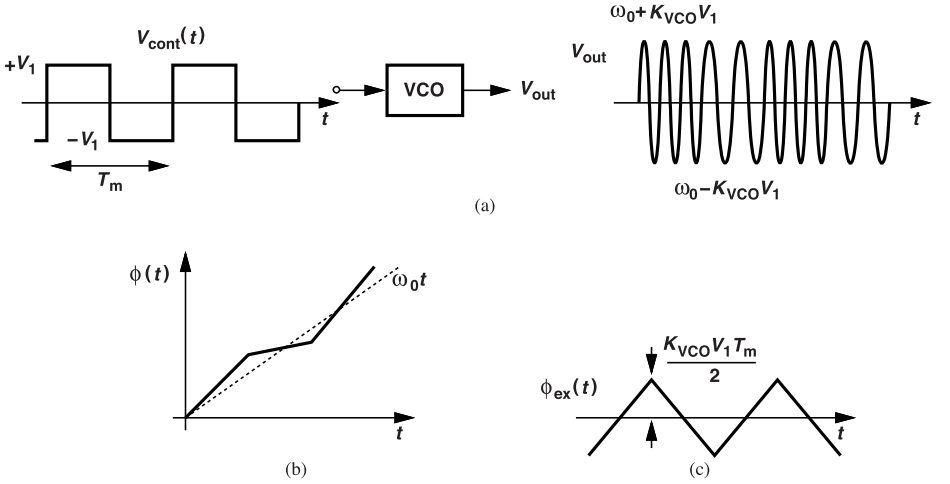


Figure 1.35 (a) VCO sensing a square wave, (b) output waveform, (c) total output phase, and (d) excess output phase .

Solution

When $V_{cont}(t)$ is equal to $+V_1$ or $-V_1$, the VCO operates at $\omega_0 + K_{VCO}V_1$ or $\omega_0 - K_{VCO}V_1$, respectively, thus producing the output waveform shown in Fig. 1.35(b). The total phase, $\phi(t) = \omega_0 t + K_{VCO} \int V_{cont} dt$, consists of a ramp, $\omega_0 t$, and the integral of a square wave, i.e., a triangular waveform [Fig. 1.35(c)]. Viewing the instantaneous frequency as $d\phi/dt$, we also observe that the slope of $\phi(t)$ toggles between $\omega_0 + K_{VCO}V_1$ and $\omega_0 - K_{VCO}V_1$. The excess phase is simply the triangular waveform, $\phi_{ex}(t) = K_{VCO} \int V_{cont}(t) dt$ [Fig. 1.35(d)], exhibiting a peak value of $K_{VCO}V_1T_m/2$ (why?).

VCO Models In summary, VCOs can be represented by one of *three* models:

1. A static system characterized by $\omega_{out} = \omega_0 + K_{VCO}V_{cont}$ if the output frequency is of interest.
2. A system generating $V_{out} = V_0 \cos[\omega_0 t + K_{VCO} \int V_{cont}(t)dt]$ if the output waveform is of interest.
3. An ideal integrator characterized by

$$\frac{\phi_{ex}}{V_{cont}}(s) = \frac{K_{VCO}}{s} \tag{1.79}$$

if the excess phase is of interest.

The last representation plays a central role in our analysis of PLLs. The integrator property implies a *dynamic* system, wherein the present value of ϕ_{ex} depends on the past values of V_{cont}. The key point here is that to change the output excess phase of a VCO, we must change V_{cont} and *wait* until $K_{VCO} \int V_{cont}dt$ reaches the desired value. In other words, the phase of a VCO cannot change instantaneously.

Example 1.20

The control voltage of a VCO jumps from zero to ΔV at $t = t_1$ and back to zero at $t = t_2$. Plot the output waveform, total phase, and the excess phase as a function of time.

Solution

As shown in Fig. 1.36(a), the output frequency jumps from ω_0 to $\omega_0 + K_{VCO}\Delta V$ and back to ω_0. The total

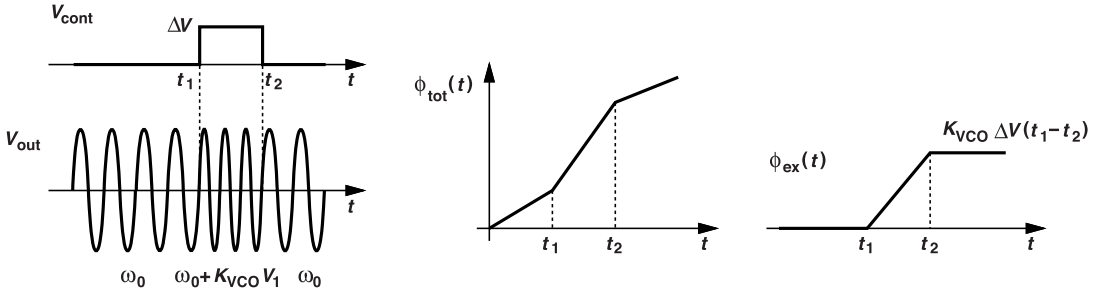


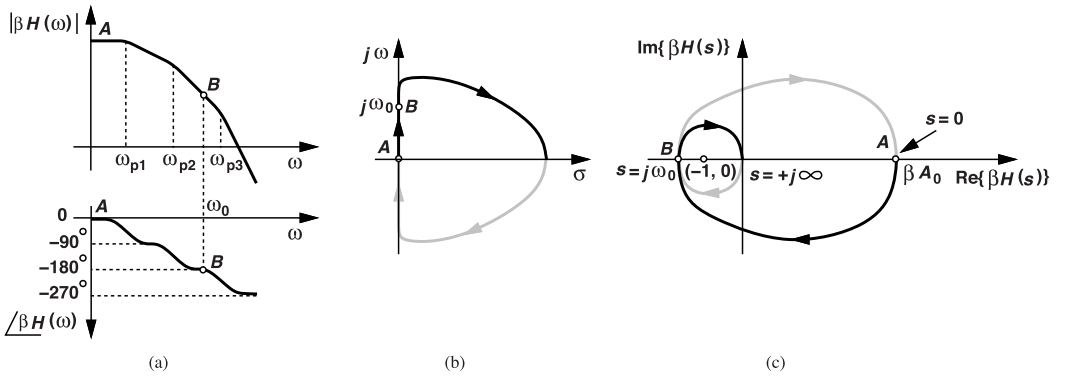
Figure 1.36 (a) Input and output waveforms of a VCO, (b) total phase, and (c) excess phase.

phase rises linearly with a slope equal to ω_0 before t_1, a slope equal to $\omega_0 + \Delta\omega$ between t_1 and t_2, and a slope equal to ω_0 again after t_2 [Fig. 1.36(b)]. The excess phase begins from zero at t_1, rises with a slope of $K_{VCO}\Delta V$ until t_2, and remains at $K_{VCO}\Delta V(t_2 - t_1)$ thereafter [Fig. 1.36(c)]. We say the VCO has accumulated an excess phase of $K_{VCO}\Delta V(t_2 - t_1)$ in this experiment. As mentioned earlier, to change the VCO excess phase, we must change the control voltage for some time. There is no other mechanism for adjusting ϕ_{ex}.

1.7 Appendix I

We wish to determine exactly what happens if $|H(j\omega_0)| > 1$ and $\angle H(j\omega_0) = -180^\circ$ in Fig. 1.10(a). Unfortunately, Bode plots do not reveal the intricacies of this case as they are confined to $s = j\omega$. We must therefore resort to Nyquist's stability criterion. The reader is referred to [2, 1] for a detailed understanding of Nyquist's approach.

Suppose $H(s)$ has three real poles, exhibiting the magnitude and phase responses shown in Fig. 1.37(a). Let us plot the values of $H(s)$ in the complex plane; $H(s)$ can be constructed in polar coordinates by plotting $|H(s)|$ vs. $\angle H(s)$. To this end, we allow s to move up on the $j\omega$ axis from 0 to ∞ [Fig. 1.37(b)], move to



(a) (b) (c)

Figure 1.37 (a) Loop transmission with three poles, (b) s-contour in the complex plane, and (c) Nyquist plot of the response.

the right half plane at very large values of $j\omega$, travel at a very large radius to $-j\infty$ (so that it reaches the $j\omega$ axis), and return to the origin on the negative $j\omega$ axis. The contour of $H(s)$ in the complex plane begins at point A [Fig. 1.37(c)], reaches point B at $s = j\omega_0$, and approaches the origin for $s \to +j\infty$ at an angle of -270°. The other (gray) half of the contour, which corresponds to $\omega < 0$, is constructed symmetrically.

We observe that if $|H(j\omega_0)| > 1$, then the H contour encircles the point $(-1, 0)$ clockwise twice, implying that $1 + H(s)$ has two zeros in the right half plane (RHP). These zeros assume a form such as $\sigma_1 \pm j\omega_1$ and become the poles of the closed-loop system, creating instability.

We recognize that poles of the form $\sigma_1 \pm j\omega_1$ signify a *growing* sinusoid. In other words, as shown in Fig. 1.38, a growing sinusoid can travel around the loop. The quantity $H(j\omega_0)$ is not equal to -1, but $H(\sigma_1 \pm j\omega_1)$ is.

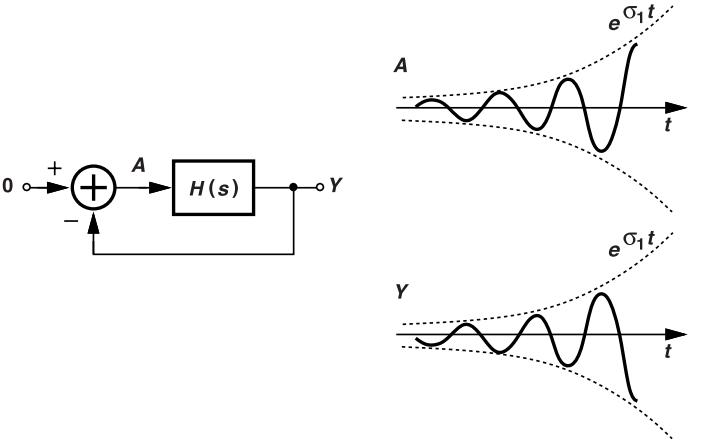


Figure 1.38 Feedback loop sustaining a growing sinusoid.

It is tempting to generalize this result and conclude that $|H(j\omega_0)| > 1$ and $\angle H(j\omega_0) = -180^\circ$ always translate to closed-loop poles in the right half plane. But this is not generally correct. As an example, consider an $H(s)$ with three poles and two zeros [Fig. 1.39(a)]. Here, $\angle H(j\omega)$ crosses -180° twice while $|H(j\omega)| > 1$ (at points A and B). The H contour appears as shown in Fig. 1.39(b), crossing the real axis at A and B and approaching the origin at an angle of -90°. Interestingly, the contour encircles $(-1, 0)$ once in the clockwise

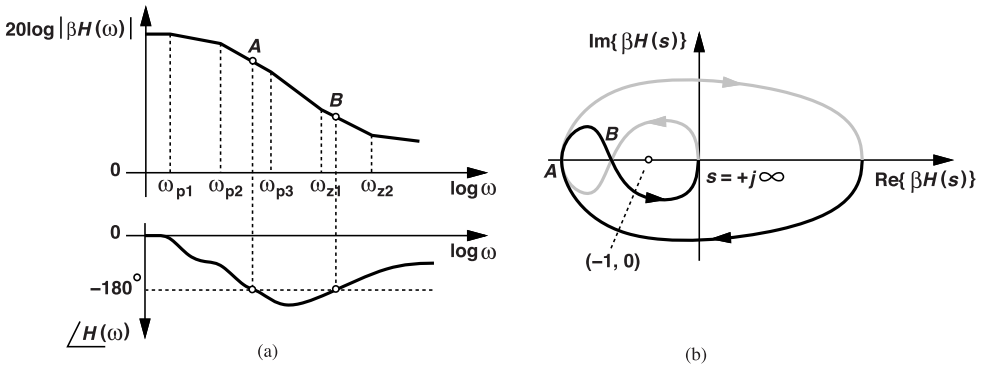


Figure 1.39 (a) Loop transmission with three poles and two zeros, and (b) Nyquist plot of frequency response.

direction and once in the counterclockwise direction. Thus, $1 + H(s)$ has no zeros in the RHP, suggesting a stable closed-loop system.

These results can be summarized as follows. If $\angle H(j\omega)$ crosses -180° an odd number of times while $|H(j\omega)| > 1$, then the system is unstable. Conversely, if $\angle H(j\omega)$ crosses -180° an even number of times while $|H(j\omega)| > 1$, then the system is stable.

References

[1]. B. Razavi, *Design of Analog CMOS Integrated Circuits,* Boston: McGraw-Hill, Second Edition, 2017.
[2]. R. C. Dorf and R. H. Bishop, *Modern Control Systems,* Reading, Massachusetts: Addison-Wesley, 1995.

Problems

1.1. Suppose I_X in Fig. 1.7(c) is an impulse, $I_0\delta(t)$. Compute V_X as a function of time, assuming small-signal operation.

1.2. In Fig. 1.7(a), V_{DD} jumps by ΔV at $t = 0$. Does the circuit begin to oscillate? Explain why or why not.

1.3. In Fig. 1.7(a), we place a voltage source in series with the source of M_1 and step it from 100 mV to zero at $t = 0$. Does the circuit begin to oscillate? Explain why or why not.

1.4. In Fig. 1.9, determine the impedance seen between nodes X and Y and therefrom obtain the oscillation condition.

1.5. In Fig. 1.12(a), compute the (closed-loop) impedance seen at node Y and determine whether it can go to infinity.

1.6. A ring oscillator employs N stages, each configured as shown in Fig. 1.40. Note that M_2 provides local positive feedback. Determine the oscillation condition and frequency.

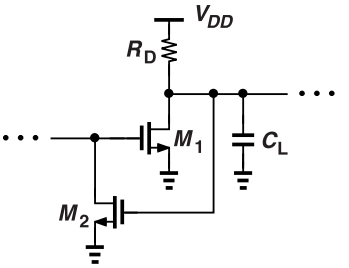


Figure 1.40 CS stage with local positive feedback.

1.7. The widths of M_1-M_6 in Fig. 1.14(a) are doubled. Explain why the frequency of oscillation remains unchanged.

1.8. For the tank in Fig. 1.16(a), derive and plot $|Z_1|$ and $\angle Z_1$ from low frequencies to high frequencies.

1.9. If the inductance value in Fig. 1.17(a) is doubled, what happens to the slope of $\angle Z_1$ at ω_0? Consider two cases: the series resistance of L_1 remains constant, or it is also doubled. Assume $R_s^2/(L_1\omega_0^2) \ll L_1$.

1.10. Repeat Example 1.11 if L_1 has a series resistance of 10 Ω and M_1 has a bias current of 2 mA.

1.11. Compute the amplitude of the third harmonic in V_X or V_Y in Fig. 1.22 normalized to that of the first harmonic.

1.12. In the circuit of Fig. 1.20(b), one inductor is slightly different from the other, $L_1 = L_2 + \Delta L$, and $\Delta L \ll L_2$. Assuming that the circuit is otherwise symmetric, compute the oscillation frequency.

1.13. Compute the (closed-loop) impedance seen between the drain of M_3 and the ground in the circuit of Fig. 1.24 and therefrom find the oscillation condition.

1.14. We insert two delay lines in series with the gates of M_1 and M_2 in Fig. 1.26(a). How much delay is necessary to allow oscillation at the resonance frequency of the tank?

1.15. Repeat the analysis of Example 1.18 for the oscillator shown in Fig. 1.30(b) by breaking the loop at the gate of M_1.

1.16. Repeat the analysis of Example 1.18 for the oscillator shown in Fig. 1.30(c) by breaking the loop at the drain of M_1 and injecting a test current, I_t, into the common terminal of L_1 and C_2. In this case, the loop transmission is given by $-I_{D1}/I_t$.

1.17. Compute the impedance seen at the source of M_1 in Fig. 1.30(b) and determine whether it goes to infinity at the oscillation frequency.

1.18. Compute the impedance seen at the gate of M_1 in Fig. 1.30(b) and determine whether it goes to infinity at the oscillation frequency.

1.19. Compute the impedance seen at the drain of M_1 in Fig. 1.30(c) and determine whether it goes to infinity at the oscillation frequency.

1.20. Consider the differential Colpitts oscillator of Fig. 1.32. Explain what happens if the two sides oscillate in the common mode. (Hint: for CM oscillations, the two sides merge into one because all of their corresponding voltages are equal. Consider the role of R_b in such a situation.)

2

Introduction to Jitter and Phase Noise

The design of phase-locked loops must deal with jitter and/or phase noise. As we will see in this chapter, oscillators bear a direct trade-off between phase noise and power consumption, requiring that we take both into account from the beginning of the design effort. In other words, it is not meaningful to design an oscillator without considering phase noise. For this reason, we study jitter and phase noise before embarking upon transistor-level design of ring and LC oscillators. The reader is encouraged to first review the VCO section in Chapter 1. This chapter serves an entry point into the interesting and complex subject of phase noise. Almost every remaining chapter of this book will deal with phase noise mechanisms.

2.1 Brief Review of Noise

The reader is expected to be familiar with the phenomenon of noise and its representation in circuits. In this section, we briefly review some of these concepts.

2.1.1 Noise in Time and Frequency Domains

Noise is a random process. For our purposes, this means that the instantaneous value of the noise in the time domain cannot be predicted accurately. The sound of traffic in a busy street, the voices of people in a crowded restaurant, and the sound of a river all exemplify random processes. We envision random waveforms such as that in Fig. 2.1(a) representing these phenomena in the time domain.

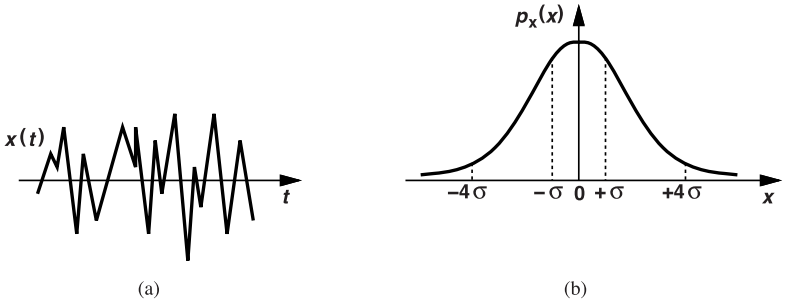


Figure 2.1 (a) Noise waveform in time domain, and (b) distribution of noise amplitude.

One property of noise that can be computed in the time domain is the probability density function (PDF) (or distribution) of its amplitude. For example, if the noise is generated by many people independently, we expect its PDF to be a Gaussian distribution [Fig. 2.1(b)]. The standard deviation of the distribution, σ, is then equal to the root mean square (rms) value of the noise. A useful rule of thumb for Gaussian noise is that

its peak value in the time domain rarely exceeds $\pm 4\sigma$. This is because the area under $p_x(x)$ for $|x| > 4\sigma$ is very small.

Most of our noise analysis occurs in the frequency domain; but, owing to its random nature, noise does not directly lend itself to the Fourier transform. That is, if we take the Fourier transform of $x(t)$ in Fig. 2.1(a), we cannot obtain a "smooth" frequency domain plot regardless of how long the $x(t)$ waveform is. We must therefore turn to a different frequency-domain characteristic that applies to random phenomena.

The concept of the "power spectral density" or simply the "spectrum" proves extremely useful in our work. We define the spectrum as the amount of power (or energy) that a signal carries in a 1-Hz bandwidth centered around each frequency. For example, we know that our voice contains frequency components from 20 Hz to 20 kHz. To construct the voice spectrum, therefore, we consider the conceptual arrangement shown in Fig. 2.2(a). A microphone converts the voice to an electric signal, which is then applied to a band-pass

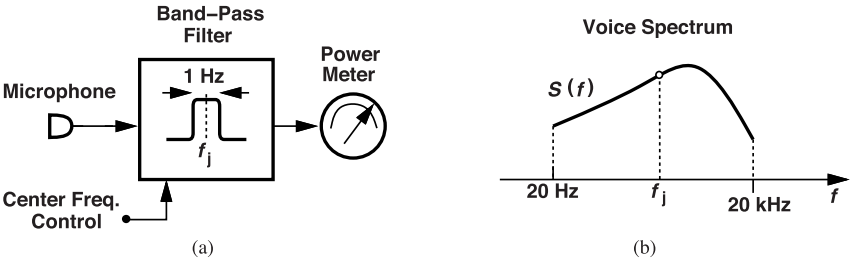


Figure 2.2 (a) Measurement, and (b) shape of spectrum.

filter having a 1-Hz bandwidth around the frequency of interest, f_j. The resulting band-limited signal drives a power meter, which displays the average power carried by the voice in a 1-Hz bandwidth. The filter's center frequency control sweeps f_j from 20 Hz to 20 kHz, yielding the profile shown in Fig. 2.2(b). We denote the spectrum by $S(f)$.

In practice, it is not easy to realize a 1-Hz filter having a high center frequency, e.g., $f_j = 1$ GHz. Instead, we use a filter with a bandwidth of, say, 1 MHz, and divide the measured output power by 10^6 to obtain the power in 1 Hz.

2.1.2 Device Noise

In CMOS design, we are primarily interested in the noise produced by resistances and MOSFETs. We review these noise sources here.

A resistor of value R converts the ambient thermal energy to a random voltage across it. The spectrum of the voltage is given by

$$S(f) = 4kTR \quad \mathrm{V}^2/\mathrm{Hz}, \tag{2.1}$$

where $k = 1.38 \times 10^{-23}$ J·K^{-1} is the Boltzmann constant and T the absolute temperature. Note that the unit of $S(f)$ is V^2/Hz rather than W/Hz. In fact, we often write this equation as

$$\overline{V_n^2} = 4kTR \quad \mathrm{V}^2/\mathrm{Hz}, \tag{2.2}$$

so as to emphasize that it represents a voltage quantity. The noise is modeled as a voltage source in series with the resistor [Fig. 2.3(a)] and its spectrum is called "white" because, like white light, it carries equal power at all frequencies [Fig. 2.3(b)].[1] In some analyses, we prefer to use the "two-sided" representation shown in Fig. 2.3(c), where the signal power is distributed on both positive and negative frequencies. Also, in some cases

[1] The actual spectrum is not flat and begins to fall above approximately 6 THz.

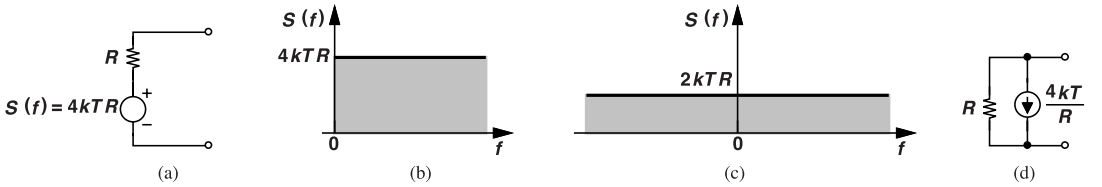


Figure 2.3 (a) Resistor noise model, (b) noise spectrum, (c) two-sided spectrum, and (d) alternative noise model.

the analysis is simpler if we model the resistor noise by a parallel current source [Fig. 2.3(d)] with a spectrum equal to $S(f) = \overline{I_n^2} = 4kT/R \quad \mathrm{A}^2/\mathrm{Hz}$. We should remark that the polarity of the voltage or current source is unimportant.

Example 2.1

We wish to compute the output noise of the network shown in Fig. 2.4(a). A student constructs the equivalent

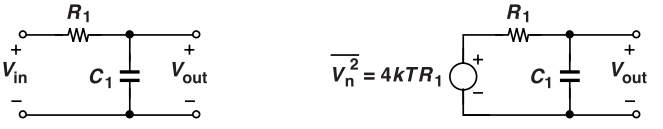


Figure 2.4 (a) Simple low-pass filter, and (b) its noise model.

circuit in Fig. 2.4(b) and writes a KVL as $V_{out} + V_{out}R_1C_1s = V_n = \sqrt{4kTR_1}$. Explain why this equation is incorrect.

Solution

The expression $\overline{V_n^2} = 4kTR$ simply means that the average noise power measured in 1 Hz is equal to $4kTR$; it does not provide any information regarding V_n itself. Thus, we cannot write $V_n = \sqrt{4kTR}$; this expression is meaningless in the frequency domain and in the time domain. The proper method of finding the output is described later in this chapter.

MOS transistors generate "flicker noise" (also called "$1/f$ noise") and thermal noise. The former is modeled as a voltage source in series with the gate [Fig. 2.5(a)] having a spectrum of

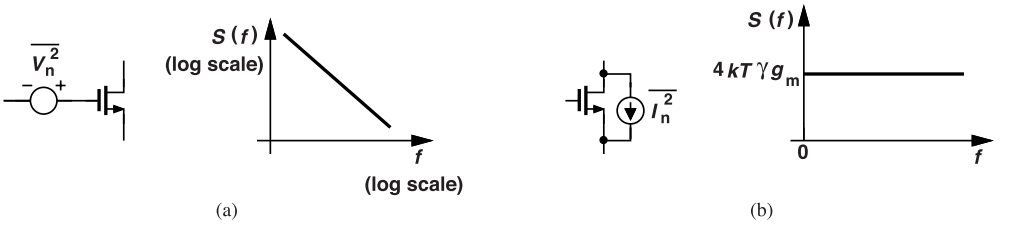


Figure 2.5 (a) MOS flicker noise model, and (b) MOS thermal noise model.

$$S_{1/f}(f) = \frac{K}{WLC_{ox}}\frac{1}{f} \quad \mathrm{V}^2/\mathrm{Hz}, \tag{2.3}$$

where K is a process-dependent parameter, and W, L, and C_{ox} denote the width, the length, and the oxide capacitance per unit area, respectively. Flicker noise becomes serious for small values of f. It also trades with the total gate capacitance WLC_{ox}.

The thermal noise of MOSFETs in the saturation region is modeled as a current source tied between the drain and source terminals [Fig. 2.5(b)] and having a white spectrum given by

$$\overline{I_n^2} = 4kT\gamma g_m \quad \mathrm{A}^2/\mathrm{Hz}, \tag{2.4}$$

where γ is called the excess noise coefficient (ideally equal to $2/3$) and is about unity in short-channel devices, and g_m is the device transconductance.

2.1.3 Propagation of Noise

A common problem in circuit and system analysis is to determine how the noise produced by a given source "propagates" to a given output [Fig. 2.6(a)]. In other words, we wish to find the transfer function experienced

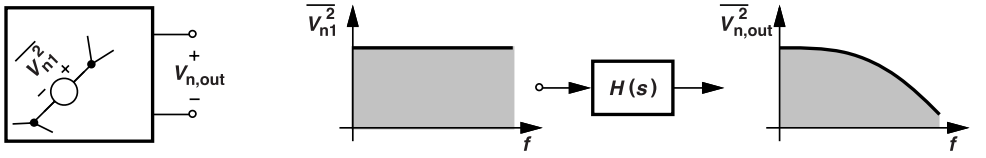


Figure 2.6 (a) Propagation of a noise source to the output, (b) shaping of noise by the transfer function.

by the noise. This task is carried out in three steps: (1) compute the transfer function, $H(s)$, from V_{n1} to $V_{n,out}$ as if V_{n1} were a deterministic source, (2) replace s with $j\omega = j2\pi f$ and determine the magnitude squared of $H(s)$ [i.e., $|H(f)|^2$], and (3) express the output spectrum as

$$S_{n,out}(f) = S_{n1}(f) \cdot |H(f)|^2, \tag{2.5}$$

where $S_{n1}(f)$ denotes the spectrum of V_{n1}. We say the noise spectrum is "shaped" by $|H(f)|^2$ as it propagates to the output. Figure 2.6(b) illustrates an example if $S_{n1}(f)$ is white and $H(s)$ a low-pass filter. This equation is similar to what we write for linear, time-invariant systems in the form of $Y(s) = X(s) \cdot H(s)$, but it applies to the power spectra of the signals.

Example 2.2

Repeat Example 2.1 using Eq. (2.5).

Solution

Since $H(s) = (R_1C_1s + 1)^{-1}$, we have[2]

$$S_{n,out}(f) = 4kTR_1 \frac{1}{R_1^2C_1^2(2\pi f)^2 + 1}. \tag{2.6}$$

Note that we have written the spectrum in terms of f rather than ω because we wish to express it in V^2/Hz rather than $\mathrm{V}^2/\mathrm{rad/s}$.

[2] The magnitude of a complex number $a + jb$ is given by $\sqrt{a^2 + b^2}$.

Example 2.3

In the common-source stage of Fig. 2.7(a), I_1 is an ideal, noiseless current source and M_1 operates in saturation. Determine the output noise spectrum.

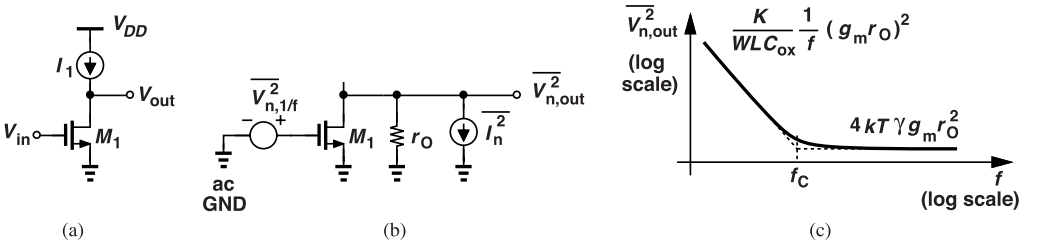


Figure 2.7 (a) Simple CS stage, (b) its noise model, and (c) output noise voltage spectrum.

Solution

We first redraw the circuit, including the noise sources. Each noise source must be multiplied by its corresponding transfer function. From Fig. 2.7(b), we observe that the $1/f$ noise spectrum experiences a gain of $(g_m r_O)^2$ as it reaches the output, and the thermal noise, $\overline{I_n^2}$, is simply multiplied by r_O^2. We then have

$$\overline{V_{n,out}^2} = \frac{K}{WLC_{ox}}\frac{1}{f}(g_m r_O)^2 + 4kT\gamma g_m (r_O)^2. \tag{2.7}$$

Figure 2.7(c) plots this spectrum on a log-log scale so that the flicker noise falls linearly with frequency. The extrapolated intersection of the $1/f$ and thermal components, f_C, is called the flicker noise "corner frequency," serving as a measure of how high this noise is. In nanometer CMOS processes, f_C reaches 50 to 100 MHz unless the transistor channel area is very large.

2.1.4 Average Power of Noise

For a periodic signal, $x_1(t)$, we define the average power as

$$P_{x1} = \frac{1}{T}\int_{-T/2}^{+T/2} x_1^2(t)dt, \tag{2.8}$$

where T denotes the period. Similarly, for a random signal, $x_2(t)$, we can compute the average power if we observe the signal for a long time:

$$P_{x2} = \overline{x_2^2(t)} = \lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{+T/2} x_2^2(t)dt. \tag{2.9}$$

The average power can also be obtained from the spectrum of the signal. Parseval's theorem states that

$$P_{x2} = \int_{-\infty}^{+\infty} S_{x2}(f)df. \tag{2.10}$$

In other words, the total average power carried by noise is equal to the area under its spectrum. Note that each frequency component in the spectrum, whether "of interest" or not, contributes to the time-domain fluctuations of the amplitude. For example, if a video signal with a bandwidth of 5 MHz is applied to a low-pass filter with a bandwidth of 10 MHz, then *all* noise frequencies produced by the filter corrupt the signal (Fig. 2.8)—unless the result undergoes further filtering so as to remove the noise components above 5 MHz.

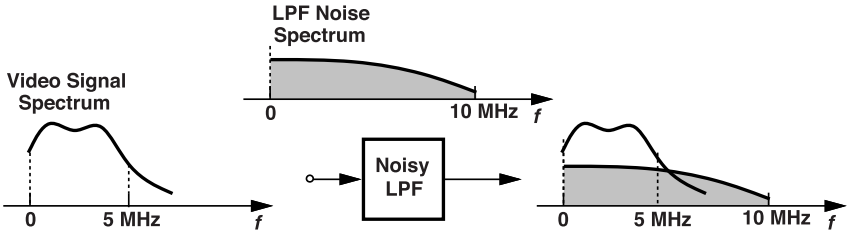


Figure 2.8 Propagation of a video signal through a noisy low-pass filter.

2.1.5 Approximation of Noise Spectrum

In this section, we describe an approximation of continuous spectra that proves useful in visualizing the effect of noise. Introduced by Rice [1], the approximation proceeds as follows. We first decompose the spectrum into narrowband slivers as depicted in Fig. 2.9(a). Next, we concentrate each sliver's area into an impulse [Fig. 2.9(b)]. Since each impulse represents a sinusoid in the time domain, we conclude that the noise can be

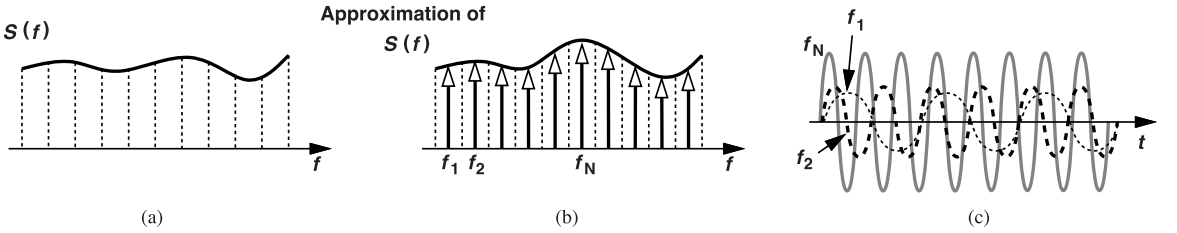


Figure 2.9 (a) Arbitrary spectrum, (b) approximation by impulses, and (c) visualization in time domain.

modeled by a large number of sinusoids [Fig. 2.9(c)].

How narrow should the slivers in Fig. 2.9(a) be? Before answering this question, we must ask another: how can the *deterministic* sinusoids in Fig. 2.9(c) represent a *random* signal? We intuitively expect that the time-domain approximation in Fig. 2.9(c) is valid only for a time scale commensurate with the slivers' bandwidths. For example, if each sliver is 1 kHz wide, the waveforms in Fig. 2.9(c) are valid for roughly 1 ms. After all, the sum of deterministic sinusoids cannot indefinitely represent a random signal. We therefore choose the slivers' bandwidth according to the time scale of interest. This view of noise readily leads to an interesting perspective that is described in the next section.

2.1.6 Accumulation of Noise with Time

Consider the noise spectrum shown in Fig. 2.10(a) along with its approximation and let us examine the noise components at f_1 and f_N in the time domain. We assume a signal of interest (not shown) also resides in this bandwidth. As depicted in Fig. 2.10(b), the former changes much more slowly than the latter. Thus, if we wish to see how the noise, $n(t)$, varies over an "observation window" of 0 to t_a, we can say that the component at f_1 changes very little in this time period and corrupts the signal of interest negligibly. Of course, the longer t_a is, the greater the effect of the noise at f_1.

The key conclusion here is that, for a given observation window, only frequency components that can change substantially in that time scale contribute to the overall noise accumulation. As a rule of thumb, we may say a component at f_1 is significant if $t_a > (1/f_1)/50$.

To arrive at another important point, let us rephrase the foregoing result as "low-frequency noise components contribute less than high-frequency components for a given (finite) observation window." This is equivalent to subjecting the noise to a *high-pass* filter (HPF) so as to accentuate the effect of high-frequency

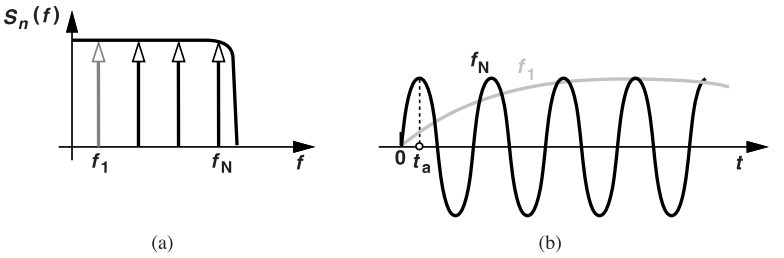


Figure 2.10 (a) Approximation of flat band-limited noise by impulses, and (b) time-domain behavior.

noise. It can be shown that the HPF transfer function is approximately equal to $\pi t_a f$ (a simple differentiator) for frequencies well below $1/(\pi t_a)$ [2]. In summary, the accumulation of noise in an observation period of t_a seconds is equivalent to subjecting all of the components to an HPF having a response given by $\pi t_a f$. This point proves useful in our discussion of cycle-to-cycle jitter later in this chapter.

2.2 Basic Jitter and Phase Noise Concepts

A noiseless oscillator generates a perfectly periodic output, e.g., in the form of a sinusoid or a square wave [Fig. 2.11(a)]. We can view such an output as (a) a waveform whose period does not change with time, (b) a

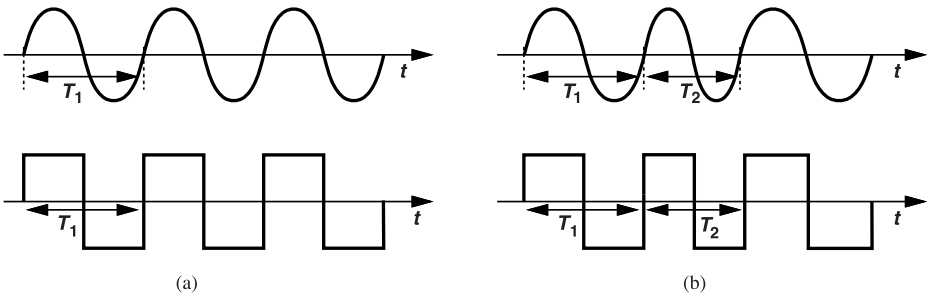


Figure 2.11 (a) Perfectly periodic signals, and (b) signals containing jitter.

waveform whose instantaneous frequency does not change with time, or (c) a waveform whose zero crossings occur uniformly and at $t = nT_1/2$, where n is an integer. These three views are equivalent.

In reality, noise sources in the oscillator circuit disturb both the amplitude and the phase of these waveforms. Let us ignore amplitude variations and observe that the noise (a) makes the periods unequal, (b) causes the instantaneous frequency to change randomly, and (c) produces zero crossings that do not occur at $nT_1/2$ [Fig. 2.11(b)]. We can say the noise randomly modulates the frequency and the phase of the output.

We know from basic communication theory that phase modulation can be expressed as $V_{out}(t) = V_0 \cos[\omega_0 t + \phi_n(t)]$, where ω_0 denotes the "carrier" frequency and, in our studies, is equal to the oscillation frequency in the absence of noise. Of course, $\phi_n(t)$ represents any of the three views mentioned above. For example, the instantaneous frequency, $\omega_0 + d\phi_n/dt$, is modulated if $d\phi_n/dt \neq 0$. (The reader is encouraged to review the VCO modeling section in Chapter 1.)

Example 2.4

Plot the total phase and excess phase for the waveform $V_{out}(t) = V_0 \sin[\omega_0 t + \phi_n(t)]$ if $\phi_n(t) \ll 1$ rad.

Solution

Depicted in Fig. 2.12, $\omega_0 t + \phi_n(t)$ generally follows the $\omega_c t$ trajectory but with small, random perturbations.

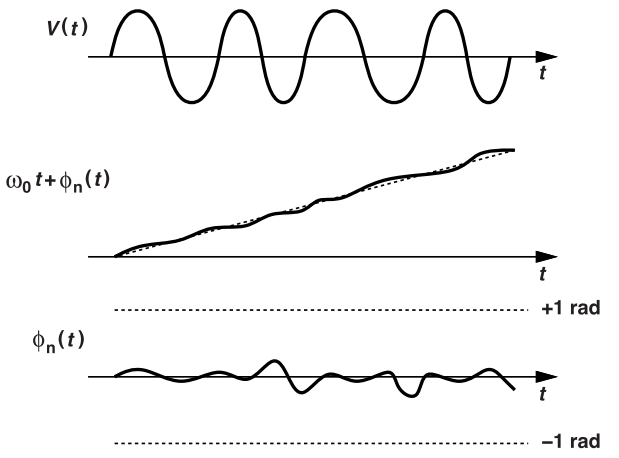


Figure 2.12 Phase noise in time domain.

As a result, the zero crossings of V_{out} are displaced from the ideal time points. Figure 2.12 also plots the excess phase, $\phi_n(t)$, whose fluctuations remain well below ± 1 rad.

2.2.1 Jitter

The departure of the zero crossings of a nominally periodic waveform from their ideal time points is called "jitter." But, in practice, the term jitter has several other definitions that will be described in Section 2.2.5. The key point to bear in mind is that jitter generally refers to a time-domain phenomenon.

Example 2.5

A VCO senses at its control voltage a small, deterministic perturbation of the form $V_{cont}(t) = V_m \cos \omega_m t$. Does the VCO output exhibit jitter?

Solution

Yes, it does. From Chapter 1, the output can be expressed as

$$V_{out}(t) = V_0 \cos(\omega_0 t + K_{VCO} \int V_m \cos \omega_m t) \tag{2.11}$$

$$= V_0 \cos \left(\omega_0 t + \frac{K_{VCO} V_m}{\omega_m} \sin \omega_m t \right). \tag{2.12}$$

We observe that some zero crossings do not occur uniformly in time. Figure 2.13 plots the total phase, ϕ_{tot}, and the excess phase, ϕ_{ex}. The maximum departure of the zero crossings from their ideal points in time occurs when the excess phase reaches a maximum or minimum. Thus, the peak-to-peak jitter is equal to $2K_{VCO}V_m/\omega_m$ radians or $[(2K_{VCO}V_m/\omega_m)/(2\pi)](2\pi/\omega_0) = 2K_{VCO}V_m/(\omega_m\omega_0)$ seconds. We can also normalize the jitter to the (average) VCO period and write the result as $[2K_{VCO}V_m/(\omega_m\omega_0)]/(2\pi/\omega_0) = K_{VCO}V_m/(\pi\omega_m)$.

Called "deterministic jitter" (DJ), this phenomenon arises if an oscillator's frequency is periodically disturbed. Such disturbance can come from the explicit control voltage or from the supply or substrate terminals.

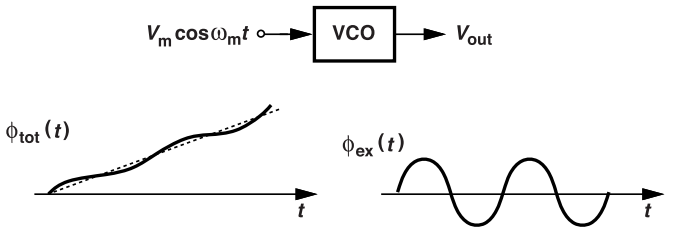


Figure 2.13 VCO modulated by a sinusoid and its phase behavior.

In order to visualize and quantify the jitter of a waveform, we can construct a cumulative diagram. Consider the jittery clock, CK_{jit}, shown in Fig. 2.14(a), where the nominal period is T_0. Using an ideal clock with the

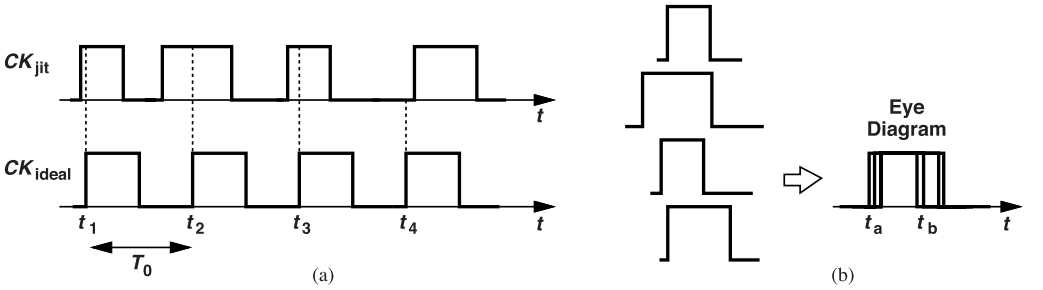


Figure 2.14 (a) Jittery waveform, and (b) construction of eye diagram.

same period, we take snapshots of CK_{jit} from t_1 to t_2, from t_2 to t_3, etc., and superimpose these cycles [Fig. 2.14(b)]. (This is how an oscilloscope displays the waveform.) The zero-crossing aberrations thus manifest themselves at times t_a and t_b. The result is called an "eye diagram."

The time-domain view of an oscillator's phase modulation does not readily reveal whether the disturbance is random or deterministic. For example, if the oscillator is disturbed periodically, as depicted in Fig. 2.13, the cumulative output waveform still appears as shown in Fig. 2.14(b). For this reason, we often resort to the frequency domain so as to distinguish between different types of disturbance.

Example 2.6

A VCO senses at its control voltage a small disturbance of the form $V_{cont}(t) = V_m \cos\omega_m t$. Determine the output spectrum.

Solution

In Eq. (2.12), we assume $K_{VCO}V_m/\omega_m \ll 1$ rad and, using the approximation $\sin\theta \approx \theta$ and $\cos\theta \approx 1$ for $\theta \ll 1$ rad, write

$$V_{out}(t) \approx V_0 \cos\omega_0 t - V_0 \frac{K_{VCO}V_m}{\omega_m} \sin\omega_0 t \sin\omega_m t \tag{2.13}$$

$$\approx V_0 \cos\omega_0 t - V_0 \frac{K_{VCO}V_m}{2\omega_m} [\cos(\omega_0 - \omega_m)t - \cos(\omega_0 + \omega_m)t]. \tag{2.14}$$

This simplification is called the "narrowband FM approximation."

Shown in Fig. 2.15, the output spectrum consists of a main component at ω_0 (called the "carrier") and two "sidebands" (also called "spurs") symmetrically disposed around it. Of particular interest to us is the ratio of

the sidebands' magnitude and the carrier magnitude, which amounts to $K_{VCO}V_m/(2\omega_m)$. Note that, as the modulating frequency, ω_m, increases, the ratio falls, and so does the jitter.

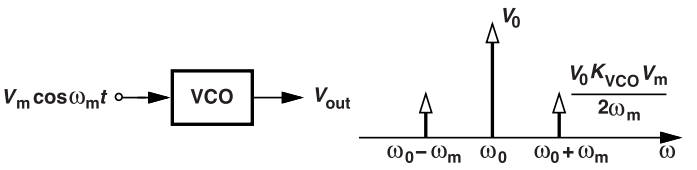


Figure 2.15 Narrowband approximation for a VCO modulated by a sinusoid.

The foregoing example demonstrates that the spectrum of the signal can reveal much about the nature of the jitter. In fact, from the relative magnitude of the sidebands, we can predict the peak-to-peak jitter. As derived in Example 2.5, the peak-to-peak deviation of the zero crossings from their ideal points in time is equal to $2K_{VCO}V_m/\omega_m$ radians. Thus, the normalized sidebands' magnitude multiplied by 4 yields the peak-to-peak jitter in radians.

Example 2.7

We wish the FM sidebands produced by an oscillator to contribute a peak-to-peak jitter less than 1% of the period. What is the maximum tolerable sideband magnitude?

Solution

For the normalized output peak-to-peak jitter to be less than 1%, we have

$$\frac{K_{VCO}V_m}{\pi\omega_m} = 1\%. \tag{2.15}$$

This implies that the relative sideband magnitude is

$$\frac{K_{VCO}V_m}{2\omega_m} = \frac{\pi}{200} \tag{2.16}$$

$$= -36 \text{ dB}. \tag{2.17}$$

This is a useful rule of thumb relating the time- and frequency-domain characteristics of deterministic jitter.

Example 2.8

The LC oscillator shown in Fig. 2.16 experiences supply noise in the form of $V_{DD} = V_{DD0} + V_m \cos \omega_m t$. Determine the resulting output sideband magnitude.

Solution

In this circuit, changes in V_{DD} modulate the oscillation frequency. This is because the output common-mode level is approximately equal to V_{DD} and its variation modulates the drain-bulk junction capacitance, C_{DB}, of M_1 and M_2. We thus compute a gain, K_{VCO}, from V_{DD} to the output frequency—as if V_{DD} were the control voltage. The capacitance at X (or Y) consists of a constant component, C_1, and a voltage-dependent

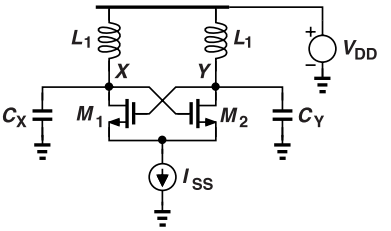


Figure 2.16 LC VCO with supply noise.

component arising from C_{DB}. The latter can be expressed as

$$C_{DB} = \frac{C_{DB0}}{(1 + \frac{V_{DD}}{\phi_B})^m}, \tag{2.18}$$

where m is around 0.3 and ϕ_B is the junction built-in potential and approximately equal to 0.8 V. Since the oscillation frequency, $\omega_1 = 1/\sqrt{L_1(C_1 + C_{DB})}$, we have

$$K_{VCO} = \frac{\partial \omega_1}{\partial V_{DD}} \tag{2.19}$$

$$= \frac{\partial \omega_1}{\partial C_{DB}} \frac{\partial C_{DB}}{\partial V_{DD}}. \tag{2.20}$$

Taking the derivative of ω_1 and C_{DB} gives

$$K_{VCO} = \frac{1}{2\sqrt{L_1(C_1 + C_{DB})}(C_1 + C_{DB})} \frac{mC_{DB0}}{\phi_B(1 + \frac{V_{DD}}{\phi_B})^{m+1}}. \tag{2.21}$$

The normalized magnitude of each sideband is equal to $K_{VCO}V_m/(2\omega_m)$.

2.2.2 Phase Noise

Recall that the aberrations in the zero crossings or the period can be viewed as phase modulation and expressed as $V_{out}(t) = V_0 \cos[\omega_0 t + \phi_n(t)]$. We call $\phi_n(t)$ the phase noise. The reader may then wonder whether there is any difference between phase noise and jitter. In general, phase noise refers only to the *random* fluctuations of the phase or the zero crossings, whereas jitter may contain a deterministic (periodic) component, as explained in Example 2.5. It is also more common to characterize phase noise in the frequency domain.

In order to examine the spectrum of a waveform having phase noise, we first make a qualitative observation: since $\phi_n(t)$ randomly modulates the phase and frequency of V_{out}, the spectrum is no longer a single impulse at $\omega = \omega_0$. That is, the instantaneous frequency randomly departs from ω_0, spilling some of the signal's energy to the vicinity of the impulse. As a result, the spectrum is broadened (Fig. 2.17).

Let us assume $\phi_n(t) \ll 1$ rad and use the narrowband FM approximation (Example 2.6) to write

$$V_{out}(t) = V_0 \cos[\omega_0 t + \phi_n(t)] \tag{2.22}$$

$$\approx V_0 \cos \omega_0 t - V_0 \phi_n(t) \sin \omega_0 t. \tag{2.23}$$

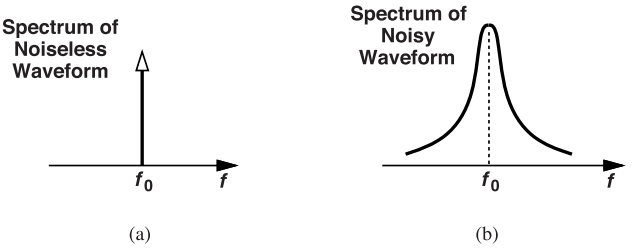


Figure 2.17 Spectra of (a) a noiseless sinusoidal waveform, and (b) a noisy sinusoidal waveform.

We recognize a carrier with an amplitude of V_0 along with the product of $\phi_n(t)$ and $V_0 \sin \omega_0 t$, which corresponds to a shift in the center frequency of the spectrum of ϕ_n by an amount equal to ω_0. Thus, if the spectrum of ϕ_n is centered around zero, the output spectrum appears as shown in Fig. 2.18, confirming that

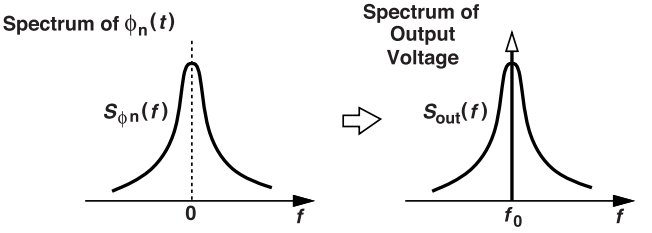


Figure 2.18 Relationship between phase noise spectrum and output voltage spectrum.

the frequency randomly deviates from f_0.

We intuitively expect that larger random frequency errors occur with *lower* probabilities because the oscillator prefers to maintain $f = f_0$ most of the time. This explains the declining shape of the spectrum as $|f - f_0|$ increases. Of course, we must eventually derive an expression for this phase noise "profile."

The reader may observe a discrepancy between the output voltage spectra shown in Figs. 2.17(b) and 2.18: the former does not contain an impulse but the latter does. We return to this point in Section 2.2.3 and conclude that the former is correct.

Since the phase noise spectrum is not flat, we must specify its value at different "frequency offsets," i.e., at different offsets from f_0 in Fig. 2.18. Of course, if the shape of $S_{out}(f)$ is unique and fully defined by *one* point, then we can simply specify the phase noise at one frequency offset. Illustrated in Fig. 2.19, the procedure consists of four steps: (1) select a frequency offset, Δf (e.g., 100 kHz), (2) calculate the noise

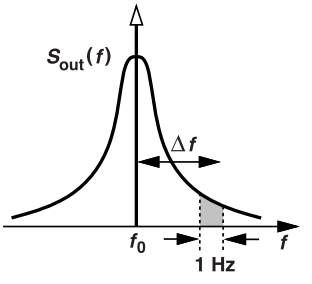


Figure 2.19 Spectrum showing how phase noise is specified.

power in a 1-Hz bandwidth at Δf, (3) normalize this power to the carrier power, given by $V_0^2/2$ in Eq. (2.23), and (4) take $10 \log$ of this ratio. The result is expressed in dBc/Hz, where the letter c indicates normalization

to the carrier power and per hertz signifies the phase noise power in a 1-Hz bandwidth. For example, a VCO in a 900-MHz GSM transmitter must achieve a phase noise of about -145 dBc/Hz at 25-MHz offset.

Example 2.9

Referring to Fig. 2.18, explain how $S_{\phi n}$ and the normalized $S_{out}(f)$ are related.

Solution

The output contains the upconverted spectrum, $V_0\phi_n \sin\omega_c t$. Since ϕ_n is multiplied by $V_0 \sin\omega_c t$, its spectrum is scaled by a factor of $V_0^2/4$ as it lands at $\pm f_0$. The one-sided spectrum is thus given by $2(V_0^2/4)S_{\phi n}(f - f_0)$, which, upon normalization to a carrier power of $V_0^2/2$, yields $S_{\phi n}(f - f_0)$. Thus, $S_{\phi n}(f)$ and the normalized one-sided output spectrum are the same except for a frequency shift. These transformations are depicted in Fig. 2.20.

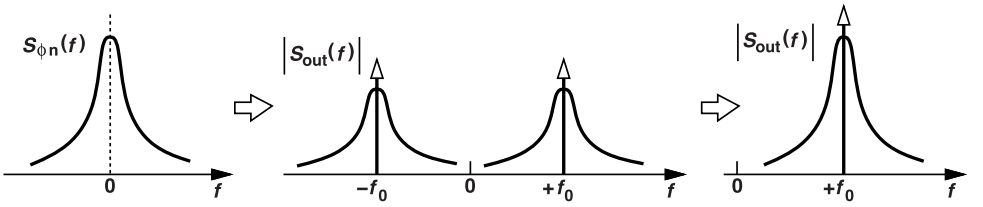


Figure 2.20 Phase noise and output voltage spectra. (As explained in Section 2.2.3, the actual voltage spectra do not contain impulses at f_0.)

Example 2.10

We connect the output of a 900-MHz oscillator to a spectrum analyzer and observe the spectrum shown in Fig. 2.21(a). Apparently, there is no impulse corresponding to $V_0 \cos\omega_0 t$ in Eq. (2.23). How should we compute the phase noise at 100-kHz offset?

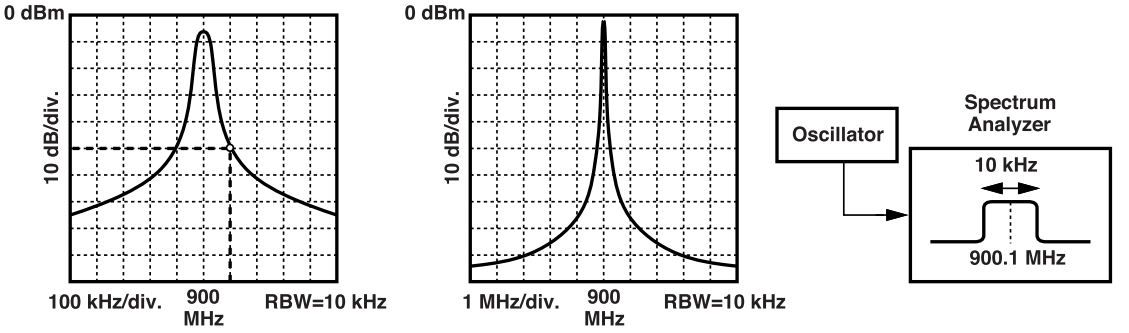


Figure 2.21 Spectrum measured on a spectrum analyzer with (a) 100 kHz/div. scale, (b) 1 MHz/div. scale, and (c) illustration of RBW.

Solution

To measure the carrier power, we must zoom out by increasing the frequency span, i.e., by increasing the scale on the horizontal scale. For example, with a scale of 1 MHz/div., the spectrum appears as in Fig.

2.21(b), resembling an impulse and yielding a carrier power of -3 dBm.[3] This peak is slightly higher than that in Fig. 2.21(a). To understand the reason, note that slow frequency fluctuations tend to shift the spectrum to the left and to the right, an effect that is more visible with a scale of 100 kHz/div. For example, the center frequency may slowly wander to 900 MHz$\pm$50 kHz. As a result, the peak is "smeared" more in Fig. 2.21(a) than in Fig. 2.21(b).[4]

The phase noise at 100-kHz offset is measured in Fig. 2.21(a) to be equal to -50 dBm, i.e., 47 dB below the carrier power of -3 dBm. However, due to the high center frequency, this noise power is not measured in a 1-Hz bandwidth, but, rather, in a "resolution bandwidth" (RBW) of 10 kHz. This means that the spectrum analyzer places at $900\ \text{MHz} + 100$ kHz a band-pass filter having a bandwidth of 10 kHz [Fig. 2.21(c)] and measures the noise power carried by the signal. The noise power in 1-Hz is thus equal to $-50\ \text{dBm} - 10\log(10\ \text{kHz}) = -90$ dBm, yielding a phase noise of $-3\ \text{dBm} + (-90\ \text{dBm/Hz}) = -87$ dBc/Hz at 100-kHz offset. It is important not to confuse the RBW with the frequency offset.

Example 2.11

A resistor R_1 is tied between the control voltage of a VCO and ground [Fig. 2.22(a)]. Determine (a) the output phase noise spectrum, and (b) the output frequency noise spectrum.

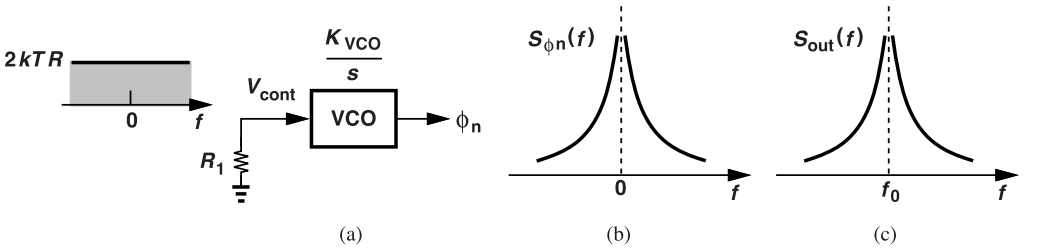


Figure 2.22 (a) VCO modulated by thermal noise, (b) resulting phase noise spectrum, and (c) output voltage spectrum.

Solution

(a) Since phase noise is the excess phase at the output, the transfer function of interest is that from V_{cont} to $\phi_{ex} = \phi_n$, i.e., K_{VCO}/s. From Eq. (2.5),

$$S_{\phi n}(f) = 2kTR_1 \left(\frac{K_{VCO}}{2\pi f}\right)^2 \tag{2.24}$$

$$= \frac{kTR_1K_{VCO}^2}{2\pi^2 f^2}. \tag{2.25}$$

Note that the resistor noise is represented by a two-sided spectrum so that, upon multiplication by $K_{VCO}^2/(2\pi f)^2$, it yields a two-sided spectrum for $S_{\phi n}$. Sketched in Fig. 2.22(b) is this spectrum. As predicted in Example 2.9, the normalized spectrum of the output waveform, $V_0\cos[\omega_0 t + \phi_n(t)]$, is the same as $S_{\phi n}(f)$ but translated to a center frequency of f_0 [Fig. 2.22(c)]:

$$S_{out}(f) = \frac{kTR_1K_{VCO}^2}{2\pi^2(f - f_0)^2}. \tag{2.26}$$

[3] The unit dBm is used for power quantities and is obtained as $10\log P$, where P is expressed in mW.

[4] The choice of the horizontal scale depends on how noisy the oscillator is.

The quantity $f - f_0$ is the frequency offset and typically denoted by Δf. This result is not exactly correct and will be revisited in Section 2.2.3.

A mistake found in some papers and books is that they multiply the one-sided spectrum of the noise preceding the VCO by $K_{VCO}^2/(2\pi f)^2$. One must bear in mind that using a two-sided spectrum for the noise applied to V_{cont} is more straightforward.

(b) The output frequency is related to the control voltage by the VCO characteristic, $\omega_{out} = K_{VCO}V_{cont} + \omega_0$. Thus, changes in V_{out} directly modulate ω_{out}; i.e., the frequency also contains white noise. From Eq. (2.5), we multiply the input noise spectrum by the magnitude squared of the transfer function, K_{VCO}^2, to obtain the spectrum of frequency noise:

$$S_{\omega out}(f) = 2kTR_1K_{VCO}^2. \tag{2.27}$$

An important observation here is that any mechanism that modulates an oscillator's *frequency* by white noise yields a phase noise profile proportional to $1/f^2$. This is to be expected as frequency and phase are related by a factor of $1/s$.

The foregoing example leads to two important results. First, the phase noise of a VCO can be represented by two different models: (1) as a voltage quantity added to the control voltage [Fig. 2.23(a)], or (2) as a phase quantity added to the output phase [Fig. 2.23(b)]. Of course, the two additive quantities have different spectra.

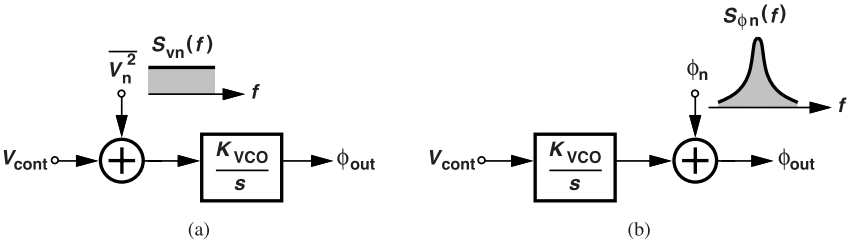


Figure 2.23 VCO phase noise model (a) using a noise voltage added to V_{cont}, or (b) using additive phase at the output.

Second, Eq. (2.25) reveals that arbitrarily low frequencies in the noise applied to V_{cont} produce arbitrarily large phase fluctuations at the output because $S_{\phi n} \propto 1/f^2$. We say the (excess) phase of a stand-alone VCO is unbounded. Illustrated in Fig. 2.24, this means that the phase difference, $\Delta\phi$, between a noisy VCO and a

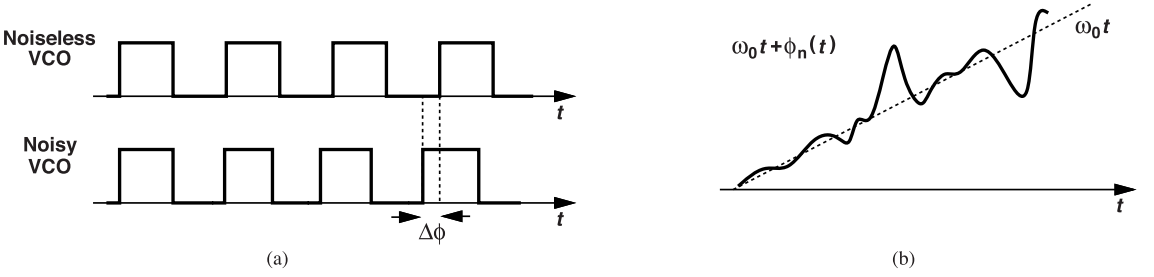


Figure 2.24 (a) Noiseless and noisy VCO output waveforms, and (b) unbounded phase fluctuations in noisy VCO.

noiseless VCO having the same nominal oscillation frequency can grow indefinitely. We can also say that the white-noise-induced jitter of a stand-alone ("free-running") VCO can become arbitrarily large with respect to an ideal clock. By contrast, as shown in Chapter 7, the jitter and phase noise of a phase-locked VCO are bounded.

The scenario depicted in Fig. 2.24(a) also exemplifies how we compute the "absolute jitter," defined as the phase difference between the oscillator of interest and an ideal oscillator as time progresses. (In the general case, absolute jitter can contain both random and periodic components.) We say the absolute jitter of a stand-alone, noisy oscillator is unbounded.

2.2.3 Limitations of Narrowband FM Approximation

Although mathematically convenient, the narrowband FM approximation, $V_0 \cos[\omega_0 t + \phi_n(t)] \approx V_0 \cos \omega_0 t - \phi_n(t) V_0 \sin \omega_n t$, does lead to some inconsistencies. First, we have surmised that frequency noise or phase noise tends to broaden the spectral impulse at ω_0, but this equation still contains an impulse (Fig. 2.18). Second, while we have assumed $\phi_n(t) \ll 1$ rad, Fig. 2.24 suggests otherwise: the excess phase can become arbitrarily large in the time domain. We typically use the two spectra, $S_{\phi n}(f)$ and $S_{out}(f)$, interchangeably, but we must resolve these inconsistencies.

Suppose the frequency of an oscillator is modulated by white noise, as in Example 2.11. Modeling the resulting frequency noise by $S_\omega = \eta/2$, where S_ω is a two-sided spectrum and η a constant, we know that $S_{\phi n}(f) = S_\omega/\omega^2 = (\eta/2)/(4\pi f)^2$. The question is, how do we determine the spectrum of the output voltage, $V_0 \cos[\omega_0 t + \phi_n(t)]$? This is accomplished by computing the autocorrelation of the waveform and taking the Fourier transform of the result [3]:

$$S_{out}(f) = \frac{V_0^2(\eta/4)}{(\omega_0 - \omega)^2 + \eta^2/16}. \tag{2.28}$$

The quantity $\eta/4$ is called the "phase diffusion constant" and denoted by D_ϕ. Sketched in Fig. 2.25, $S_{out}(f)$ is called a "Lorentzian" spectrum, reaching a peak of V_0^2 at $\omega = \omega_0 = 2\pi f_0$. Note that there is no impulse at f_0.

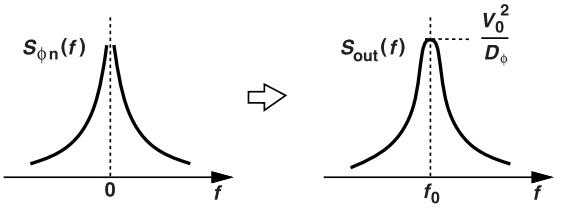


Figure 2.25 Lorentzian output voltage spectrum corresponding to white-noise-induced phase noise.

In reality, D_ϕ^2 is small enough that it can be neglected with respect to $(\omega - \omega_0)^2$ for $(\omega - \omega_0)/(2\pi)$ above a few hundred hertz. In other words, the narrowband FM approximation still serves us well in most cases, obviating the need for distinction between $S_{\phi n}(f)$ and $S_{out}(f)$.

The foregoing studies have established that white noise driving the control voltage of a VCO generates white frequency noise and a phase noise profile proportional to $1/\Delta f^2$. We must also examine the effect of flicker noise, as illustrated by the following example.

Example 2.12

Recall from Chapter 1 and Example 2.8 that oscillators generally are sensitive to their supply voltage, displaying a finite gain from V_{DD} to ω_{out}. For this reason, VCOs are often fed from a dedicated on-chip voltage regulator [Fig. 2.26(a)] so as to isolate their supply from the noise produced by other circuitry on the chip. Unfortunately, regulators themselves suffer from flicker noise. Modeling the noise on V_{DD} as α/f, where α is a constant, determine the output phase noise spectrum.

Solution

The gain from V_{DD} to ω_{out}, K_{VCO}, can be computed as outlined in Example 2.8. With this K_{VCO} known,

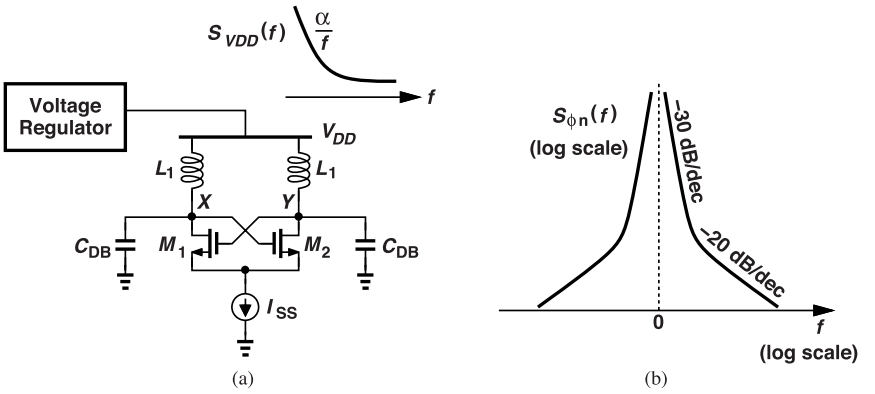


Figure 2.26 (a) VCO with $1/f$ noise on its supply, and (b) resulting phase noise spectrum.

we have

$$S_{\phi n}(f) = \frac{\alpha}{f}\frac{K_{VCO}^2}{(2\pi f)^2} \tag{2.29}$$

$$S_{\phi n}(f) = \frac{\alpha K_{VCO}^2}{4\pi^2 f^3}. \tag{2.30}$$

We say flicker-noise-induced phase noise follows a $1/f^3$ profile. Illustrated in Fig. 2.26(b), this behavior manifests itself as a -30-dB/dec slope on a log-log scale,[5] followed by a -20 dB/dec slope corresponding to white-noise-induced phase noise. In this case, too, the phase fluctuations at low offset frequencies (e.g., below 10 Hz) become very large, violating the narrowband FM approximation and yielding a Lorentzian shape, but we are typically interested in much higher offsets and can assume S_{out} (the output voltage spectrum) is a shifted copy of $S_{\phi n}$.

2.2.4 Relationship between Jitter and Phase Noise

We have seen that jitter in the time domain can arise from (deterministic) sidebands or (random) phase noise. We learned in Section 2.2.1 that the peak-to-peak deterministic jitter in radians is equal to 4 times the normalized magnitude of the output sidebands. In this section, we derive the relation between random jitter and the phase noise spectrum. The randomness suggests that we cannot specify a peak-to-peak jitter and must seek an rms value.[6]

Consider the waveform $V_0 \cos[\omega_0 t + \phi_n(t)]$. We know that $\phi_n(t)$ represents the fluctuations in the zero crossings and, as with any other time-domain quantity, exhibits an average "power" given by

$$\phi_n^2(t) = \lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{+T/2} \phi_n^2(t)dt. \tag{2.31}$$

(Recall the definition of average power in Section 2.1.4.) This quantity is difficult to compute directly, but we

[5] Since the spectrum represents power (or squared) quantities, we take $10\log$ of $S(f)$.

[6] For Gaussian jitter, we can say that the peak-to-peak jitter rarely exceeds 8 times the rms value.

also know from Parseval's theorem that the average power is equal to the area under the spectrum:

$$\overline{\phi_n^2(t)} = \int_{-\infty}^{+\infty} S_{\phi n}(f)df. \tag{2.32}$$

The rms jitter, J_{rms}, is equal to the square root of $\overline{\phi_n^2(t)}$:

$$J_{rms} = \sqrt{\int_{-\infty}^{+\infty} S_{\phi n}(f)df}. \tag{2.33}$$

This relationship proves extremely useful in computing the rms (random) jitter from the phase noise profile. Note that J_{rms} is in radians and must be divided by 2π and multiplied by the period if we wish to express it in seconds.

Example 2.13

The output phase noise of a phase-locked loop can be approximated by the profile shown in Fig. 2.27, where $S_{\phi n}(f)$ is flat for $|f| < f_1$ and falls in proportion to $1/f^2$ for $|f| > f_1$. Compute the rms jitter.

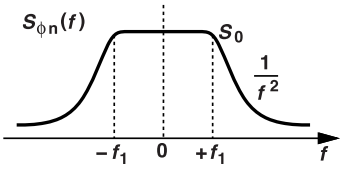


Figure 2.27 Spectrum of phase-locked VCO.

Solution

We have

$$\overline{\phi_n^2(t)} = 2\int_0^{f1} S_0 df + 2\int_{f1}^{+\infty} S_{\phi n}(f)df, \tag{2.34}$$

where the factor of 2 accounts for the power symmetrically distributed over positive and negative frequencies. Since $S_{\phi n}(f) = S_0$ at $|f| = f_1$, the profile beyond this frequency is given by $S_0 f_1^2/f^2$. It follows that

$$\overline{\phi_n^2(t)} = 2S_0 f_1 + 2\int_{f1}^{+\infty} \frac{S_0 f_1^2}{f^2} df \tag{2.35}$$

$$= 2S_0 f_1 + 2S_0 f_1 \tag{2.36}$$

$$= 4S_0 f_1. \tag{2.37}$$

Interestingly, the rms jitter is equal to 4 times the area under the spectrum from 0 to $f = +f_1$. It is useful to memorize this result.

Example 2.14

The control line of a VCO is driven by a resistor R_1. Determine the output rms jitter.

Solution

As explained in Example 2.11, in response to $V_{n,R1}^2 = 2kTR_1$, the VCO produces a phase noise profile given by $S_{\phi n}(f) = 2kTR_1K_{VCO}^2/(2\pi f)^2$. The area under this profile is infinite, which agrees with our observation that the phase fluctuations can accumulate indefinitely [Fig. 2.24(b)]. A similar situation arises if the VCO is modulated by flicker noise. We say the "absolute" jitter of a stand-alone oscillator is unbounded and its rms value is infinite.

2.2.5 Types of Jitter

We have thus far seen two types of jitter: (a) deterministic jitter, resulting from periodic frequency modulation (manifested by sidebands in the spectrum), and (b) absolute jitter, defined as the phase difference between a noisy oscillator and a noiseless oscillator running at the same nominal frequency.

That the absolute jitter is unbounded presents a quandary. How do we specify the jitter of a stand-alone oscillator? Fortunately, we can consider the "cycle-to-cycle" jitter, J_{cc}, i.e., the difference between consecutive periods, $T_k - T_{k-1}$. Of course, this measurement must be performed for many cycles, with the rms value defined as

$$J_{cc,rms} = \lim_{n\to\infty} \sqrt{\frac{1}{n}\sum_{k=1}^{n}(T_k - T_{k-1})^2}. \tag{2.38}$$

We expect the cycle-to-cycle jitter to be very small (but not negligible!) because phase noise components do not have much time to accumulate from one output period to the next (Section 2.1.6). As an example, consider the white-noise-induced phase noise profile depicted in Fig. 2.28(a). We assume $S_{\phi n}(f) = \alpha/f^2$, where α is a constant. This profile contains all noise frequencies but with declining power levels as $|f|$ increases. Now,

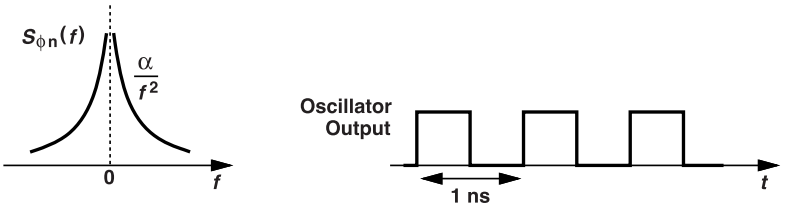


Figure 2.28 (a) Phase noise resulting from white noise, and (b) a 1-GHz oscillator output waveform.

suppose this phase noise appears at the output of a 1-GHz oscillator [Fig. 2.28(b)]. Since each cycle is 1 ns long, we surmise that only high-frequency phase noise components can affect the zero crossings 1 ns later. For example, a phase noise component around 1 MHz can be approximated by $\phi_0 \cos[2\pi(1\text{ MHz})t]$, which changes negligibly in 1 ns.

The foregoing thoughts are similar to our study of noise accumulation in Section 2.1.6, suggesting that phase noise experiences high-pass filtering if observed for a limited time, i.e., for one period. It can be shown that for white-noise-induced phase noise, the cycle-to-cycle jitter of an oscillator running at ω_0 is equal to [3]

$$J_{cc,rms}^2 = \frac{4\pi}{\omega_0^3}\omega^2 S_{\phi n}(\omega), \tag{2.39}$$

as if $S_{\phi n}(\omega)$ has traveled through a differentiator. If $S_{\phi n}(\omega)$ is of the form α/f^2, we have

$$J_{cc,rms} = \sqrt{\frac{2\alpha}{f_0^3}}. \tag{2.40}$$

Normalized to the oscillation period, $T_0 = 1/f_0$, this result emerges as

$$\frac{J_{cc,rms}}{T_0} = \sqrt{\frac{2\alpha}{f_0}}. \tag{2.41}$$

Example 2.15

A 1-GHz oscillator exhibits a phase noise of -90 dBc/Hz at 100-kHz offset. (a) Determine the cycle-to-cycle jitter if only white noise sources are considered. (b) How much jitter do we expect after 50 cycles?

Solution

(a) From the phase noise value, we can compute α:

$$10 \log S_{\phi n}(f = 100 \text{ kHz}) = -90 \text{ dBc/Hz} \tag{2.42}$$

and, thus,

$$\frac{\alpha}{(100 \text{ kHz})^2} = 10^{-9}. \tag{2.43}$$

It follows that $\alpha = 10$ Hz and $J_{cc,rms} = \sqrt{20/(1 \text{ GHz})^3} = 0.14$ ps.

(b) The jitter accumulates randomly with time. If the cycle-to-cycle jitters of consecutive cycles are uncorrelated, we can add their squares:

$$J_N^2 = J_{1,2}^2 + J_{2,3}^2 + \cdots + J_{N-1,N}^2, \tag{2.44}$$

where J_N^2 denotes the jitter (with respect to a noiseless oscillator) after N cycles, and $J_{m,n}$ is the jitter between the mth cycle and the nth cycle. We then have

$$J_N = \sqrt{N} J_{cc,rms}. \tag{2.45}$$

In this example, $J_{50} \approx 1$ ps. It is interesting to note that it takes roughly 50 million cycles (50 ms) for the jitter to accumulate to one period (1 ns).

Another type of jitter that remains bounded is the "period jitter" (also called the "cycle jitter"). Representing the random departures of the period from its *average* value, T_{avg}, this type is expressed as

$$J_c = \lim_{n \to \infty} \sqrt{\frac{1}{n} \sum_{k=1}^{n} (T_k - T_{avg})^2}. \tag{2.46}$$

It can be shown that $J_c = J_{cc}/\sqrt{2}$ for white-noise-induced jitter [3].

Table 2.1 summarizes the three common types of jitter and their characteristics. The deterministic jitter can be obtained from the magnitude of the output FM sidebands and the absolute jitter is unbounded for a stand-alone oscillator. The rms cycle-to-cycle jitter resulting from white noise sources is computed from Eq. (2.40).

2.3 Trade-Off Between Phase Noise and Power

In this section, we show that (random) phase noise directly trades with power dissipation; hence, in most cases, the power drawn by an oscillator is dictated by the phase noise specification.

Let us take N nominally identical oscillators and add their output voltages (Fig. 2.29). Using Eq. (2.23),

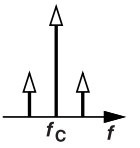
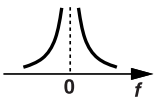
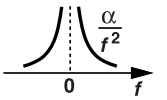
Deterministic Jitter Due to Periodic Modulation		$J_{pp} = \frac{2K_{VCO}V_m}{\omega_m}$
Absolute Jitter Due to White Noise		$J_{rms} = \infty$
Cycle–to–Cycle Jitter Due to White Noise		$J_{rms} = \sqrt{\frac{2\alpha}{f_0^3}}$

Table 2.1 Common types of jitter and their values for stand-alone oscillators.

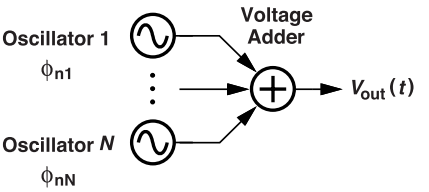


Figure 2.29 Addition of output voltages of N oscillators.

we can express the output as

$$V_{out}(t) = V_0 \cos\omega_1 t - V_0\phi_{n1}\sin\omega_1 t + \cdots + V_0\cos\omega_1 t - V_0\phi_{nN}\sin\omega_1 t, \tag{2.47}$$

where ω_1 is the average oscillation frequency and $\phi_{n1}, ..., \phi_{nN}$ denote the phase noise of the individual oscillators. It follows that

$$V_{out}(t) = NV_0\cos\omega_1 t - V_0(\phi_{n1} + \cdots + \phi_{nN})\sin\omega_1 t. \tag{2.48}$$

The normalized phase noise is equal to $(\phi_{n1} + \cdots + \phi_{nN})/N$ and exhibits a spectrum given by

$$S_\phi(f) = \frac{S_{\phi n1} + \cdots + S_{\phi N}}{N^2}, \tag{2.49}$$

because the phase noises of the N oscillators are uncorrelated. Since the oscillators are nominally identical, their phase noise spectra are equal and

$$S_\phi(f) = \frac{S_{\phi n1}}{N}. \tag{2.50}$$

That is, the overall phase noise is reduced by a factor of N in return for an N-fold increase in power dissipation (if the voltage adder incurs none). The reader can show that this proof and this trade-off do not apply to *deterministic* jitter.

While intuitive, the approach illustrated in Fig. 2.29 is difficult to implement in practice. We introduce other methods of trading power for phase noise in Chapters 3 and 5.

Example 2.16

The "thought experiment" illustrated in Fig. 2.29 appears to suffer from a flaw: even if they begin at the same value, the oscillators' phases drift with time, possibly causing *destructive* addition of the output voltages. Is this a serious flaw?

Solution

For our purpose, it is not. Recall from Example 2.15 that it takes millions of cycles for each oscillator's jitter to accumulate to a fraction of the period. Thus, the outputs remain "coherent" for at least thousands of cycles, and our experiment is valid for such time scales.

2.4 Basic Phase Noise Mechanisms

2.4.1 Phase Noise versus Frequency Noise

In this section, we make two observations that offer useful insights. First, consider a chain of inverters acting as a simple delay line [Fig. 2.30(a)], noting that the inverters' delays vary with their supply voltage, V_{DD}. We can say that a small change, ΔV, in V_{DD} causes a change, Δt_D, in the line's delay, or, equivalently, a change, $\Delta\phi$, in the line's input-output phase shift. We can then define the supply sensitivity (or gain) as $K_{DD} = \Delta\phi/\Delta V$. With a supply noise voltage, V_n, the output of the delay line is expressed as

$$V_{out} = \frac{V_{DD}}{2} + \frac{V_{DD}}{2}\cos(\omega_{in}t + \phi_0 + \Delta\phi) \tag{2.51}$$

$$= \frac{V_{DD}}{2} + \frac{V_{DD}}{2}\cos(\omega_{in}t + \phi_0 + K_{DD}V_n), \tag{2.52}$$

where ϕ_0 denotes the phase shift in the absence of supply perturbation. We thus conclude that supply noise *directly* modulates the output phase of a delay line; e.g., white supply noise translates to white phase noise.

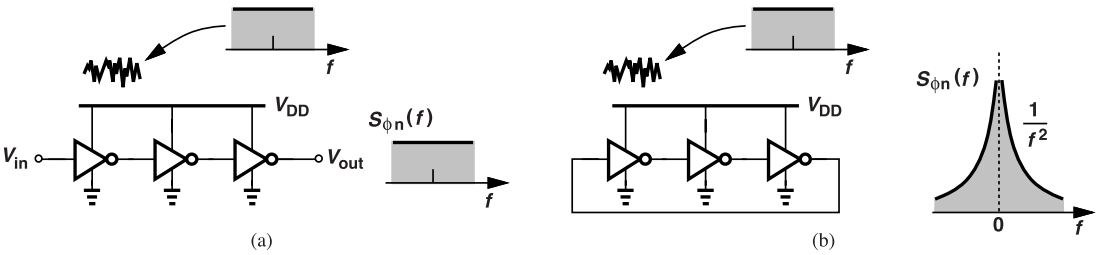


Figure 2.30 Effect of supply noise on (a) delay line, and (b) ring oscillator.

Next, let us repeat the supply noise experiment for a ring oscillator [Fig. 2.30(b)]. We know that, for example, a higher V_{DD} leads to a greater oscillation frequency. That is, supply perturbation directly modulates the *frequency* of the oscillation. For example, white supply noise produces white frequency noise. We can define a gain from V_{DD} equal to $K_{DD} = \partial\omega_{out}/\partial V_{DD}$, which is similar to K_{VCO}. (Note that the delay line and oscillator K_{DD}'s have different units.) The oscillator output is equal to

$$V_{out} = \frac{V_{DD}}{2} + \frac{V_{DD}}{2}\cos(\omega_{osc}t + \int K_{DD}V_n dt). \tag{2.53}$$

In this case, the spectrum of the phase noise is proportional to $1/f^2$ if V_n is white. Similarly, in the presence of flicker noise, the phase noise profiles produced by the delay line and the oscillator are proportional to $1/f$ and $1/f^3$, respectively.

Our study of phase noise in the previous sections has considered only noise sources *outside* the oscillator circuit, for example, noise on the control voltage or on the supply. But we are also interested in the "intrinsic" phase noise, i.e., that generated by the oscillator's constituent devices. In the following sections, we familiarize ourselves with the latter at a qualitative level. We formulate these mechanisms in subsequent chapters.

Example 2.17

A student surmises that the phase noise of a ring oscillator should be the same as that of an infinitely-long delay line because the circulation of an edge around the ring is equivalent to passing the edge through an infinite number of inverters. Is the student's thinking correct?

Solution

No, it is not. Recall that the noise sources in a delay line directly modulate its delay and hence its output phase. That is, the phase noise spectrum itself has a shape given by $\alpha/f + \beta$, where α and β are constant values. As we increase the number of stages, the output phase noise spectrum simply rises proportionally, but its shape remains unchanged. On the other hand, the phase noise of an oscillator has the form $\alpha/f^2 + \beta/f$. Thus, the infinite delay line does not predict the oscillator phase noise correctly.

2.4.2 Ring Oscillators

Let us first study ring oscillators. Consider the inverter in Fig. 2.31(a) and assume that the input rises from

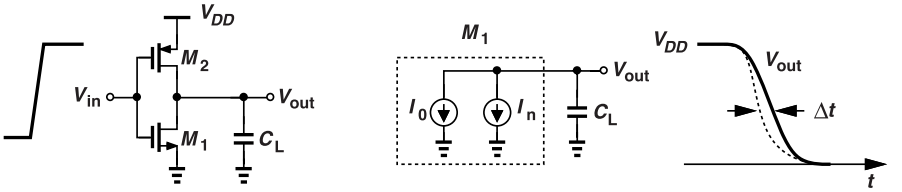


Figure 2.31 (a) CMOS inverter sensing a step, and (b) ouput response.

zero to V_{DD}. Transistor M_1 thus turns on and begins to discharge the load capacitance, C_L, from V_{DD} to zero. Modeling M_1 as shown in Fig. 2.31(b), where I_n represents its noise, we observe that the flow of I_n through C_L leads to a certain jitter, Δt. It can be shown that if I_n has a white spectrum, so does Δt, i.e., white current noise translates to white phase noise for an inverter.

What can we predict about the jitter and phase noise of N cascaded inverters? Since the inverters contain uncorrelated noise sources, we expect that the phase noise rises by a factor of N, and the rms jitter by a factor of $\sqrt{N}$. This is indeed how we compute the phase noise of a delay line (Fig. 2.32).

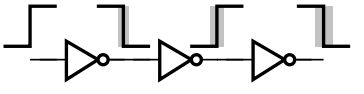


Figure 2.32 Accumulation of jitter of inverters along a delay line.

Example 2.18

Consider the flicker noise of M_2 in Fig. 2.31(a). Modeling this noise by a gate-referred voltage, explain how it modulates the inverter.

Solution

We first study the single PMOS device shown in Fig. 2.33(a), where V_n denotes its flicker noise. We can view

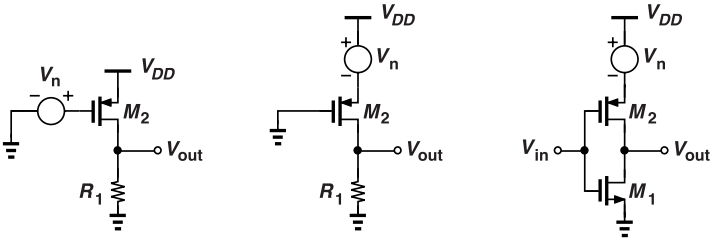


Figure 2.33 (a) Noise voltage in series with gate of a MOS device, (b) noise voltage moved to source terminal, and (c) PMOS noise voltage in inverter.

M_2 as a common-source device that amplifies V_n. If channel-length modulation is negligible, we can move V_n to the transistor's source terminal, as depicted in Fig. 2.33(b), and consider the device a common-gate stage, which also amplifies V_n by the same amount. Thus, the two topologies are equivalent.

Now, let us perform the same transformation in an inverter [Fig. 2.33(c)], arriving at the important conclusion that V_n appears in series with V_{DD} and, therefore, modulates the *delay* of the circuit. In other words, the phase noise inherits the spectrum of V_n and falls at a rate of 10 dB/dec—a notable difference with respect to the oscillator phase noise profile shown in Fig. 2.26.

What happens to the phase noise if the cascade of inverters in Fig. 2.32 is turned into a ring oscillator? As a simple case, we repeat the previous example for the flicker noise of one of the PMOS devices [Fig.

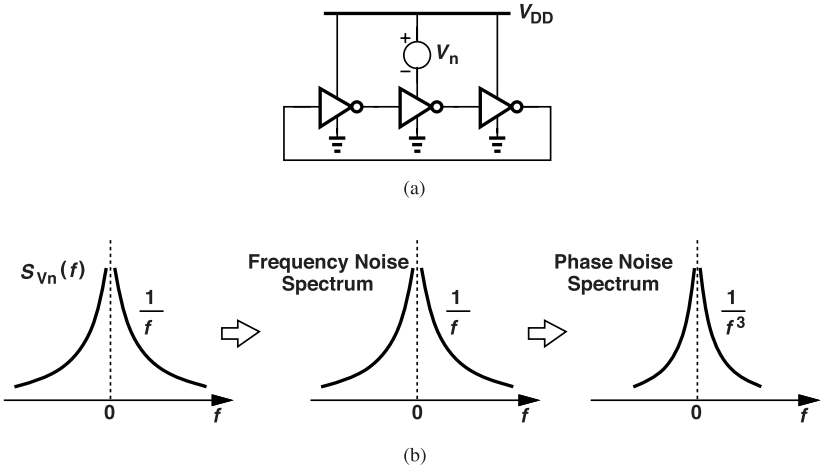


Figure 2.34 (a) Noise modulating delay of one stage in a ring, and (b) resulting spectra.

2.34(a)]. Here, too, V_n modulates the delay of the second inverter, but this means that the *frequency* of the

oscillation, f_0, is directly modulated because f_0 is inversely proportional to the sum of the delays. That is, V_n translates to *frequency noise* here and the resulting spectrum must be divided by $(2\pi f)^2$ so as to obtain the phase noise.[7] Figure 2.34(b) summarizes these results.

2.4.3 LC Oscillators

The simple cross-coupled LC oscillator studied in Chapter 1 entails several phase noise mechanisms (Chapters 5 and 6). We briefly consider one here.

Shown in Fig. 2.35 is the circuit with I_{n1} modeling the thermal noise current of M_1. We note that I_{n1} is

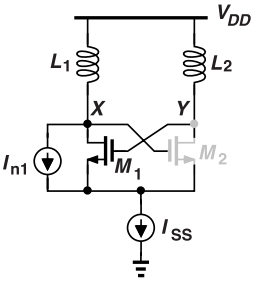


Figure 2.35 Effect of transistor noise current in an LC oscillator.

injected into node X, causing some jitter, Δt. But we also observe that (1) I_{n1} is zero when M_1 is off and (2) I_{n1} does *not* flow from the tank when M_2 is off (because I_{n1} is degenerated by I_{SS}). Thus, the analysis must view I_{n1} as a noise source whose "strength" changes with time. The step-by-step phase noise analysis in Chapter 5 determines how I_{n1} translates to phase noise in the output waveform.

The flicker noise of M_1 and M_2 can cause phase noise in the output under certain conditions. We study such phenomena in Chapter 6. Also studied in that chapter are mechanisms that convert the thermal and flicker noise of I_{SS} to phase noise.

2.5 Effect of Jitter on Performance

Jitter primarily manifests itself in "timing" applications, wherein random data is sampled by a clock. For example, the transfer of high-speed data and clock between a microprocessor and a memory chip must deal with jitter.

We intuitively expect that a jittery clock waveform degrades the performance of a system. Specifically, the displacement of the clock edges from their ideal positions in time can lead to erroneous sampling of data. To understand this point, let us begin with the idealized situation depicted in Fig. 2.36(a), where a random sequence of binary data, D_{in}, is applied to a master-slave flipflop (FF)[8] and sampled by a clock, CK. On each rising edge of the clock, the value of D_{in} is read and sent to the output, D_{out}. What happens if CK has jitter [Fig. 2.36(b)]? If one of the rising edges of CK incurs so much jitter as to sample the wrong point, e.g., at t_b rather than t_a, then D_{out} contains an error. In this case, the maximum tolerable clock jitter is one-half of the data period, T_{in}, provided that D_{in} itself has no jitter. [In data communications, we call T_{in} the "unit interval" (UI).]

In practice, however, the data waveform suffers from imperfections, too. For example, the data transitions take a finite time, creating the scenario shown in Fig. 2.36(c). It appears that CK can still have a jitter of $< T_{in}/2$ ($= 0.5$ UI) for the data to be sampled correctly, but if the rising edge of CK occurs at $t = t_1$, the data swing sensed by the flipflop is degraded to V_1. The FF therefore takes a longer time to generate a proper

[7] Recall that the phase is the integral of the frequency with respect to time.

[8] In this book, we denote master-slave flipflops by two concentric rectangles and latches by one rectangle.

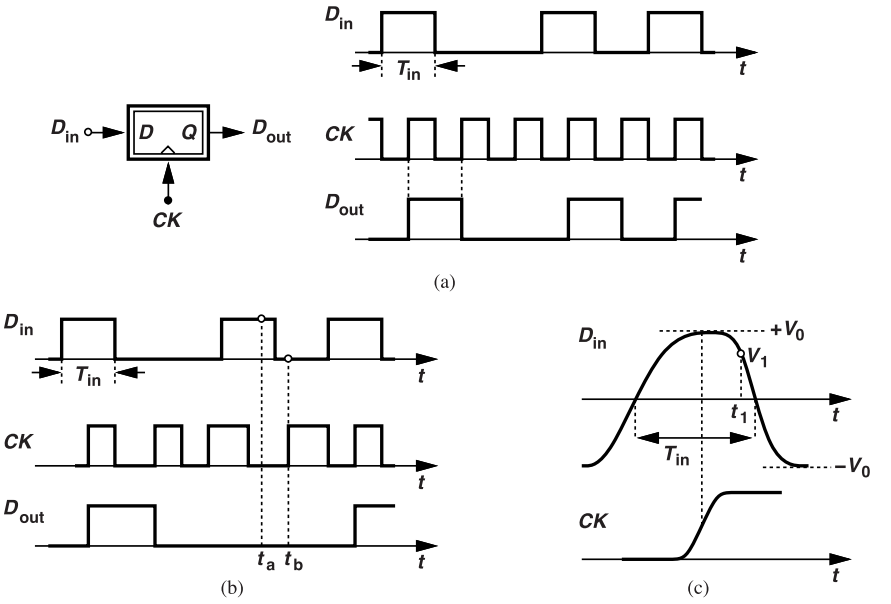


Figure 2.36 (a) A flipflop sensing random data, (b) effect of clock jitter on sampling, and (c) more realistic waveforms.

logical level at its output. We say the FF is "metastable." As a result of metastability, the FF output may be interpreted incorrectly by subsequent stages. In addition, if V_1 is not substantially greater than the offset or electronic noise of the flipflop, the data level may be interpreted as a logical ZERO rather than a logical ONE.

The problem of clock jitter becomes more serious in the presence of two other imperfections: (1) if the data suffers from jitter that is uncorrelated with the clock jitter, and (2) if D_{in} and CK in Fig. 2.36(a) incur a *skew*, i.e., if the rising edge of CK is (deterministically) offset to the left or to the right. The skew stems from asymmetries between the D_{in} and CK paths within the FF or from external sources, such as different interconnect lengths or capacitances seen by D_{in} or CK. These jitter and skew components further constrain the maximum allowable jitter in the clock.

2.6 Effect of Phase Noise on Performance

The frequency-domain view of phase noise is invoked in RF and wireless communication systems, wherein a "desired" modulated signal may be accompanied by undesired interferers (also called "blockers"). In this section, we examine how an oscillator's phase noise affects the performance in such systems.

Let us consider the simple RF transmitter (TX) depicted in Fig. 2.37(a). Here, the baseband signal (e.g.,

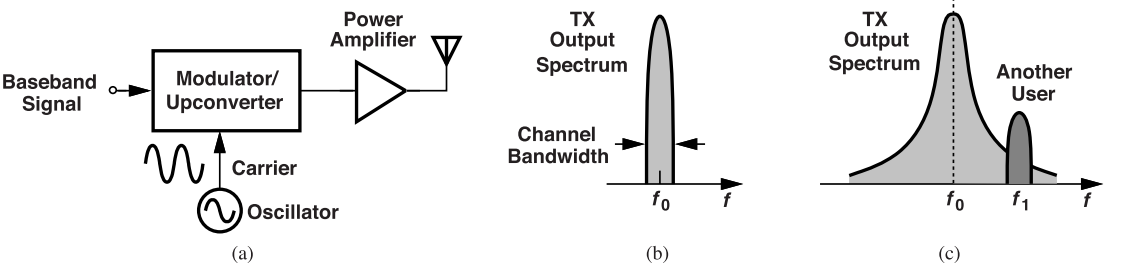


Figure 2.37 (a) Simple RF transmitter, (b) ideal output spectrum, and (c) output spectrum with noisy oscillator.

voice, video, or data) is applied to a modulator and impressed upon a carrier generated by the oscillator. The modulated signal then travels through a power amplifier (PA) and is transmitted by an antenna. In the ideal case, the carrier has no phase noise, leading to a TX output expressed as $A(t)\cos[\omega_0 t+\theta(t)]$, where $A(t)$ and $\theta(t)$ carry the baseband information. Depicted in Fig. 2.37(b), the spectrum is thus confined to the specified channel bandwidth, e.g., 1 MHz in Bluetooth. On the other hand, if the oscillator output contains phase noise, the TX output assumes the form $A(t)\cos[\omega_c t+\theta(t)+\phi_n(t)]$, exhibiting the spectrum shown in Fig. 2.37(c). Now, suppose another user transmits a weak signal at f_1. We observe that the noise skirts of the high-power signal transmitted at f_0 corrupt the signal at f_1. That is, a receiver interested in the latter signal may not be able to detect it properly. We must therefore ensure that the TX oscillator phase noise, integrated across this other user's channel, is sufficiently small.

Another adverse effect of TX phase noise relates to the transmitted signal itself. After all, in the above expression, $\phi_n(t)$ corrupts the information carried by $\theta(t)$. For example, if $\theta(t)$ is equal to integer multiples of 90°, then the total integrated phase noise (from very low to very high offset frequencies) must be much less than 90°. More complex modulation schemes, such as those used in WiFi, require an integrated rms phase noise of about 1°.

Phase noise also manifests itself in RF receivers. Illustrated in Fig. 2.38(a) is a common scenario: a desired, weak signal centered around f_1 is accompanied by a strong interferer centered around f_2. A receiver

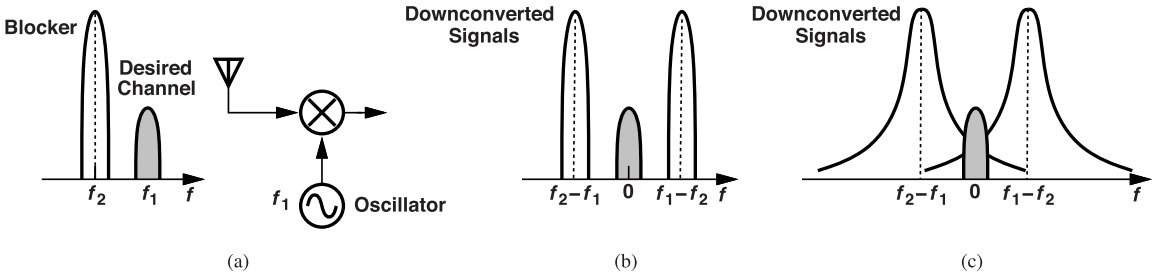


Figure 2.38 (a) Simple RF receiver sensing a desired channel and a blocker, (b) downconverted spectra for an ideal oscillator, and (c) downconverted spectra for a noisy oscillator.

interested in the former signal multiplies, in the time domain, these signals by the output of an oscillator running at f_1. In the frequency domain, this "mixing" operation translates the center of the desired channel to zero, with the interferer appearing at $\pm(f_1-f_2)$ [Fig. 2.38(b)]. We say the mixer "downconverts" the RF inputs. We now repeat this experiment with a noisy oscillator. As depicted in Fig. 2.38(c), the oscillator phase noise appears as skirts extending to the desired channel. Called "reciprocal mixing," this phenomenon places stringent phase noise requirements on cellphone receivers.

Another impact of phase noise upon the received signal relates to the direct corruption of the phase, as explained above for transmitters. In other words, when deciding on the tolerable integrated phase noise for a given modulation scheme, we must bear in mind that both the TX and the RX corrupt the phase.

References

[1]. S. O. Rice, "Mathematical analysis of random noise," *Bell Sys. Tech. J.*, vol. 23, pp. 282-332, July 1944.

[2]. C. C. Enz and G. C. Temes, "Circuit techniques for reducing the effects of op-amp imperfections: autozeroing, correlated double sampling, and chopper stabilization," *Proc. of IEEE*, vol. 84, pp.1584-1614, Nov.1996.

[3]. F. Herzel and B. Razavi, "A study of oscillator jitter due to supply and substrate noise," *IEEE Trans. Circuits and Systems, Part II*, vol. 46, pp. 56-62, Jan. 1999.

Problems

2.1. Determine the output noise spectrum of the circuit shown in Fig. 2.39.

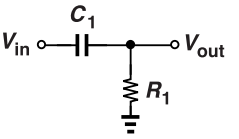
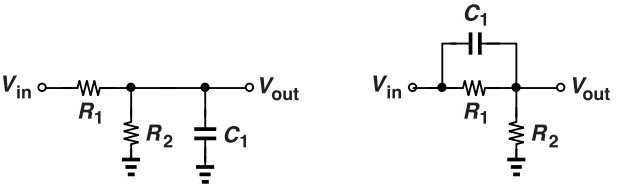


Figure 2.39 Circuit for output noise calculation.

2.2. Determine the output noise spectrum of the circuits shown in Fig. 2.40.



(a) (b)

Figure 2.40 Circuit for output noise calculation.

2.3. In the circuit of Fig. 2.7(a), we double the transistor's transconductance. What happens to f_C?

2.4. In the circuit of Fig. 2.7(a), the transistor's width and bias current are doubled. What happens to f_C?

2.5. Consider a resistor R_1 whose ambient temperature toggles periodically between 0 K and T_1 with 50% duty cycle. Explain intuitively why the resistor's noise spectrum is halved, i.e., it is given by $2kTR_1$.

2.6. Consider the circuit shown in Fig. 2.41, where M_1 operates in saturation and channel-length modulation and flicker noise are neglected.

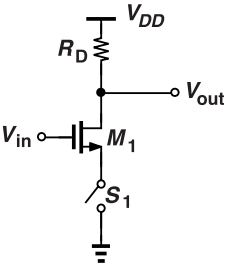


Figure 2.41 Periodically-switched CS stage.

(a) Determine the output thermal noise spectrum if S_1 remains on.
(b) Suppose S_1 turns on and off with a duty cycle of 50%. Can you intuitively predict the output noise spectrum?

2.7. In Fig. 2.13, we double ω_0. What happens to the output jitter if it is expressed in seconds?

2.8. In Example 2.6, we choose $\omega_m = \omega_0$. Plot the output spectrum. Does the output waveform contain jitter in this case?

2.9. The control voltage of a VCO carries a noise spectrum given by $\alpha + \beta/f$. Determine the output phase noise.

2.10. A "frequency doubler" doubles the output phase, e.g., it converts $V_0 \cos \omega_0 t$ to $V_0 \cos(2\omega_0 t)$. Suppose the VCO in Example 2.5 is followed by a doubler. Determine the output jitter in seconds and as a percentage of the period.

2.11. Repeat the above problem for Example 2.6 and determine the nomalized magnitude of the sidebands. Plot the doubler's output spectrum. Does the spacing between the sidebands and the carrier double?

2.12. Suppose in Example 2.8, the supply voltage contains white noise with a one-sided spectrum given by $S_V(f) = \eta$, where η is a constant. Determine the output phase noise.

2.13. In Fig. 2.23(b), $S_{\phi n}(f)$ is equal to -90 dBc/Hz at 1-MHz offset. Determine the equivalent thermal noise in Fig. 2.23(a) if $K_{VCO} = 100$ MHz/V.

2.14. If in Fig. 2.23(b), $S_{\phi n}(f)$ contains both flicker and thermal components, e.g., in the form of $S_{\phi n}(f) = \alpha/f^2 + \beta/f^3$, determine the equivalent voltage spectral density in Fig. 2.23(a).

2.15. An oscillator exhibits a (two-sided) phase noise profile given by $S_{\phi n}(f) = \alpha/f^2$. If the phase noise is equal to -100 dBc/Hz at 100-kHz offset, determine the value of α.

2.16. An oscillator exhibits a (two-sided) phase noise profile given by $S_{\phi n}(f) = \alpha/f^2 + \beta/f^3$. If the phase noise is equal to -90 dBc/Hz at 100-kHz offset and -115 dBc/Hz at 1-MHz offset, determine the values f α and β.

2.17. For the oscillator in the previous problem, compute the total rms phase noise (in radians) from 50 kHz to 500 kHz.

2.18. Repeat Example 2.11 for the circuit shown in Fig. 2.42. Assume the VCO control input presents an infinite impedance.

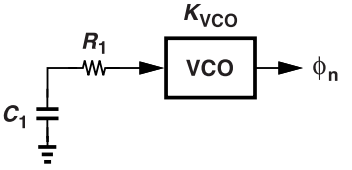


Figure 2.42 VCO sensing noise produced by a resistor.

2.19. Repeat Example 2.11 for the circuit shown in Fig. 2.43.

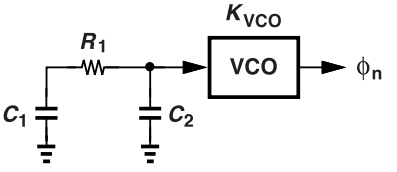


Figure 2.43 VCO sensing band-pass noise.

2.20. Consider the cascade of VCOs shown in Fig. 2.44, where we have for the two VCOs

$$V_{out1} = V_m \cos\left[\omega_{01} t + K_{VCO1} \int V_{cont1} dt + \phi_{n1}(t)\right] \tag{2.54}$$

$$V_{out2} = V_0 \cos\left[\omega_{02} t + K_{VCO2} \int V_{cont2} dt\right], \tag{2.55}$$

where $\phi_{n1}(t)$ denotes the phase noise of VCO_1 and is given by α/f^2. If the narrowband FM approximation holds, determine the output spectrum of VCO_2. Assume the output frequency of VCO_1 is much lower than that of VCO_2.

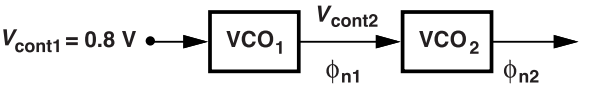


Figure 2.44 Cascaded VCOs.

2.21. The oscillators in Fig. 2.29 share the same supply voltage. If the supply contains a small sinusoidal perturbation, $V_m \cos \omega_m t$, determine $V_{out}(t)$ and show that the resulting jitter is the same as that of a single oscillator experiencing this supply noise. Assume the oscillators begin with the same phase.

3

Design of Inverter-Based Ring Oscillators

Having studied the fundamentals of oscillators in Chapter 1, we are now ready to embark upon their design. This chapter deals with ring oscillators that incorporate CMOS inverters as their delay stages. We first extend our phase noise study in Chapter 2 to develop simple equations for ring oscillators. Next, we methodically design, simulate, and analyze inverter-based rings. Finally, we introduce frequency tuning techniques.

3.1 Phase Noise in Ring Oscillators

As mentioned in Chapter 2, phase noise directly trades with power dissipation and its analysis forms an essential step in the design of oscillators. In other words, without a phase noise specification, the design is not meaningful. In this section, we develop simple phase noise equations for ring oscillators so that our subsequent design efforts can readily optimize for this parameter.

3.1.1 General Equation

In this section, we wish to establish a relation between the phase noise of a ring oscillator (the "closed-loop" phase noise) and the phase noise of the delay line that is obtained if the loop is broken (the "open-loop" phase noise). Figure 3.1 illustrates the two cases. Such a relation proves useful in both analysis and design. It is important to note that these two quantities are *not* the same! As a transition travels through a delay line, e.g.,

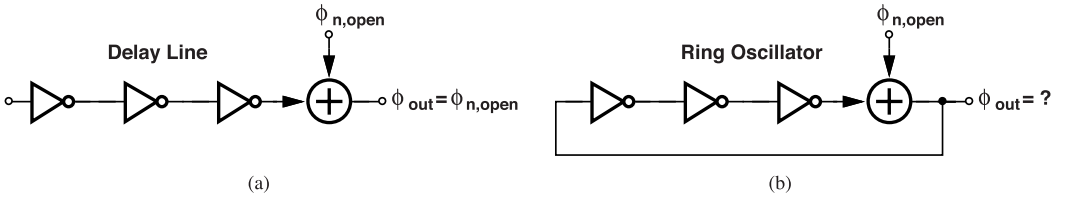


Figure 3.1 Phase noise of (a) a delay line, and (b) a ring oscillator.

through three cascaded inverters, it is corrupted by each stage only *once* [Fig. 3.2(a)]. In a ring oscillator, on the other hand, the edge continues to circulate around the loop, accumulating jitter indefinitely [Fig. 3.2(b)].

Consider the ring shown in Fig. 3.3(a), assuming that it is designed for an oscillation period of $T_0 = 1/f_0$. We observe that (1) T_0 is equal to six gate delays, (2) the delay line consisting of the three inverters exhibits a delay, T_{DL}, equal to three gate delays, i.e., $T_0/2$, and (3) T_{DL} corresponds to a phase shift of $2\pi(T_{DL}/T_0) \equiv 180^\circ$. The other 180° phase shift necessary for oscillation is provided by an odd number of inversions in the chain. Let us now add a small, *constant* phase, ϕ_1, to the output of the delay line, as depicted in Fig. 3.3(b). (For example, we can insert an additional stage after the third inverter to create ϕ_1.) What happens to the

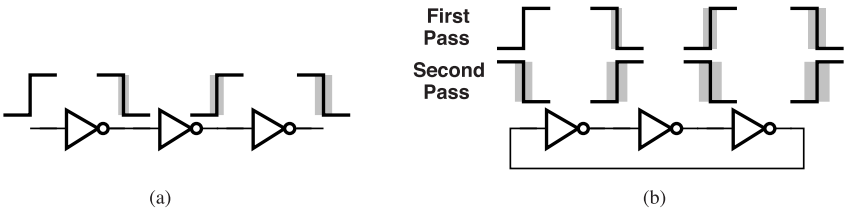


Figure 3.2 Accumulation of phase noise in (a) a delay line, and (b) a ring oscillator.

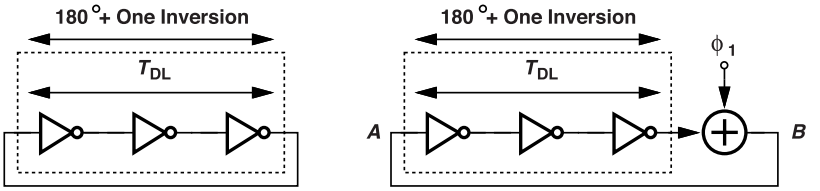


Figure 3.3 (a) Ring oscillator with a total delay of T_{DL}, and (b) addition of a constant phase within the loop.

oscillation frequency? The circuit adjusts the frequency so that the total phase shift from A to B (excluding the net inversion) returns to 180°. In other words, we must have

$$2\pi \frac{T_{DL}}{T_{osc}} + \phi_1 = \pi, \tag{3.1}$$

where $T_{osc} = 1/f_{osc}$ denotes the new oscillation period.

It is interesting to note that, in this experiment, the gate delays do not change,[1] but the oscillation frequency does. That is, T_{DL} is constant if measured in seconds but changes if measured in degrees relative to the new oscillation frequency. Since $T_{DL} = T_0/2 = 1/(2f_0)$, we obtain from Eq. (3.1)

$$f_{osc} = f_0 - f_0 \frac{\phi_1}{\pi}. \tag{3.2}$$

Thus, the oscillation frequency departs from the original value by an amount equal to $f_0\phi_1/\pi$. This result applies to both small-signal and large-signal operation so long as f_0 and f_{osc} are computed accordingly.

Example 3.1

Figure 3.4(a) shows an inverting delay line, where the amplifier has no phase shift. Plot the phase shift as a function of the input frequency and therefrom derive Eq. (3.2).

Solution

As mentioned earlier, the delay, T_{DL}, of a circuit such as cascaded inverters is relatively constant and independent of the input frequency, f_{in}. Excluding the net inversion, we write the phase shift, ϕ_{DL} (in radians), as

$$\phi_{DL} = 2\pi \frac{T_{DL}}{T_{in}} \tag{3.3}$$

$$= 2\pi T_{DL} f_{in}. \tag{3.4}$$

[1] The delay of each stage is given by its physical properties (the inverter's strength and parasitic capacitances) and is relatively independent of the operation frequency.

Plotted in Fig. 3.4(b), the phase shift crosses π at $f_{in} = f_0 = 1/(2T_{DL})$. Now consider a loop utilizing the delay line and oscillating at f_0 [Fig. 3.4(c)]. This situation corresponds to the ring shown in Fig. 3.3(a). We

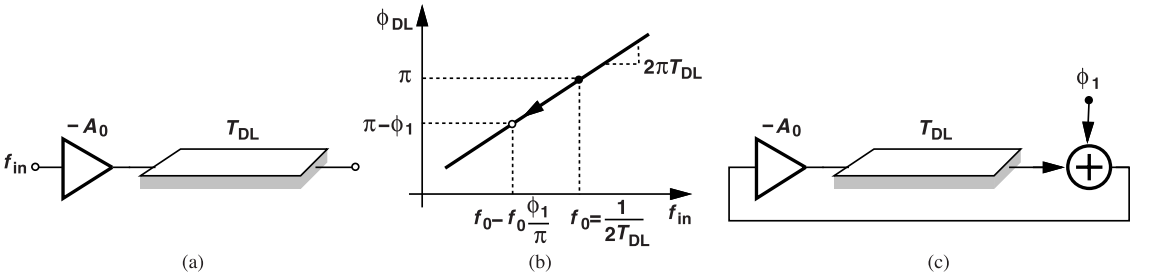


Figure 3.4 (a) Inverting delay line, (b) phase response, and (c) an oscillator formed from the delay line.

say the open-loop "phase slope" is equal to $2\pi T_{DL}$.

If we insert a positive phase shift of ϕ_1 as in Fig. 3.3(b), the oscillation frequency of the loop must *decrease* so that T_{DL} corresponds to *less* than $180°$. The slope of $2\pi T_{DL}$ thus requires a frequency change to $f_0 - \phi_1/(2\pi T_{DL}) = f_0 - f_0\phi_1/\pi$.

A key point offered by the above example is that, if ϕ_1 varies, then the phase shift of the delay line travels on the characteristic shown in Fig. 3.4(b) so as to satisfy the overall phase shift required for oscillation; consequently, the oscillation frequency changes. Now suppose ϕ_1 in fact represents the phase noise, $\phi_{n,DL}$, of the delay line (the open-loop phase noise), i.e., the edge displacement that a periodic signal experiences as it propagates through the three inverters in Fig. 3.2(a). We recognize that, as $\phi_{n,DL}$ increases and decreases, the oscillation frequency also varies according to $f_0 - f_0\phi_{n,DL}/\pi$.

In summary, the open-loop phase noise of the delay line translates to a closed-loop *frequency* noise of $f_0\phi_{n,DL}/\pi$ Hz in the ring oscillator. The *excess* output phase of the oscillator can therefore be expressed as

$$\phi_{n,osc}(t) = \int 2\pi \frac{f_0}{\pi}\phi_{n,DL}(t)dt, \tag{3.5}$$

where the factor of 2π converts the unit from hertz to radians per second. We note that, in general, if

$$g(t) = \int w(t)dt, \tag{3.6}$$

then the spectra of $g(t)$ and $w(t)$ are related as follows (Chapter 2):

$$S_g(f) = \frac{1}{4\pi^2 f^2} S_w(f) \tag{3.7}$$

because the transfer function of an integrator is given by $H(s) = 1/s$. Equation (3.5) thus yields

$$S_{\phi n,osc}(f) = \frac{4f_0^2}{4\pi^2 f^2} S_{\phi n,DL}(f) \tag{3.8}$$

$$= S_{\phi n,DL}(f)\left(\frac{f_0}{\pi f}\right)^2. \tag{3.9}$$

Derived in [1] using a different method, this simple result suggests that the phase noise of a ring oscillator running at f_0 is equal to that of its open-loop circuit multiplied by a "shaping function" $[f_0/(\pi f)]^2$ (Fig. 3.5).

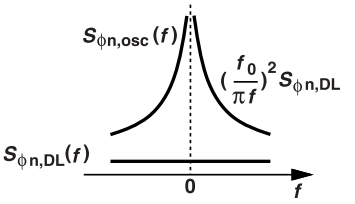


Figure 3.5 Shaping of a delay line phase noise as a result of closing the loop.

Since it is generally easier to compute the open-loop phase noise, this relation reduces the mathematical labor. It is important to note that our derivation has *not* assumed small-signal operation and thus applies to both linear or nonlinear circuits.

Example 3.2

A single-loop oscillator can be represented as shown in Fig. 3.6(a), where ϕ_{open} denotes the phase shift from

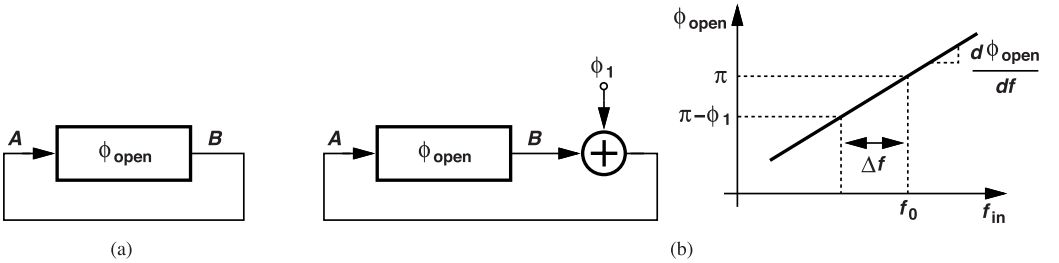


Figure 3.6 (a) Loop consisting of a circuit with a phase shift of ϕ_{open}, and (b) effect of phase noise.

A to B. Assuming negative feedback at low frequencies, derive a general relation between open-loop and closed-loop phase noise.

Solution

For oscillation at f_0, we must have $\phi_{open}(f_0) = 180^\circ$ (for small-signal or large-signal operation).[2] If we add a small phase, ϕ_1, to the output, the circuit adjusts the oscillation frequency so that $\phi_{open} + \phi_1 = 180^\circ$ [Fig. 3.6(b)]. To compute the change in the frequency, we note from Example 3.1 that ϕ_1 must be divided by the phase slope:

$$\Delta f = \frac{\phi_1}{d\phi_{open}/df}. \tag{3.10}$$

If ϕ_1 represents the phase noise of the open-loop circuit, $\phi_{n,open}$, then Δf is the frequency noise in the oscillator output. The closed-loop phase noise is given by

$$\phi_{n,closed}(t) = \int 2\pi \frac{\phi_{n,open}}{d\phi_{open}/df} dt, \tag{3.11}$$

[2] We have excluded from ϕ_{open} the signal inversion due to negative feedback.

and its spectrum emerges as

$$S_{\phi n,closed}(f) = \frac{4\pi^2}{|d\phi_{open}/df|^2} \cdot \frac{1}{4\pi^2 f^2} \cdot S_{\phi n,open}(f) \tag{3.12}$$

$$= \frac{1}{|d\phi_{open}/df|^2} \cdot \frac{1}{f^2} \cdot S_{\phi n,open}(f). \tag{3.13}$$

This result proves the $1/f^2$ shaping function for the phase noise of any (single-loop) oscillator regardless of the circuit implementation and regardless of small-signal or large-signal operation. Note that $d\phi_{open}/df$ is simply the slope of the open-loop phase profile at f_0, a constant value if the phase perturbations are small. Of course, if $S_{\phi n,open}(f)$ contains flicker noise, then $S_{\phi n,closed} \propto 1/f^3$ because $S_{\phi n,open} \propto 1/f$.

The foregoing developments allow us to calculate the phase noise of inverter-based ring oscillators from the phase noise of a single inverter. The latter is obtained by examining how the noise of the NMOS and PMOS transistors displaces the output edge. As illustrated in Fig. 3.7, when V_{in} rises to V_{DD}, M_2 turns off (like a switch), depositing kT/C noise on C_L, and M_1 turns on, injecting both thermal and flicker noise into

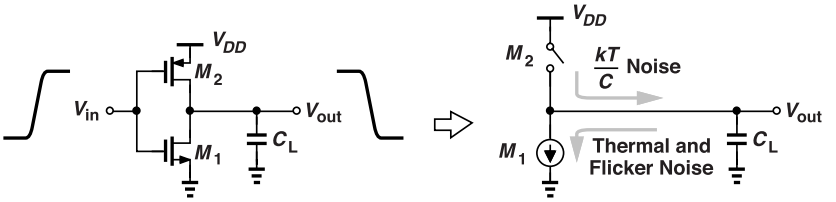


Figure 3.7 Noise mechanisms in an inverter.

the output. Consequently, the falling edge in V_{out} is corrupted. The analyses in [2] and [1] yield the following phase noise equations for white and flicker noise sources in a closed-loop N-stage ring oscillator operating at a frequency of f_0:

$$S_{\phi n,white}(f) = \frac{f_0^2}{f^2}\left[\frac{1}{2I_D^2}(S_{I,NMOS} + S_{I,PMOS}) + \frac{2kT}{I_D V_{DD}}\right] \tag{3.14}$$

$$S_{\phi n,flicker}(f) = \frac{f_0^2}{f^2}\frac{1}{4NI_D^2}(S_{1/f,NMOS} + S_{1/f,PMOS}). \tag{3.15}$$

Here, f denotes the frequency offset and I_D and S_I correspond, respectively, to the drain current and the drain white noise current spectrum at a single operating point, namely, at $|V_{GS}| = V_{DD}$ and $|V_{DS}| \approx V_{DD}/2$.[3] Similarly, $S_{1/f}$ is the drain flicker noise current spectrum at the same operating point and, of course, has a $1/f$ dependence. We should emphasize that I_D is the drain current under the above conditions and *not* equal to the average supply current per stage. Both S_I and $S_{1/f}$ are two-sided spectra, e.g., $S_I(f) = 2kT\gamma g_m$.

Equations (3.14) and (3.15) provide useful insights into the design of ring oscillators. The former suggests that the white-noise-induced phase noise is independent of (a) the capacitance seen at each node, and (b) the number of stages in the ring. These two points are revisited in Section 3.3. Noting that $S_I = 2kT\gamma g_m$ (Chapter 2) and $g_m = 2I_D/(V_{DD} - |V_{TH}|)$ (because $|V_{GS}| = V_{DD}$),[4] we can simplify (3.14) to

$$S_{\phi n,white}(f) = \frac{f_0^2}{f^2} \cdot \frac{2kT}{I_D}\left(\frac{2\gamma}{V_{DD} - |V_{TH}|} + \frac{1}{V_{DD}}\right), \tag{3.16}$$

[3] It is assumed that I_D is the same for the NMOS and PMOS transistors under these conditions.

[4] This g_m expression assumes that the transistors operate in saturation even though $|V_{DS}| \approx V_{DD}/2$.

where we have assumed $g_{m,N} = g_{m,P}, \gamma_N = \gamma_P, I_{DN} = I_{DP}$, and $V_{THN} = |V_{THP}| = V_{TH}$. If $4\gamma/(V_{DD} - |V_{TH}|) \gg 1/V_{DD}$, then

$$S_{\phi n,white}(f) = \frac{f_0^2}{f^2} \cdot \frac{4kT\gamma}{I_D(V_{DD} - |V_{TH}|)}. \tag{3.17}$$

Example 3.3

A 2-GHz ring oscillator is designed with $I_D = 0.5$ mA for both NMOS and PMOS devices under the bias conditions $|V_{GS}| = V_{DD}$ and $|V_{DS}| = V_{DD}/2$. If $V_{DD} = 1$ V and $|V_{TH}| = 0.3$ V, estimate $S_{\phi n,white}$ at 1-MHz offset. Assume $T = 300$ K and $\gamma = 1$.

Solution

From (3.17), we have

$$S_{\phi n,white}(f) = 9.46 \times 10^{-17} \frac{f_0^2}{f^2} \quad \text{Hz}^{-1} \tag{3.18}$$

$$= \frac{189}{f^2} \quad \text{Hz}^{-1}. \tag{3.19}$$

It follows that

$$S_{\phi n,white}(f = 1 \text{ MHz}) = 1.89 \times 10^{-10} \quad \text{Hz}^{-1} \tag{3.20}$$

$$= -97 \text{ dBc/Hz}. \tag{3.21}$$

Equation (3.15) implies that the flicker-noise-induced phase noise is inversely proportional to the number of stages but independent of the load capacitance. Since $S_{1/f} = g_m^2 K/(2WLC_{ox}f)$, where the factor of 2 signifies a two-sided spectrum, we have

$$S_{\phi n,flicker} = \frac{f_0^2}{f^3} \cdot \frac{g_m^2}{4NI_D^2} \left[\frac{K_N}{(2WL)_{\text{NMOS}} C_{ox}} + \frac{K_P}{(2WL)_{\text{PMOS}} C_{ox}} \right], \tag{3.22}$$

observing the $1/f^3$ dependence. It follows that

$$S_{\phi n,flicker} = \frac{f_0^2}{f^3} \frac{1}{(V_{DD} - |V_{TH}|)^2} \left[\frac{K_N}{2N(WL)_{\text{NMOS}} C_{ox}} + \frac{K_P}{2N(WL)_{\text{PMOS}} C_{ox}} \right]. \tag{3.23}$$

Interestingly, this phase noise is inversely proportional to the total NMOS gate capacitance, $N(WL)_{\text{NMOS}} C_{ox}$, and the total PMOS gate capacitance, $N(WL)_{\text{PMOS}} C_{ox}$.

Example 3.4

We double the widths of the NMOS and PMOS transistors in a 2-GHz ring oscillator. Explain what happens to the oscillation frequency and the phase noise.

Solution

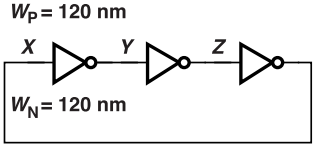
Since both the strength and the gate and drain capacitances of the inverters are doubled, the circuit still oscillates at 2 GHz.[5] In Eq. (3.14), I_D is doubled and so is $S_I = 2kT\gamma g_m$ (why?). Thus, $S_{\phi n,white}$ is halved. In Eq. (3.22), g_m and I_D are doubled and so are the gate areas, $(W/L)_N$ and $(W/L)_P$, suggesting a twofold

[5] Alternatively, we can say that the fanout seen by each stage is still equal to unity, concluding that the delay does not change.

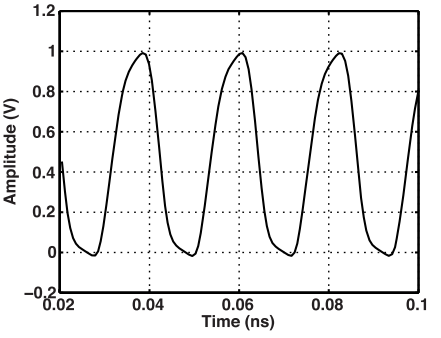
drop in $S_{\phi n,1/f}$ as well. This "linear scaling" also confirms the direct trade-off between phase noise and power dissipation (Chapter 2). We return to this point later in this chapter.

3.2 Preliminary Design Ideas

We wish to design an inverter-based ring oscillator running at 2 GHz. Where do we begin? Our approach to circuit design is as follows: we always begin with minimum-size devices and the minimum number of stages, unless there is a compelling reason not to do so. In 40-nm technology, $L_{min} = 40$ nm and $W_{min} = 120$ nm. We thus consider a three-stage ring with $W/L = 120\ \text{nm}/40$ nm for both NMOS and PMOS devices [Fig. 3.8(a)], where the values shown above and below the inverter denote the PMOS and NMOS widths, respectively. We simulate this circuit with $V_{DD} = 1$ V at the room temperature in the typical-typical



(a)



(b)

Figure 3.8 (a) Simple ring oscillator design, and (b) its output waveform.

corner[6] of the process, obtaining the output plotted in Fig. 3.8(b).

As explained in Chapter 2, with only three stages, the voltage waveform has no time to settle near V_{DD} or ground because the feedback signal returns quickly.

Example 3.5

Sketch the supply current, I_{DD}, of the above ring oscillator as a function of time.

Solution

Each time a PMOS device turns on, it draws a current from V_{DD} to charge its load capacitance. (The current that flows from V_{DD} to ground through the PMOS and NMOS transistors in each inverter is negligible.) As illustrated in Fig. 3.9, I_{DD} peaks for every rising edge within the ring, thereby swinging at a frequency of $3f_0$.

The oscillation frequency observed in the above simulation is around 45 GHz, implying a gate delay of about 3.7 ps. This frequency seems impressive for a ring oscillator but, in reality, we must take into account several other factors: (1) the interconnect capacitance, (2) the input capacitance of the subsequent circuit, e.g., a buffer, and (3) the worst-case conditions, for example, $V_{DD} = 0.95$ V, a temperature of $75°$, and operation in the slow-slow (SS) corner. We therefore revise the circuit to that shown in Fig. 3.10(a), where an

[6]The typical-typical corner means both the NMOS and the PMOS devices have their typical characteristics.

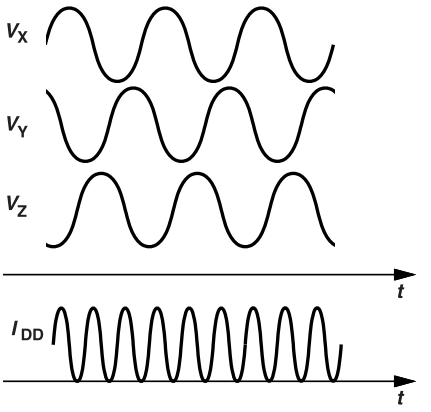
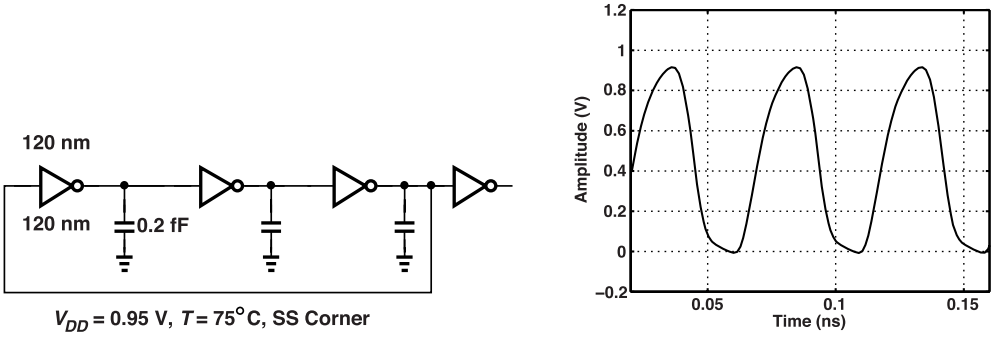


Figure 3.9 Three-stage ring oscillator voltage and current waveforms.



(a) (b)

Figure 3.10 (a) Ring oscillator design in extreme conditions, and (b) its output waveform.

estimated interconnect capacitance of 0.2 fF is added to each node and the output is buffered with a fanout of 1. Simulating the circuit with the above worst-case conditions, we obtain the output plotted in Fig. 3.10(b), where the oscillation frequency has fallen to 21 GHz.

Example 3.6

In Fig. 3.10(a), the wire connecting the output of the last gate to the input of the first is quite long, presenting a larger capacitance. How do we alleviate this issue?

Solution

We rearrange the layout as shown in Fig. 3.11, where one inverter is placed below the other two. If all three phases are utilized and their precision is important, we also "meander" the wire between the top inverters so as to make its length equal to the other two wires.

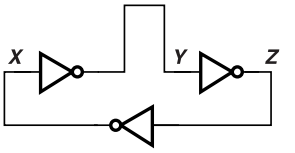
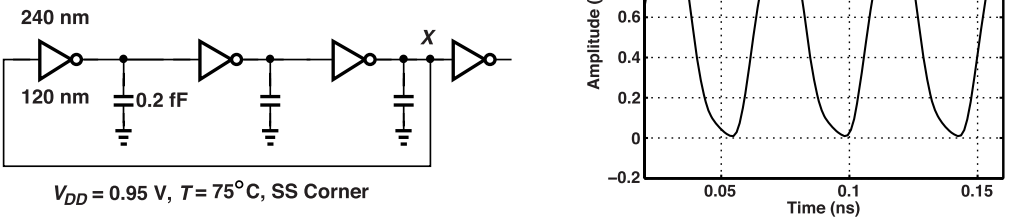


Figure 3.11 Layout technique for equalizing delays.

In the circuit of Fig. 3.10(a), the weak PMOS devices cannot lift the voltage at each node to V_{DD}. To remedy this issue, we double their widths, arriving at the design shown in Fig. 3.12(a) and the output plotted in Fig. 3.12(b). The oscillation frequency is now 22.6 GHz. We shall refer to this circuit as our "reference



(a) (b)

Figure 3.12 (a) Ring oscillator design with stronger PMOS devices, and (b) its output waveform.

design." The frequency increases slightly because the stronger PMOS pull-up more than compensates for the greater capacitance.

In addition to the oscillation frequency and the output waveform, we are also interested in the power consumption and supply sensitivity of the ring oscillator. The former is given by

$$P = 3f_0 C_{tot} V_{DD}^2, \tag{3.24}$$

where C_{tot} denotes the total capacitance seen from each node to ac ground (including the interconnect capacitance). According to simulations, the ring in Fig. 3.12(a) draws an average current of 60 μA, thus consuming 60 $\mu\text{A} \times 0.95\ \text{V} = 57\ \mu\text{W}$. (This means that $C_{tot} \approx 0.93$ fF.[7]) To compute the supply sensitivity, K_{VDD}, we change V_{DD} by a small amount (10 mV) and measure the change in f_0 (502 MHz), obtaining $K_{VDD} = 50.2$ GHz/V. As explained in Chapter 2, such an extremely high sensitivity converts supply noise to a large amount of phase noise.

We hereafter abandon the idealized ring in Fig. 3.8(a) and continue with the more realistic arrangement in Fig. 3.12(a). Since a practical design must ultimately take the worst-case conditions into account, there is little value in expending effort on the typical-typical corner.

3.3 Obtaining the Desired Frequency

The reference design in Fig. 3.12(a) runs at 22.6 GHz, well above the specified value of 2 GHz. To reach the desired frequency, we have four options:

1. Add capacitance to each node.
2. Increase the number of stages.
3. Increase the length of the transistors.
4. Divide the output frequency by a factor of 10 or 11.

Figure 3.13 depicts these methods conceptually. Of course, two or more of these techniques can be combined as well. Let us explore each option.

[7] Owing to the buffer, the capacitance is higher at X, but we neglect this for now.

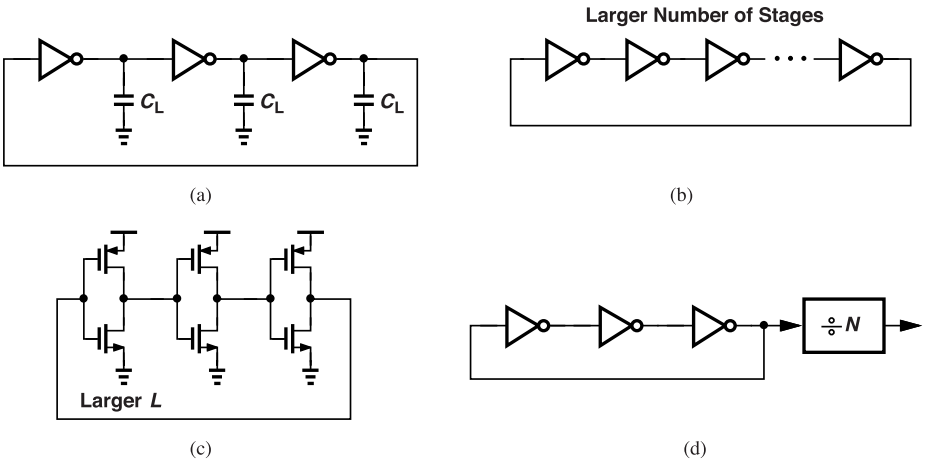


Figure 3.13 Methods of obtaining a lower output frequency: (a) adding capacitances to the nodes, (b) a longer ring, (c) longer transistors, and (d) use of a frequency divider.

3.3.1 Greater Node Capacitances

In order to determine the amount of additional capacitance that is necessary to decrease f_0 to 2 GHz, we first recall that C_{tot} in Fig. 3.12(a) is equal to 0.93 fF. If the inverters themselves do not change, the frequency scales directly with C_{tot}, requiring a new value of $(22.6\ \text{GHz}/2\ \text{GHz}) \times 0.93\ \text{fF} = 10.5\ \text{fF}$. We therefore add $10.5\ \text{fF} - 0.93\ \text{fF} \approx 9.6\ \text{fF}$ to each internal node. Figure 3.14 shows the revised design and its output waveform. The oscillation frequency is 1.93 GHz, close to the target. We must point out that the rise and fall

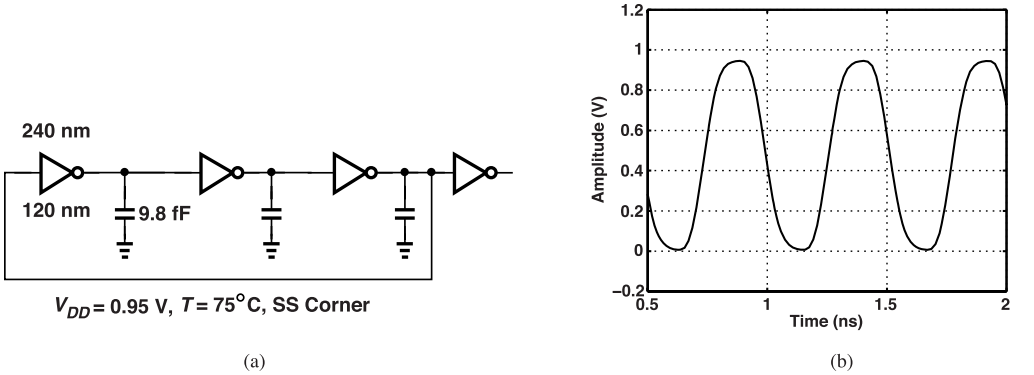


Figure 3.14 (a) Ring oscillator with additional capacitances, and (b) its output waveform.

times (and hence the gate delay) have increased proportionally.

It is interesting to examine the power dissipation, P, of the revised design. Since $P = 3f_0C_{tot}V_{DD}^2$ and since C_{tot} and f_0 have changed in opposite directions and by about the same factor, we recognize that the power dissipation is equal to that of the reference design.

The supply sensitivity of the 1.93-GHz design is obtained from simulations to be around 4.3 GHz/V. For a fair comparison with the reference design, we normalize K_{VDD} to f_0 and note that K_{VDD}/f_0 has not changed. This is simply because the strength of the inverters determines both K_{VDD} and f_0.

3.3.2 Greater Number of Stages

For a frequency of 2 GHz, we can increase the number of stages in the ring from 3 to about 31. Figure 3.15 shows such a design along with its output, revealing an oscillation frequency of 2.3 GHz. Interestingly, this

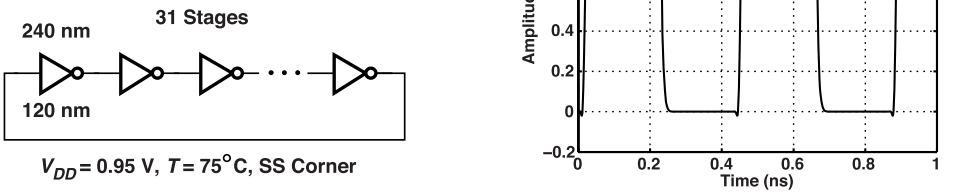


Figure 3.15 (a) Long ring oscillator, and (b) its output waveform.

waveform exhibits substantially sharper edges than that in Fig. 3.14(b); this is because the inverters in the 31-stage ring still see a unity fanout and retain their original rise and fall times.

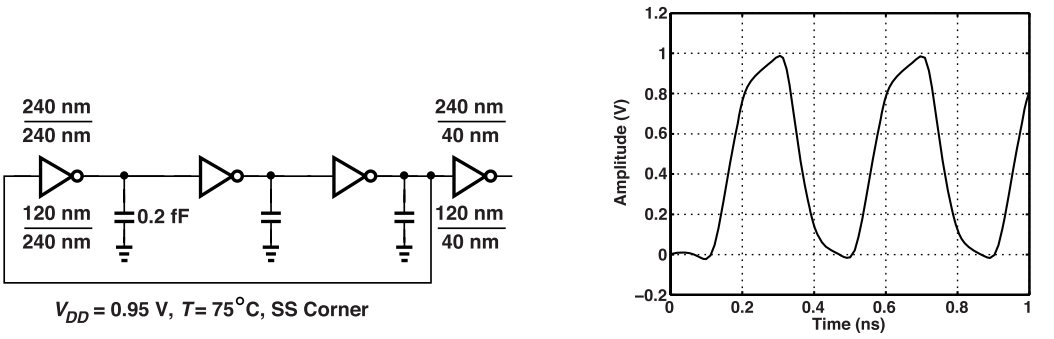
An N-stage ring operating at a frequency of f_0 consumes a power equal to $P = Nf_0C_{tot}V_{DD}^2$, where C_{tot} denotes the total load capacitance seen by one stage. Also, $f_0 = (2NT_D)^{-1}$, where T_D is the gate delay. Thus, $P = C_{tot}V_{DD}^2/(2T_D)$, a quantity independent of N. In other words, we expect the three-stage design in Fig. 3.12(a) and the 31-stage design in Fig. 3.15(a) to draw approximately equal powers.

The supply sensitivity of the 31-stage ring oscillator is obtained from simulations to be equal to 4.9 GHz/V. The normalized sensitivity, K_{VDD}/f_0, is about the same as those of the previous designs.

3.3.3 Greater Transistor Lengths

How should we increase the transistor channel lengths to go from 22.6 GHz to 2 GHz? We know that as the lengths of the transistors in an inverter-based ring increase, three effects occur: (1) the inverters become weaker, (2) their input capacitance rises, and (3) the flicker noise of the transistors decreases. The first two imply that, if the (effective) lengths are increased by a factor of m, then the gate delay goes up by roughly a factor of m^2. Thus, to the first order, $m = \sqrt{22.6\ \text{GHz}/2\ \text{GHz}} = 3.4$. However, since the 0.2-fF parasitic capacitance and the drain junction capacitances in the reference design are comparable with the gate capacitances, the lengths have to be increased by a factor greater than 3.4 for f_0 to fall to 2 GHz. In fact, simulations indicate an oscillation frequency of 2.5 GHz for a channel length of 6×40 nm $= 240$ nm. Depicted in Fig. 3.16 are the design and its output waveform. Note the long rise and fall times.

An interesting aspect of this modification is its lower power consumption, about $14\ \mu\text{A} \times 0.95\ \text{V} = 13.3\ \mu\text{W}$ from simulations. To understand the cause, we compare this design with the 1.93-GHz-ring in Fig. 3.14(a) in terms of the inverters' strength and the node capacitances. We note that the inverters are weakened by a factor of 6 here. For the circuit to operate at the same frequency, these inverters must see a total load capacitance that is lower by a factor of 6 than those in Fig. 3.14(a). It follows that $P = 3f_0C_{tot}V_{DD}^2$ falls by a factor of 6. In fact, if we account for the different oscillation frequencies (2.5 GHz versus 1.93 GHz), we predict a power reduction factor of $6 \times (1.93/2.5) = 4.6$, obtaining $I_{DD} = 60\,\mu\text{A}/4.6 = 13\,\mu\text{A}$ for the new design, which is close to the simulated value of 14 μA. Of course, the lower power consumption may lead to higher phase noise. We return to this point later in this chapter. The supply sensitivity of this design is about 4.96 GHz/V, translating to a slightly smaller K_{VDD}/f_0 than those of the previous designs.



(a) (b)

Figure 3.16 (a) Ring oscillator using long-channel devices, and (b) its output waveform.

3.3.4 Frequency Division

The output of the reference design in Fig. 3.12(a) can be applied to a frequency divider (a counter) so as to obtain $f_0 = 2$ GHz. How does the performance differ from that of the foregoing topologies? The divider consumes additional power, and the normalized supply sensitivity, K_{VDD}/f_0, does not change. We therefore see no advantage in this approach [except for the possibility of generating quadrature phases (Chapter 15)].

In summary, the four approaches to reducing the oscillation frequency do not change the supply sensitivity. The use of longer transistors lowers the power consumption. Ultimately, however, the pros and cons of these methods should be studied in terms of phase noise.

3.4 Phase Noise Considerations

The design efforts in the previous section have focused on the oscillation frequency, power consumption, and supply sensitivity. Having developed a basic understanding of these design principles, we can now take phase noise into account as well. We begin with the 22.6-GHz reference oscillator in Fig. 3.12(a) and then investigate the phase noise behavior of the 2-GHz topologies derived using our four methods. All of the simulations use the SS model at 75 °C with $V_{DD} = 0.95$ V.

It is important to bear in mind that phase noise directly trades with power dissipation (Chapter 2) and, as exemplified by Eqs. (3.14) and (3.15), with the oscillation frequency. The following example illustrates the former trade-off.

Example 3.7

An oscillator is designed for a frequency of f_0. Explain how its phase noise can be reduced by 3 dB.

Solution

One approach is to add the output voltages of two nominally identical oscillators running at f_0 [Fig. 3.17(a)]. For Oscillator 1, we have $V_1(t) = V_0 \cos(\omega_0 t + \phi_{n1})$ and for Oscillator 2, $V_2(t) = V_0 \cos(\omega_0 t + \phi_{n2})$. Note that ϕ_{n1} and ϕ_{n2} are uncorrelated because the noise sources in the two oscillators know nothing about each other. It follows that

$$V_{out}(t) = V_1(t) + V_2(t) \tag{3.25}$$

$$\approx V_0 \cos\omega_0 t - V_0 \phi_{n1} \sin\omega_0 t + V_0 \cos\omega_0 t - \phi_{n2} V_0 \sin\omega_0 t \tag{3.26}$$

$$\approx 2V_0 \cos\omega_0 t - V_0(\phi_{n1} + \phi_{n2}) \sin\omega_0 t. \tag{3.27}$$

The amplitude of the second term must be normalized to that of the first, yielding a phase noise of $(\phi_{n1} + \phi_{n2})/2$. Since the two oscillators are designed identically, they have the same phase noise spectra, $S_{\phi n1} =$

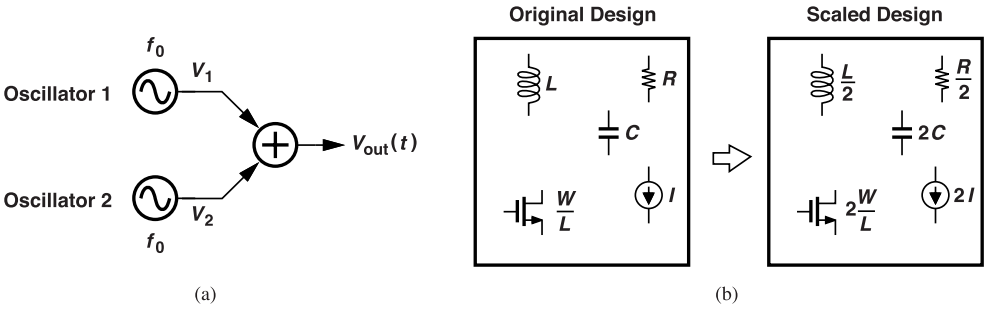


Figure 3.17 Trading power for phase noise by (a) adding outputs of nominally identical oscillators, or (b) linear scaling.

$S_{\phi n2}$, and hence

$$S_{\phi,out} = \frac{S_{\phi n1} + S_{\phi n2}}{4} \tag{3.28}$$

$$= \frac{S_{\phi n1}}{2}. \tag{3.29}$$

In other words, the overall phase noise spectrum is halved for a doubling of the power dissipation.

The above arrangement proves impractical because the oscillators inevitably exhibit some frequency mismatch. A better approach to exploiting the phase noise-power trade-off involves "linear scaling." Depicted in Fig. 3.17(b), the idea is to scale all of the transistor widths, bias currents, and capacitances up by a factor of, say, 2, and all of the inductors and resistors down by the same factor. The voltage swings remain unchanged because we have kept constant the product of the currents and the impedances through which they flow. So does the oscillation frequency because the LC and RC products have not changed. Now, we also observe that all noise *current* spectra are doubled (why?), which, upon multiplication by the magnitude *squared* of the impedances, eventually lead to half as much phase noise (Example 3.4). We often apply linear scaling to an optimized oscillator so as to achieve lower phase noise—albeit at the cost of power—without repeating the entire design effort.

3.4.1 Transistor Noise Simulations

Recall from Eqs. (3.14) and (3.15) that the phase noise is proportional to $S_I(f)$ and $S_{1/f}(f)$ (the power spectral densities of the drain current). For this reason, we must first compute these spectra for the transistors used in the reference design of Fig. 3.12(a). For the 120 nm/40 nm NMOS device, simulations yield the drain noise current spectrum plotted in Fig. 3.18 if $V_{GS} = V_{DD}$ and $V_{DS} = V_{DD}/2$. Here, the vertical axis represents $10 \log I_n^2$, including both flicker and thermal noise. This plot provides a wealth of information. First, the flicker noise drops at a rate of 10 dB/dec from 100 kHz to several tens of megahertz and, more gradually, to several hundred megahertz. Second, we can compute the factor K_N in $S_{1/f}(f) = g_m^2 K_N/(WLC_{ox}f)$ by substituting for g_m, WLC_{ox}, and f.[8] To this end, we use simulations to compute g_m and I_D under the foregoing bias conditions, obtaining 0.22 mS and 92 μA, respectively. Since $WL_{\text{eff}}C_{ox} \approx 120\ \text{nm} \times 36\ \text{nm} \times (17\ \text{fF}/\mu\text{m}^2)$ and since Fig. 3.18 indicates that $S(f = 1\ \text{MHz}) = -207\ \text{dB} = 2 \times 10^{-21}\ \text{A}^2/\text{Hz}$, we have $K_N = 3 \times 10^{-24}\ \text{FV}^2$. (The use of dB to denote A^2/Hz is not quite correct, but we proceed with the understanding that $10 \log S(f)$ is expressed in dB.)

[8] Note that L denotes the effective length here, about 36 nm in this technology.

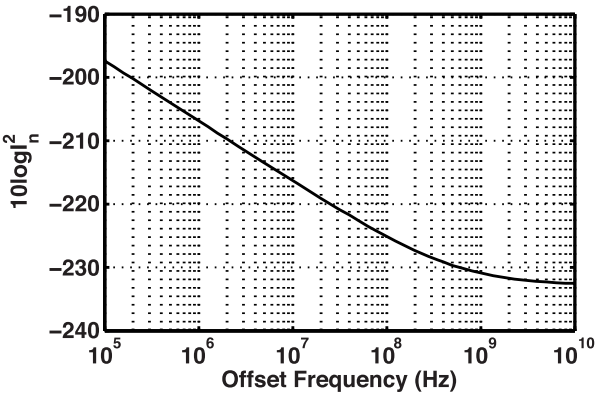


Figure 3.18 Noise current of a 120 nm/40 nm NMOS device with $V_{GS} = V_{DD}$ and $V_{DS} = V_{DD}/2$.

Second, if we extrapolate the low-frequency section of the plot at a slope of -10 dB/dec, it intercepts the thermal noise floor (≈ -233 dB) around $f \approx 500$ MHz,[9] the flicker noise corner frequency of the transistor. With $\gamma = 1$, we have $4kT\gamma g_m = 3.64 \times 10^{-24}\ \mathrm{A}^2/\mathrm{Hz} = -234\,\mathrm{dB}$, a value close to the simulated result.

Example 3.8

For a $1/f$ corner of 500 MHz in the transistor noise current, do we expect the phase noise of the reference ring oscillator of Fig. 3.12(a) to have the same corner frequency?

Solution

Not quite. We must compute the intersection frequency of (3.14) and (3.15). Consider a simpler case where $2kT/(I_D V_{DD})$ is negligible and let us consider only the contribution of NMOS devices. Equating the two yields

$$S_{I,\mathrm{NMOS}} = \frac{1}{2N} S_{1/f,\mathrm{NMOS}}, \tag{3.30}$$

indicating that the corner frequency is lower by a factor of $2N$ (Fig. 3.19). For a three-stage ring, $2N = 6$ and $4kT\gamma g_m = [g_m^2 K/(WLC_{ox} f_{c,osc})]/(2N)$ yields $f_{c,osc} \approx 500\ \mathrm{MHz}/6 \approx 83$ MHz. Of course, to obtain a more accurate result, we must also include the PMOS noise contributions.

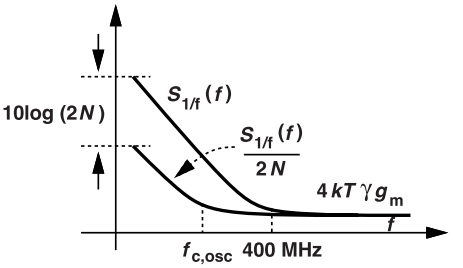


Figure 3.19 Flicker noise corner frequency for a single transistor and for the ring oscillator.

[9] Alternatively, we can write $g_m^2 K/(WLC_{ox} f) = 4kT\gamma g_m$ and find f.

We now repeat the foregoing simulations for a 240 nm/40 nm PMOS device. Plotted in Fig. 3.20 is the drain noise current spectrum, revealing a flicker noise corner frequency of about 650 MHz. Under the bias

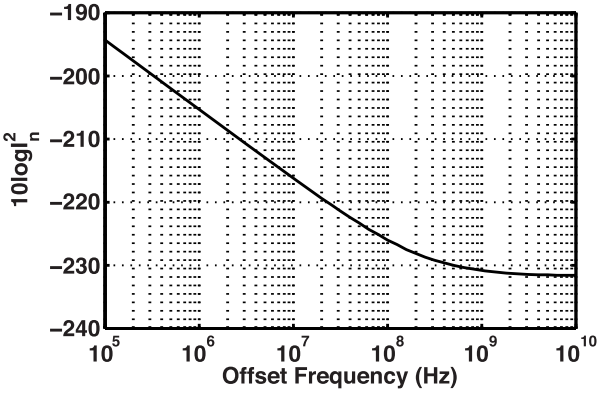


Figure 3.20 Noise current of a 240 nm/40 nm PMOS device with $|V_{GS}| = V_{DD}$ and $|V_{DS}| = V_{DD}/2$.

conditions $|V_{GS}| = V_{DD}$ and $|V_{DS}| = V_{DD}/2$, we have $I_D = 95\ \mu$A and $g_m = 0.29$ mS. Interestingly, in this technology, PMOS transistors exhibit a higher flicker noise than do NMOS devices. This can be seen more clearly if the drain current spectra of Figs. 3.18 and 3.20 are normalized to the gate areas. In fact, we have $K_P = 5.5 \times 10^{-24}$ FV^2.

3.4.2 Reference Oscillator Phase Noise

We begin with the phase noise of the 22.6-GHz reference design in Fig. 3.12(a). Shown in Fig. 3.21 is the simulated phase noise, with the horizontal axis representing the frequency offset from the carrier. We observe

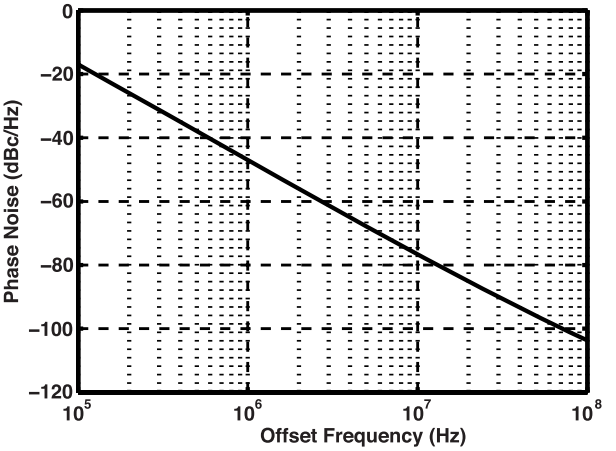


Figure 3.21 Phase noise of 22.6-GHz reference oscillator in Fig. 3.12(a).

that the phase noise falls by about 30 dB from 100 kHz offset to 1 MHz and also from 1 MHz to 10 MHz—as is characteristic of flicker noise upconversion. Extrapolation from these decades implies a corner frequency of about 90 MHz. The discrepancy with respect to the corner frequency calculated above (67 MHz) is due to the noise contribution of the PMOS devices.

The reference oscillator suffers from a very high phase noise, e.g., -47 dBc/Hz at 1-MHz offset. Of course, this value must be normalized to both the power dissipation and the square of the oscillation frequency for meaningful comparisons.

Example 3.9

Define a figure of merit (FOM) for oscillators based on the oscillation frequency, phase noise, and power dissipation.

Solution

Based on our derivations, we define

$$\text{FOM} = \frac{1}{\text{phase noise} \times \text{power dissipation}} \left(\frac{f_0}{f}\right)^2, \tag{3.31}$$

where f is the offset at which the phase noise is measured and is often denoted by Δf. We typically report $10 \log \text{FOM}$. Note that this FOM improves as we go from the $1/f$-noise regime (low f) to the white-noise regime (high f) (why?). The figure of merit provides a consistent normalization procedure for evaluating the performance of oscillators. It is common in the literature to express P in milliwatts rather than in watts.

It is instructive to compare the simulated phase noise to the theoretical values predicted by Eqs. (3.14) and (3.15). For thermal noise, we have from Figs. 3.18 and 3.20: $S_{I,\text{NMOS}} = -233 \text{ dB} = 5 \times 10^{-24} \text{ A}^2/\text{Hz}$ and $S_{I,\text{PMOS}} = -232 \text{ dB} = 6.3 \times 10^{-24} \text{ A}^2/\text{Hz}$, respectively. Since the two are uncorrelated, we add their spectral densities, obtaining $S_{I,NMOS} + S_{I,PMOS} = 1.13 \times 10^{-23} \text{ A}^2/\text{Hz}$. With $|I_D| \approx 93 \ \mu\text{A}$ and $V_{DD} = 0.95$ V, we have

$$S_{\phi n,white}(f) = 7.45 \times 10^{-16} \left(\frac{f_0}{f}\right)^2 \text{ Hz}^{-1} \tag{3.32}$$

$$= \frac{3.8 \times 10^5}{f^2} \text{ Hz}^{-1}. \tag{3.33}$$

For example, at $f = 100$ MHz, the phase noise is predicted to be around -104 dBc/Hz, close to that in Fig. 3.21.

Similarly, for flicker noise, we compute from Figs. 3.18 and 3.20 at, say, 1 MHz, $S_{1/f,\text{NMOS}} = -207 \text{ dB} = 2 \times 10^{-21} \text{ A}^2/\text{Hz}$ and $S_{1/f,\text{PMOS}} = -205 \text{ dB} = 3.2 \times 10^{-21} \text{ A}^2/\text{Hz}$. The corresponding noise current spectra are equal to $(2 \times 10^{-21})(1 \text{ MHz})/f$ and $(3.2 \times 10^{-21})(1 \text{ MHz})/f$, respectively. We add these two values and multiply the result by $(f_0^2/f^2)/(4NI_D^2)$, obtaining

$$S_{\phi n,flicker}(f) = 5 \times 10^{-8} \frac{f_0^2}{f^3} \text{ Hz}^{-1}. \tag{3.34}$$

It follows that

$$S_{\phi n,flicker}(f) = \frac{2.55 \times 10^{13}}{f^3} \text{ Hz}^{-1}. \tag{3.35}$$

For example, at $f = 1$ MHz, the phase noise reaches -46 dBc/Hz, close to the simulated value in Fig. 3.21.

Example 3.10

Compute the FOM given by (3.31) for the 22.6-GHz oscillator at $f = 1$ MHz and $f = 100$ MHz.

Solution

At 1-MHz offset, simulations yield $S_{\phi n} = -47$ dBc/Hz $= 2 \times 10^{-5}$ Hz^{-1}. With $P = 60\ \mu\text{A} \times 0.95\ \text{V} = 57\ \mu\text{W}$, we have $\text{FOM} = 4.5 \times 10^{17}$ Hz/W $= 177$ dB. At 100-MHz offset, $S_{\phi n}$ drops to -104 dBc/Hz, yielding $\text{FOM} = 194$ dB. To remain consistent with the literature, we express P in milliwatts, obtaining $\text{FOM} = 147$ dB and 164 dB, respectively.

Example 3.11

How do we modify the reference oscillator to obtain a phase noise of -120 dBc/Hz at 100-MHz offset with the same oscillation frequency?

Solution

For a 16-dB reduction in phase noise, we can scale the power consumption up by a factor of $10^{1.6} \approx 40$. That is, we increase the widths of all transistors by this factor. The parasitic capacitances in Fig. 3.12(a) also rise by approximately a factor of 40 due to the proportionally larger layout dimensions.The entire phase noise profile, including the $1/f$ regime, shifts down by 16 dB while the power rises to $40 \times 60\ \mu\text{A} \times 0.95\ \text{V} = 2.3$ mW.

3.4.3 First 2-GHz Oscillator Phase Noise

Recall from Section 3.3 that we can add capacitance to the internal nodes of a ring to lower the oscillation frequency without changing the power consumption. For the 1.93-GHz design in Fig. 3.14(a), we obtain the phase noise depicted in Fig. 3.22. As predicted by Eqs. (3.14) and (3.15), the phase noise should fall by

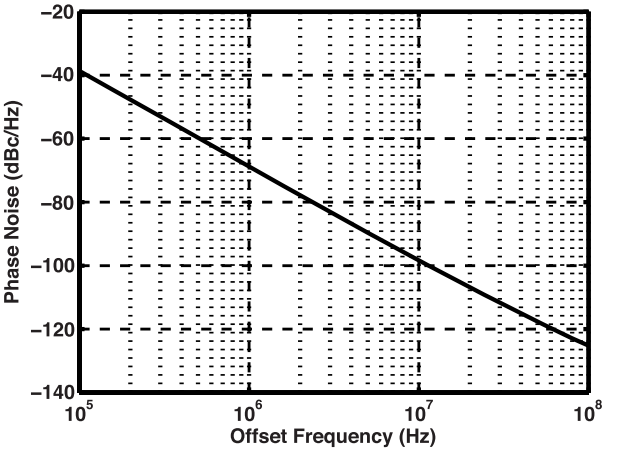


Figure 3.22 Phase noise of ring oscillator in Fig. 3.14(a).

$10 \log(22.6\ \text{GHz}/1.93\ \text{GHz})^2 = 21.5$ dB. Indeed, the plots in Figs. 3.21 and 3.22 differ by almost exactly this amount. That is, the oscillator figure of merit has not changed.

3.4.4 Second 2-GHz Oscillator Phase Noise

In the second type of modification in Section 3.3, we increased the number of stages to 31 and arrived at $f_0 = 2.3$ GHz. Plotted in Fig. 3.23 is this ring's phase noise, exhibiting substantial improvement in the flicker-noise

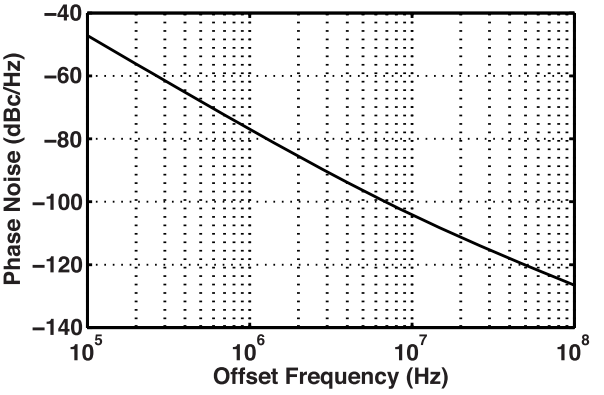


Figure 3.23 Phase noise of 31-stage ring oscillator in Fig. 3.15(a).

regime. Equation (3.15) implies the same trend as N increases. In fact, if we calculate $10\log(31/3) \approx 10$ dB and account for the frequency difference between the designs in Figs. 3.14(a) and 3.15(a) by subtracting $10\log(2.3\text{ GHz}/1.93\text{ GHz})^2 = 0.8$ dB, we expect a phase noise reduction of 9.2 dB, which is close to the difference between the plots in Figs. 3.22 and 3.23. The second oscillator thus has a better FOM.

The key point emerging from this study is that, for a given oscillation frequency, it is preferable to increase the number of stages than to add capacitance to the internal nodes.

3.4.5 Third 2-GHz Oscillator Phase Noise

The third design incorporates three stages but with longer NMOS and PMOS transistors. As explained in Section 3.3, such an arrangement draws less power, about 13 μW, because its node capacitances are smaller. Figure 3.24 shows the phase noise of this 2.5-GHz ring.

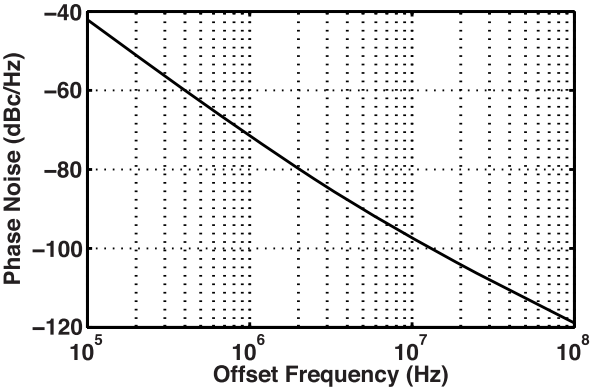


Figure 3.24 Phase noise of ring oscillator in Fig. 3.16(a).

Example 3.12

Compare the $1/f^3$ regimes of the phase noise profiles in Figs. 3.23 and 3.24 and determine whether they agree with Eq. (3.23).

Solution

Equation (3.23) suggests that, for a given overdrive voltage, $V_{DD} - V_{TH}$, the phase noise simply depends on the total NMOS and PMOS channel capacitances. We can compute $K/[N(WL)_{NMOS}] +$

$K/[N(WL)_{PMOS}]$ for both oscillators and see whether the phase noise scales accordingly. Assuming an effective length of 236 nm for the devices in the third oscillator, we obtain this sum as $4.11 \times 10^{-11}\ \mathrm{FV}^2/\mathrm{m}^2$ and $6.47 \times 10^{-11}\ \mathrm{FV}^2/\mathrm{m}^2$ for the second and third designs, respectively. We calculate $10\log$ of the ratio $6.47/4.11$ and add $10\log(2.5\ \mathrm{GHz}/2.3\ \mathrm{GHz})^2 = 0.7$ dB to account for the frequency difference. That is, Eq. (3.23) predicts that the phase noise of the third oscillator should be higher by 2.7 dB at 100-kHz offset.

Why is the actual difference between Figs. 3.23 and 3.24 close to 5 dB? According to device noise simulations, the K_N and K_P factors in fact *vary* with the channel length, leading to this 2.7-dB discrepancy.

It is instructive to determine which strategy results in higher performance: increasing the number of stages or increasing the transistor lengths. Taking into account the power difference between the two designs as $10\log(57\ \mu\mathrm{W}/13\ \mu\mathrm{W}) = 6.4$ dB and adding another 0.7 dB for the frequency difference, we expect the phase noise of the third design to be higher by 7.1 dB if the two strategies are equally good. This is indeed the case at 100-MHz offset. At 1-MHz offset, on the other hand, the third design's phase noise (-74 dBc/Hz) is only 4 dB higher than the second, indicating a 3-dB FOM advantage. In other words, increasing the channel lengths is generally preferred.

Example 3.13

How should we modify the third oscillator to obtain a phase noise of -100 dBc/Hz at 1-MHz offset with the same oscillation frequency?

Solution

For the phase noise to fall from -72 dBc/Hz to -100 dBc/Hz, we can sacrifice power by a factor of $10^{2.8} \approx 631$. We thus choose $(W/L)_{NMOS} = 76\ \mu\mathrm{m}/240$ nm and $(W/L)_{PMOS} = 152\ \mu\mathrm{m}/240$ nm. The circuit now draws 8.2 mW, a large amount, and also occupies a large area.

3.4.6 Fourth 2-GHz Oscillator Phase Noise

The fourth approach to achieving the desired frequency is through the use of a frequency divider. Let us neglect the divider's phase noise and power dissipation. If the output frequency of a oscillator is divided by M, the result can be expressed as

$$V_{div}(t) = V_0 \cos\left[\frac{\omega_0 t + \phi_n(t)}{M}\right]. \tag{3.36}$$

That is, the phase is also divided by M because phase and frequency are simply related by a (linear) differential operator. We thus conclude that the output phase noise is lower by a factor of M (by $20\log M$ in dB), just as in the first 2-GHz oscillator design. Frequency division, therefore, does not improve a ring oscillator's FOM.

Example 3.14

Explain what happens to the output jitter when the frequency is divided.

Solution

The jitter, expressed in seconds, does not change even though the phase noise falls. This can be seen if we visualize at the divider input a random edge displacement equal to ΔT seconds (Fig. 3.25): the divider output is displaced by the same amount. This illustration also shows why phase noise decreases: since the output period, T_{out}, is *longer*, when we express the phase in radians as $(\Delta T/T_{out}) \times (2\pi)$, we obtain a smaller value than $(\Delta T/T_{in}) \times (2\pi)$.

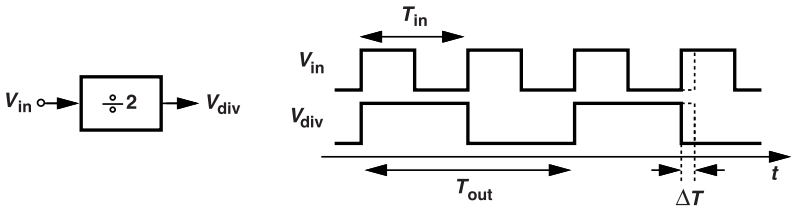


Figure 3.25 Effect of input jitter in a frequency divider .

In summary, if we wish to reduce the oscillation frequency of a ring to a desired value, the best solution is to increase the channel lengths of the transistors; the next best solution is to increase the number of stages. Table 3.1 summarizes the performance of our four designs. The topology incorporating long-channel transistors exhibits the highest FOM in both $1/f^2$ and $1/f^3$ regimes and will be the focus of our study in the next section.

	Design 1 Higher Node Capacitances	Design 2 Greater Number of Stages	Design 3 Longer Transistors	Design 4 Frequency * Division
f_0 (GHz)	1.93	2.3	2.5	2.26
P (uW)	57	68	13	60
Supply Sensitivity K_{VDD}(GHz/V)	4.3	4.9	5	5
Phase Noise (dBc/Hz) at				
1–MHz Offset	–69	–78	–74	–67
100–MHz Offset	–126	–126	–118	–124
FOM (dB) at				
1–MHz Offset	147	158	161	147
100–MHz Offset	164	164	165	164

* Divider's power consumption and phase noise are neglected.

Table 3.1 Summary of four 2-GHz oscillator designs.

3.5 Frequency Tuning

We described the concept of voltage-controlled oscillators in Chapter 2 and constructed mathematical models for them. In this section, we devise frequency tuning techniques for inverter-based ring oscillators and apply them to the optimized oscillator of Fig. 3.16(a).

3.5.1 Tuning Considerations

We must study five aspects of each tuning mechanism:

1. The frequency range that can be achieved; as explained in Chapter 1, this range must be wide enough to cover both PVT variations and the desired frequency range(s) necessary in a given application.

2. The value of the VCO gain, K_{VCO}. In general, we wish to maintain a relatively "low" K_{VCO} so that any noise coupled to the control line does not translate to large frequency and phase fluctuations. This issue necessitates a *wide* range for the control voltage, V_{cont}; as depicted in Fig. 3.26, if the allowable range for V_{cont}—dictated by the VCO itself or by the preceding circuit—is limited to $[V_1\ V_2]$ rather than to $[0\ V_{DD}]$,

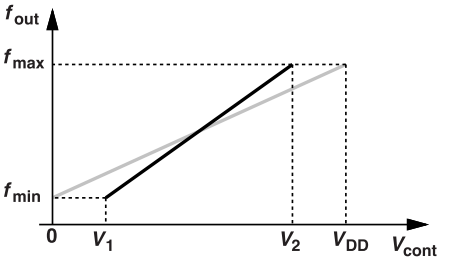


Figure 3.26 Undesirable scaling of K_{VCO} due to narrower control voltage range.

then K_{VCO} inevitably increases. As a rule of thumb, we keep K_{VCO} below roughly $0.1 f_m$ Hz/V, where $f_m = (f_{max} + f_{min})/2$. For example, if $f_m = 2$ GHz, we prefer a K_{VCO} below 200 MHz/V. In some cases, even lower K_{VCO} values are required.

3. The "linearity" of the tuning characteristics, loosely viewed as how much the slope, K_{VCO}, varies (Fig. 3.27); it is generally desirable to limit this variation to no more than 20%. If, for example, the operation frequency changes from f_1 to f_2, we wish the VCO's gain to remain relatively constant. As explained in Chapter 7, this ensures that a phase-locked loop employing such an oscillator incurs negligible change in its static and dynamic behavior.

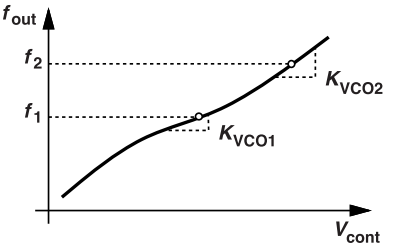


Figure 3.27 Nonlinearity in VCO characteristic.

4. The change in the oscillator internal voltage swings; we prefer to keep the swings relatively constant across the tuning range, primarily to avoid degrading the phase noise.
5. The change in the phase noise across the tuning range.

3.5.2 Continuous and Discrete Tuning

The direct relation between K_{VCO} and the tuning range, depicted in Fig. 3.26, poses a difficulty: a wide frequency tuning range inevitably makes the VCO more sensitive to noise on its control line. In most modern systems, we avoid this issue through the use of both continuous (analog) and discrete (digital) tuning. Illustrated in Fig. 3.28(a), the idea is to choose the path controlled by V_{cont} weak enough—and hence K_{VCO} low enough—while also providing discrete jumps in the frequency by means of a digital control, D_{cont}. For example, V_{cont} can vary a resistance and D_{cont} can switch more or fewer capacitors into the oscillator nodes. We observe that, with $D_{cont} = 11$, f_{out} varies on the lowest curve as V_{cont} goes from V_1 to V_2. If this frequency range is too low, we change D_{cont} to 10, etc. The voltage range $[V_1\ V_2]$ may be limited by the preceding stage or by the oscillator itself.

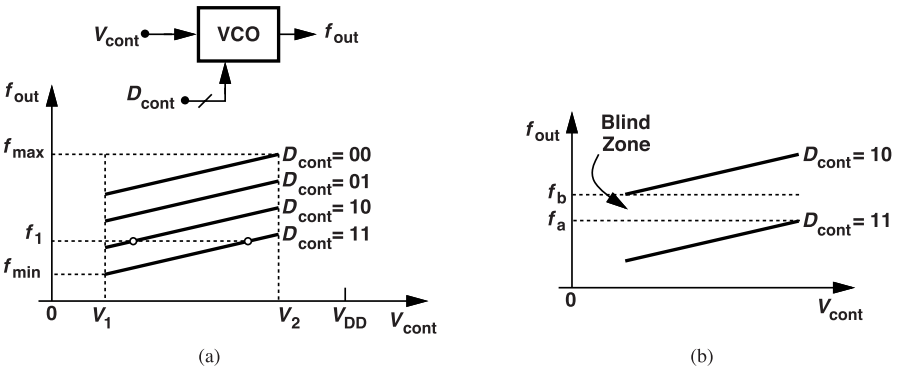


Figure 3.28 (a) Discrete tuning of a VCO, and (b) problem of blind zone.

The overall tuning characteristics depicted in Fig. 3.28(a) must cover the entire desired frequency range, $[f_{min}\ f_{max}]$, as V_{cont} varies from V_1 to V_2. It is important to note that each continuous characteristic must have enough *overlap* with those immediately above and below it to ensure no "blind zone" can occur. As illustrated in Fig. 3.28(b), if two adjacent curves do not have overlap, the VCO fails to cover a certain range of frequencies, $[f_a\ f_b]$. In Fig. 3.28(a), frequency f_1 is covered by two characteristics, ensuring that the VCO can be tuned to this frequency even if the analog curve corresponding to $D_{cont} = 10$ shifts up with PVT.

Example 3.15

From the characteristics in Fig. 3.28(a), plot f_{out} versus D_{cont} with V_{cont} as a parameter.

Solution

Figure 3.29 depicts this behavior, showing the overlaps more clearly.

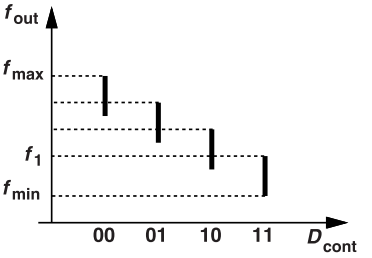


Figure 3.29 Illustration of overlap between discrete tuning characteristics.

As a rule of thumb, we design a ring VCO such that V_{cont} provides a tuning range of ± 10 to $\pm 20\%$ of the center frequency, and then select the resolution of the digital control so as to cover the entire desired range with no blind zones. In some cases, this choice still leads to a high K_{VCO}, necessitating finer discrete steps.

3.5.3 Tuning by Variable Resistance

Since the gate delay in a ring is given by the inverter strength and the load capacitance, we surmise that the frequency can be tuned by varying one of the two. We begin with the former.

Figure 3.30(a) shows an example of frequency tuning by adjusting the strength of the NMOS transistors.

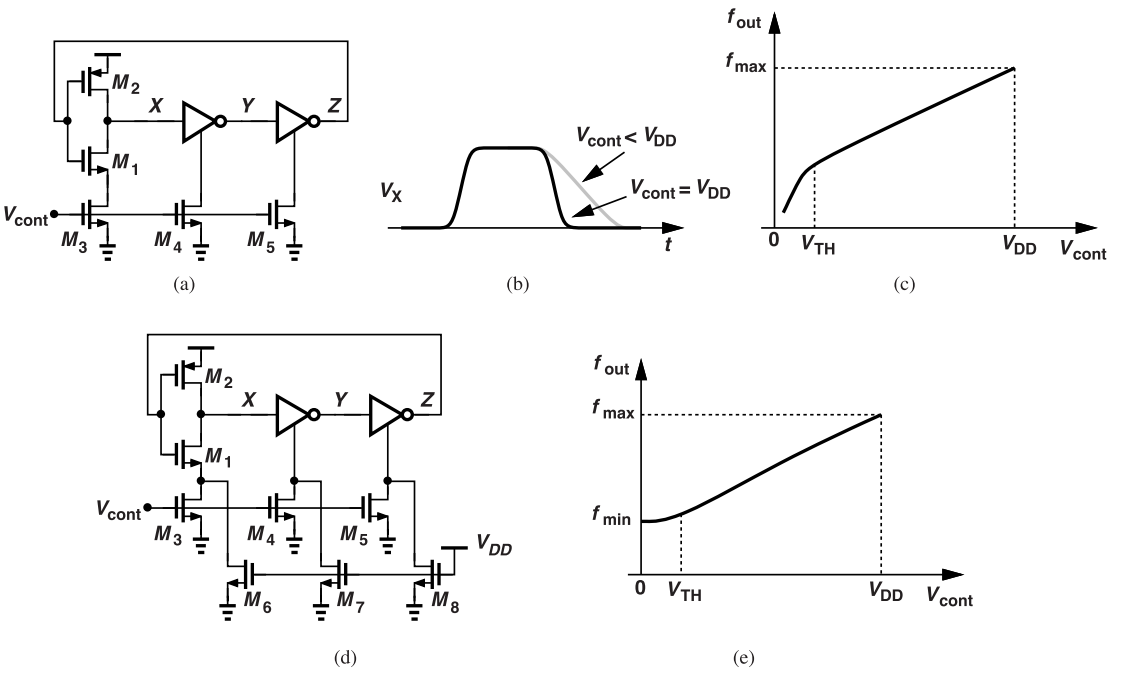


Figure 3.30 (a) Frequency control by variable resistance, (b) effect on output waveform, (c) resulting characteristic, (d) modification to avoid the sharp drop, and (e) new characteristic.

Here, M_3-M_5 act as voltage-dependent resistors. If V_{cont} starts at V_{DD} and decreases, so does the pull-down strength of the inverters, yielding a longer fall time [Fig. 3.30(b)]. The increase in the gate delay thus lowers the oscillation frequency.

This arrangement faces two issues. First, the disparate rise and fall times translate to a nonoptimum design because the inverters produce more phase noise on the falling edges than on the rising edges. This asymmetry becomes more pronounced as the frequency is reduced. Second, as V_{cont} approaches the NMOS threshold voltage, the on-resistance of M_3-M_5 rises sharply, thereby creating a very nonlinear tuning characteristic [Fig. 3.30(c)]. In fact, for $V_{cont} \approx 0$, the circuit fails to oscillate. This issue can be ameliorated by placing a constant resistance in parallel with M_3-M_5 as realized by M_6-M_8 in Fig. 3.30(d). Now, even for $V_{cont} = 0$, the fall time is limited by the series combination of the NMOS on-resistances (e.g., those of M_1 and M_6) and f_{out} reaches a finite minimum [Fig. 3.30(e)]. Nonetheless, the problem of asymmetric transitions remains. Another less critical issue is that M_3-M_8 inevitably reduce the *maximum* oscillation frequency (why?).

Let us apply these ideas to our third 2-GHz oscillator design. As mentioned previously, the continuous control should provide a tuning range equal to ± 10 to $\pm 20\%$ of the center frequency (for ring oscillators). We therefore select M_6-M_8 in Fig. 3.30(d) weak enough to obtain $f_{max} - f_{min} \approx 400$ to 500 MHz in Fig. 3.30(e). According to simulations, we need $(W/L)_{6-8} = 120\,\text{nm}/180\,\text{nm}$, arriving at the design depicted in Fig. 3.31(a) and the phase noise profile in Fig. 3.31(b) at 2 GHz (for $V_{cont} \approx V_{THN}$).

We should make two remarks. First, M_3-M_5 are chosen strong enough not to limit the maximum frequency significantly. Nonetheless, f_{max} has dropped from 2.5 GHz to 2.4 GHz. Second, the phase noise at 100-kHz and 1-MHz offsets is about 3 dB less than that in Fig. 3.24. Why? Let us denote the NMOS and PMOS contributions by S_N and S_P, respectively, and assume they are equal. We first recognize that, for $V_{cont} \approx V_{THN}$, M_6-M_8, which operate in the triode region, degenerate the corresponding NMOS devices within the inverters [Fig. 3.32(a)]. Since degeneration is equivalent to making the inverters' NMOS transistors longer, we roughly view M_6 as shown in Fig. 3.32(b). Equivalently, the NMOS channel area has risen from $120\,\text{nm}\times$

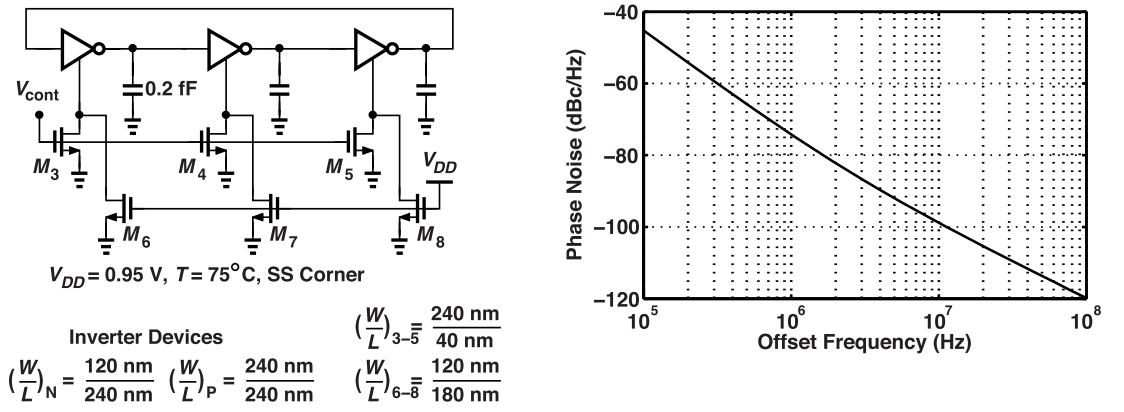


Figure 3.31 (a) Ring VCO, and (b) its phase noise.

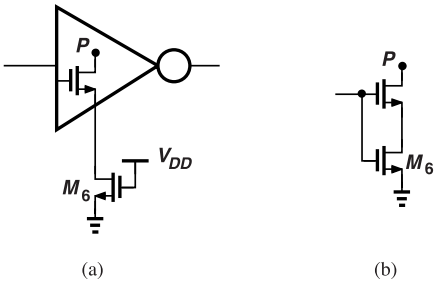


Figure 3.32 (a) Transistor M_6 acting a degenerating device, and (b) approximate equivalent structure.

236 nm to 120 nm × 236 nm + 120 nm × 176 nm.[10] That is, the total noise drops from $S_N + S_P = 2S_P$ to $S_N/1.75 + S_P = 1.57S_P$, which translates to a 1-dB reduction. In addition, f_0 has fallen from 2.5 GHz to 2 GHz (for this phase noise plot), giving another 1.94 dB. Thus, we expect a phase noise reduction of 2.94 dB, close to what simulations predict.

Example 3.16

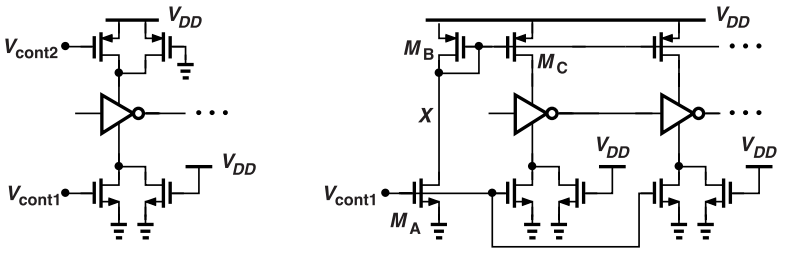
How do we tune the strength of the inverters symmetrically?

Solution

We can insert a PMOS device in series at the top so as to adjust both NMOS and PMOS strengths as depicted in Fig. 3.33(a). The difficulty, however, is that the two control voltages must vary in opposite directions. We therefore add an inverting (common-source) stage to produce V_{cont2} from V_{cont1} [Fig. 3.33(b)]. Here, as V_{cont1} rises from V_{THN} to V_{DD}, V_{cont2} falls from $V_{DD} - |V_{THP}|$ to nearly zero if M_A is much stronger than M_B. Proper ratioing of devices therefore yields approximately equal rise and fall times across the tuning range.

Unfortunately, the common-source stage in Fig. 3.33(b) introduces considerable noise, corrupting the rising edges within the ring oscillator. To quantify this effect, we first compute the noise voltage produced by M_A at X. We multiply the gate-referred $1/f$ noise voltage by $g_{m,A}$ to obtain the corresponding drain current,

[10] Even though the gate of M_6 is not driven by the input, this transistor appears in series with the inverter's NMOS device, reducing the $1/f$ noise injection into the output transitions.



(a) (b)

Figure 3.33 (a) Controlling both PMOS and NMOS strengths, and (b) current-mirror implementation.

add the thermal noise current to it, and let the result flow through the impedance of the diode-connected transistor, M_B:

$$V_{n,A}^2 = \left[\frac{K}{(WL)_A C_{ox}}\frac{1}{f}g_{m,A}^2 + 4kT\gamma g_{m,A}\right]\frac{1}{g_{m,B}^2}, \tag{3.37}$$

where channel-length modulation is neglected. The $1/f$ and thermal noise of M_B must also be added to this result, which we leave as an exercise for the reader.

The noise current of M_B is not directly mirrored onto M_C because the latter operates in the triode region part of the time. Nevertheless, we can still determine the phase noise due to $V_{n,A}$. Since the total noise voltage at node X, $V_{n,X}$, appears in series with V_{cont2} (why?), we can readily express the output phase noise if we know the gain from V_{cont2} to the output frequency, K_{VCO2}. Specifically, we have

$$V_{out}(t) = V_0 \cos\left(\omega_0 t + K_{VCO2}\int V_{cont2}dt\right) \tag{3.38}$$

and hence

$$\phi_n(t) = K_{VCO2}\int V_{n,X}dt. \tag{3.39}$$

It follows that

$$S_{\phi n}(f) = K_{VCO2}^2\frac{\overline{V_{n,X}^2}}{4\pi^2 f^2}. \tag{3.40}$$

Note that the oscillator's internal phase noise adds to this result. The phase noise obtained above tends to dominate, making this technique less attractive.

We now construct the tuning characteristic of the VCO shown in Fig. 3.31 to examine K_{VCO} and its variation. Plotted in Fig. 3.34, the curve shows a flat region for $V_{cont} < V_{THN}$ because the controlled NMOS devices are off, and another flat region for $V_{cont} > 0.85$ because the on-resistance of these devices varies negligibly with V_{cont}. If we select the frequency range to be from 2 GHz to 2.4 GHz, then K_{VCO} varies from 300 MHz/V near the ends to 1.25 GHz/V in the middle, a large ratio. This issue arises primarily because control voltages below 0.35 V have little effect on the frequency and remain unutilized.

Design Modification In order to remove the flat zone in Fig. 3.34, we must modify the tuning mechanism such that the frequency decreases further as V_{cont} falls below 0.35 V. This means that V_{cont} must also drive the gate of some PMOS devices, but how? We wish the frequency to drop as a PMOS branch becomes *stronger*. This is possible if we apply positive feedback around the inverters through a controlled PMOS

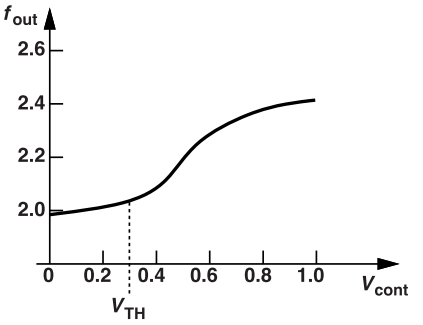


Figure 3.34 Tuning characteristic of the ring in Fig. 3.31.

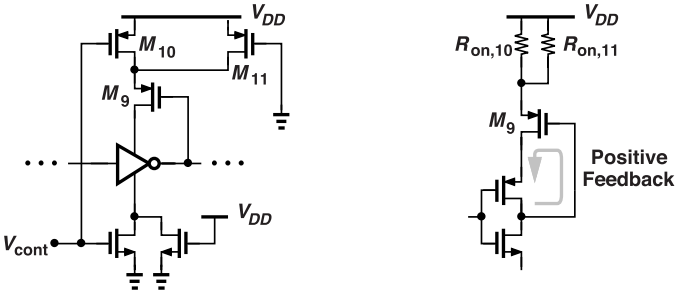


Figure 3.35 (a) Tuning by positive feedback, and (b) simplified circuit.

transistor. Illustrated in Fig. 3.35, the idea is to provide greater feedback through M_9 as V_{cont} decreases; as the on-resistance of M_{10} falls, so does the degeneration of M_9, making the feedback around it stronger. Since positive feedback tends to raise the output impedance of the inverter in this case (why?), the oscillation slows down. Acting as a contant resistor, M_{11} ensures that the inverter is not disabled as V_{cont} approaches $V_{DD} - |V_{TH10}|$.

The topology of Fig. 3.35 entails two issues. First, M_9 must have small dimensions so as to present a negligible load to the inverter. Second, M_{10} must be a weak device so as to lower the frequency by only a moderate amount. (Recall that V_{cont} should provide a moderate tuning range so that K_{VCO} is not inordinately large.)

Rather than attempt to make M_{10} weak, we increase the strength of the main signal path. After all, we know that the present design's phase noise is unacceptably high and can only be reduced by linear scaling (Section 3.4). For example, if we quadruple the widths of all of the transistors in Fig. 3.35, except for M_9 and M_{10}, then the phase noise power spectral density falls by $10 \log 4 = 6$ dB. Shown in Fig. 3.36(a), the resulting circuit is based on that in Fig. 3.31(a) and exhibits the tuning behavior plotted in Fig. 3.36(b). (Note that the interconnect capacitance has also been quadrupled because the larger transistors inevitably lead to wider cells and hence longer wires.) The tuning range is about 550 MHz and the maximum K_{VCO} around 1.4 GHz/V.

Figure 3.36(c) plots the phase noise of the new design at $f_0 = 2$ GHz. As expected, this profile is 6 dB lower than that in Fig. 3.31(b), albeit at the cost of a supply current of 42 μA.

Example 3.17

The maximum K_{VCO} in Fig. 3.36(b) is still quite large. How can we reduce both this value and the variation of K_{VCO}?

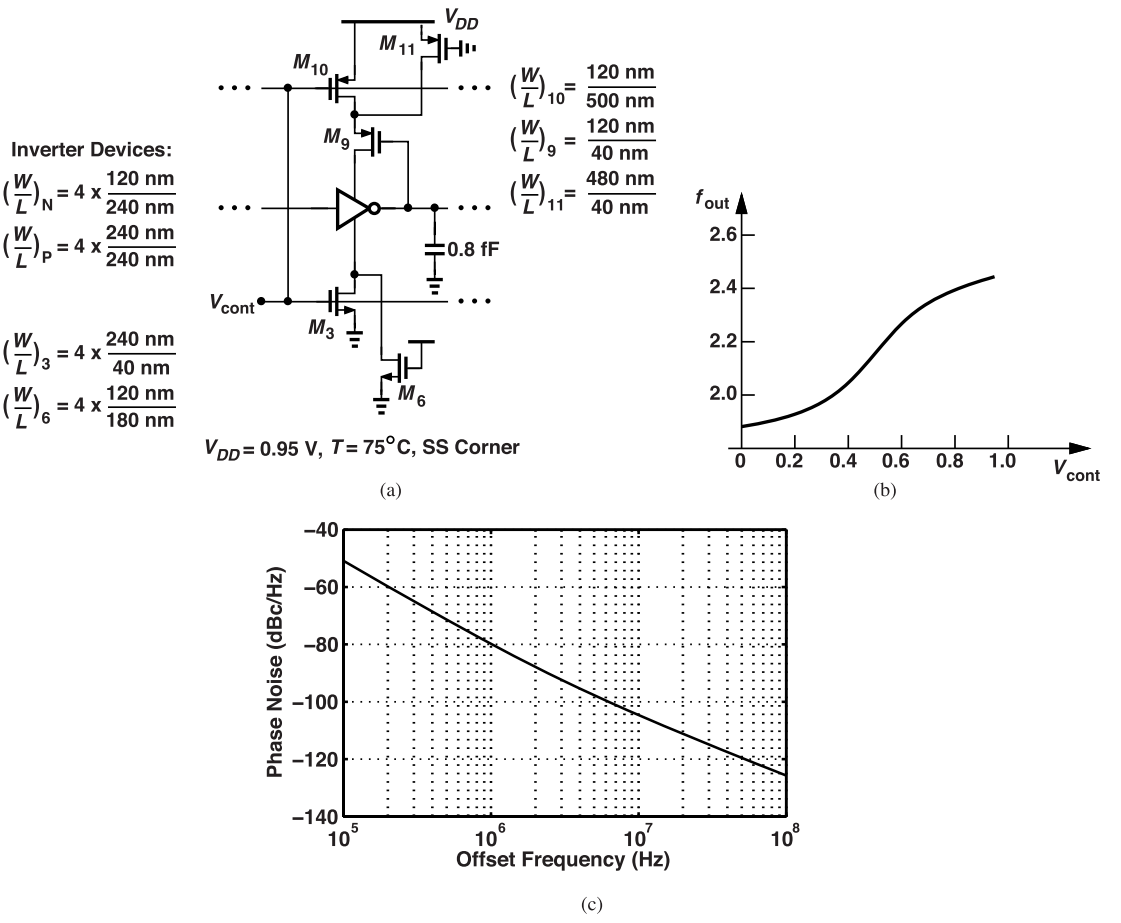


Figure 3.36 (a) Modified VCO using positive feedback for tuning, (b) resulting characteristic, and (c) its phase noise.

Solution

We can weaken the effect of the NMOS control path, e.g., M_3, but this lowers the maximum oscillation frequency. To compensate, we make the NMOS device within the inverter slightly shorter. Figure 3.37 shows the resulting design along with its tuning characteristic and phase noise at $f_0 = 2.07$ GHz. The phase noise is about 1 dB higher than that in Fig. 3.36(c), partially due to $10\log(2.07\text{ GHz}/2.0\text{ GHz})^2 = 0.3$ dB and partially because the NMOS within the inverter has a smaller area. The maximum K_{VCO} is about 800 MHz/V.

3.5.4 Tuning by Variable Capacitance

It is possible to realize continuous tuning by means of voltage-dependent capacitors, also called "varactors." Depicted in Fig. 3.38(a) is a varactor structure commonly used in CMOS technology. It consists of an NMOS transistor placed inside an n-well, with the source, drain, and n-well serving as one terminal and the gate as the other. If V_{AB} is negative, electrons are repelled from the oxide-silicon interface, leaving behind a depletion region [Fig. 3.38(b)]. The total capacitance, C_{AB}, is thus equal to the series combination of the gate oxide capacitance, C_G, and the depletion region capacitance, C_{dep}, $C_{AB} = C_G C_{dep}/(C_G + C_{dep})$. As $|V_{AB}|$ decreases, so does the thickness of the depletion region, yielding a higher C_{AB}. For a sufficiently positive

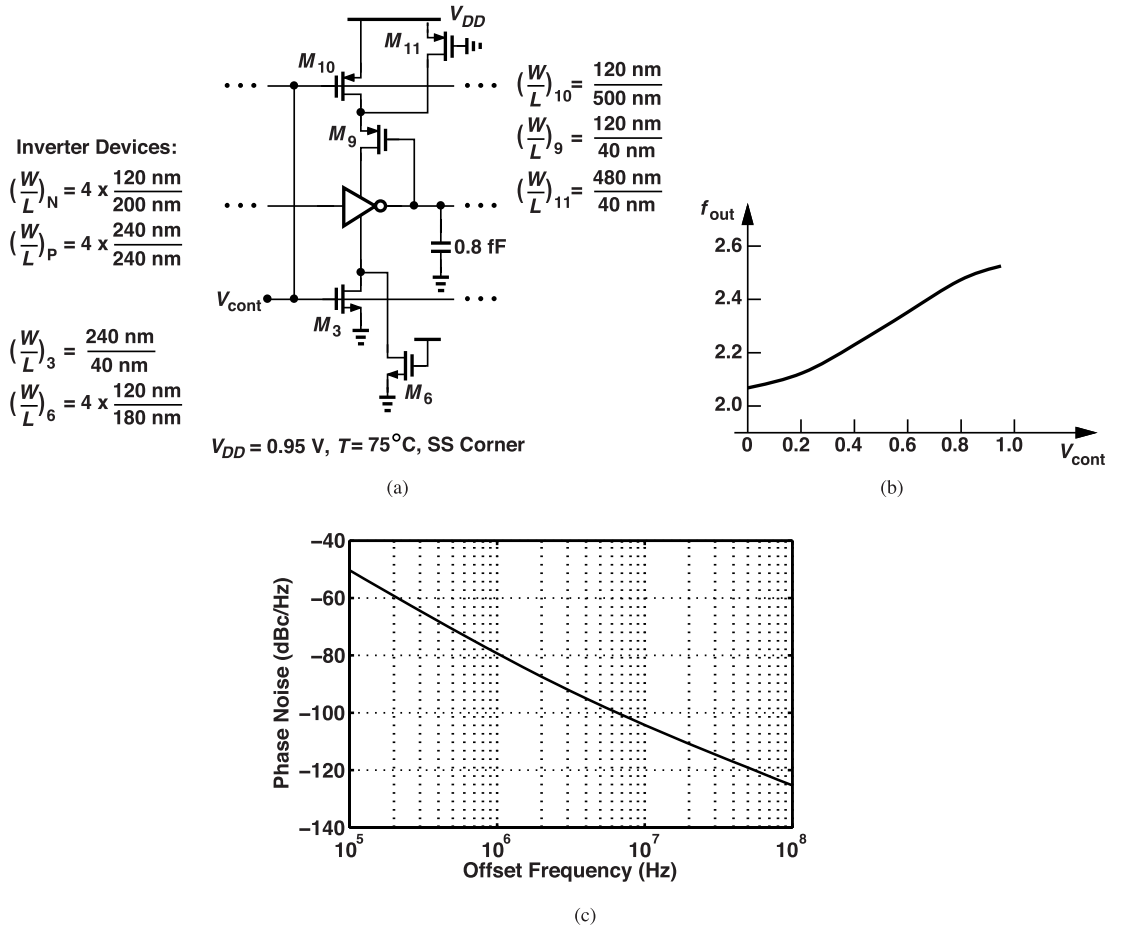


Figure 3.37 (a) VCO with weaker NMOS devices in control path, (b) resulting characteristic, and (c) its phase noise.

V_{AB} [Fig. 3.38(c)], a channel of electrons forms under the oxide, the device operates in the "accumulation mode," and C_{AB} becomes equal to C_G. The C-V characteristic thus emerges as shown in Fig. 3.38(d). We should remark that the varactor suffers from parasitic capacitances between the n-well and the substrate.

The C-V behavior of varactors must be properly modeled in circuit simulations. We approximate the curve in Fig. 3.38(d) by a hyperbolic tangent:

$$C_{AB}(V_{AB}) = \frac{C_{max} - C_{min}}{2} \tanh\left(\alpha + \frac{V_{AB}}{V_0}\right) + \frac{C_{max} + C_{min}}{2}, \tag{3.41}$$

where α and V_0 are fitting parameters for the intercept and the slope, respectively. As $|V_{AB}|$ becomes large enough, C_{AB} approaches C_{min} or C_{max}. It is important to note that the range $[C_{min}\ C_{max}]$ is available only if V_{AB} can assume both negative and positive values. We return to this point later.

In order to tune the three-stage ring of Fig. 3.16(a), we attach the gates of three varactors to the internal nodes and use their source/drain/well terminal as the control voltage [Fig. 3.39(a)]. The varactor dimensions are chosen so as to provide the necessary continuous tuning range. Varactor tuning introduces negligible noise in the control path, an advantage over transistor tuning, e.g., the topology shown in Fig. 3.36(a). However, this method requires accurate varactor models.

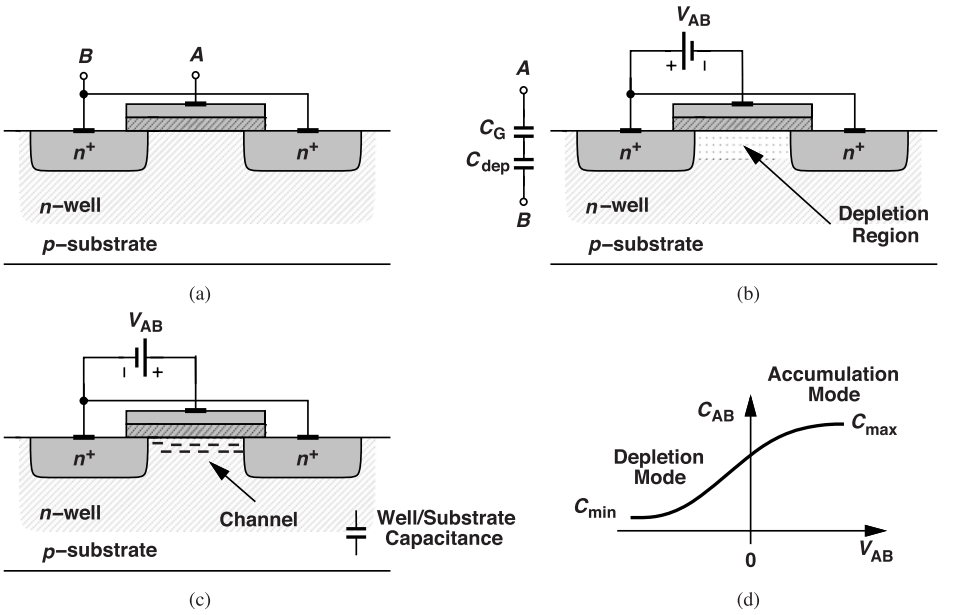


Figure 3.38 (a) MOS varactor structure, (b) illustration of oxide and depletion region capacitances, (c) operation in accumulation mode, and (d) C/V characteristic.

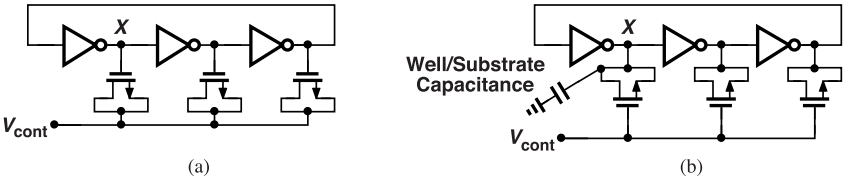


Figure 3.39 Ring oscillator using varactors with (a) $K_{VCO} > 0$, and (b) $K_{VCO} < 0$.

Example 3.18

Sketch as a function of time the voltage across the varactors in Fig. 3.39(a) if $V_{cont} = 0$. How does their capacitance vary with time?

Solution

Each varactor's gate voltage swings between 0 and V_{DD}. As shown in Fig. 3.40, such an excursion causes C_{AB} to go from C_{int} to C_{max}. Thus, the varactor capacitances are modulated periodically. Note that $C_{AB}(t)$ can be expanded as a Fourier series, an observation that proves useful later. In this case, the range [C_{min} C_{int}] remains unutilized because the varactors' voltages cannot become negative.

If the capacitances loading an oscillator's nodes vary with time, how do we determine the oscillation frequency, f_0? As a first-order approximation, we use the "average" value of these capacitances, defined as the first term in their Fourier series expansion. In Fig. 3.40,

$$C_{AB}(t) = C_0 + C_1 \cos \omega_0 t + C_2 \cos(2\omega_0 t) + \cdots, \tag{3.42}$$

where C_0 is the average value. If the C_{AB} waveform is relatively symmetric, we can write $C_0 \approx (C_{max} + C_{int})/2$.

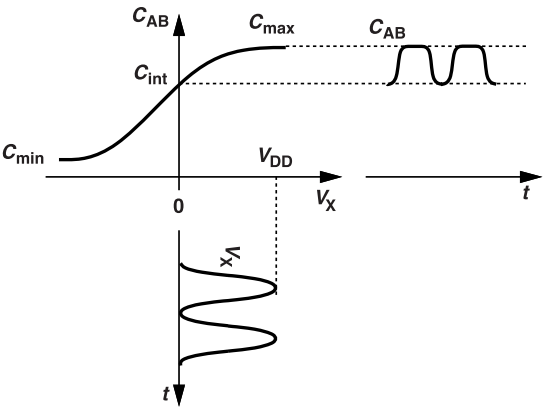


Figure 3.40 Variation wth time of varactor capacitance in an oscillator.

Let us make two observations. First, as V_{cont} in Fig. 3.39(a) rises from zero to V_{DD}, so does f_0 (why?), i.e., $K_{VCO} > 0$. If we had connected the gates of the varactors to V_{cont} [Fig. 3.39(b)], then $K_{VCO} < 0$. But this choice would load the ring with the n-well/substrate capacitances of the varactors. Second, even for $V_{cont} = V_{DD}$, the ring is loaded by the varactors, experiencing a drop in its maximum oscillation frequency. Nonetheless, since varactor tuning covers a narrow frequency range (for K_{VCO} to be small), the varactors' fixed capacitance is small and this drop is tolerable.

3.6 Discrete Frequency Tuning

We wish to add discrete tuning to the design shown in Fig. 3.37(a) so as to obtain a wide frequency range. We face two questions. (1) Should we change the inverter strength or the load capacitance in discrete steps? (2) How much discrete tuning is necessary? For the former, we prefer to simply switch load capacitors into or out of the output node. For the latter, we should think more.

Let us assume that the VCO must always operate at 2 GHz in a given application. The discrete (and continuous) tuning is therefore expected to bring the frequency to this value at any PVT extreme. Since our design has thus far assumed SS devices at 75 °C with $V_{DD} = 0.95$ V, we now turn to another extreme defined by FF transistors at 0 °C with $V_{DD} = 1.05$ V. As shown in Fig. 3.41, the circuit runs at a higher frequency, necessitating that discrete tuning slow it down to 2 GHz. With $V_{cont} \approx 0$ V, we have, from simulations, $f_{out} = 3.45$ GHz. Other PVT corners yield frequencies between these two extremes. We should

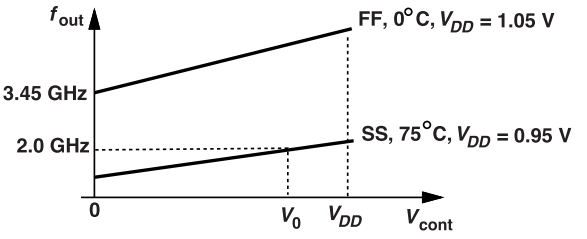


Figure 3.41 PVT extremes of tuning characteristics.

add sufficient capacitance to each node to obtain $f_{out} \approx 2$ GHz. Given the continuous tuning range of Fig. 3.37(b), about 400 MHz, we surmise that four or five partially-overlapping tuning characteristics suffice to bring f_{out} from 3.45 GHz to 2 GHz.

Figure 3.42(a) shows one stage of the ring along with programmable capacitors C_1-C_4. The switches are

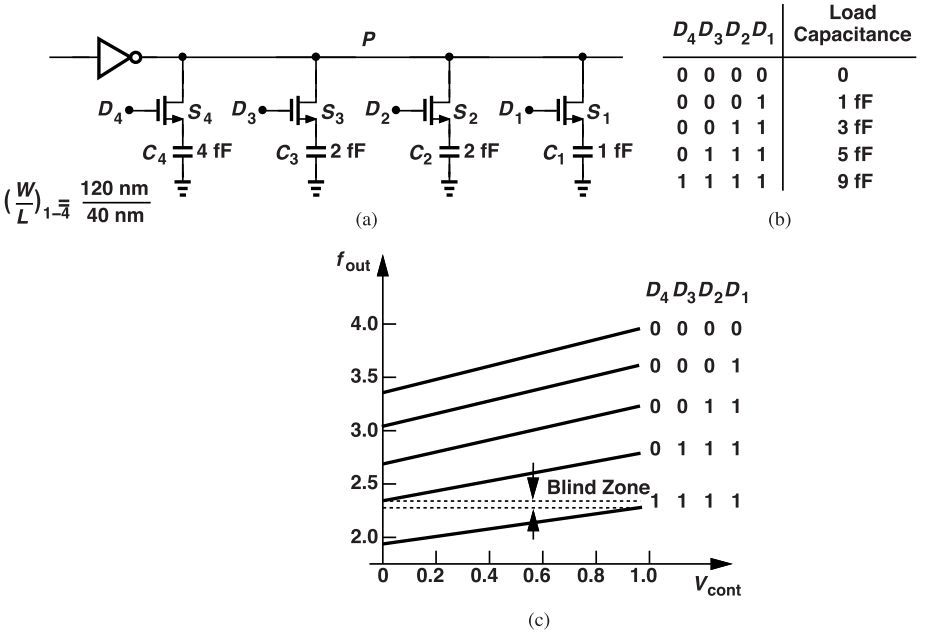


Figure 3.42 (a) One stage of a ring oscillator with discrete tuning, (b) digital control in thermometer code form, and (c) tuning characteristics.

driven by a "thermometer" code.[11] That is, to reduce f_{out}, we turn on S_1 first, and S_2 next while S_1 is on, and then S_3 while S_1 and S_2 are on, etc. Figure 3.42(b) depicts the logical values assumed by $D_4D_3D_2D_1$: as the number of ONEs in the thermometer code increases, so does the load capacitance. The thermometer-code scheme guarantees that f_{out} monotonically rises with $D_4D_3D_2D_1$, allowing the coarse tuning logic to converge easily (Chapters 8 and 9). We wish to choose a small width for the switches so as to minimize their loading upon the circuit when they are off. However, their on-resistance must be low enough for the inverter to "see" C_1-C_4 when they are on.

Example 3.19

If S_4 in Fig. 3.42(a) is on, does the voltage on C_4 track V_P?

Solution

No, when V_P reaches V_{DD}, the capacitor voltage approaches $V_{DD} - V_{TH}$. This means that node P does not see C_4 for the entire cycle, a minor issue.

How do we select the capacitor values in Fig. 3.42(a)? Should they be equal? Suppose the total capacitance at node P when the switches are off is denoted by C_P. In the first downward frequency step, $D_4D_3D_2D_1$ goes from 0000 to 0001 and we turn on only S_1, causing a frequency change proportional to $(C_1 + C_P)/C_P$. In the last step, $D_4D_3D_2D_1$ changes from 0111 to 1111 and we turn on S_4 while C_1-C_3 already load the output. Now the frequency change is proportional to $(C_1 + \cdots + C_4 + C_P)/(C_1 + \cdots + C_3 + C_P)$. Thus, if $C_1 = \cdots = C_4$, the percentage change in the load capacitance is greater in the first step than in the last,

[11] For a decimal value of, say, 5 in a 3-bit system, the thermometer code is given by 00011111. As the value increases, so does the number of ONEs in the code.

leading to very different jumps in the frequency. We therefore expect that C_1 must be the smallest, and C_4 the largest.

The procedure for finding the necessary capacitor values is as follows. With $V_{cont} \approx 0$, we turn on only S_1 and choose C_1 to reduce the frequency by an amount equal to the continuous tuning range minus some overlap. For example, we target a change of about 300 MHz. Next, we turn on S_2 while S_1 is on and choose C_2 accordingly, etc. Finally, we test the other ends of the continuous tuning range by measuring f_{out} when $V_{cont} \approx 1$ V.

The foregoing procedure yields $C_4 = 4$ fF, $C_3 = C_2 = 2$ fF, and $C_1 = 1$ fF with a switch size of 120 nm/40 nm. The rough tuning characteristics are shown in Fig. 3.42(c). Minimum-size switches suffice in this case.[12] We observe a blind zone between the two lowest curves. This can be avoided by decomposing the S_4-C_4 branch into two, each using about 3 fF.

Example 3.20

Compare the topology of Fig. 3.42(a) to that shown in Fig. 3.43(a), where the switches are tied to the bottom plates of the capacitors.

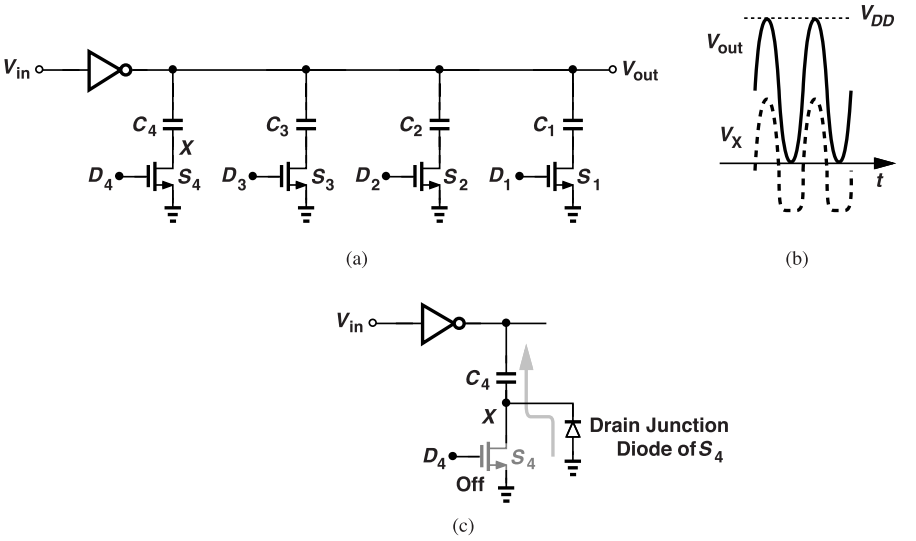


Figure 3.43 (a) Discrete tuning with bottom-plate switching, (b) internal waveforms, and (c) turn-on of the drain junction diode.

Solution

Suppose the switches in Fig. 3.43(a) are off. How much is the dc voltage at the drains of the switches, e.g., at node X? Switch S_4 suffers from drain junction leakage and subthreshold leakage, thereby producing a zero dc voltage at this node. We also recognize that, as V_{out} swings between zero and V_{DD}, V_X must follow it. Since V_X has a zero average, it must swing above zero and *below* zero [Fig. 3.43(b)], turning on the drain-substrate junction of S_4 [Fig. 3.43(c)]. That is, the capacitors do load the output for part of the period even though the switches are off. For this reason, we prefer the topology of Fig. 3.42(a), at least for ring oscillators.

[12] Strictly speaking, we must scale the switch widths according to their corresponding capacitors, but the effect on the characteristics is minimal here.

3.7 Problem of Supply Noise

In the 2-GHz ring oscillator designs described in Section 3.3, we noted the high supply sensitivity, around 4.5 to 5 GHz/V. This means that the circuits convert deterministic perturbations or random noise on V_{DD} to sidebands (spurs) or phase noise, respectively. For example, as illustrated in Fig. 3.44(a), a microprocessor draws a large, data-dependent current from the supply. In the presence of a parasitic inductance, L_B, in the

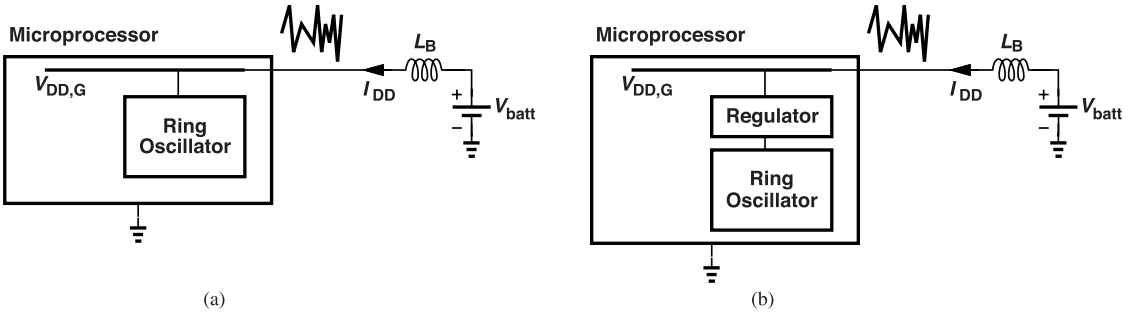


Figure 3.44 (a) Supply noise due to transient currents, and (b) use of a regulator to reduce the noise.

supply path, this current change translates to a large voltage change, $L_B dI_{DD}/dt$, thereby disturbing the oscillator.

To alleviate this issue, we often interpose a regulator between the global on-chip supply, $V_{DD,G}$, and the ring oscillator, as shown in Fig. 3.44(b). In this section, we study two examples of regulation. We should remark, however, that the regulator output may contain substantial flicker noise, thus modulating the ring and producing phase noise. Furthermore, the regulator's output voltage is lower than the global supply,[13] leaving less headroom for the ring.

3.7.1 Voltage Regulation

Figure 3.45(a) depicts an example of a "low-dropout" (LDO) regulator feeding a ring oscillator. A

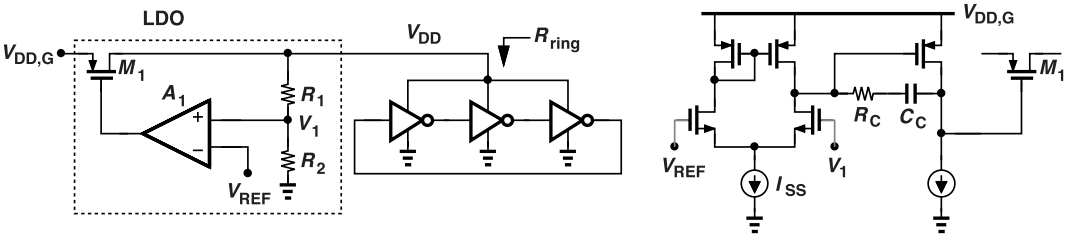


Figure 3.45 (a) Simple LDO implementation, and (b) realization of op amp.

high-gain servo loop consisting of M_1, the resistor divider, and A_1 ensures that $V_1 \approx V_{REF}$ and hence $V_{DD} \approx V_{REF}(1 + R_1/R_2)$. That is, despite variations in $V_{DD,G}$, the oscillator supply, V_{DD}, remains regulated. Operating in the deep-triode region, M_1 acts as a voltage-dependent resistor whose value is controlled by A_1. To minimize the "dropout," $V_{DD,G} - V_{DD} = |V_{DS}|$, we choose a wide transistor here. The reference voltage, V_{REF}, can be generated using a bandgap circuit. The transconductance of M_1 in the triode region is given by $g_m = \mu_p C_{ox}(W/L)|V_{DS}|$ (why?), a low value because we wish to keep $|V_{DS}|$ small, e.g., around

[13] Unless we use a "switching" regulator, in which case an off-chip inductor is necessary.

50 mV. The loop gain associated with the the feedback network is equal to

$$\text{Loop Gain} = \frac{R_2}{R_1 + R_2} A_1 g_m [(R_1 + R_2)||R_{ring}], \tag{3.43}$$

where R_{ring} denotes the small-signal resistance seen at the supply pin of the ring oscillator. This resistance is computed in Problem 3.17. If $R_2/(R_1 + R_2) \approx 0.5$ and, as an example, $g_m[(R_1 + R_2)||R_{ring}] \approx 0.2$, then A_1 must reach several hundred to ensure an accurate value for V_{DD}. Alternatively, M_1 can operate in saturation so as to provide a higher transconductance, relaxing the gain required of the op amp. In this case, M_1 must be very wide to consume only a small headroom.

Example 3.21

Can we choose $R_1 = 0$ and $R_2 = \infty$ in Fig. 3.45(a) to raise the loop gain?

Solution

If we do, then V_{REF} must be chosen equal to the desired V_{DD} and A_1 must accommodate an input common-mode level equal to V_{DD}. These requirements become problematic in the design of the V_{REF} generator and A_1.

Figure 3.45(b) shows a possible realization of A_1, where a simple two-stage op amp provides a gain in the range of 100-200. The circuit must incorporate wide, long transistors to exhibit low flicker noise and a high voltage gain. The series branch consisting of R_C and C_C performs frequency compensation. One can instead consider operating M_1 in the saturation region and eliminating the second stage of the op amp, thereby improving the high-frequency regulation.

Example 3.22

Explain how the voltage regulator of Fig. 3.45 behaves at high noise frequencies entering from $V_{DD,G}$.

Solution

Since the gain of the amplifier falls at high frequencies, the feedback loop has less ability to keep V_{DD} constant in the presence of fast fluctuations in $V_{DD,G}$. For this reason, we typically tie a capacitor from V_{DD} to ground so that it manifests itself as the loop gain begins to drop.

The LDO output noise voltage in Fig. 3.45(a) is also of interest. In Problem 3.15, we prove that, if the loop gain is large, the gate-referred noise voltage of M_1 is divided by $R_2 A_1/(R_2 + R_1)$ as it travels to V_{DD}. The op amp's input-referred noise is simply added to V_{REF} and hence multiplied by $1 + R_1/R_2$.

3.7.2 Current Regulation

Another approach to isolating ring oscillators from the supply noise is to feed the circuit by a *current source* rather than a voltage source. Illustrated in Fig. 3.46, the idea is to interpose a current source between the global supply, $V_{DD,G}$, and V_{DD}, and ensure that I_{DD} itself does not change significantly with $V_{DD,G}$. Of course, I_{DD} consumes some voltage headroom, limiting the maximum value of V_{DD} and hence the oscillator's internal swings.

Example 3.23

Compute the value of V_{DD} in Fig. 3.46.

Solution

The average power consumed by the ring is equal to $I_{DD}V_{DD}$ and also equal to $3f_0C_{tot}V_{DD}^2$, where C_{tot}

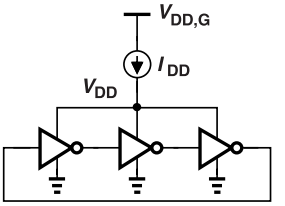


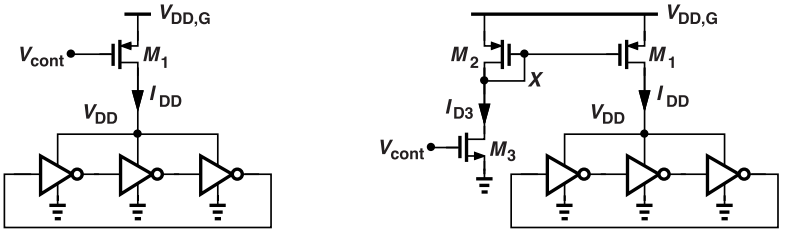
Figure 3.46 Current-controlled ring oscillator.

denotes the total capacitance at each inverter output. It follows that

$$V_{DD} = \frac{I_{DD}}{3f_0 C_{tot}}. \tag{3.44}$$

As expected, if I_{DD} changes, so does f_0. To see this, we invoke proof by contradiction: if we reduce I_{DD} and f_0 remains constant, then V_{DD} must fall. But we know that the delay of the inverters then increases, concluding that f_0 must decrease. In practice, f_0 and, to some extent, V_{DD}, change with I_{DD}.

The dependence of f_0 upon I_{DD}, as implied by Eq. (3.44), suggests that the oscillation frequency can be tuned by adjusting this current. Figure 3.47(a) depicts a simple implementation, where M_1 operates in the



(a) (b)

Figure 3.47 (a) VCO using current control, and (b) realization by a current mirror.

saturation region. We can readily identify several serious issues here: (1) noise on $V_{DD,G}$ modulates V_{GS1} and I_{DD} unless V_{cont} tracks this noise, (2) the $1/f$ and thermal noise of M_1 directly translates to phase noise, (3) V_{cont} cannot reach $V_{DD,G}$ as it would turn off M_1, and (4) V_{cont} cannot reach ground as it would drive M_1 into the triode region (why?), making I_{DD} supply-dependent. The last two issues severely limit the control voltage range, dictating a large K_{VCO}.

We now consider the modified topology shown in Fig. 3.47(b). Which one of the above issues is resolved here? Noise on $V_{DD,G}$ has negligible effect on I_{DD} if the channel-length modulation of M_1 and M_3 is small: as $V_{DD,G}$ rises, for example, so does V_X, but I_{D3} remains constant; similarly, if $|V_{DS1}|$ increases, I_{D1} does not change much. We thus prefer long devices here. Unfortunately, however, all three transistors contribute noise, and V_{cont} has both a lower and an upper bound. In Problem 3.18, we explain why we cannot filter the noise of M_1-M_3 by tying a capacitor from V_{DD} to ground. For these reasons, the regulator of Fig. 3.45 is generally superior, albeit at the cost of an op amp.

Example 3.24

Determine the total noise current in I_{DD} in Fig. 3.47(b). Neglect channel-length modulation.

Solution

The flicker and thermal noise voltage contributed by M_2 and M_3 at node X is expressed as

$$\overline{V_{n,X}^2} = \overline{V_{1/f,3}^2}\left(\frac{g_{m3}}{g_{m2}}\right)^2 + 4kT\gamma\frac{g_{m3}}{g_{m2}^2} + \overline{V_{1/f,2}^2} + \frac{4kT\gamma}{g_{m2}}. \tag{3.45}$$

This noise is multiplied by g_{m1}^2 as it appears in I_{DD}.

Example 3.25

In the VCO of Fig. 3.47(b), how do we choose I_{DD}/I_{D3}?

Solution

In order to save power, we prefer a large ratio, in the range of 5 to 10. However, the noise currents of M_2 and M_3 are also multiplied by this ratio as they travel to I_{DD}. For this reason, we must often limit this ratio to unity or even *less*.

We have observed that the regulation of the supply voltage or current introduces flicker noise, raising oscillators' phase noise at low offset frequencies. In some applications, this issue becomes particularly troublesome.

References

[1]. A. Homayoun and B. Razavi, "Relation between delay line phase noise and ring oscillator phase noise," *IEEE J. Solid-State Circuits*, vol. 49, pp. 384-391, Feb. 2014.

[2]. A. Homayoun and B. Razavi, "Analysis of phase noise in phase/frequency detectors," *IEEE Trans. Circuits and Systems - Part I*, vol. 60, pp. 529-539, Mar. 2013.

Problems

3.1. Suppose the number of stages in a ring oscillator is approximately doubled. Explain exactly what happens to the oscillation frequency, the power consumption, and the phase noise values given by Eqs. (3.14) and (3.15).

3.2. In Fig. 3.48, the corresponding nodes in two identical rings are shorted to each other. Explain what happens to the

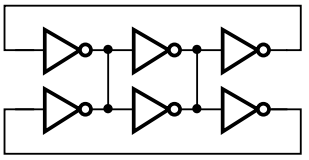


Figure 3.48 Two ring oscillators connected in parallel.

oscillation frequency, power consumption, and phase noise. [This is an alternative view of the linear scaling method depicted in Fig. 3.17(a).]

3.3. The widths of the PMOS and NMOS transistors in a ring oscillator are doubled. Explain how the plots in Fig. 3.19 change.

3.4. Consider the 22.6-GHz reference oscillator in Section 3.4.2 and suppose we break its loop to form a delay line. Determine the phase noise of this delay line at input frequency of 22.6 GHz at 100-kHz, 10-MHz, and 100-MHz offsets.

3.5. Repeat the above problem for the first 2-GHz oscillator in Section 3.4.3.

3.6. We must design a ring oscillator for $f_0 = 5$ GHz with a phase noise of -100 dBc/Hz at 10-MHz offset. If $V_{DD} = 1$ V and $|V_{TH}| = 0.3$ V, what is the value of I_D in Eq. (3.16)?

3.7. The transistor capacitances do not directly appear in Eqs. (3.14) and (3.15). Suppose we double the capacitance seen at each node of a ring oscillator while other parameters remain unchanged. Explain what happens to the phase noise and the power consumption.

3.8. We wish to reduce the oscillation frequency of a ring oscillator by adding equal capacitances from its nodes to ground. Is there an upper bound on the capacitance value? In other words, does the oscillator fail if the capacitance is very large?

3.9. Consider the ring shown in Fig. 3.49. If C_1 is increased, is it possible the oscillator fails? Explain.

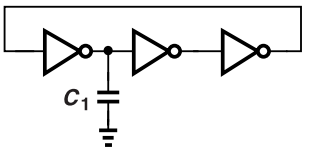


Figure 3.49 Ring oscillator loaded at one node.

3.10. Repeat the above problem for the circuit shown in Fig. 3.50.

3.11. Shown in Fig. 3.51 is a ring oscillator employing feedback capacitors. Does the circuit fail to oscillate if C_F is very large?

3.12. Consider the ring oscillator in Fig. 3.51. Student A surmises that the Miller effect of the capacitors at the input of each inverter is simply given by $2C_F$ because, when the left plate of C_F swings from 0 to V_{DD}, the right plate falls from V_{DD} to 0. Student B disagrees and notes that the voltage waveforms at the input and output of each inverter are 120° apart and hence C_F must be multiplied by a *complex* factor. Who is right?

3.13. The op amp in Fig. 3.45(a) exhibits an input-referred two-sided noise spectrum $\overline{V_n^2} = \eta\ \mathrm{V}^2/\sqrt{\mathrm{Hz}}$. If we denote the supply sensitivity of the ring oscillator by $K_{VCO} = \partial f_{out}/\partial V_{DD}$, determine the resulting phase noise.

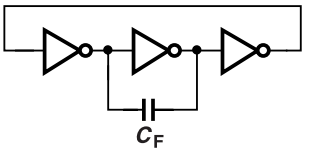


Figure 3.50 Ring oscillator loaded by a Miller capacitor.

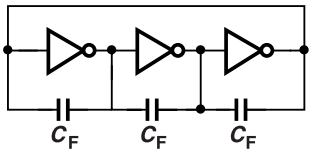


Figure 3.51 Ring oscillator loaded by Miller capacitors.

3.14. Repeat the above problem if transistor M_1 in Fig. 3.45(a) operates in saturation and has a gate-referred two-sided noise spectrum $\overline{V_n^2} = 2kT\gamma g_m$.

3.15. Prove that the gate-referred noise voltage of M_1 in Fig. 3.45(a) is divided by $R_2A_1/(R_2 + R_1)$ as it reaches V_{DD} if the loop gain is large.

3.16. We define a gain from I_{DD} to f_{out} in Fig. 3.47(a) as $K = \partial f_{out}/\partial I_{DD}$. We model the two-sided noise spectrum of M_1 by a gate-referred voltage according to

$$\overline{V_n^2} = 2kT\gamma g_m + \frac{1}{2}\frac{K_F}{WLC_{ox}f}. \tag{3.46}$$

Determine the resulting phase noise.

3.17. The power consumption of a CMOS inverter can be calculated by an interesting method. First, consider the arrangement shown in Fig. 3.52(a), where C_1 is periodically (and completely) charged to V_{in} and discharged to ground. Noting that the charge taken from V_{in} in one clock cycle is equal to C_1V_{in}, prove that the network in the dashed box

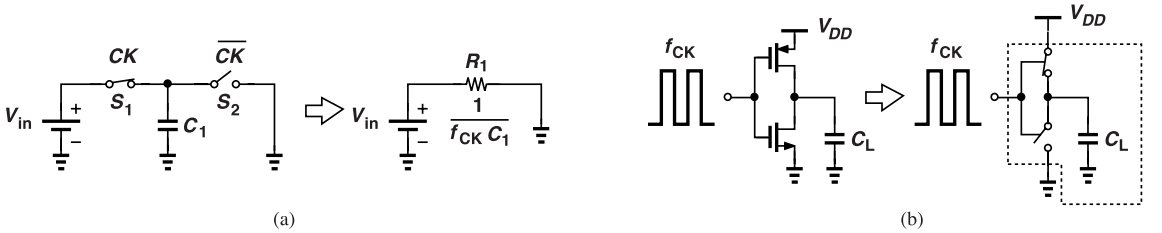


Figure 3.52 (a) Equivalence of a switched capacitor to a resistor, and (b) use of (a) to find power consumption of an inverter.

has an "average" resistance equal to $1/(f_{CK}C_1)$, where f_{CK} denotes the clock frequency. Now, let us study the equivalent inverter circuit shown in Fig. 3.52(b), recognizing that the dashed box acts as an average resistor equal to $1/(f_{CK}C_L)$ tied between V_{DD} and ground. Calculate the power consumed by the dashed box.

3.18. Suppose a large capacitor is tied between V_{DD} and the ground in Fig. 3.47(b) so as to filter the noise of M_1-M_3. The pole created by this capacitor appears in the control path of the oscillator. That is, we can no longer express the transfer function as K_{VCO}/s. This additional pole can degrade the stability of phase-locked loops. Neglecting channel-length modulation, estimate the pole frequency by first computing the equivalent resistance seen at V_{DD}. (Hint: use the results of the previous problem.)

3.19. Assuming square-law devices in Fig. 3.47(b) and neglecting channel-length modulation, determine the range of V_{cont} so that M_3 remains in saturation.

3.20. For the oscillator shown in Fig. 3.47(b), sketch the oscillation frequency as a function of V_{cont}.

4

Design of Differential and Multiphase Ring Oscillators

The inverter-based rings studied in Chapter 3 suffer from poor supply rejection. We therefore explore in this chapter differential rings as an alternative that potentially exhibits much less sensitivity to the supply. Our design work parallels that in Chapter 3: we begin with a reference topology and, through simulations, evaluate its performance. Next, we modify the design so as to obtain a desired frequency of 2 GHz, and then introduce methods of tuning the frequency. We also study rings providing multiple evenly-spaced output phases. The reader is encouraged to review Chapter 1 before proceeding with this chapter.

4.1 General Considerations

The problem of supply noise can be greatly alleviated through the use of differential rings. The basic design proceeds as outlined for single-ended rings described in Chapter 3. For example, the three-stage topology shown in Fig. 4.1 must provide, in each stage, a frequency-dependent phase shift of 60° and a low-frequency

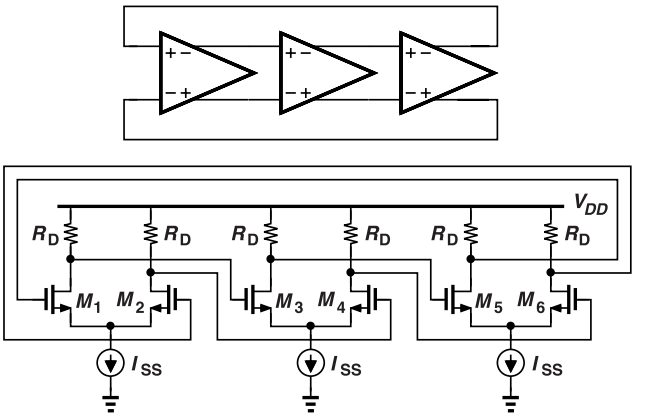


Figure 4.1 Three-stage differential ring oscillator.

voltage gain of 2 (Chapter 1). Note that negative feedback at low frequencies ensures the loop does not latch up. The circuit generates six output phases in 60° steps.

The primary supply dependence in the circuit of Fig. 4.1 arises from the drain-substrate junction capacitances of the transistors. If V_{DD} fluctuates, so does the common-mode level at each output and hence these capacitances. This effect is studied in Problem 4.1. In contrast to inverters, the small-signal *resistance* seen at the drains is relatively independent of V_{DD} here. The supply sensitivity is therefore much less than that of single-ended rings.

Example 4.1

A student reasons that the tail current sources in Fig. 4.1 are not necessary and reduces the circuit to that shown in Fig. 4.2(a) so as to save voltage headroom. Explain how this arrangement operates.

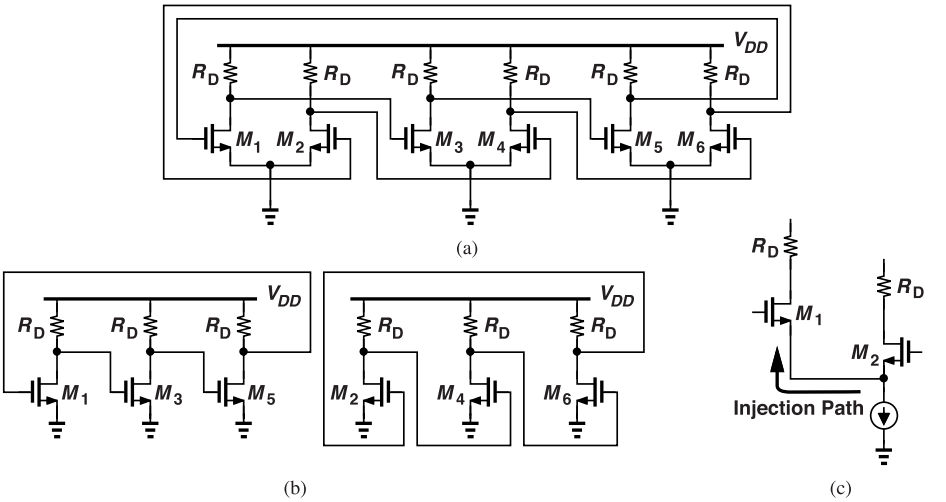


Figure 4.2 (a) Ring oscillator without tail current sources, (b) equivalent circuit showing two single-ended loops, and (c) illustration of mutual injection in the presence of tail current source.

Solution

We can identify two *independent*, single-ended ring oscillators here [Fig. 4.2(b)]. Each ring then operates according to its own frequency and initial phase. In other words, if the tail current sources in Fig. 4.1 are replaced with resistors and the value of these resistors is decreased, at some point differential operation ceases and two single-ended rings appear. In more advanced studies, we say that the differential ring with tail current sources can be viewed as two single-ended loops that are "injection-locked" to each other through the source terminals of their transistors [Fig. 4.2(c)]. In the presence of the tail current source, M_1 acts as a source follower and M_2 as a common-gate stage, and vice versa, allowing each oscillator to inject a current into the other. For the two single-ended loops to synchronize to each other and operate differentially, the injection must have enough "strength," hence the need for a high-impedance tail device. (If the tail current source is replaced by a small resistance, the injection from one oscillator is mostly shunted to ground.)

The second advantage of differential rings over their single-ended counterparts is the ease with which they can generate multiple desired output phases. For example, a four-stage loop yields eight phases with $45°$ separations (Fig. 4.3). As explained in Section 4.9, such a phase spacing is more difficult to obtain using single-ended rings.

The third advantage of differential topologies is their higher speed. The delay per stage in Fig. 4.1 is smaller than that of a CMOS inverter delay for two reasons: (1) the single-ended voltage swing, $I_{SS}R_D$, is less, e.g., around 300 mV, requiring less transition time, and (2) compared to a PMOS transistor in an inverter, the simple resistive load contributes no capacitance to the input and less capacitance to the output.

One disadvantage of differential rings is that they do not provide rail-to-rail voltage swings, necessitating a level converter (an amplifier). A common method is to use self-biased inverters that sense the outputs capacitively (Fig. 4.4), but at the cost of greater power consumption.

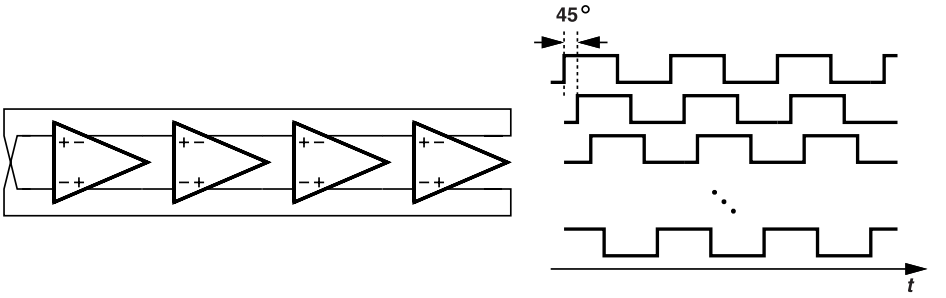


Figure 4.3 Four-stage differential ring oscillator and its waveforms.

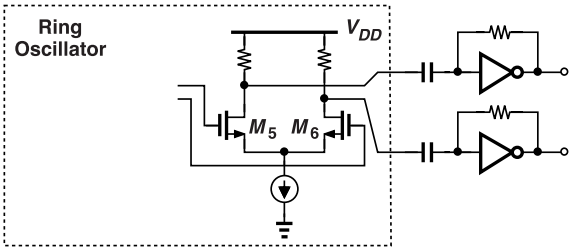


Figure 4.4 Generating rail-to-rail swings by means of self-biased inverters.

4.2 Phase Noise Considerations

The principal drawback of differential ring oscillators is their higher phase noise. The relation expressed by Eq. (3.13) still applies, but we turn to published results: [1] expresses the phase noise in the thermal regime in a form that can be reduced to

$$S_{\phi n}(f) = \frac{8kT}{3I_{SS}} \left(\frac{\gamma}{V_{GS} - V_{TH}} + \frac{1}{R_D I_{SS}} \right) \frac{f_0^2}{\Delta f^2}, \tag{4.1}$$

where the tail current source's noise is neglected and $V_{GS} - V_{TH}$ denotes the overdrive of the transistors when each carries $I_{SS}/2$. This equation assumes that the single-ended voltage swing is equal to $R_D I_{SS}$, which, as seen later, may not be accurate in some cases. Let us examine this equation from different angles.

Example 4.2

In a three-stage ring, how do the two terms $\gamma/(V_{GS} - V_{TH})$ and $1/(R_D I_{SS})$ in (4.1) compare? Assume $\gamma = 1$.

Solution

For start-up, we must have $g_m R_D \geq 2$. Writing g_m as $2I_D/(V_{GS} - V_{TH}) = I_{SS}/(V_{GS} - V_{TH})$, we obtain

$$\frac{1}{V_{GS} - V_{TH}} \geq \frac{2}{R_D I_{SS}}. \tag{4.2}$$

Those two terms signify the relative contributions of the transistors and the load resistors to $S_{\phi n}$.

Example 4.3

Equation (4.1) appears independent of the number of stages, N. Explain what happens as N increases.

Solution

For a fair comparison, we must keep the power consumption, the voltage swings, and the oscillation frequency constant as N increases. We therefore reduce I_{SS} and W/L and increase R_D proportionally. All of the transistor capacitances scale down, while $V_{GS} - V_{TH}$ and $R_D I_{SS}$ remain unchanged. Thus, the phase noise rises in proportion to $8kT/(3I_{SS})$. For example, if N is doubled and I_{SS} is halved, $S_{\phi n}$ rises by 3 dB.

The foregoing example implies that the best phase noise-power trade-off occurs if the number of stages is *minimized*. As we will see in Section 4.3, this is not quite correct because the voltage swings increase as N goes from 3 to 4 to 5.

The flicker-noise-induced phase noise is more complex. As discussed in Chapter 6, the flicker noise of the differential pair transistors contributes to phase noise if the single-ended drain voltage waveforms do not have symmetric rise and fall transitions [1]. Illustrated in Fig. 4.5, this scenario can be viewed as unequal slopes for the rising and falling edges, especially at the crossing points.

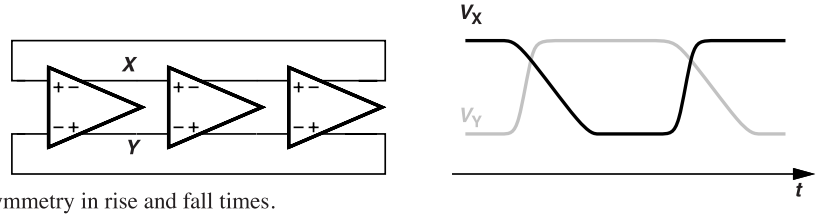


Figure 4.5 Asymmetry in rise and fall times.

Example 4.4

Explain how the presence of the second harmonic in V_X and V_Y leads to unequal slopes.

Solution

Suppose $V_X(t) = V_1 \sin \omega_0 t + V_2 \sin(2\omega_0 t) + V_{CM}$. As shown in Fig. 4.6, the second harmonic speeds up the rising edges and slows down the falling edges. We observe that the slope around V_{CM} alternates between $V_1\omega_0 + 2V_2\omega_0$ and $-V_1\omega_0 + 2V_2\omega_0$. We will use the ratio of the first and second harmonic amplitudes as a measure of the waveform asymmetry.

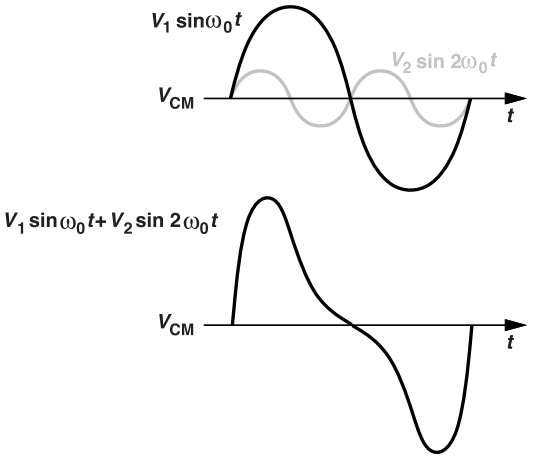


Figure 4.6 Asymmtery due to the second harmonic.

Let us examine the origins of the second harmonic. Considering the differential pair in Fig. 4.7(a), we can identify a mechanism related to V_P and the tail capacitance, C_T. As depicted in Fig. 4.7(b), V_P drops

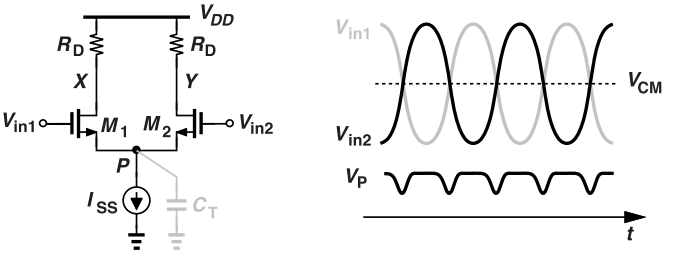


Figure 4.7 (a) Differential pair with tail capacitance, and (b) its waveforms.

around the crossing points of V_{in1} and V_{in2}, exhibiting a frequency of $2\omega_0$. From another perspective, each time V_{in1} reaches its maximum value, M_2 turns off and M_1 acts as a source follower, lifting V_P. Thus, V_P is lifted twice per cycle. If we assume $V_{in1} = V_{CM} + V_0 \sin\omega_0 t$ and $V_{in2} = V_{CM} - V_0 \sin\omega_0 t$ and neglect the current drawn by C_T, we can prove that [1]

$$V_P = V_{CM} - V_{TH} - \sqrt{(V_{GS} - V_{TH})^2 - V_0^2 \sin^2 \omega_0 t}, \tag{4.3}$$

where $V_{GS} - V_{TH}$ denotes the equilibrium overdrive. Note that this equation is valid only if both transistors are on, whereas in a ring oscillator we choose V_0 large enough to ensure that they turn on and off. Nevertheless, this result provides some insight if we assume $V_0^2 \sin^2 \omega_0 t \ll (V_{GS} - V_{TH})^2$, factor $(V_{GS} - V_{TH})^2$, assume $\sqrt{1+\epsilon} \approx \epsilon/2$, write $\sin^2 \omega_0 t = (1 - \cos 2\omega_0 t)/2$, and hence obtain

$$V_P \approx V_{CM} - V_{TH} - (V_{GS} - V_{TH}) + \frac{V_0^2[1 - \cos(2\omega_0 t)]}{4(V_{GS} - V_{TH})}. \tag{4.4}$$

The voltage component $[-V_0^2 \cos(2\omega_0 t)]/[4(V_{GS} - V_{TH})]$ generates in C_T a current, $C_T dV/dt$, of the form $[V_0^2 C_T \omega_0 \sin(2\omega_0 t)]/[2(V_{GS} - V_{TH})]$, which flows from X and Y, thereby altering the positive and negative slopes in V_X and V_Y. This phenomenon is studied from another perspective in Chapter 6. An interesting point here is that, if we increase the width of the transistors in the ring oscillator, (a) V_0 rises (Section 4.4), (b) C_T increases, and (c) $V_{GS} - V_{TH}$ falls, all causing the second harmonic to rise.

If the ratio of the positive and negative slopes is denoted by α, e.g., $\alpha = (V_1 + 2V_2)/(V_1 - 2V_2)$ in Example 4.4, then it can be shown that the flicker noise corner frequency of the phase noise profile is given by [1]

$$f_{1/f^3} \approx f_{1/f} \frac{3}{2N} \frac{(1-\alpha)^2}{1 - \alpha + \alpha^2}, \tag{4.5}$$

where $f_{1/f}$ is the flicker noise corner frequency of the transistors themselves.

Example 4.5

A student reasons that transistor M_1 in Fig. 4.7(a) turns on and off at a rate of f_0, operating as a mixer. If the second harmonic current drawn by the tail is mixed with f_0, it produces only $2f_0 + f_0 = 3f_0$ and $2f_0 - f_0 = f_0$. Thus, the tail-induced second harmonic does not create asymmetry in the rise and fall transitions. Explain the flaw in this argument.

Solution

Acting as a mixer, M_1 multiplies the tail current by a square wave, $S(t)$, toggling between 0 and 1 (Fig. 4.8).

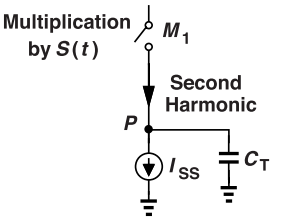


Figure 4.8 Multiplication of tail current by $S(t)$ as a result of switching.

Considering the Fourier series of $S(t)$, we recognize that it has a dc level of 0.5, by which the tail-induced second harmonic is multiplied: $0.5 \times \sin 2\omega_0 t$. Consequently, I_{D1} does carry half of this harmonic. That is, I_{D1} contains components at $f_0, 2f_0, 3f_0$, etc.

We should also consider the flicker noise of the tail current sources. Recall from Chapter 1 that the oscillation frequency of N-stage differential rings is approximately equal to $\omega_0 = (R_D C_L)^{-1} \tan(180°/N)$. Since this value is independent of the tail current, we expect that fluctuations in this current do not contribute significant frequency or phase modulation. However, if C_L contains a nonlinear component, the tail flicker noise can translate to phase noise. This effect is studied in the context of LC oscillators in Chapter 6.

4.3 Basic Differential Ring Design

4.3.1 Initial Design

Let us design a three-stage differential ring oscillator and examine its behavior. As with the developments in Section 3.2, we begin with relatively small transistors, bearing in mind that linear scaling (Section 3.1) eventually may be necessary to reduce the phase noise. The choice of the tail current, the load resistance, and the transistor dimensions is dictated by the following requirements: (1) the small-signal low-frequency voltage gain of one stage must be at least 2 to guarantee start-up (Chapter 1); (2) the voltage swings must be large enough to reduce the phase noise; we aim for a single-ended peak-to-peak swing of about 300 mV for now; and (3) it is desirable that the differential pairs steer their tail current almost completely so as to maximize the voltage swings for a given I_{SS} (and hence for a given power dissipation); in Fig. 4.7(a), the swing at X or Y is equal to $I_{SS}R_D$ if the transistors switch completely. Unfortunately, this is not possible for only three stages in the loop, as shown in the following example.

Example 4.6

Prove that the single-ended voltage swings in the three-stage ring of Fig. 4.7(a) cannot reach $I_{SS}R_D$ even if the tail currents are steered completely.

Solution

Consider the stage shown in Fig. 4.9(a). We expect the output waveforms to resemble those in Fig. 4.9(b), swinging between V_{min} and V_{max}. Using the small-signal estimate given by Eq. (1.22) for the oscillation frequency, we write the period as

$$T_0 = \frac{1}{f_0} = \frac{2\pi}{\sqrt{3}} R_D C_L. \tag{4.6}$$

As M_1 turns on, we model the left side of the circuit as depicted in Fig. 4.9(c), and approximate $I_1(t)$ by a

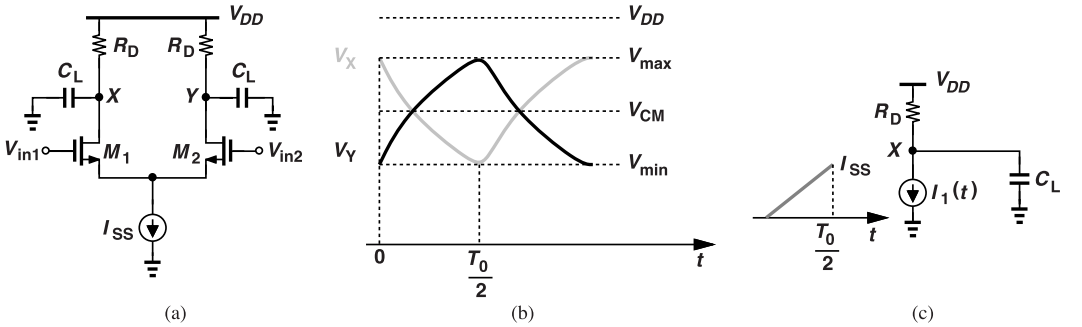


Figure 4.9 (a) Differential pair with load capacitance, (b) approximate waveforms, and (c) simplified model.

ramp going from 0 to I_{SS} in $T_0/2$ seconds:

$$I_1(t) = I_{SS}\frac{t}{T_0/2}u(t). \tag{4.7}$$

The ramp response of the circuit can be obtained by integrating the step response with respect to time:

$$V_X(t) = V_{max} - \frac{R_D I_{SS}}{T_0/2}\left(t + R_D C_L \exp\frac{-t}{R_D C_L} - R_D C_L\right), \tag{4.8}$$

where $-R_D C_L$ is an integration constant that ensures $V_X(0) = V_{max}$. At $t = T_0/2 = (\pi/\sqrt{3})R_D C_L = 1.8 R_D C_L$, we have $V_{min} = V_{max} - 0.54 R_D I_{SS}$. In other words, the peak-to-peak single-ended voltage swing is roughly equal to $0.54 R_D I_{SS}$. This is because the delay through the loop is too short to allow V_X and V_Y to reach V_{DD}. Note that V_{CM} is still equal to $V_{DD} - R_D I_{SS}/2$ (why?). Of course, as the number of stages in the ring increases, the swing approaches $R_D I_{SS}$.

Is our small-signal estimate of T_0 justified here? Yes, it is, because Eq. (1.22) is based on the 60° phase shift criterion rather than the small-signal gain requirement. As we increase the tail current and depart from the small-signal regime, we expect that $\omega_0 = \sqrt{3}/(R_D C_L)$ still holds approximately because the phase shift is equal to 60° even in large-signal operation.

We continue our design effort, selecting $I_{SS} = 100\ \mu$A and $R_D = 6\ \text{k}\Omega$ for a single-ended swing of approximately $0.54 I_{SS} R_D \approx 320$ mV. For the start-up condition, we consider the small-signal regime and write $g_m \approx 2I_D/(V_{GS} - V_{TH}) = I_{SS}/(V_{GS} - V_{TH})$, and enforce the condition $g_m R_D \geq 2$. If $W/L = 400$ nm/40 nm, this condition is satisfied, leading to the design shown in Fig. 4.10(a). As explained in Section 3.2, we simulate the circuit in the slow-slow corner at 75 °C and with $V_{DD} = 0.95$ V. We also estimate an interconnect capacitance of 0.6 fF,[1] and use ideal tail current sources for now. The circuit's simulated waveforms are plotted in Fig. 4.10(b).

These waveforms provide a wealth of information about the oscillator. The single-ended voltage swings are about 150 mV$= 0.25 R_D I_{SS}$ due to the short delay around the loop *and* incomplete steering of the tail current. (In fact, I_{D1} and I_{D2} swing between 30 μA and 70 μA.) We observe an oscillation frequency of 42 GHz, around twice that of the inverter-based ring in Section 3.2. But we also note that $V_P \approx 80$ mV because the small transistor widths result in a large overdrive, $V_{GS} - V_{TH} \approx 0.3$ V, leaving little headroom for the tail current source. We address these issues in the next section.

[1] This capacitance is greater than that in the basic inverter-based design in Section 3.2 because the 6-kΩ resistors occupy a large area, necessitating long interconnects between the stages.

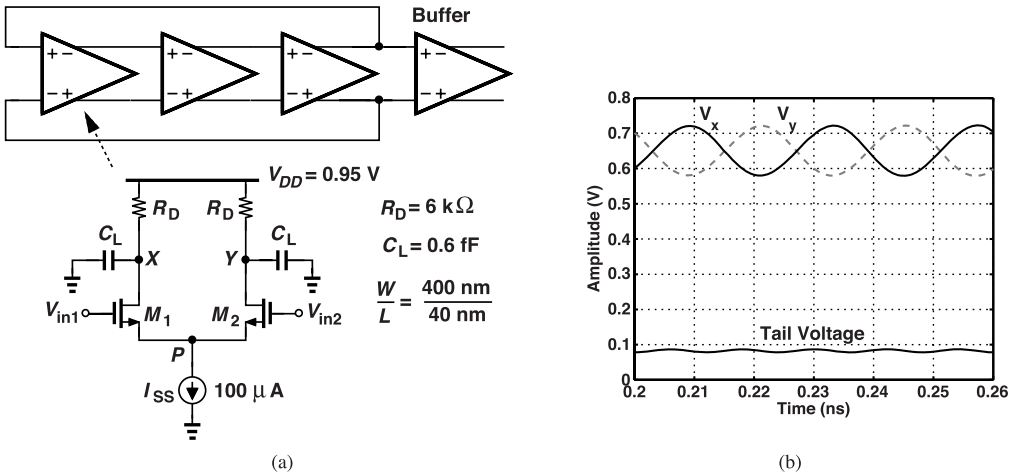


Figure 4.10 (a) Three-stage differential ring design, and (b) its waveforms.

Figure 4.11 plots the phase noise, revealing flicker noise upconversion for offsets below approximately 1 MHz. (The phase noise drops by about 26 dB from 100 kHz to 1 MHz.) Let us examine the phase noise at 100 MHz, which is predominantly produced by thermal noise, in view of Eq. (4.1). With $T = 348$ K (75 °C), $I_{SS} = 100$ μA, and $V_{GS} - V_{TH} = 300$ mV, Eq. (4.1) yields $S_{\phi n}(f = 100\text{ MHz}) \approx -99.5$ dBc/Hz, about 9.5 dB lower than observed in Fig. 4.11. This discrepancy arises primarily because (4.1) assumes complete

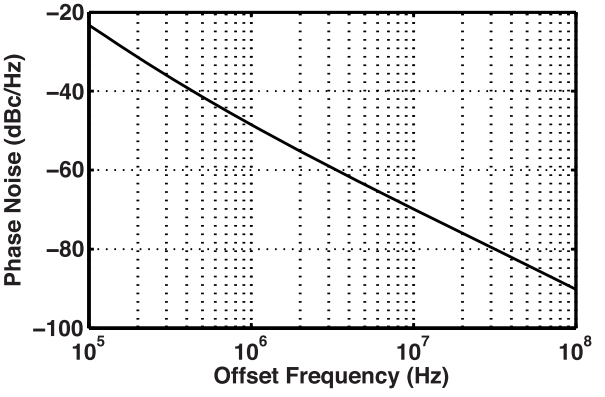


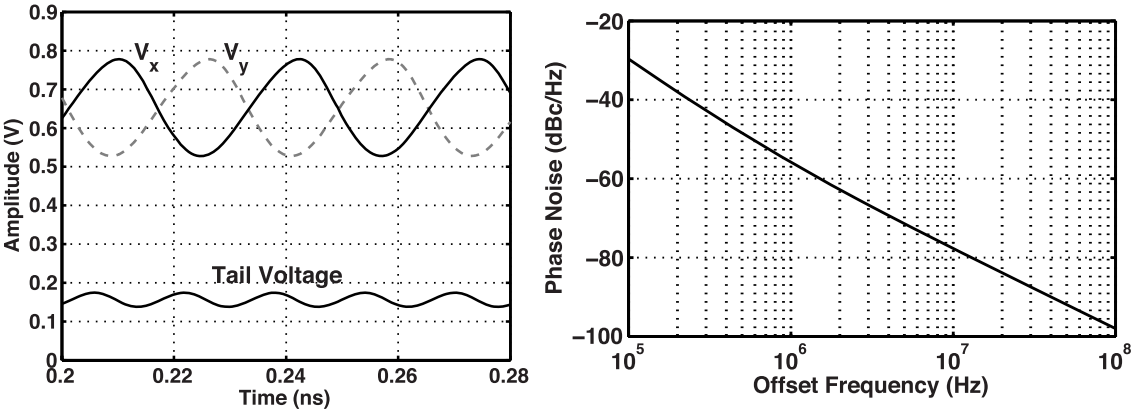
Figure 4.11 Phase noise of design in Fig. 4.10.

current switching and hence a single-ended voltage swing of $I_{SS}R_D$. We also note that the small voltage swings and the large $V_{GS} - V_{TH}$ lead to low $1/f$ noise upconversion. (The second harmonic in the single-ended outputs is at -37 dB.)

It is instructive to compare the performance of this design to that of the 22.6-GHz single-ended ring in Section 3.2. We note that the latter exhibited flicker noise upconversion for offsets as high as 100 MHz. We summarize our results as follows: (1) for the differential ring: $f_0 = 42$ GHz, $S_{\phi n}(100\text{ MHz}) = -90$ dBc/Hz, $P = 3 \times 100\ \mu\text{A} \times 0.95\text{ V} = 285\ \mu$W; and (2) for the inverter-based ring: $f_0 = 22.6$ GHz, $S_{\phi n}(100\text{ MHz}) = -104$ dBc/Hz, $P = 60\ \mu\text{A} \times 0.95\text{ V} = 57\ \mu$W. Thus, the figure of merit, $f_0^2/(\Delta f^2 \cdot P \cdot S_{\phi n})$, is about 15.6 dB higher (better) for the latter in the thermal-noise regime. At 100-kHz offset, this advantage narrows to 7.6 dB owing to less flicker noise upconversion in the differential design.

4.3.2 Design Improvements

We must modify the differential ring to provide a greater voltage headroom for the tail current sources and also to allow more complete switching. To this end, we double the widths of the transistors, arriving at the waveforms shown in Fig. 4.12(a). The single-ended drain voltage swings increase to 250 $\mathrm{mV_{pp}}$, the drain



(a) (b)

Figure 4.12 (a) Waveforms, and (b) phase noise of design in Fig. 4.10(a) with $W/L = 800$ nm$/40$ nm.

currents swing from 17.5 μA to 87 μA, and the circuit oscillates at 31.3 GHz. Note that the tail node voltage swing has risen to about 35 $\mathrm{mV_{pp}}$, causing a higher displacement current at the second harmonic in the tail capacitance. The single-ended drain voltages now exhibit a second harmonic level of -23 dB. The tail voltage is barely sufficient for a tail current source.

Figure 4.12(b) plots the phase noise of the modified design. We expect the phase noise at 100-MHz offset to drop by $10\log(42\text{ GHz}/31.3\text{ GHz})^2 = 2.6$ dB and $10\log(250\text{ mV}/150\text{ mV})^2 = 4.4$ dB, i.e., a total of 7 dB. The actual drop is about 8 dB. With $V_{GS} - V_{TH} = 220$ mV, Eq. (4.1) predicts a phase noise of about -101 dBc/Hz at 100-MHz offset, about 4 dB lower than the simulated value.

For the flicker-noise-induced phase noise around 100-kHz offset, we guess a drop of 7 dB plus another 3 dB due to the doubling of the transistors' channel area. But we still observe only a 7-dB reduction. This is because of the higher second harmonic and the greater asymmetry between rise and fall times of V_X and V_Y.

Continuing our exploration, we again double the transistor widths, reaching the design shown in Fig. 4.13(a) and the waveforms in Fig. 4.13(b). The ring oscillates at 20.8 GHz with a single-ended swing of 270 $\mathrm{mV_{pp}}$, a value somewhat close to our prediction of $0.54 I_{SS} R_D$ in Example 4.6. The second harmonic in V_X or V_Y is around -19 dB.

Figure 4.13(c) plots the phase noise. We estimate the reduction at 100-MHz offset to be equal to $10\log(31.3\text{ GHz}/20.8\text{ GHz})^2 = 3.55$ dB plus $10\log(270\text{ mV}/250\text{ mV})^2 = 0.7$ dB. The actual reduction is 4.5 dB. With $V_{GS} - V_{TH} = 165$ mV, Eq. (4.1) gives a phase noise of -103.7 dBc/Hz, about 1.2 dB lower than the simulated value. Owing to the higher second harmonic, the phase noise at 100-kHz offset falls by 6 dB rather than by $4.25\text{ dB} + 3\text{ dB} = 7.25$ dB. Note from the waveforms in Fig. 4.13(b) that the rise and fall times are approximately equal to 21 ps and 28 ps, respectively. While it is desirable to maintain symmetry between the rising and falling transitions, the circuit does not offer such a flexibility in its design.

The 21-GHz oscillator in Fig. 4.13(a) will serve as our reference design. The supply sensitivity is about 1.82 GHz/V, a factor of 28 times lower than that of the 22.6-GHz inverter-based design in Section 3.2.

Table 4.1 summarizes the performance of the single-ended and differential reference designs developed in this and the previous chapters. The latter exhibits a much lower supply sensitivity and an 11-dB improvement in the figure of merit at 100-kHz offset, but suffering a 10-dB disadvantage at 100-MHz offset as a result of its higher power consumption.

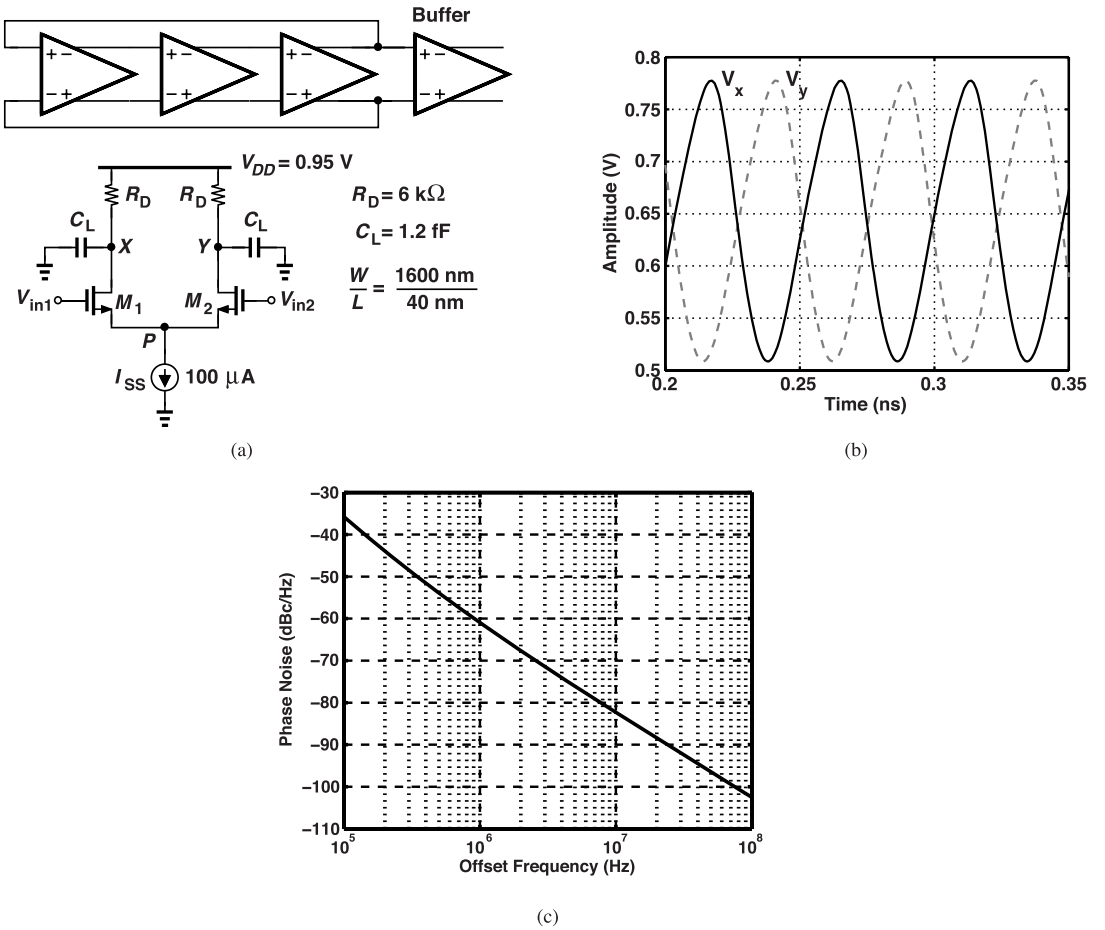


Figure 4.13 (a) Ring design of Fig. 4.10(a) but with $W/L = 1600\text{ nm}/40\text{ nm}$, (b) its waveforms, and (c) its phase noise.

$$\text{FOM} = 10\log\frac{f_0^2}{\Delta f^2 P S_{\phi n}(\Delta f)}$$

	f_0	P	K_{VDD}	Δf = 100 kHz	Δf = 100 MHz
Inverter–Based Ring	22.6 GHz	0.057 mW	50.2 GHz/V	137 dB	164 dB
Differential Ring	20.8 GHz	0.285 mW	1.82 GHz/V	148 dB	154 dB

Table 4.1 Comparison of inverter-based and differential ring oscillators.

4.4 Obtaining the Desired Frequency

In this section, we repeat the design efforts of Section 3.3 for differential ring oscillators, aiming for $f_0 \approx 2$ GHz. We have the following options:

1. Add capacitance to each node.

2. Increase the length of the transistors. The width should also be increased to ensure sufficient voltage headroom for the tail current source.
3. Increase the number of stages.

(We observed in Section 3.3 that dividing the frequency does not improve the figure of merit.)

4.4.1 Method 1: Greater Node Capacitances

With $R_D = 6$ kΩ and $f_0 = 20.8$ GHz in Fig. 4.13(a), we can estimate from $f_0 = \sqrt{3}/(2\pi R_D C_L)$ a node capacitance of 2.21 fF. For f_0 to drop to 2 GHz, we must raise C_L by about a factor of 10 [Fig. 4.14(a)]. Simulations reveal that an extrinsic capacitance of 25 fF yields $f_0 = 2.14$ GHz. The waveforms are plotted

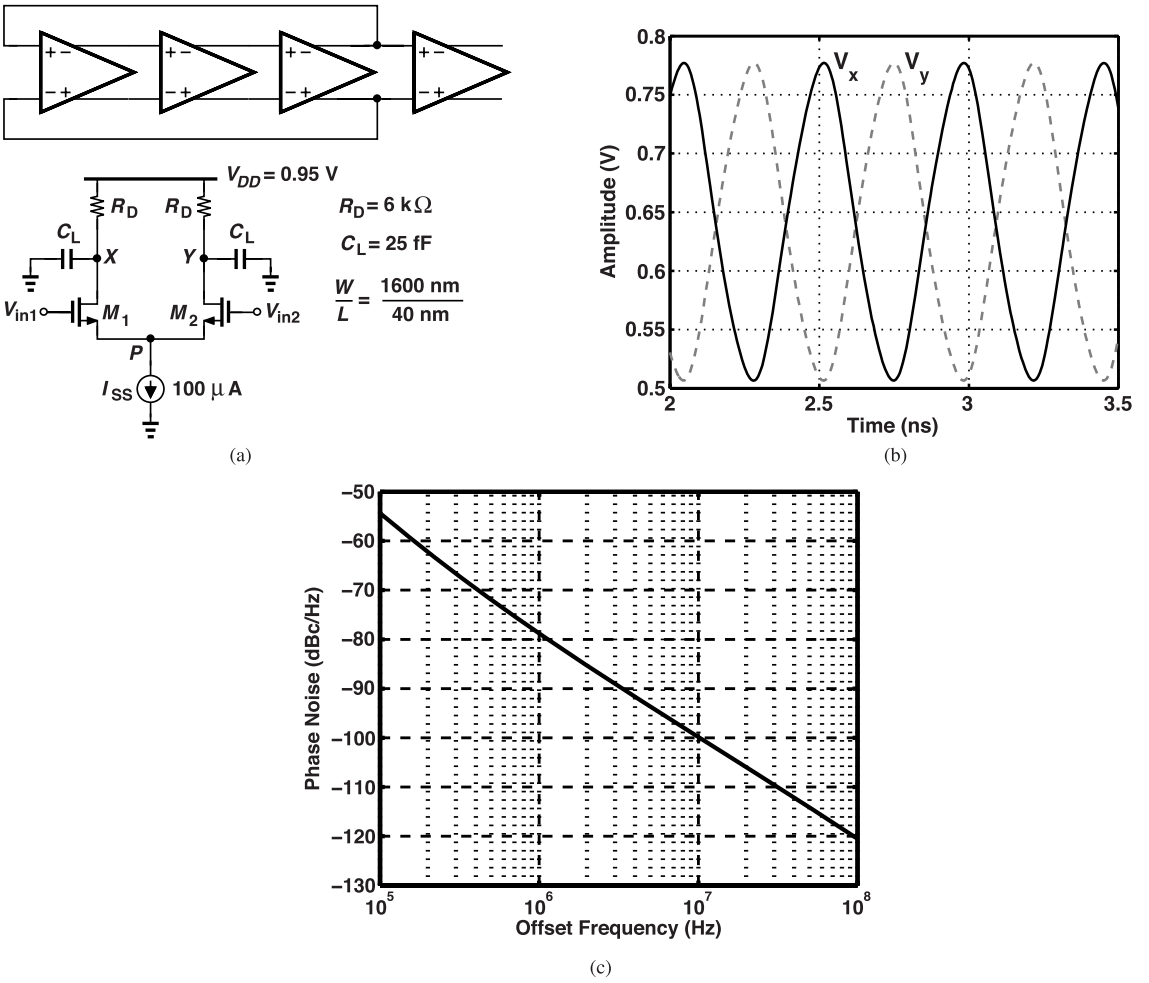


Figure 4.14 Oscillator of Fig. 4.13(a) redesigned for 2 GHz using large load capacitances, (b) its waveforms, and (c) its phase noise.

in Fig. 4.14(b) and the phase noise in Fig. 4.14(c).

Equation (4.1) implies that the phase noise in the thermal noise regime falls by $10\log(20.8\ \text{GHz}/2.14\ \text{GHz})^2 = 19.8$ dB. The actual drop is about 18 dB for both thermal and flicker noise regimes.

An important benefit that accrues from adding constant, grounded capacitances to the drains is the lower supply sensitivity, $K_{VDD} = 74$ MHz/V. This is because $C_L = 25$ fF is much greater than the transistors' voltage-dependent capacitances (C_{DB}). If normalized to the oscillation frequency, this value is 2.5 times less than that of the reference design. We formulate this improvement in Problem 4.3.

4.4.2 Method 2: Larger Transistors

We increase the transistors' dimensions from $W/L = 1600$ nm/40 nm to 8 μm/240 nm, obtaining $f_0 = 2.27$ GHz. Figure 4.15 shows the design and the simulation results. The interconnect capacitance is doubled

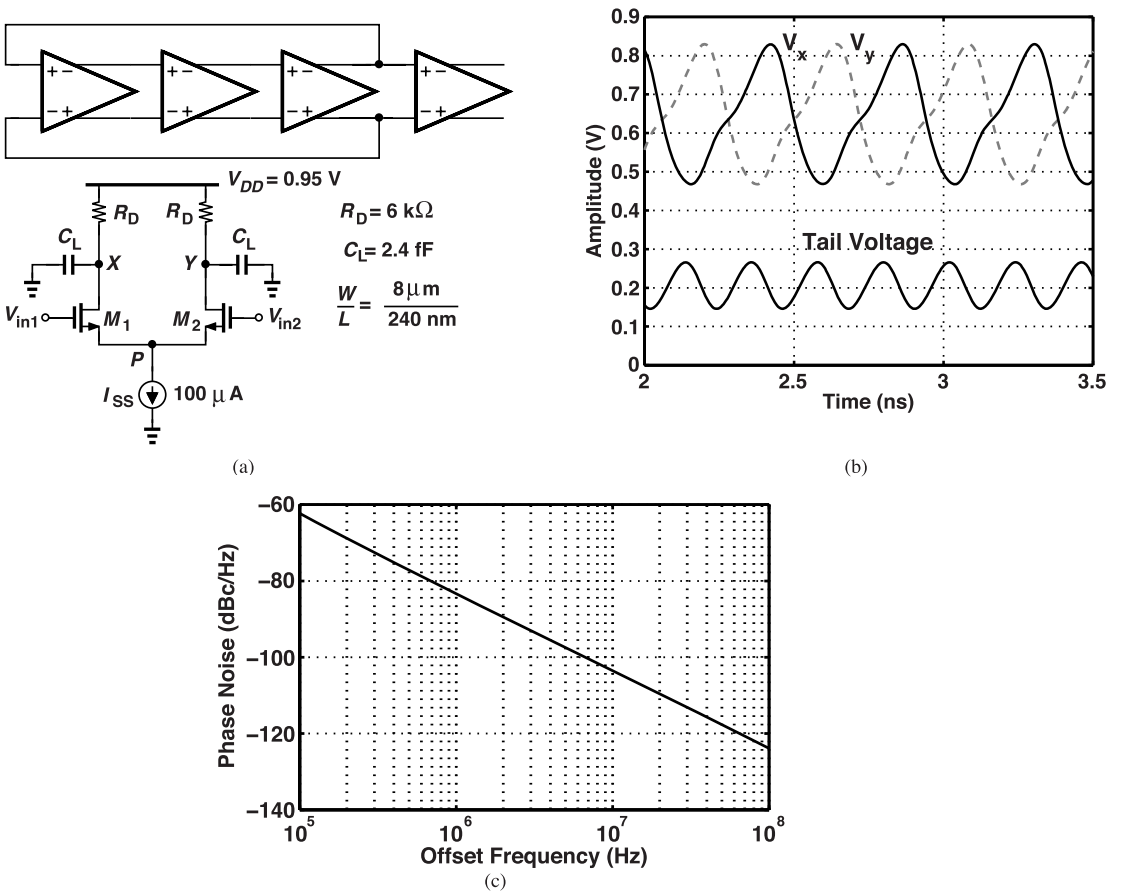


Figure 4.15 Oscillator of Fig. 4.13(a) redesigned with large transistor dimensions, (b) its waveforms, and (c) its phase noise.

to account for the larger cell size and hence longer wires.

Let us make several observations. First, the voltage swings at X and Y have increased from 270 $\mathrm{mV_{pp}}$ in the reference design to 360 $\mathrm{mV_{pp}}$, a value close to the $0.54R_DI_{SS}$ estimate found in Example 4.6. Indeed, the transistors now experience complete switching. Second, the phase noise at 100-MHz offset has fallen by 21.5 dB, that is, 2 dB more than $20\log(20.8\ \mathrm{GHz}/2.27\ \mathrm{GHz})$. This can be attributed to the larger swings, as we have $10\log(360\ \mathrm{mV}/270\ \mathrm{mV})^2 = 2.5$ dB. Third, the phase noise at 100-kHz offset has dropped by 26 dB, of which 19.2 dB is due to frequency scaling. Is the remaining 6.8 dB expected? Since the transistor

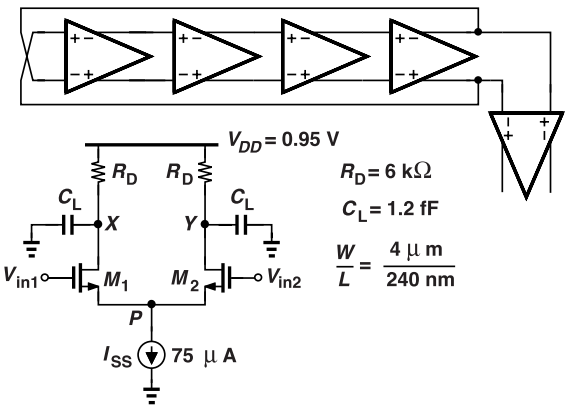
channel areas are increased by a factor of $(8000\ \text{nm} \times 236\ \text{nm})/(1600\ \text{nm} \times 36\ \text{nm}) \approx 33$,[2] we predict a reduction of $10 \log 33 = 15$ dB. But, due to the greater tail capacitance and the larger voltage swings across this capacitance, the second harmonic in V_X and V_Y has risen to -12 dB, leading to significant asymmetry in the waveforms of Fig. 4.15(b). Thus, the effect of flicker noise decreases by 6.8 dB rather than by 15 dB.

This design is still superior to that in Fig. 4.14(a) in terms of phase noise, with an 8-dB advantage at 100-kHz offset and a 4-dB advantage at 100-MHz offset. Its supply sensitivity, however, is about 228 MHz/V.

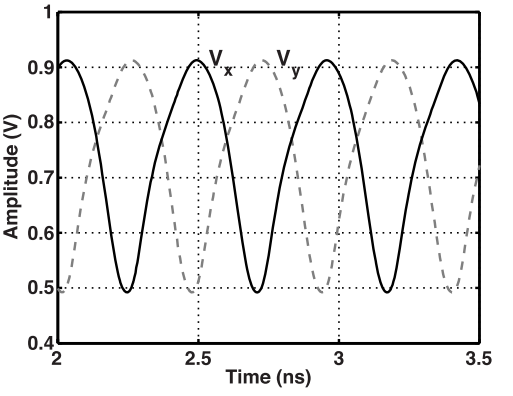
4.4.3 Method 3: Greater Number of Stages

Our third approach would be to increase the number of stages for a tenfold reduction in f_0. But we recognize two opposing trends: (1) as explained in Example 4.3, a larger N translates to higher phase noise if the power dissipation is constant, and (2) as N goes from 3 to 4 to 5, the voltage swings increase, lowering the normalized phase noise. We therefore raise N by only a moderate amount and scale the frequency further through the use of larger transistors.

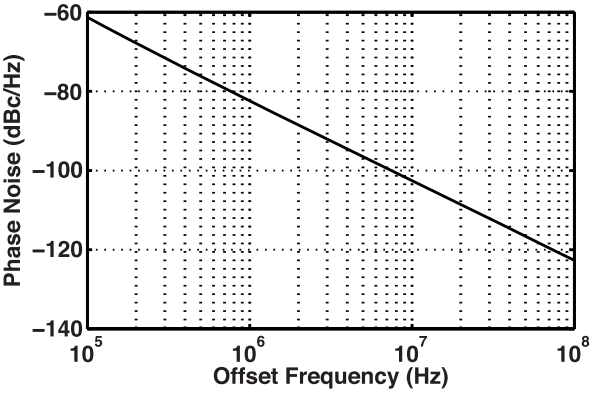
Figure 4.16(a) shows a four-stage loop, in which I_{SS} is reduced to 75 μA nd $W/L = 4$ μm/240 nm. Note the crossover between the feedback lines, necessary to produce negative feedback at low frequencies. Plotted



(a)



(b)



(c)

Figure 4.16 (a) Four-stage ring, (b) its waveforms, and (c) its phase noise.

[2] The effective channel length is assumed 4 nm shorter than the drawn length.

in Fig. 4.16(b) are the waveforms, revealing that the swings now reach 420 mV despite the reduction in the tail current. Oscillating at $f_0 = 2.16$ GHz, the circuit exhibits a supply sensitivity of 194 MHz/V and the phase noise plotted in Fig. 4.16(c), which is about 1 dB higher than that in Fig. 4.15(c). In other words, the two opposing trends mentioned above partialy cancel each other as N goes from 3 to 4.

Table 4.2 summarizes the performance of the three designs. Our conclusion here is that three-stage rings with large transistors achieve the best FOM. Four-stage rings, on the other hand, offer the advantage of generating larger voltage swings and phases that are integer multiples of 45°.

	Design 1 Higher Node Capacitances	Design 2 Longer Transistors	Design 3 4 Stages
f_0 (GHz)	2.14	2.27	2.16
P (uW)	285	285	285
Supply Sensitivity K_{VDD}(MHz/V)	74	228	194
Phase Noise (dBc/Hz) at			
1–MHz Offset	–79	–83	–82
100–MHz Offset	–120	–123	–122
FOM (dB) at			
1–MHz Offset	151	156	154
100–MHz Offset	152	156	154

Table 4.2 Performance summary of three differential ring oscillators.

4.5 Two-Stage Ring Oscillators

4.5.1 Basic Idea

Our study of differential rings began in Section 4.3 with the tacit assumption that the minimum number of stages is three. We surmise that, if successfully implemented, a two-stage differential ring operates faster and provides quadrature outputs. Of course, if a ring consists of only two simple differential pairs, no oscillation occurs because the loop contains only two poles and the frequency-dependent phase shift around the loop does not reach 180° for $\omega < \infty$.

Let us consider a two-stage-ring and its negative-feedback system model, Fig. 4.17(a). Suppose each stage is realized as a differential pair, Fig. 4.17(b), but with a *negative* resistance, $-R_p$, tied from each output to ground. We examine the possibility of oscillation for such a ring. The transfer function of one stage is given by

$$\frac{V_{out}}{V_{in}} = -g_{m1,2}\left[R_D||\frac{1}{C_L s}||(-R_p)\right] \tag{4.9}$$

$$= -g_{m1,2}\frac{-R_D R_p}{R_D - R_p}\frac{1}{\dfrac{-R_D R_p}{R_D - R_p}C_L s + 1}. \tag{4.10}$$

We assume that $R_p < R_D$ so that the parallel combination of R_D and $-R_p$ is negative. The two-stage ring

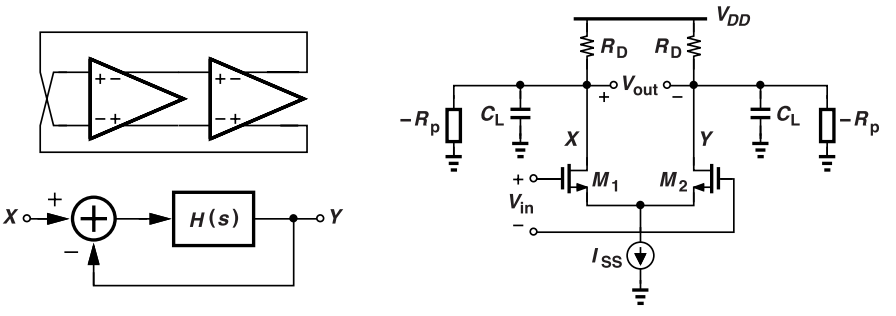


Figure 4.17 (a) Two-stage ring and its model, and (b) differential pair with negative load resistance.

with negative feedback exhibits the following loop transmission (the square of the above transfer function):

$$H(s) = \left(\frac{g_{m1,2}R_DR_p}{R_D - R_p}\right)^2 \frac{1}{\left(\dfrac{-R_DR_p}{R_D - R_p}C_Ls + 1\right)^2}. \tag{4.11}$$

For the transfer function $Y/X = H/(1+H)$ in Fig. 4.17(a) to go to infinity, we seek a value of s that yields $H(s) = -1$. Denoting $R_DR_p/(R_D - R_p)$ by R_L (> 0), we have

$$(g_{m1,2}R_L)^2 \frac{1}{(-R_LC_Ls + 1)^2} = -1 \tag{4.12}$$

and hence

$$g_{m1,2}R_L = \pm j(-R_LC_Ls + 1). \tag{4.13}$$

That is,

$$s = \frac{1}{R_LC_L} \pm j\frac{g_{m1,2}}{C_L}. \tag{4.14}$$

Interestingly, the closed-loop system contains two complex poles in the *right* half plane. This means that the loop is unstable. In the small-signal regime, the circuit does not produce a simple sinusoidal oscillatory waveform because the poles do not reside on the $j\omega$ axis. Writing the above values of s as $\sigma_1 \pm j\omega_1$, we note that the solution satisfying $H(s) = -1$ is of the form $\exp(\sigma_1 \pm j\omega_1)$, i.e., a sinusoid at ω_1 but with an exponentially growing amplitude. Figure 4.18 illustrates how this waveform is multiplied by -1 as it propagates through $H(s)$. Of course, the growth only continues until the negative resistance and/or the differential pairs begin to saturate.

Example 4.7

What does Eq. (4.14) predict if (a) $R_p = \infty$, or (b) $R_p = R_D$?

Solution

For $R_p = \infty$, we have $R_L = -R_D$ and $s = -1/(R_DC_L) \pm jg_{m1,2}/C_L$. The closed-loop system contains two complex poles in the left half plane and fails to oscillate. This is also true if $R_p > R_D$. If $R_p = R_D$, then $R_L = \infty$ and $s = \pm jg_{m1,2}/C_L$. In this case, the ring reduces to two ideal integrators (why?), oscillating at $\omega = g_{m1,2}/C_L$.

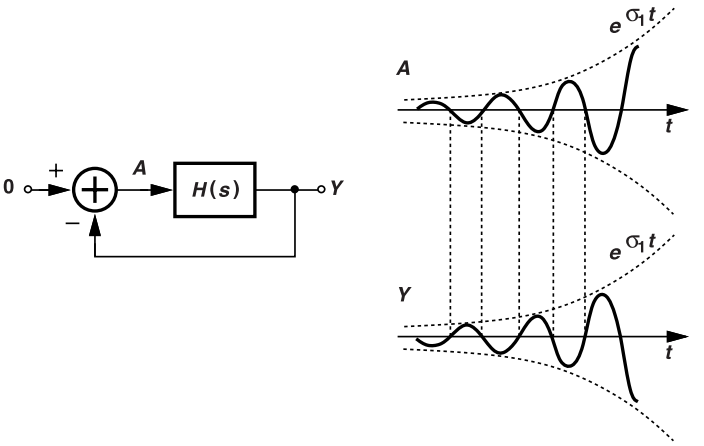


Figure 4.18 Growth of oscillation waveforms.

The negative resistances necessary in Fig. 4.17(b) can be realized by means of a cross-coupled pair, as shown in Fig. 4.19. Here, M_3 and M_4 introduce a resistance equal to $-2/g_{m3,4}$ between X and Y, yielding

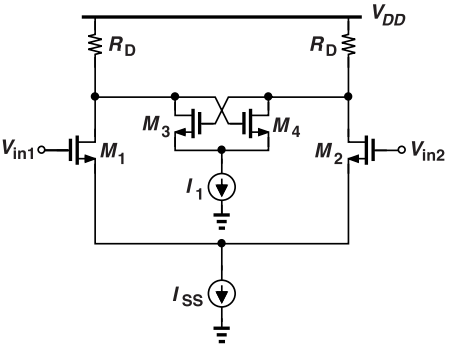


Figure 4.19 Differential pair with cross-coupled pair acting as negative resistance.

$R_p = 1/g_{m3,4}$. That is, $R_L = R_D R_p/(R_D - R_p) = R_D/(g_{m3,4}R_D - 1)$. Note that M_3 and M_4 must be strong enough to ensure $1/g_{m3,4} < R_D$, i.e., $g_{m3,4}R_D > 1$.

Example 4.8

Can the value of $g_{m1,2}$ be chosen arbitrarily?

Solution

We intuitively expect that $g_{m1,2}$ must play a role: if $g_{m1,2} = 0$, the two-stage ring reduces to two independent cross-coupled pairs, failing to oscillate.

To understand the actual situation, recall that $g_{m3,4}R_D$ must exceed unity to create poles in the right half plane. Now, consider the cross-coupled pair shown in Fig. 4.20, noting that positive feedback around the loop pushes V_X and V_Y in opposite directions because the loop gain, $(g_{m3,4}R_D)^2$, is greater than unity. This divergence continues until one transistor turns off, the other carries all of I_1, and, for example, $V_X = V_{DD}$ and $V_Y = V_{DD} - I_1 R_D$. The regenerative loop thus latches up, presenting no negative resistance anymore. In other words, the oscillation condition inevitably leads to latch-up, unless M_1 and M_2 in Fig. 4.19 are strong enough to bring M_3 and M_4 back to equilibrium in every cycle, allowing the cross-coupled pair to act as a negative resistance again. This means that both $g_{m1,2}$ and I_{SS} should exceed a minimum.

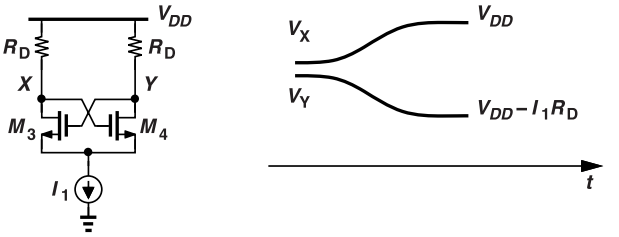


Figure 4.20 Growth of node voltages in a cross-coupled pair.

Two-stage rings are attractive for two reasons. (1) They can potentially achieve a higher oscillation frequency than three-stage topologies, and (2) they provide quadrature waveforms, i.e., four phases that are integer multiples of $90°$. Such rings, however, suffer from smaller swings due to the short delay around the loop.

Example 4.9

A student reasons that I_1 in Fig. 4.19 lowers the output CM level, limiting the voltage headroom for itself and for I_{SS}. The student then proposes the topology shown in Fig. 4.21 to avoid this issue. Is this a good idea?

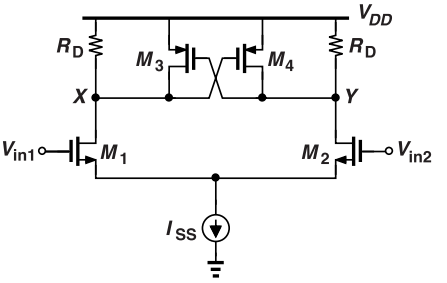


Figure 4.21 Use of a PMOS cross-coupled pair in a differential pair.

Solution

No, it is not. Consider the point in time at which $V_X = V_Y = V_{DD} - |V_{GS3,4}|$. The supply voltage must accommodate $|V_{GS3,4}| + V_{DS1,2} + V_{ISS}$, where V_{ISS} denotes the voltage necessary for I_{SS}. By contrast, in Fig. 4.19, we need $V_{DD} = (I_{SS} + I_1)R_D/2 + V_{DS1,2} + V_{ISS}$, which is not constrained by a gate-source voltage. Thus, the PMOS devices require a higher supply voltage.

Example 4.10

A student draws the two-stage ring of Fig. 4.22(a) as shown in Fig. 4.22(b) for small signals, reasoning that the circuit suffers from *positive* feedback and hence latches up. Explain the flaw in this claim.

Solution

Suppose the circuit in Fig. 4.22(b) tends to latch up with the logical states $A = 1, C = 0, B = 1$, and $D = 0$. This means that $A = B = 1$ and $C = D = 0$, an impossible state. To understand why, note from Fig. 4.22(a) that these states would imply that M_1 and M_2 are off and M_3 and M_4 are on, which cannot happen for differential pairs in a loop. In other words, the conditions $A = B$ and $C = D$ hold only if all four node voltages are equal to $V_{DD} - R_D I_{SS}/2$, which can be true at the startup. (The circuit still oscillates due to

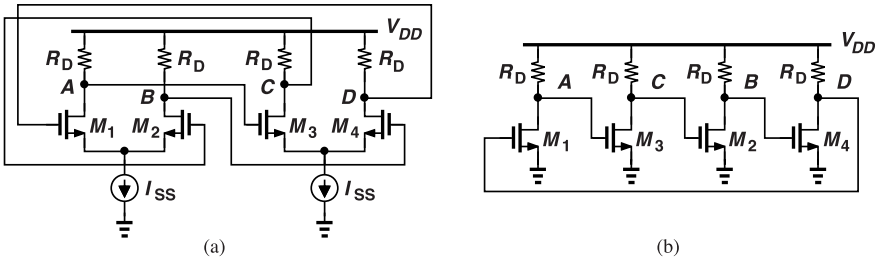


Figure 4.22 (a) Simple two-stage ring, and (b) its small-signal differential model.

the noise of the devices.) This point stands in contrast to the situation where the common source nodes of the differential pairs are grounded, in which case the circuit does latch up.

4.5.2 Design Example

Let us begin with the stage shown in Fig. 4.19 and design it for $f_0 \approx 2$ GHz and $P = 300\ \mu\text{A} \times 0.95$ V, the same parameters as those of the three-stage ring depicted in Fig. 4.15(a). Since each stage can draw 150 μA from V_{DD}, we allocate 100 μA to the input pair and 50 μA to the cross-coupled pair. Figure 4.23(a) shows the design; R_D is 4 kΩ to maintain the same voltage swing, and $(W/L)_{3,4}$ is so chosen to provide sufficient

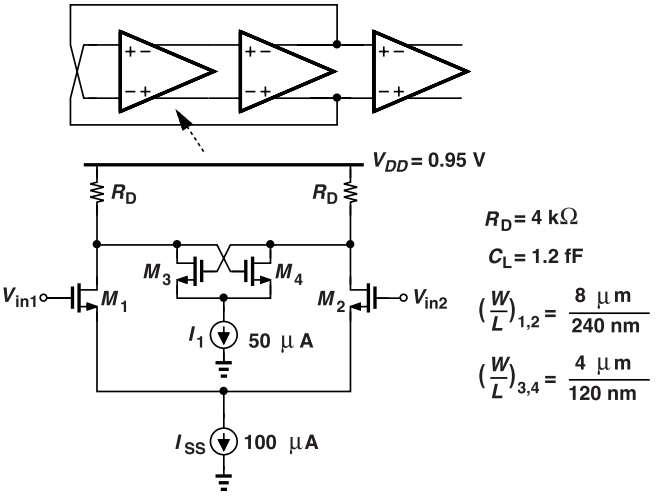


Figure 4.23 Two-stage ring design.

transconductance.

According to simulations, $f_0 = 2.14$ GHz but the swings are less than 250 mV_{pp}. The phase noise is 3 to 4 dB higher than that of the three-stage design in Fig. 4.15, which can be attributed to the ratio of the voltage swings, $10\log(360\ \text{mV}/250\ \text{mV})^2 = 3.2$ dB.

The stage of Fig. 4.23 can be reconfigured as shown in Fig. 4.24, where the nodes of the differential and cross-coupled pairs are shared. For proper splitting of I_{SS} between the two pairs, all transistor lengths must be equal. Then, since the input and output common-mode levels are equal in a ring oscillator environment, the differential pair carries a current of $I_{SS}W_{1,2}/(W_{1,2} + W_{3,4})$, and the cross-coupled pair, $I_{SS}W_{3,4}/(W_{1,2} + W_{3,4})$ (why?).

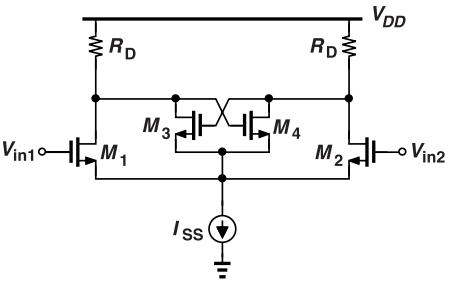


Figure 4.24 Differential and cross-coupled pairs with their tails shorted.

A two-stage ring using the topology of Fig. 4.24 oscillates with larger voltage swings than that in Fig. 4.19 but also exhibits 10 to 15 dB higher flicker noise upconversion! The following example investigates this point.

Example 4.11

Examine the tail node waveforms of the differential and cross-coupled pairs in Figs. 4.23 and 4.24.

Solution

In a two-stage ring, the input and output waveforms of one stage are 90° out of phase. Thus, as illustrated in Fig. 4.25, the separate tail nodes carry voltage waveforms (at $2f_0$) having a phase difference of 180°.

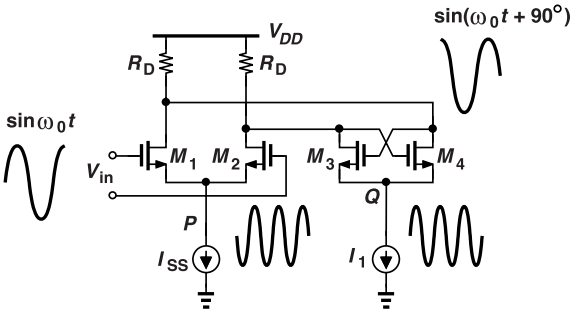


Figure 4.25 Differential waveforms appearing at tail nodes.

This effect can be understood by viewing the two pairs as frequency doublers and noting that they convert $A\sin\omega_0 t$ and $A\sin(\omega_0 t + 90^\circ)$ at their drains to $A\sin(2\omega_0 t)$ and $A\sin(2\omega_0 t + 180^\circ)$, respectively, at their sources.

If we now short the tail nodes, as in Fig. 4.24, the two waveforms partially cancel each other. If V_P and V_Q exactly cancel, the shared tail node voltage remains constant, as if the sources of the four transistors were grounded. In this case, the second-order nonlinearity of each transistor produces a strong second harmonic in its drain current. The resulting asymmetry between the rise and fall times translates to higher flicker noise upconversion. Note that this mechanism is different from the previous one related to the tail capacitance and occurs even without such a capacitance.

4.6 Linear Scaling

As explained in Section 3.1, the phase noise of any oscillator can be reduced by exploiting its fundamental trade-off with power dissipation. For the previous 2-GHz differential rings, $P \approx 285\ \mu$W, a small value. For

example, we can scale up the transistor widths and the tail currents by a factor of 10 and scale down the resistors by the same factor, obtaining 10 dB less phase noise with $P = 2.85$ mW. An additional penalty in this endeavor is the proportionally larger area.

4.7 Tuning Techniques

In this section, we repeat the design efforts of Section 3.5 for the 2-GHz differential ring oscillator of Fig. 4.15(a). Our objective is to create a relatively linear continuous tuning characteristic and subsequently add discrete tuning. The reader is encouraged to review the design criteria outlined in Section 3.5.

Recall from Eq. (1.25) that

$$\omega_0 = \frac{1}{R_D C_L} \tan \frac{180^\circ}{N} \tag{4.15}$$

for an N-stage ring. This small-signal estimate suggests that the oscillation frequency is relatively independent of the tail current and should be tuned by varying R_D, C_L, or N. We thus envision three tuning techniques.

4.7.1 Resistive Tuning

It is tempting to replace the load resistors of the differential pairs with PMOS devices that act as controlled resistances, i.e., operate in the deep triode region [Fig. 4.26(a)]. However, as the on-resistance of the PMOS loads varies, so do the voltage swings. Moreover, V_{cont} has an upper bound well below $V_{DD} - |V_{THP}|$ if M_3 and M_4 are to remain in the triode region.[3] This issue translates to substantial nonlinearity in the frequency tuning profile.

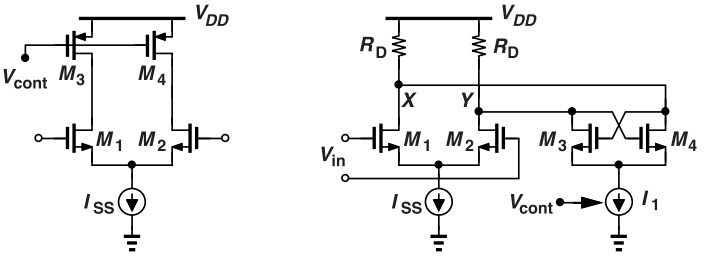


Figure 4.26 Frequency tuning by (a) PMOS loads, or (b) a variable negative resistance.

Another approach to adjusting the load resistance is to place in parallel with it a variable *negative* resistance. Illustrated in Fig. 4.26(b), the idea is to introduce a resistance equal to $-2/g_{m3,4}$ between X and Y by means of cross-coupled transistors M_3 and M_4, with I_1 serving as the tuning element. As pointed out in Section 3.5, the tuning should negligibly degrade the phase noise, a possibility here if the continuous tuning range is limited to 10-20% of f_0.

Figure 4.27(a) shows a three-stage ring employing negative-resistance tuning and Fig. 4.27(b) plots its tuning characteristic. The dimensions of M_3-M_5 are so chosen as to provide the necessary tuning range and avoid degrading the phase noise. As V_{cont} rises from 250 mV (slightly below V_{TH}) to 900 mV, f_0 falls from 2.2 GHz to 1.83 GHz and the voltage swings increase from 360 mV to 480 mV. The maximum K_{VCO} is about 1 GHz/V. Unfortunately, we cannot allow V_{cont} to reach 900 mV if M_5 must remain in saturation.

[3] In other words, we must ensure that $2|V_{GS} - V_{THP}| \gg |V_{DS}|$.

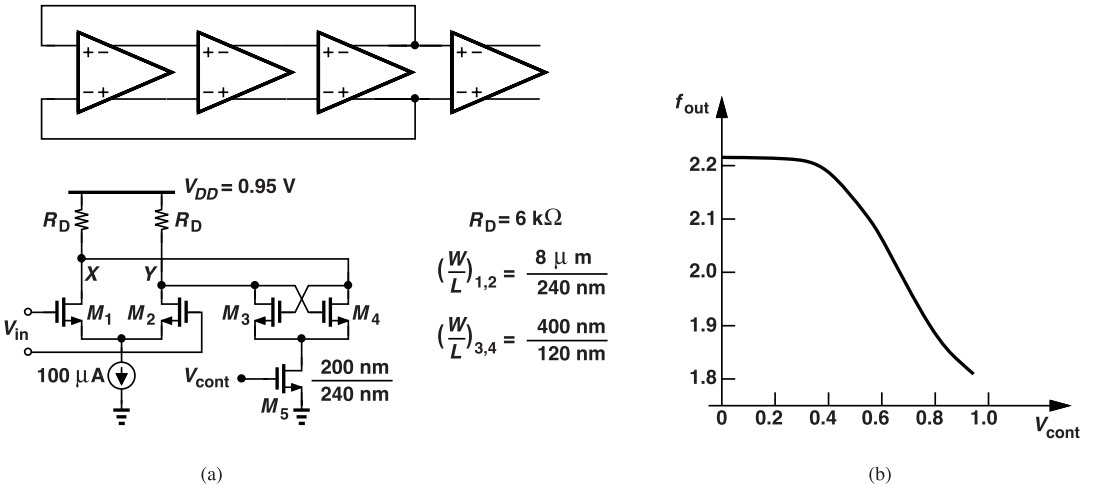


Figure 4.27 (a) Three-stage VCO, and (b) its tuning characteristic.

Example 4.12

Explain why the supply rejection of the foregoing oscillator degrades as M_5 enters the triode region.

Solution

If M_5 operates in the triode region, the bias current of the cross-coupled pair becomes a function of V_{DD} because, at the instant when $V_X = V_Y$, we have $V_{DS5} = V_{DD} - V_{RD} - V_{GS3,4}$. Thus, V_{DD} fluctuations cause $I_{D3,4}$ and hence $g_{m3,4}$ to change, modulating the negative resistance and f_0.

The phase noise at $f_0 = 1.83$ GHz is expected to be lower than at $f_0 = 2.2$ GHz by at least $10\log(2.2/1.83)^2 = 1.6$ dB. In fact, the phase noise at 100-MHz offset drops by 3 dB to -127 dBc/Hz because the swings have also become larger. At 100-kHz offset, however, the phase noise remains unchanged and equal to -62 dBc/Hz. Why? According to simulations, the tuning device, M_5 in Fig. 4.27(a), is now a major phase noise contributor; after all, the flicker noise voltage of this transistor directly adds to V_{cont}. This is the principal disadvantage of negative-resistance tuning. If we double both the width and the length of M_5, the phase noise falls by 1 dB.

The flat zone between $V_{cont} = 0$ and $V_{cont} \approx 400$ mV in Fig. 4.27(b) results in a high K_{VCO} for $400\ \text{mV} < V_{cont} < 900$ mV. Unlike inverter-based rings, this differential topology does not lend itself to techniques that combine PMOS and NMOS devices to remove the flat zone (Section 3.5). The difficulty stems from the low supply voltage and hence the lack of headroom for PMOS transistors. We can reduce K_{VCO} by scaling down M_3-M_5. The continuous tuning range also decreases, requiring finer discrete steps when we add switched capacitors to the drain nodes (as we did in Chapter 3).

An important advantage of resistively-loaded differential rings is that their PVT-induced frequency variation is less than that of inverter-based loops. As $\omega_0 = (R_D C_L)^{-1} \tan(180°/N)$ suggests, the oscillation frequency varies with R_D and C_L, exhibiting little dependence on the strength of the *transistors* in the signal path. With a typical variation of $\pm 15\%$ for R_D and $\pm 5\%$ for C_L, the overall tuning range can be about $\pm 20\%$ for narrowband applications.

4.7.2 Varactor Tuning

The MOS varactor tuning method described in Section 3.5 can be applied to differential rings as well. Illustrated in Fig. 4.28, this approach negligibly adds phase noise through the tuning network, but it relies on varactor models to establish the desired tuning range. Varactor tuning is generally the preferred technique

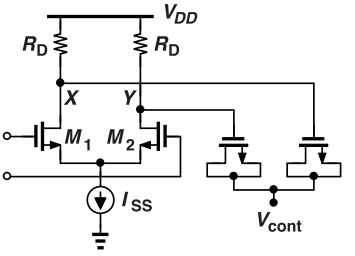


Figure 4.28 Differential pair with varactor tuning.

as it introduces negligible noise in the control path. By comparison, any transistor in the path injects flicker noise, modulating the oscillation frequency. We return to this point in the next section.

Example 4.13

A MOS varactor exhibits the C-V characteristic shown in Fig. 4.29. If V_{cont} in Fig. 4.28 varies from V_1 (≈ 0)

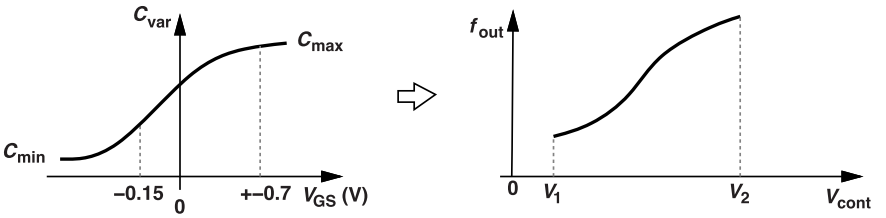


Figure 4.29 Varactor C/V characteristic and resulting tuning behavior.

to V_2 ($\approx V_{DD}$), which part of the characteristic is exercised?

Solution

The average voltage across the varactors in Fig. 4.28 is approximately equal to the common-mode level at X and Y minus V_{cont}. Thus, as V_{cont} goes from V_1 to V_2, the varactors' voltage varies from $V_{DD} - I_{SS}R_D/2 - V_1$ to $V_{DD} - I_{SS}R_D/2 - V_2$. For example, if $V_1 = 50$ mV, $V_2 = 0.9$ V, $I_{SS}R_D/2 = 200$ mV, and $V_{DD} = 0.95$ V, we have a range of -150 mV to $+700$ mV for V_{GS} in Fig. 4.28. If C_{var} does not flatten significantly as V_{GS} approaches $+700$ mV, then neither does f_{out} as V_{cont} falls to V_1.

4.7.3 Tuning the Number of Stages

It is possible to vary the number of stages in a ring oscillator *continuously*. Particularly suited to differential configurations, this method operates as depicted in Fig. 4.30(a) for a three-stage example. Each stage incorporates a slow path and a fast path, with their outputs summed and applied to the next stage. The strength of the fast path is adjusted by V_{cont}. We observe that the ring reaches the minimum oscillation frequency if the fast path is off because the stages in the loop run with the longest delay. On the other hand, if the fast path is fully on, the signal mostly bypasses the slow path in each stage, yielding the maximum f_0. For V_{cont} values between these two, we say the circuit "interpolates" between the slow path and the fast path. We call the corresponding waveforms the slow and fast signals, respectively. We see in this example that the total number of stages in the ring varies from 6 to 3.

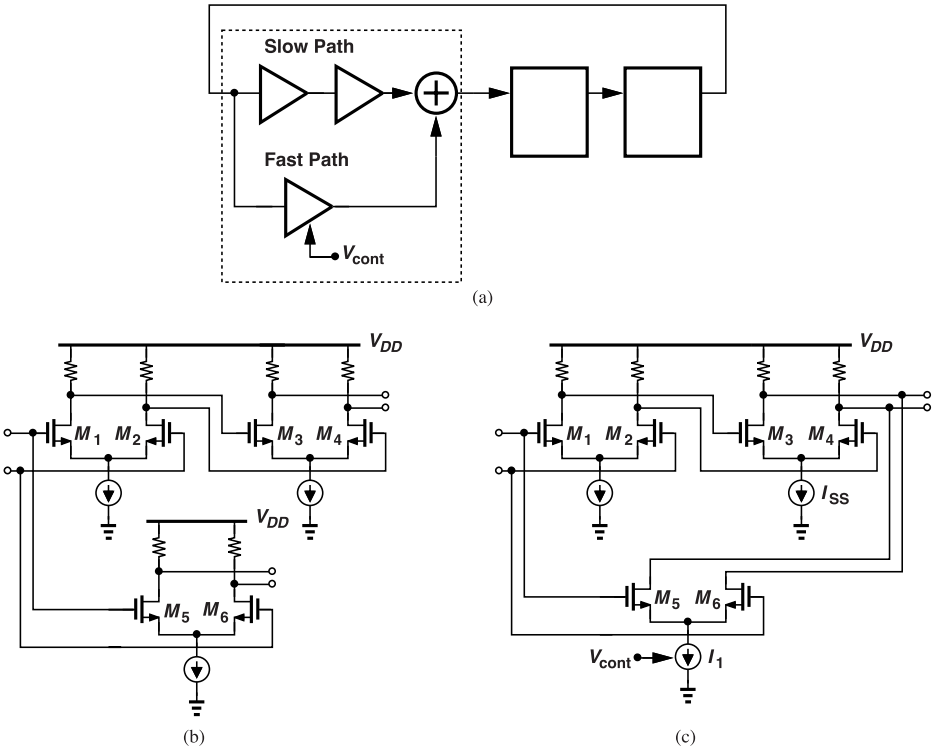


Figure 4.30 (a) Frequency tuning by interpolation, (b) implementation of slow and fast paths, and (c) interpolation circuit.

Let us implement the foregoing concept at the transistor level. We readily conceive the arrangement shown in Fig. 4.30(b) for the two paths, but how do we sum their outputs? Noting that the drain currents of M_3-M_4 and M_5-M_6 are proportional to the slow and fast signals, respectively, we can first sum these currents and then pass the result through resistors. Depicted in Fig. 4.30(c) is the outcome, with V_{cont} adjusting the fraction of the fast signal that is added to the slow signal. The output voltage swings now vary with I_1, but by a small amount because we are interested in a narrow (continuous) tuning range.

Example 4.14

Estimate the value of I_1 in Fig. 4.30(c) for the delay of the overall circuit to decrease by 20% compared to when $I_1 = 0$.

Solution

Writing the slow (current) signal as $A\cos(\omega_0 t + \omega_0 \Delta T)$, where ΔT denotes the delay, and the fast (current) signal as $a\cos\omega_0 t$, we express the sum as $A\cos(\omega_0 t + \omega_0 \Delta T) + a\cos\omega_0 t = [A\cos(\omega_0 \Delta T) + a]\cos\omega_0 t - A\sin(\omega_0 \Delta T)\sin\omega_0 t$. The phase of the resultant is given by

$$\phi = \tan^{-1}\frac{-A\sin(\omega_0 \Delta T)}{A\cos(\omega_0 \Delta T) + a}, \tag{4.16}$$

which, for $\omega_0 \Delta T \ll 1$, reduces to

$$\phi \approx \tan^{-1}\frac{-A\omega_0 \Delta T}{A + a}. \tag{4.17}$$

This phase translates to a delay of $A\Delta T/(A+a)$ seconds. Thus, for a 20% change, we must have $a \approx 0.2A$ and hence $I_1 \approx 0.2I_{SS}$.

The circuit of Fig. 4.30(c) follows the same design procedure described for the negative-resistance tuning topology shown in Fig. 4.27(a). An NMOS device acts as I_1, sensing the control voltage from low values to near V_{DD}. For this highest value of V_{cont}, the NMOS current source resides in the triode region, exacerbating the supply sensitivity (Example 4.12).

The above oscillator generally suffers from high phase noise. When the fast path is off, the requisite oscillation frequency is obtained with six differential pairs (in a three-stage ring) rather than with three pairs. As explained in Section 4.2, the phase noise tends to rise as the number of differential pairs increases while the total power consumption remains constant. Also, when the fast path is on, the noise of I_1 modulates the frequency (Problem 4.15).

In summary, among various tuning techniques, we prefer continuous varactor tuning along with switched capacitors for coarse control. As evident from Example 4.13, such a design suffers less from a flat zone in the tuning characteristic than that in Fig. 4.27.

4.8 Comparison of Inverter-Based and Differential Rings

Our studies in Chapter 3 and this chapter lead to two important differences between inverter-based and differential ring oscillators. (1) As noted in Table 4.1, the former exhibit a more favorable trade-off between phase noise and power dissipation in the thermal-noise regime. (2) The latter have much less supply sensitivity.

These observations suggest that in "hostile" applications—systems with a high supply noise—we opt for differential rings. If, on the other hand, the supply noise is small or can be suppressed through the use of low-noise LDOs, then we prefer inverter-based oscillators.

4.9 Inverter-Based Oscillators with Complementary or Quadrature Outputs

4.9.1 Coupled Oscillators

The need for negative feedback at low frequencies dictates an *odd* number of stages in simple single-ended oscillators, disallowing a phase step of the form $360°/2^M$. Let us first ask, how can we use single-ended arrangements to produce a $180°$ phase separation ($M = 1$)? Recall from Example 4.1 that two oscillators can be synchronized if they are "coupled," i.e., if a fraction of one's output is "injected" into the other. We then consider the two identical oscillators shown in Fig. 4.31 and ask whether the mutual coupling created by R_i

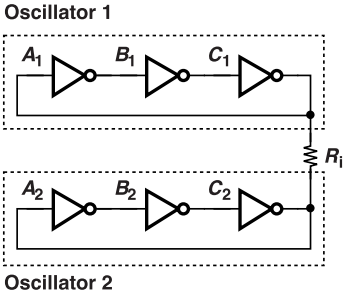


Figure 4.31 Two oscillators coupled through a resistor.

forces them to oscillate with opposite phases. If that were the case, then a current would flow from Oscillator 1 through R_i to Oscillator 2 and vice versa. That is, some energy would be transferred from one oscillator to the other. But suppose the two circuits oscillate in-phase. Then, nodes A_1 and A_2 change together, no current

flows through R_i, and no energy is exchanged. This simple view predicts that the two oscillators prefer to operate in-phase, failing to generate complementary outputs. We say the oscillators reach a point of minimum energy exchange.

We now turn to the topology depicted in Fig. 4.32(a) [2], where equal resistors R_{i1} and R_{i2} couple B_1

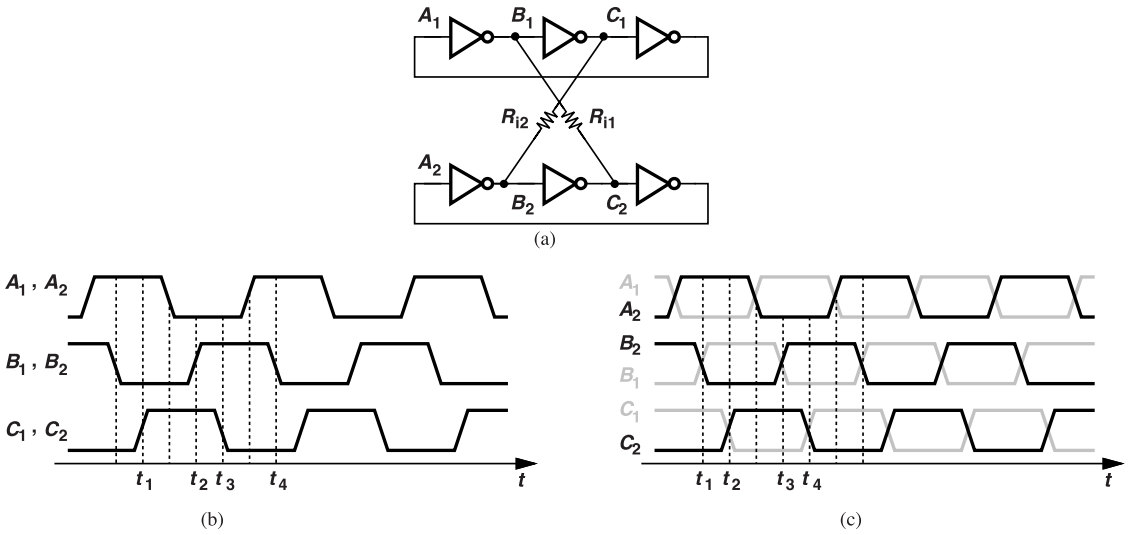


Figure 4.32 (a) Two oscillators coupled by two resistors, (b) waveforms in the case of in-phase operation, and (c) waveforms in the case of out-of-phase operation.

to C_2 and C_1 to B_2, respectively. Can the two oscillators operate in-phase here? Drawing the idealized waveforms as in Fig. 4.32(b), we recognize that energy exchange occurs from t_1 to t_2 and from t_3 to t_4 (why?). That is, energy is exchanged for two-thirds of the period.

If, on the other hand, the two oscillators operate 180° out of phase, the waveforms in Fig. 4.32(c) reveal that energy is exchanged for only one-third of the period. We therefore conclude that the circuit prefers the latter mode to the former.

Example 4.15

Can the two oscillators in Fig. 4.32(a) operate with some other phase difference between zero and 180°?

Solution

No, they cannot. The reader can show that, in such a case, the energy flow between B_1 and C_2 would last longer than between B_2 and C_1. This cannot happen as it conflicts with the symmetry between the two oscillators.

Another coupling arrangement for generating complementary waveforms is shown in Fig. 4.33. Here, two "anti-parallel" (cross-coupled) inverters, Inv_1 and Inv_2, inject a fraction of one oscillator's signal into the other. We observe that in-phase operation ($A_1 = A_2$) is discouraged here as it would mean *equal* (rather than complementary) input and output voltages for the coupling inverters, allowing, for example, Inv_1 to severely disrupt the top ring. We say anti-parallel inverters fight equal states. The oscillators therefore prefer to operate with a phase difference of 180° so that Inv_1 and Inv_2 inject only a transient current. The symmetry of the circuit prohibits other phase difference values as well.

The topology of Fig. 4.33 merits two remarks. First, Inv_1 and Inv_2 cannot be arbitrarily weak or strong. If they are weak, and the oscillators suffer from a frequency mismatch, then synchronization ("injection locking") does not occur. In such a case, the oscillators "pull" each other, exhibiting corrupted outputs. If

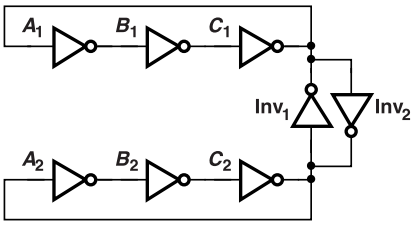


Figure 4.33 Oscillators coupled by inverters.

Inv_1 and Inv_2 are excessively strong, then they act as a regenerative latch, forcing permanently a ONE at one output and a ZERO at the other. As a rule of thumb, the strength of these inverters is chosen about 20% of that of the main inverters within the two loops.

Second, the coupling inverters present loading only at A_1 and A_2, thereby creating systematic phase mismatches around each ring. In applications requiring all of the phases, this issue is resolved by adding similar inverters between the other nodes (Fig. 4.34). Of course, the phase separations are multiples of 60° rather than $180^\circ/2^M$.

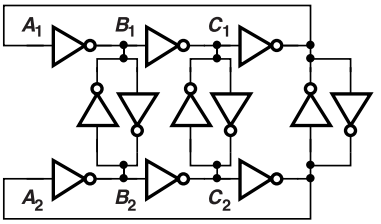


Figure 4.34 Two oscillators uniformly coupled by inverters.

4.9.2 Phase Noise Considerations

Let us design the coupled oscillators of Fig. 4.34 for a frequency of 2 GHz and examine the phase noise performance. Exploiting the 2.5-GHz three-stage ring depicted in Fig. 3.16(a), we construct the stage shown in Fig. 4.35(a). We select the coupling inverters according to those in Fig. 3.16(a), and the main inverters four times as strong. Simulations with $V_{DD} = 0.95$ V, $T = 75\,^\circ$C, and the SS corner reveal the waveforms shown in Fig. 4.35(b) and $f_0 = 2.3$ GHz. The phase noise is plotted in Fig. 4.35(c).

In order to assess the new design's phase noise, we note that the oscillator in Fig. 3.16(a) incorporates three "unit" inverters with $(W/L)_N = 120$ nm/240 nm and $(W/L)_P = 240$ nm/240 nm, whereas the circuit in Fig. 4.35(a) contains 30 such unit inverters (why?). The new design therefore consumes approximately 10 times as much power and, we hope, exhibits $10\log 10 = 10$ dB less phase noise. In reality, the supply current is 115 μA rather than 10×14 μA, but the oscillation frequency is also lower. In fact, $10\log(140\ \mu\text{A}/115\ \mu\text{A})$ happens to be about equal to $10\log(2.5\ \text{GHz}/2.3\ \text{GHz})^2$. We thus still expect a 10-dB reduction in phase noise compared to the simple three-stage ring, which is approximately the case in Fig. 4.35(c). The current therefore generates complementary outputs with little FOM penalty.

Example 4.16

It is often thought that the "differential" operation of the two rings in Figs. 4.34 and 4.35(a) considerably reduces the supply sensitivity. Is this true?

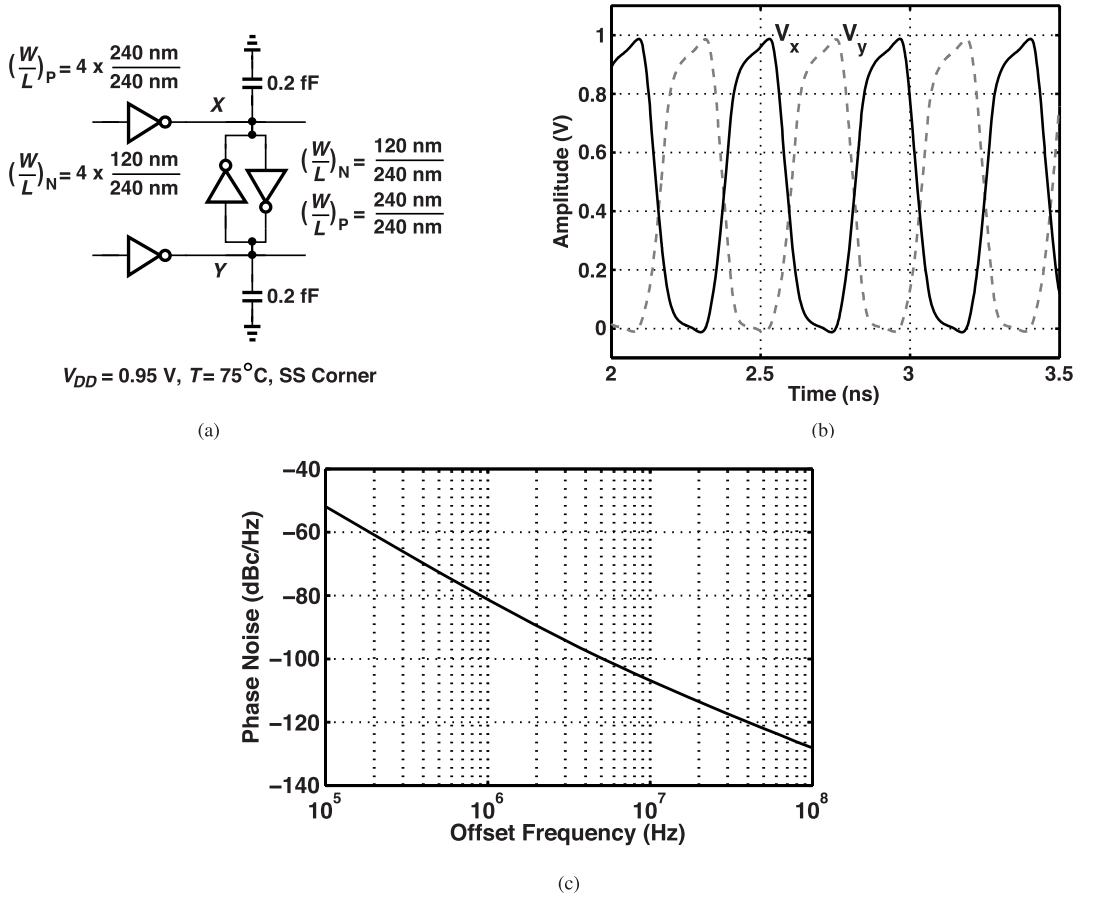


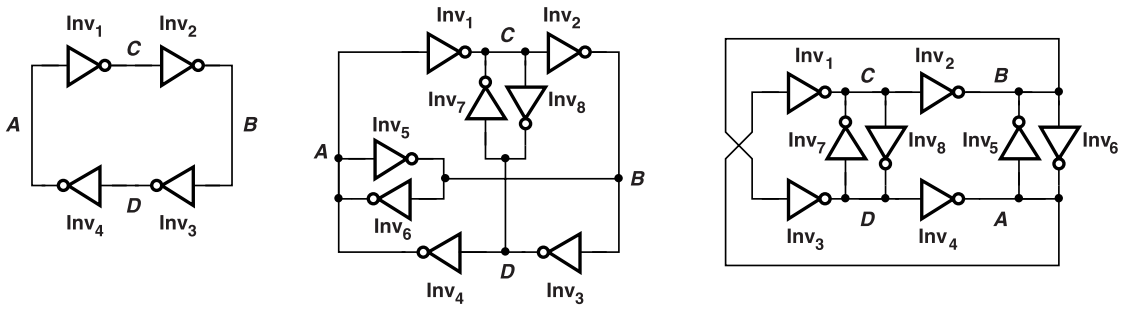
Figure 4.35 (a) Oscillator using coupling inverters, (b) its waveforms, and (c) its phase noise.

Solution

No, it is not. In contrast to the differential oscillators studied in Section 4.1, the topology of Fig. 4.35(a) still suffers from the dependence of the inverters' strength upon V_{DD}. As V_{DD} fluctuates, so do the inverter delays; that is, the oscillation frequencies of the two loops change in the *same* direction. In the foregoing design, for example, the normalized supply sensitivity of the coupled oscillators is only about 15% less than that of a single ring.

4.9.3 Direct Quadrature Generation

It is possible to obtain complementary, quadrature phases from an inverter-based ring oscillator. Let us begin with the four-stage loop shown in Fig. 4.36(a). Ignoring for the moment the inversion provided by each stage, we recognize that the circuit can oscillate at a frequency $f_0 = 1/(8T_d)$, where T_d denotes the gate delay. Thus, A and B carry complementary waveforms and so do C and D. Also, the latter two are $90°$ out of phase with respect to the former two. From another perspective, the loop consists of four one-pole stages, thereby generating $90°$ phase separations between consecutive nodes.



(a) (b) (c)

Figure 4.36 (a) Four-stage ring, (b) addition of cross-coupled inverters to avoid latchup, and (c) redrawing of the topology in (b).

Unfortunately, the circuit of Fig. 4.36(a) prefers to latch up: the loop can indefinitely maintain $A = B = 1$ and $C = D = 0$ or vice versa. We must therefore devise a mechanism that avoids such a state. Recall from Fig. 4.33 that a pair of cross-coupled inverters can fight equal states between their input and output nodes. This point leads to the structure depicted in Fig. 4.36(b), often redrawn as in Fig. 4.36(c). Here, the cross-coupled inverters are strong enough to ensure that the main loop does not latch up.

How should the cross-coupled inverters be sized with respect to those in the main loop? The former must fight the latter if Inv_1-Inv_4 tend toward latchup. From this perspective, Inv_5-Inv_8 should be strong enough. On the other hand, these inverters also fight Inv_1-Inv_4 during transitions, both draining power and injecting noise. Thus, Inv_5-Inv_8 should not be excessively strong. As a rule of thumb, we choose a ratio of 2 between the strengths of Inv_1-Inv_4 and those of Inv_5-Inv_8. Greater ratios run the risk of latchup in the presence of mismatches, and lower ratios degrade the figure of merit.

Let us design the circuit of Fig. 4.36(b) and study its performance. Returning to our reference ring design in Fig. 3.12(a) and assuming a minimum allowable width of 120 nm for the transistors, we choose for the main inverters $(W/L)_N = 240$ nm/40 nm and $(W/L)_P = 480$ nm/40 nm, and for the cross-coupled inverters $(W/L)_N = 120$ nm/40 nm and $(W/L)_P = 240$ nm/40 nm. We also attach output buffers to all of the nodes, arriving at the design shown in Fig. 4.37(a). Plotted in Fig. 4.37(b), the quadrature waveforms exhibit a slight swing degradation due to the fight between the main and cross-coupled inverters. The oscillation frequency, f_0, is equal to 26.5 GHz, about 17% *greater* than that of the reference design in Fig. 3.12(a). Interestingly, the quadrature topology runs faster even though it contains more stages. The supply current is $235\,\mu$A.

Figure 4.37(c) plots the phase noise of the quadrature oscillator. At 1-MHz offset, the phase noise is equal to -50 dBc/Hz, 3 dB lower than that of the reference design. For a fair comparison, we must take into account (a) the supply current ratio, $10\log(235\ \mu\text{A}/60\ \mu\text{A}) = 5.9$ dB, (b) the frequency ratio, $20\log(26.5\ \text{GHz}/22.6\ \text{GHz}) = 1.4$ dB, and (c) the phase noise difference of 3 dB. Thus, the quadrature design's figure of merit is worse by $5.9 - 1.4 - 3 = 1.5$ dB. This is the penalty paid for generating quadrature phases.

4.9.4 Quadrature Generation by Interpolation

It is possible to derive additional phases from a simple inverter-based ring by means of "phase interpolation." Returning to Fig. 4.34, we can add phasors B_1 and C_2 with equal weights and obtain an interpolated phasor perpendicular to A_1 [Fig. 4.38(a)]. This can be repeated for C_1 and B_2 as well [Fig. 4.38(b)], yielding A_1, D_1, A_2, and D_2 as the desired quadrature phases.

Let us study this phase interpolation technique more carefully. How do we implement phasor summation? Considering the B_1 and C_2 waveforms in Fig. 4.38(c), we recognize that $(B_1 + C_2)/2$ lies exactly half-way between the two. We thus apply B_1 and C_2 to two inverters and short their outputs [Fig. 4.38(d)]. We ignore the signal inversion for now and consider the result roughly equivalent to $(B_1 + C_2)/2$.

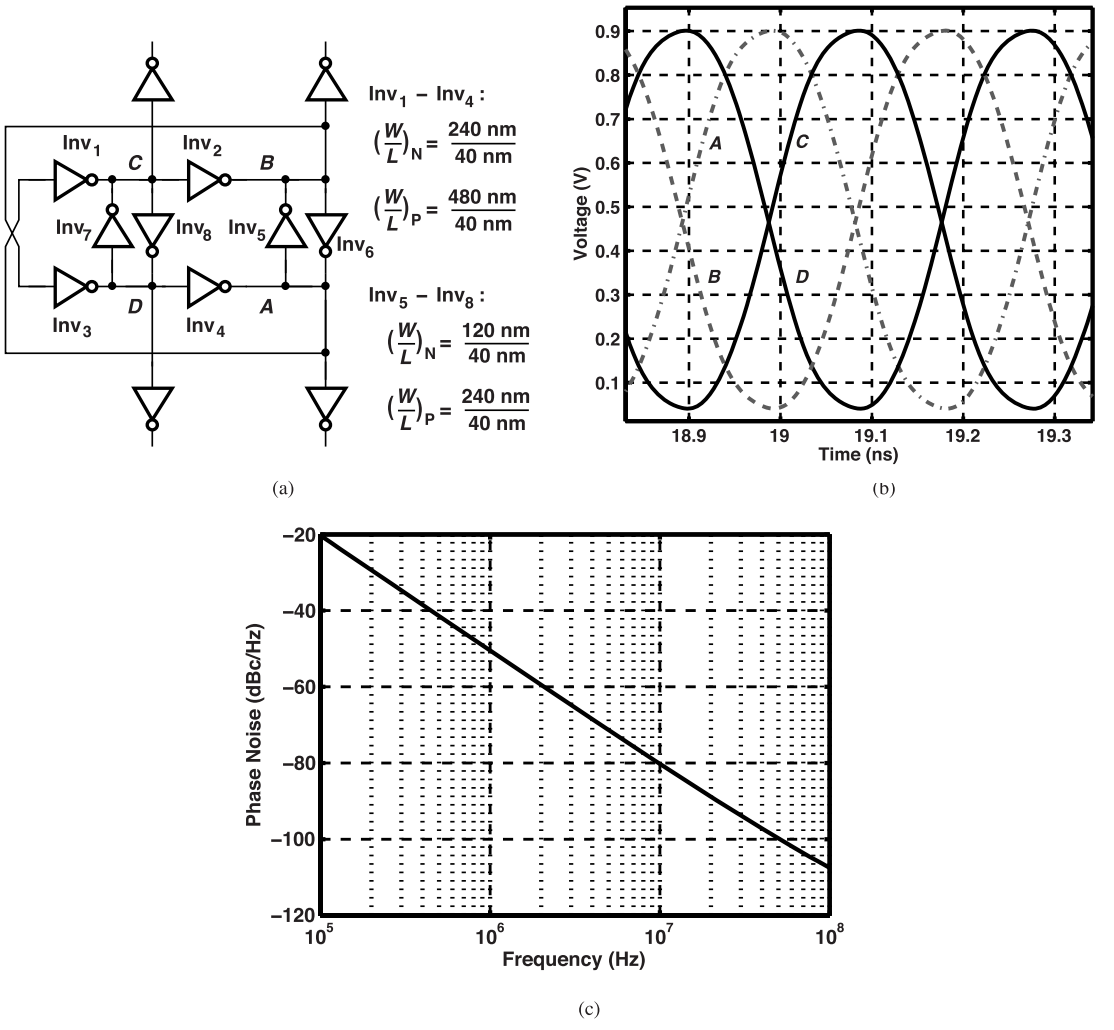


Figure 4.37 (a) Quadrature oscillator design example, (b) its output waveforms, and (c) its phase noise.

The reader may raise two questions here. First, do Inv $_1$ and Inv$_2$ fight and degrade the high and low levels of $(B_1 + C_2)/2$ in Fig. 4.38(c)? No, when both B_1 and C_2 are high or low, the two inverters do not fight and the interpolated output is close to the supply rails. Second, do the two inverters fight around $t = t_1$? Yes, in fact, they must fight so that they generate an interpolated output between B_1 and C_2. As depicted in Fig. 4.39, the NMOS device in one inverter and the PMOS transistor in the other are simultaneously on, producing an interpolated output voltage. In other words, the circuit draws static power for part of the period.

The circuit of Fig. 4.38(d) faces two more issues. First, the load capacitance introduced by the interpolating inverters creates a phase mismatch between B_1 and A_1 (and between B_2 and A_2). Second, the finite delay of Inv$_1$ and Inv$_2$ leads to another phase mismatch between D_1 and A_1 (and between D_2 and A_2). Illustrated in Fig. 4.40 are the modifications necessary to alleviate these issues. The four outputs, D_1-D_4, serve as the quadrature phases. The various interconnects in this topology tend to introduce significant phase mismatches, requiring careful layout.

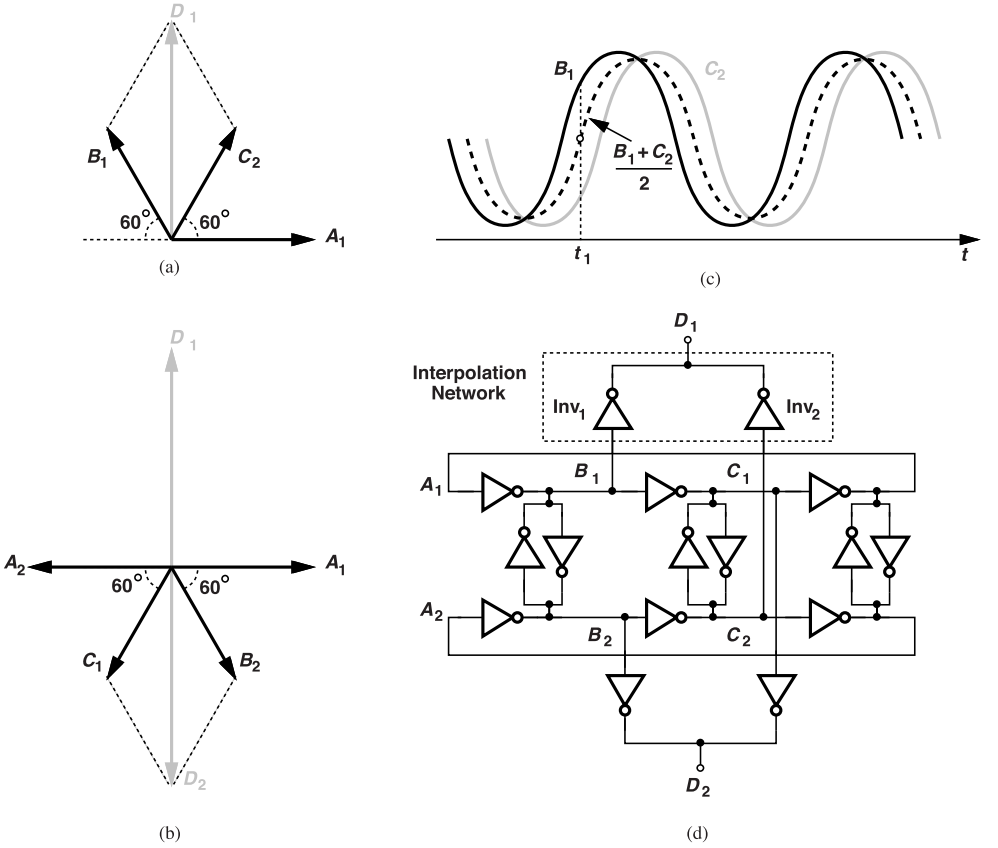


Figure 4.38 (a) Generation of a quadrature phasor, D_1, from C_2 and B_1, (b) generation of D_2 from C_1 and B_2, (c) interpolated waves, and (d) implementation of interpolation network.

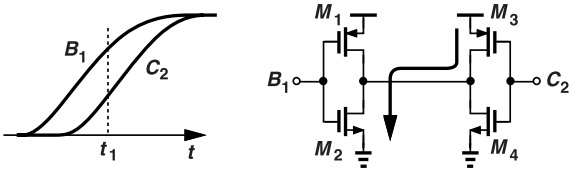


Figure 4.39 Fight between two inverters during interpolation.

Example 4.17

Suppose D_1 and D_3 in Fig. 4.40 drive equal load capacitances. Do these two signals experience a phase mismatch in this case?

Solution

Yes, they do. The driving strength at D_1 is greater than that at D_3 because D_1 is formed by two inverters. To remedy the situation, we should ensure that the fanout seen at these outputs is small, i.e., Inv_1-Inv_3 are several times *larger* than the subsequent stages.

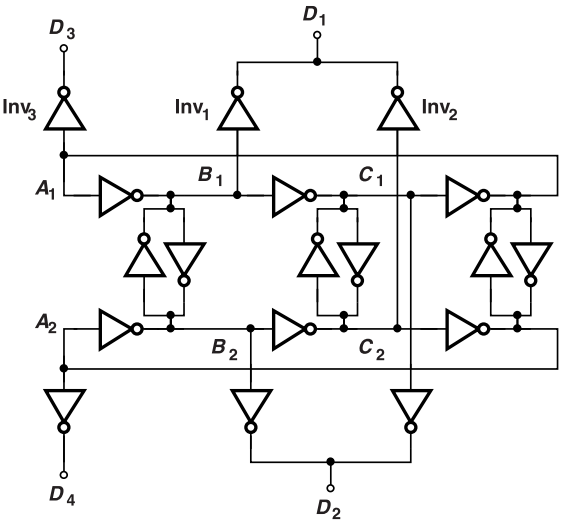


Figure 4.40 Oscillator with interpolation network.

Example 4.18

Two five-stage ring oscillators are coupled so as to operate $180°$ out of phase. How should quadrature waveforms be generated?

Solution

As shown in Fig. 4.41, the two rings produce a phase resolution of $360°/10 = 36°$. For $90°$ phase generation,

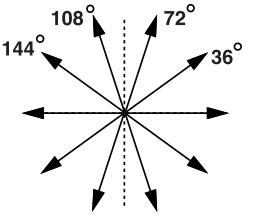


Figure 4.41 Phasor diagram for the oscillator in Fig. 4.40.

we can interpolate between the $36°$ and $144°$ outputs or between $72°$ and $108°$ outputs. We prefer the latter option as the smaller phase difference means that the interpolating inverters fight to a lesser extent.

We should also remark that the interpolation scheme introduced in Fig. 4.40 negligibly degrades the phase noise. Since the interpolating inverters lie outside the oscillation loops, their phase noise is not shaped by the $(f_0/f)^2$ function in Eq. (3.9) and hence remains small. For this reason, these inverters need not employ large transistors.

The two quadrature generation techniques introduced here are somehwat similar in performance, with the former (based on direct generation) exhibiting a 1.5-dB penalty.

4.10 Ring Oscillators with LC Loads

The ring oscillators studied in the previous sections face frequency limitations arising from both the number of stages and the load capacitance seen by each stage. In applications where the ring must provide multiple phases, these limitations become serious.

What happens if we replace the resistive loads in a differential ring with LC tanks? Consider the structure depicted in Fig. 4.42(a), where the feedback is negative at low frequencies. Barkhausen's criterion dictates

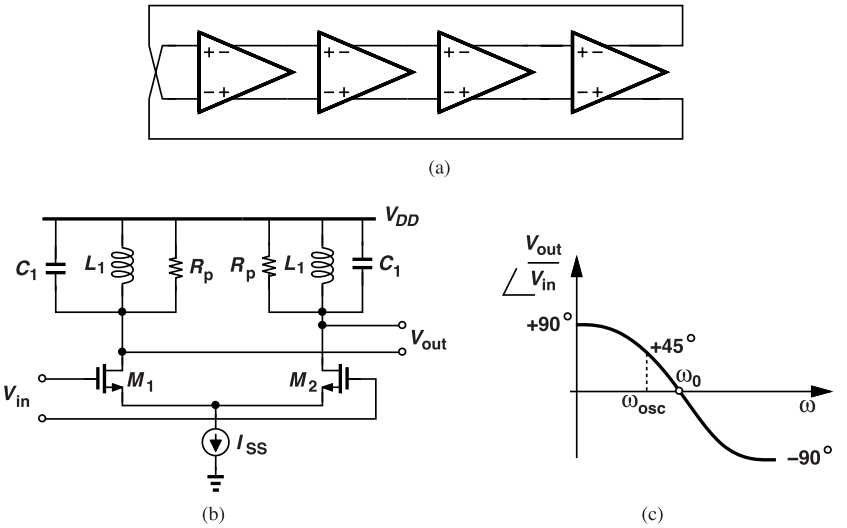


Figure 4.42 (a) Four-stage ring, (b) inductively-loaded stage, and (c) phase response of one stage.

that the four stages produce a total phase shift of 180° at the oscillation frequency. If implemented as shown in Fig. 4.42(b), each stage exhibits a resonance frequency given by $\omega_0 = 1/\sqrt{L_1 C_1}$ but no phase shift at ω_0. Thus, the oscillation frequency must depart from ω_0 so as to yield a phase shift of 45° per stage [Fig. 4.42(c)] [3]. To calculate the oscillation frequency, ω_{osc}, we equate the phase shift of the parallel RLC tank to 45°:

$$\frac{\pi}{2} - \tan^{-1} \frac{L_1 \omega_{osc}}{R_p(1 - L_1 C_1 \omega_{osc}^2)} = \frac{\pi}{4}. \tag{4.18}$$

If we approximate $R_p/(L_1\omega_{osc})$ by the Q at ω_0, we have (Problem 4.19)

$$\omega_{osc} = \sqrt{1 - \frac{1}{Q}}\, \frac{1}{\sqrt{L_1 C_1}}. \tag{4.19}$$

Alternatively, if the departure, $\Delta\omega = \omega_0 - \omega_{osc}$, is small, we can express the Q of the tank near ω_0 as $Q = (\omega_0/2)d\phi/d\omega$ and, write $d\phi \approx \pi/4$. That is,

$$\Delta\omega \approx \frac{\pi}{8Q}\omega_0. \tag{4.20}$$

For example, if $Q = 5$, then (4.19) predicts a departure of 11% and (4.20), 8%.

The LC-load ring oscillator offers three advantages over previous topologies. First, with little dc voltage drop across the inductive loads, the circuit can operate with lower supply voltages. This is because the input and output CM levels are close to V_{DD} here, i.e., $V_{GS} + V_{ISS} \approx V_{DD}$, where V_{ISS} is the voltage across I_{SS}. Second, the oscillation frequency is only a weak function of the number of stages. For N stages in the loop, $\Delta\omega$ must provide a phase shift of π/N per stage, assuming the following value:

$$\Delta\omega \approx \frac{\pi}{2QN}\omega_0. \tag{4.21}$$

Interestingly, as N increases, ω_{osc} approaches ω_0. Third, the load capacitance can be absorbed by the tanks, leading to a higher oscillation frequency than that of resistively-loaded rings. The drawback of the circuit is the need for multiple spiral inductors. This issue can be alleviated through the use of stacked structures [3].

References

[1]. A. Hajimiri, S. Limotyrakis, and T. H. Lee, "Jitter and phase noise in ring oscillators," *IEEE J. Solid-State Circuits*, vol.34, pp. 790-804, June 1999.

[2]. J. Alvarez et al., "A wide-bandwidth low-voltage PLL for PowerPC microprocessors," *IEEE J. Solid-State Circuits*, vol. 30, pp. 383-391, April 1995.

[3]. J. Savoj and B. Razavi, "A 10-Gb/s CMOS clock and data recovery circuit with a half-rate binary phase/frequency detector," *IEEE J. of Solid-State Circuits,* vol. 38, pp.13-21, Jan. 2003.

Problems

4.1. Consider the ring shown in Fig. 4.1. Suppose each node has a drain junction capacitance,

$$C_j(V) = \frac{C_0}{\left(1 + \frac{V_D}{V_0}\right)^m}, \tag{4.22}$$

where V_D denotes the drain voltage, $V_0 \approx 0.7$ V, and m is in the range of 0.3 to 0.4. For simplicity, assume the oscillator voltage swings are small so that V_D is approximately equal to the common-mode level, $V_{DD} - R_D I_{SS}/2$. Determine $\partial C_j / \partial V_{DD}$.

4.2. Using the result obtained in the previous problem and modeling the other capacitances by a constant, single-ended capacitor, C_p, to ground, determine $K_{VDD} = \partial f_0 / \partial V_{DD}$. For f_0, use the small-signal approximation obtained in Chapter 1.

4.3. Repeat the previous problem if we increase C_p to reduce f_0, e.g., if f_0 is lowered by a factor of 10. Normalize the supply sensitivity to f_0.

4.4. Assume the ring in Fig. 4.1 is designed for a certain oscillation frequency. We double the transistor widths and I_{SS} and halve R_D. Explain what happens to the supply sensitivity of the circuit.

4.5. In Fig. 4.4, the ring oscillator has low supply sensitivity, but the inverters do not. Returning to the delay-line phase noise considerations in Chapter 2, explain how the supply noise affects the inverters' outputs.

4.6. Suppose we double the transistor widths and the tail currents and halve the load resistors in a three-stage differential ring. Explain from Eq. (4.1) what happens to the phase noise.

4.7. In Fig. 4.6, we have assumed that the first and second harmonics are "in-phase," i.e., both are expressed as sines. Explain what happens if the second harmonic has the form $V_2 \cos 2\omega_0 t$.

4.8. For the circuit in Fig. 4.9(a), explain why the output common-mode level is equal to $V_{DD} - I_{SS} R_D/2$ regardless of whether the transistors switch completely or not.

4.9. Let us apply Eq. (4.1) to the ring in Fig. 4.10(a), but by replacing $R_D I_{SS}$ with the actual single-ended swing observed in Fig. 4.10(b) (about 150 mV). Does this remedy eliminate the 9.5-dB discrepancy in the phase noise calculation at 100-MHz offset?

4.10. Suppose the transistor widths in Fig. 4.10(a) are so large that $V_{GS} - V_{TH} \approx 50$ mV. We can assume that M_1 and M_2 experience *abrupt* and complete switching, and their drain currents resemble a square wave. Repeat Example 4.6 for this case.

4.11. The output waveforms of Fig. 4.15(b) exhibit asymmetric rise and fall times due to a second harmonic level of -12 dB. Using Example 4.4 and Eq. (4.5), determine f_{1/f^3} interms of $f_{1/f}$.

4.12. In Fig. 4.27(a), a charge of ΔV in V_{DD} causes approximately the same change in the source voltages of M_3 and M_4 (why?). Suppose M_5 presents a small-signal resistnace of r_{eq} from this node to ground. Determine the change in the negative resistance and the change in f_0 if V_{DD} changes by ΔV. Assume square-law devices.

4.13. We wish to repeat the previous problem if the tail current of M_1 and M_2 exhibits a small-signal resistance of r_{eq}. (Note that, in a ring oscillator environment, the input CM level of M_1 and M_2 changes by ΔV if the supply voltage of the *preceding* stage changes by the same amount. That is, the oscillator's supply variations appear at the source nodes of M_1 and M_2.) Do we expect significant frequency modulation in this case?

4.14. Modeling the flicker noise of M_5 in Fig. 4.27(a) by a gate-referred voltage source, $\overline{V_n^2} = K/(WLC_{ox}f)$, and neglecting other noise sources, compute the phase noise of the oscillator. Assume square-law devices.

4.15. Modeling the flicker noise of I_1 in Fig. 4.30(c) by a current source I_{n^2} and neglecting other noise sources, determine the phase noise of the oscillator. (Hint: Example 4.14 proves useful.)

4.16. In Fig. 4.32(a), the supply voltage of the inverters varies slowly. Which statement is true: (a) the oscillation frequency is modulated but the phase relationships in Fig. 4.32(c) remain intact, or (b) the oscillation frequency remains relatively constant but the phase relationships are distrubed.

4.17. Consider the topology shown in Fig. 4.43, where the outer inverters are chosen weaker than the inner ones. Does this circuit prefer to latch up or oscillate?

4.18. We wish to use a single three-stage ring along with interpolation so as to generate quadrature phases (Fig. 4.44). Draw the interpolation network and explain how the strengths of the interpolating inverters must be weighted.

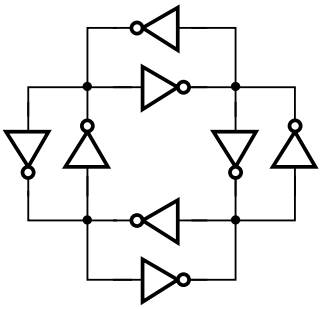


Figure 4.43 Two rings connected in anti-parallel form.

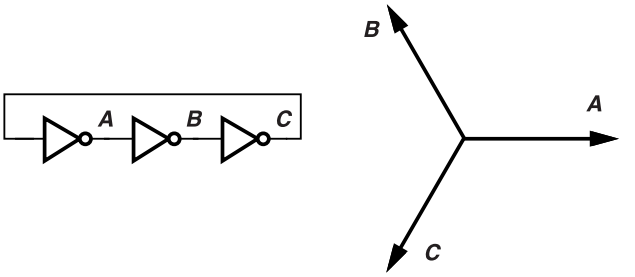


Figure 4.44 Three-stage ring for interpolation.

4.19. Equation (4.18) can be written as

$$\frac{L_1\omega_{osc}}{R_p(1-L_1C_1\omega_{osc}^2)} = 1. \tag{4.23}$$

Assuming $Q \approx R_p/(L_1\omega_{osc})$, compute ω_{osc}.

4.20. We wish to apply linear scaling to the oscillator shown in Fig. 4.42 so as to lower the phase noise by 3 dB. Explain how the devices must be scaled.

5

LC Oscillator Design

Many applications require the use of LC oscillators, primarily because these topologies exhibit less phase noise than ring oscillators do (for the same power consumption). Other advantages of LC topologies include higher oscillation frequencies—not limited by $1/(2NT_D)$—and voltage excursions beyond the supply rails. The disadvantages are fourfold: (a) LC oscillators have a much narrower tuning range, (b) integrated inductors have a large "footprint," occupying significant chip area and creating difficulty in routing of signals, (c) such inductors suffer from substantial unwanted coupling to and from the substrate and other circuits, and (d) the design of LC oscillators heavily relies on proper inductor and varactor models, often requiring margins in the design to hit the targeted oscillation frequency.

Drawing upon concepts developed in Chapters 1 and 2, our study of LC oscillators in this and the next chapters parallels that of ring oscillators in Chapters 3 and 4, which the reader is encouraged to review. We deal with phase noise in LC oscillators, describe various topologies, analyze their properties, and, using simulations, delve into their power-phase noise trade-offs. But we first take a detour to introduce a simple model for inductors.

5.1 Inductor Modeling

Integrated inductors have been studied extensively [1]-[5]. In this section, we deal with only some concepts necessary in oscillator design and refer the reader to [6] for additional details.

The most commonly used inductor geometry in oscillator design is the symmetric spiral shown in Fig. 5.1(a). If driven differentially, this structure exhibits a higher quality factor (Q) than asymmetric topologies [7], a critical benefit for reducing phase noise. In practice, the inductor is realized in an octagonal shape to minimize the series resistance for a given inductance. Our objective is to develop a circuit model that yields reasonably accurate results in oscillator design.

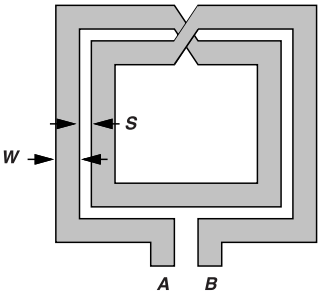


Figure 5.1 Symmetric inductor structure.

Three attributes of an integrated inductor must be reflected in the model: (1) the inductance value, (2) the parasitic capacitances surrounding the structure, and (3) the quality factor. The inevitably narrow tuning range

of LC oscillators dictates that, in $\omega_0 = 1/\sqrt{LC}$, both L and C be set accurately. A number of expressions have been reported for the inductance of spiral inductors [1, 2] but they incur 10% to 20% error in some cases. That in [1] can be rewritten as

$$L = 1.3 \times 10^{-7} \frac{l_{tot}^{5/3}}{\left[\frac{l_{tot}}{4N} + W + (N-1)(W+S)\right]^{1/3} W^{0.083}(W+S)^{0.25}} \text{ nH}, \tag{5.1}$$

where l_{tot}, N, W, and S denote the total length of the trace, the number of turns, the trace width, and the trace spacing, respectively. Such equations prove useful for only quick estimates of the inductor dimensions, but in practice, we must resort to electromagnetic field simulators such as Ansoft's HFSS to compute a given structure's inductance.

The parasitic capacitances associated with spiral inductors consist of two components: the capacitance to the substrate and the capacitance between the windings (Fig. 5.2). In the differential topology of Fig. 5.1, the

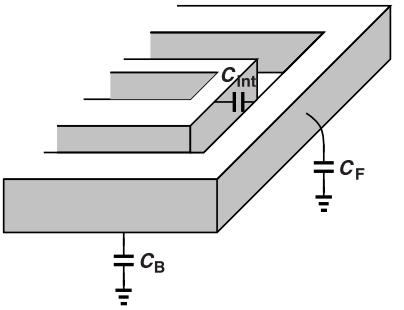


Figure 5.2 Parasitic capacitances of spiral inductors.

latter component is dominant because of the large potential difference between adjacent turns. This can be seen from Fig. 5.3, where a differentially-excited spiral is "unwrapped" in the form of four inductor segments:

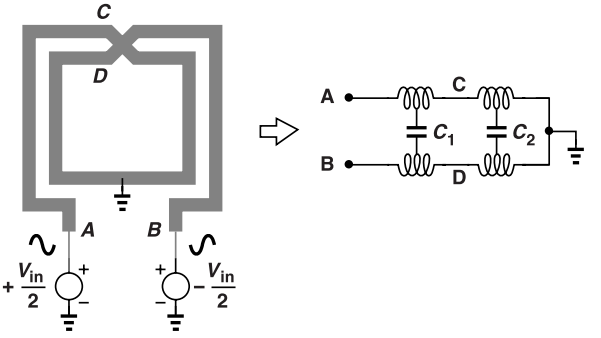


Figure 5.3 Symmetric inductor with differential excitation.

segment AC is modeled by inductor AC, etc.[1] This model reveals that C_1 sustains a larger voltage difference than C_2. Thus, C_1 and C_2 cannot be simply added to each other. A three-turn structure, for example, exhibits an equivalent capacitance given by

$$C_{eq} = \frac{5}{18} C_{tot}, \tag{5.2}$$

[1] In reality, there is also mutual coupling among these inductor segments.

where C_{tot} denotes the total capacitance present between adjacent windings [6]. The factor $5/18$ transforms this distributed capacitance to a lumped representation (Fig. 5.4).

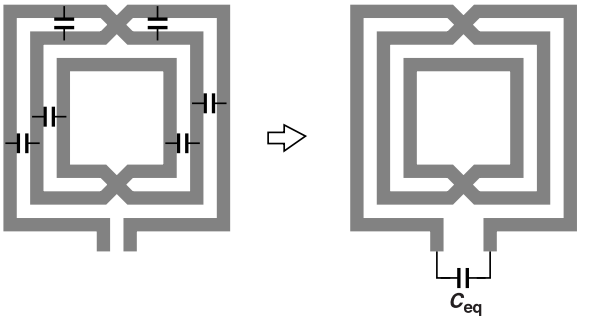


Figure 5.4 Equivalent lumped capacitance of an inductor.

Accurate modeling of the inductor's parasitic capacitances is only somewhat critical because they constitute a small fraction of the total C in $\omega_0 = 1/\sqrt{LC}$; after all, a typical design includes transistor capacitances and also adds substantial capacitance for tuning. The parasitic capacitances can also be calculated by electromagnetic field simulators. If, instead, a standard layout extraction tool is used to obtain C_{tot}, then this value must be scaled down properly—as in Eq. (5.2)—to arrive at a lumped model.

The third property of inductors, namely, the Q, plays a central role in oscillator phase noise (Section 5.2) and must be modeled with a relatively small error (e.g., $\pm 10\%$). The quality factor of inductors relates to the thermal losses that they incur upon passing a current. For example, the resistance of the metal trace comprising the spiral converts electric energy to heat, behaving as the friction effect in a pendulum (Chapter 1). In this context, we can model the inductor as shown in Fig. 5.5 and define the Q as the ratio of the desired



Figure 5.5 Simple inductor model including trace resistance.

impedance and the undesired impedance:

$$Q = \frac{L_1 \omega}{R_s}. \tag{5.3}$$

Equation (5.3) implies that the Q rises linearly with ω, but at frequencies above a few gigahertz, several other effects counteract this trend. First, the skin effect within the metal wire causes the oscillation current to crowd near the edges [Fig. 5.6(a)], thereby raising R_s (roughly in proportion to $\sqrt{\omega}$). Second, the inductor loses energy to the silicon substrate as well. Illustrated in Fig. 5.6(b), one such mechanism stems from the capacitance between the spiral and the distributed substrate resistance; the voltage excursions on the inductor couple through the capacitance to this resistance, producing loss. Another mechanism, shown in Fig. 5.6(c), is due to the magnetic coupling between the inductor and the substrate—as if the two formed the primary and secondary of a transformer, respectively. In this case, a time-varying *current* in the inductor induces another in the substrate, creating thermal loss. We remark that the coupling phenomena in Figs. 5.6(b) and (c) become more pronounced as the frequency increases. The former can be reduced through the use of a "patterned ground shield" [8], as illustrated in Fig. 5.7, where the broken metal lines prohibit the electric field lines from reaching the substrate. The cost is a greater capacitance to the substrate.

The foregoing observations provide insight into what limits the Q but do not help much with its modeling. In practice, we rely on electromagnetic field simulations for this aspect of inductor design as well.

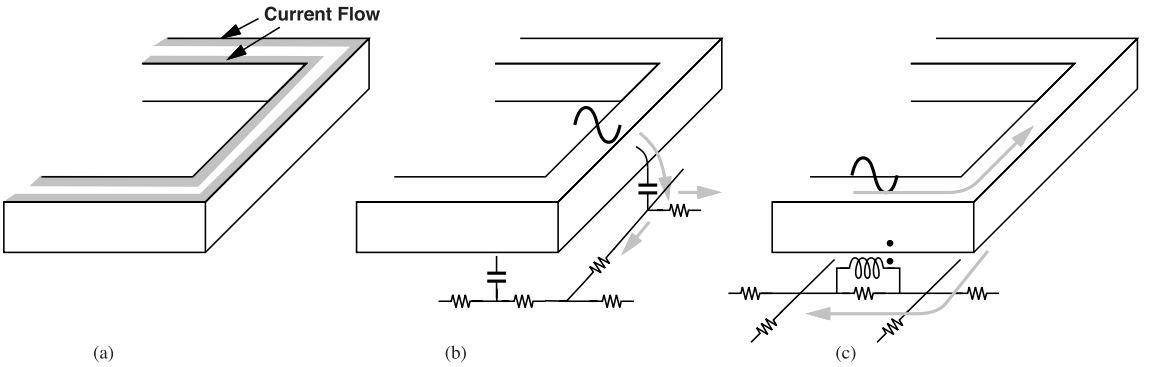


Figure 5.6 (a) Skin effect in an inductor, (b) capacitive coupling to the substrate, and (c) magnetic coupling to the substrate.

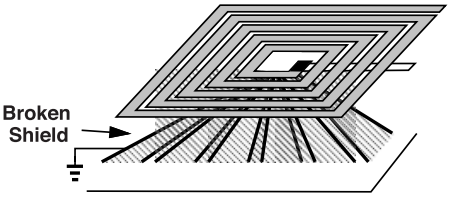


Figure 5.7 Spiral inductor with patterened ground shield.

Example 5.1

An inductor incurs negligible series resistance but significant substrate loss. How can this inductor be modeled?

Solution

For electric coupling to the substrate, we can construct the lumped model shown in Fig. 5.8(a), but how do we compute C_1, C_2, and R_{sub}? This is difficult because the original elements are distributed across the spiral

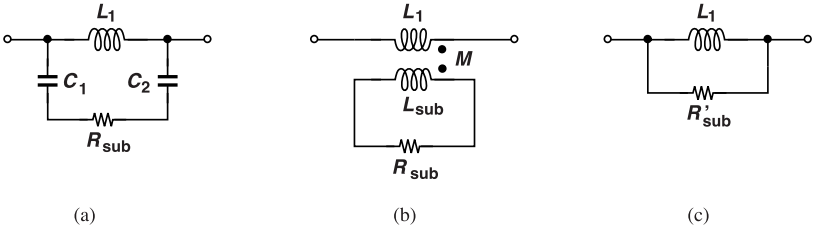


Figure 5.8 Inductor model with (a) capacitive coupling to the substrate, (b) magnetic coupling to the substrate, and (c) a single resistor representing both mechanisms .

while we wish to represent them by a lumped model. Thus, we view these values as fitting parameters, i.e., as quantities that allow us to match the model to measured or simulated results.

For magnetic coupling, we invoke the equivalent circuit depicted in Fig. 5.8(b), with M, L_{sub}, and R_{sub} unknown. Now, since the transformer presents to L_1 a scaled version of R_{sub}, we simplify the circuit as shown in Fig. 5.8(c) for some frequency range. Here, R'_{sub} includes both losses. These considerations point to inductor modeling challenges in general.

The above example suggests that substrate loss manifests itself as a resistance in *parallel* with the inductor. Let us generalize this view and define the Q for the arrangement shown in Fig. 5.9 as

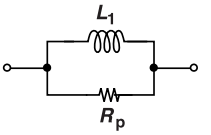


Figure 5.9 Simple parellel inductor model.

$$Q = \frac{R_p}{L_1\omega}, \tag{5.4}$$

so that $Q \to \infty$ as $R_p \to \infty$ and the inductor becomes more ideal. While called a fictitious quantity in Chapter 2, R_p in fact can represent the substrate loss for integrated inductors. Note that this Q *falls* as ω rises.

Example 5.2

A student surmises that inductor models necessary for LC oscillator design need not be valid for a wide frequency range. Is this true?

Solution

Not quite. Recall from Chapter 1 that the currents in an LC oscillator can resemble square waves, exhibiting relatively strong harmonics. Now consider the series-resistance model in Fig. 5.10(a), where I_{in} is a square wave, and note that here $Q = L_1\omega/R_S$ *increases* linearly at higher harmonics, but we know that the actual

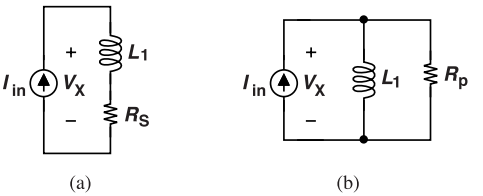


Figure 5.10 Excitation of series and parellel models by a square-wave current source.

Q begins to saturate at high frequencies. The resulting voltage, V_X, therefore exhibits unrealistically sharp edges in this model. The converse occurs for the parallel-resistance model in Fig. 5.10(b). We must therefore include both a series resistance and a parallel resistance.

From our studies emerges the simple and relatively accurate model illustrated in Fig. 5.11(a), with the assumption that the substrate is grounded. The series and parallel resistances represent physical effects and better model the harmonic behavior described in the above example. Fitted to electromagnetic field simulation results, the values are symmetric for the spiral of Fig. 5.1. Note that the model in Fig. 5.11(a) resembles that in Fig. 5.9 at very high frequencies, at which $L_s\omega$ overwhelms R_s, and C_1 and C_2 act as short circuits. That is, the Q asymptotically approaches $(R_1 + R_2)/(L_s\omega)$ as the frequency rises if we assume R_S increases negligibly.

It is difficult to define a Q for this model based on the component values, but for oscillator design, we proceed as follows. First, we add sufficient capacitance between terminals A and B to create resonance at the desired frequency [Fig. 5.11(b)]. Next, we measure the impedance $Z_{AB} = V_X/I_X$ in simulations and plot its magnitude and phase [Fig. 5.11(c)]. The overall Q for oscillator design purposes is given by $(\omega_0/2)d(\angle Z_{AB})/d\omega$ (Chapter 1). Alternatively, we can divide ω_0 by the 3-dB bandwidth of $|Z_{AB}|$ as an approximation of the Q.

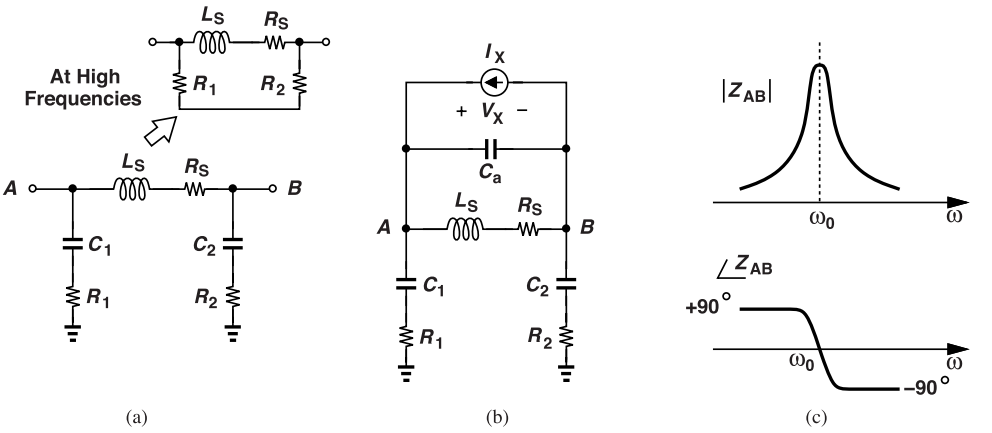


Figure 5.11 (a) Basic inductor model, (b) excitation of inductor in parallel with C_a, and (c) impedance behavior.

The Q of symmetric inductors integrated in standard CMOS technology begins around 3 or 4 at 1 GHz and rises to about 8 at 5 GHz, 10 at 10 GHz, and 15 at 20 GHz. Processes that provide thick metal layers offer higher Q values.

Example 5.3

A 5-nH inductor exhibits a Q of about 8 at 5 GHz with a lumped capacitance of 50 fF. Construct a circuit model for this inductor.

Solution

As seen earlier, the Q must be modeled by both a series resistance, R_s, and a parallel resistance, R_p. As a simple approximation, we allocate half of the loss to the former and the other to the latter. That is, we assume a Q of 16 for calculating both R_s and R_p. Since $Q = L\omega/R_s = R_p/L\omega$, we obtain $R_s = 9.8\ \Omega$ and $R_p = 2.5\ \text{k}\Omega$. The overall model is shown in Fig. 5.12.

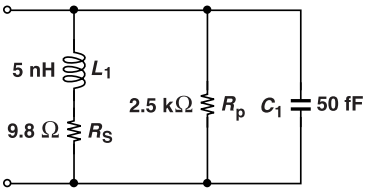


Figure 5.12 Example of inductor model.

The model developed in the foregoing example serves as a good starting point in LC oscillator design. We use this type of model for our studies in this chapter.

5.2 Phase Noise Analysis

The phase noise principles described in Chapters 2, 3, and 4 form the foundation for our study here. We wish to analyze the phase noise behavior of the simple cross-coupled LC oscillator introduced in Chapter 1, assuming, for now, that (a) the transistors do not enter the triode region, and (b) the tail current source is ideal

and noiseless. After finishing this chapter, the reader is encouraged to study another approach to phase noise analysis described in Chapter 6.

Our analysis must answer two questions. (1) How do the devices inject noise into the oscillatory waveform? (2) How does the injected noise convert to phase noise? The challenge facing us is that the transistors turn on and off in every cycle, injecting noise only some of the time. We remind the reader that a signal $x(t)$ having a spectrum $S_X(f)$ and subjected to a transfer function $H(s)$ yields an output spectrum given by $S_X(f)|H(2\pi jf)|^2$. This section is relatively long as we must build up our understanding gradually and rigorously.

5.2.1 A Simple Case

Let us begin with a simple, ideal tank, as shown in Fig. 5.13(a). Suppose the circuit produces an oscillatory

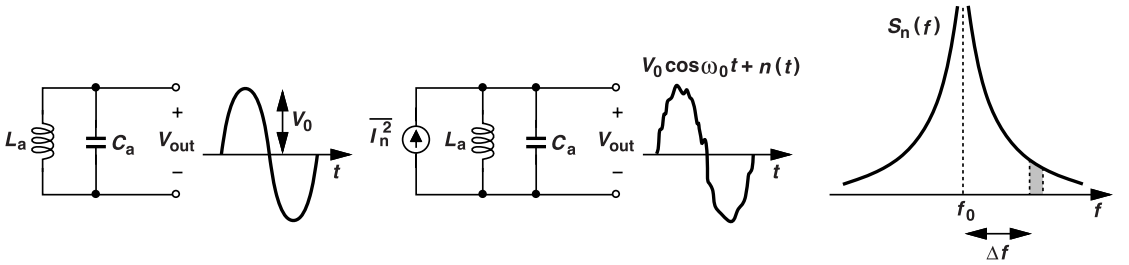


Figure 5.13 (a) Ideal LC tank released with an initial condition, (b) injection of noise into the tank, and (c) output noise spectrum.

output as a result of an initial condition. Now, we inject a noise current as in Fig. 5.13(b) and wish to determine the phase noise in $V_{out}(t)$. We intuitively expect that $\overline{I_n^2}$ creates an additive noise voltage, $n(t)$, at the output. As an example, let us assume $\overline{I_n^2}$ is white. The tank impedance is given by $L_a s/(L_a C_a s^2 + 1)$, which, for $s = j\omega$, assumes the form $jL_a\omega/(1-L_a C_a\omega^2)$. Since we are interested in frequencies near $\omega_0 = 1/\sqrt{L_a C_a}$, we write $\omega = \omega_0 + \Delta\omega$, $\Delta\omega \ll \omega_0$, simplifying the impedance to $jL_a(\omega_0 + \Delta\omega)/(-2L_a C_a\omega_0\Delta\omega - L_a C_a\Delta\omega^2) \approx -j/(2C_a\Delta\omega)$.

We compute the output noise voltage spectrum, $S_n(f)$, as the spectrum of $\overline{I_n^2}$ multiplied by the magnitude squared of the impedance:

$$S_n(\Delta f) = \overline{I_n^2} \frac{1}{4C_a^2(2\pi\Delta f)^2}, \tag{5.5}$$

where $2\pi\Delta f$ is replaced for $\Delta\omega$. Depicted in Fig. 5.13(c), this spectrum *adds* to the oscillatory output waveform as a voltage quantity.

Conversion of Additive Noise to Phase Noise The overall output voltage in Fig. 5.13(b) is given by $V_{out}(t) = V_0\cos\omega_0 t + n(t)$, where $n(t)$ has the narrowband spectrum shown in Fig. 5.13(c) and given by Eq. (5.5). We must quantify how the phase of $V_{out}(t)$ is corrupted by $n(t)$. It can be proved that band-pass noise whose spectrum is centered around $\omega_0 = 2\pi f_0$ can be expressed as

$$n(t) = n_I(t)\cos\omega_0 t - n_Q(t)\sin\omega_0 t, \tag{5.6}$$

where $n_I(t)$ and $n_Q(t)$ are called the quadrature components of $n(t)$ and have a low-pass spectrum [11]. This operation can be visualized as quadrature upconversion [Fig. 5.14(a)]. As shown in Fig. 5.14(b), $n_I(t)\cos\omega_0 t$ and $n_Q(t)\sin\omega_0 t$ exhibit quadrature envelopes. To obtain the spectra of n_I and n_Q, we simply shift the

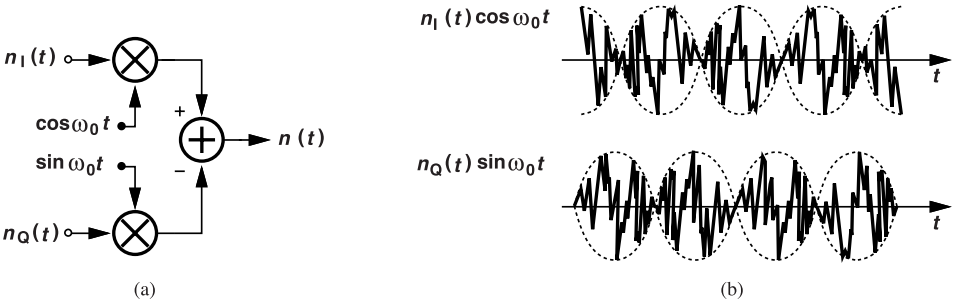


Figure 5.14 (a) Viewing $n(t)$ as a quadrature-modulated signal, and (b) quadrature envelopes of n_I and n_Q.

one-sided spectrum of $n(t)$ to zero (Fig. 5.15).[2] For the band-pass noise given by Eq. (5.5), the quadrature components have a spectrum equal to $\overline{I_n^2}/[4C_a^2(2\pi f)^2]$ around $f = 0$.

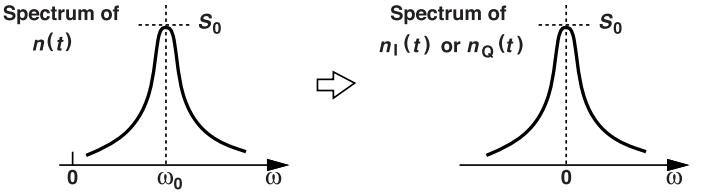


Figure 5.15 Relation between one-sided spectrum of band-pass noise and the spectrum of its quadrature components.

The output waveform in Fig. 5.13(b) is then written as

$$V_{out}(t) = V_0 \cos\omega_0 t + n(t) \tag{5.7}$$

$$= [V_0 + n_I(t)]\cos\omega_0 t - n_Q(t)\sin\omega_0 t \tag{5.8}$$

$$= \sqrt{[V_0 + n_I(t)]^2 + n_Q^2(t)}\cos\left[\omega_0 t + \tan^{-1}\frac{n_Q(t)}{V_0 + n_I(t)}\right]. \tag{5.9}$$

The output contains both amplitude and phase modulation. This can also be seen from the phasor diagram shown in Fig. 5.16, where $n_I(t)$ perturbs the length of the carrier phasor, and $n_Q(t)$, its phase.[3] The phase

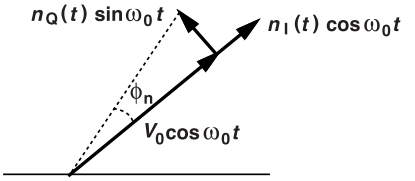


Figure 5.16 Phasor diagram showing the AM and PM components arising from additive noise.

noise emerges as

$$\phi_n(t) = \tan^{-1}\frac{n_Q(t)}{V_0 + n_I(t)} \tag{5.10}$$

[2] For clarity, we have shown these spectra with a peak of η whereas the spectrum in Fig. 5.13(c) has an infinite peak.

[3] If $n_Q(t)\sin\omega_0 t$ is drawn from the tip of $n_I(t)\cos\omega_0 t$, then we observe that n_Q also perturbs the amplitude, as suggested by Eq. (5.9).

$$\approx \frac{n_Q(t)}{V_0}. \tag{5.11}$$

The reader can show that this result can also be derived by dividing $n_Q(t)$ by the slope of the waveform at the zero crossings. The spectrum of $\phi_n(t)$ is simply that of $n_Q(t)$ scaled down by V_0^2. It follows from (5.5) that

$$S_{\phi n}(f) = \overline{I_n^2} \frac{1}{4C_a^2(2\pi f)^2} \cdot \frac{1}{V_0^2}. \tag{5.12}$$

This is the phase noise observed in the oscillatory output. Note that we have replaced $\Delta f = f - f_0$ with f because $S_{\phi n}$ is centered around the zero frequency.

Let us go one step further, neglect n_I and n_Q in the amplitude of the waveform, and write Eq. (5.9) as $V_{out} \approx V_0 \cos[\omega_0 t + n_Q(t)/V_0]$. The spectrum of V_{out} is the same as that of $n_Q(t)/V_0$ but translated from a zero center frequency to ω_0.

Figure 5.17 summarizes the steps involved in our computation. We inject the noise, I_n, as an additive

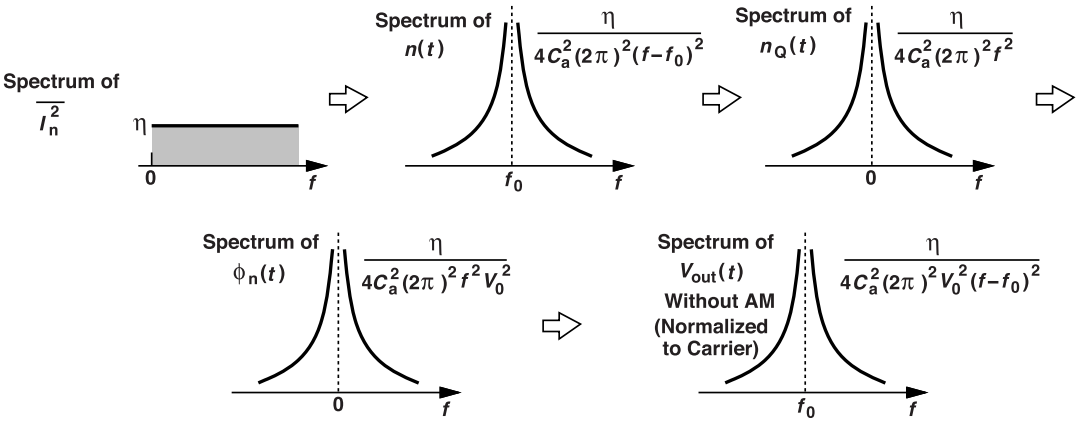


Figure 5.17 Transformation of injected noise current to output voltage spectrum.

component, convert it to voltage, decompose it to quadrature components, find its contribution to the phase of the output, and translate the phase noise spectrum to a center frequency of f_0 to obtain the spectrum of V_{out} (without AM components). Note that our calculations begin with the *one-sided* spectrum of I_n^2. Also, as explained in Chapter 2, the spectrum of V_{out} normalized to the carrier power is the same as that of ϕ_n.

In the study that follows, we wish to model by $\overline{I_n^2}$ the noise of the transistors in an LC oscillator. The difficulty is that the transistors turn on and off, generating noise whose "strength" *varies* during the oscillation period. We must therefore deal with "cyclostationary" noise.

5.2.2 Cyclostationary Noise

Let us examine how the transistors in the oscillator of Fig. 5.18(a) inject noise into the tanks, noting that their transconductance varies periodically with time. Since the drain noise current spectrum is equal to $4kT\gamma g_m$, we recognize that M_1 and M_2 produce noise only when they are on. Moreover, when, for example, M_2 is completely off, M_1 does *not* inject noise into its tank because it is heavily degenerated by the tail current source [Fig. 5.18(b)] (Problem 5.6). We can alternatively view the circuit as a differential pair receiving input swings equal to those of the oscillation waveforms [Fig. 5.18(c)].[4] As illustrated in Fig. 5.18(d), when the

[4]This view is valid only for studying the injection but not for phase noise calculations because the open-loop phase noise and closed-loop phase noise are very different (Chapter 3).

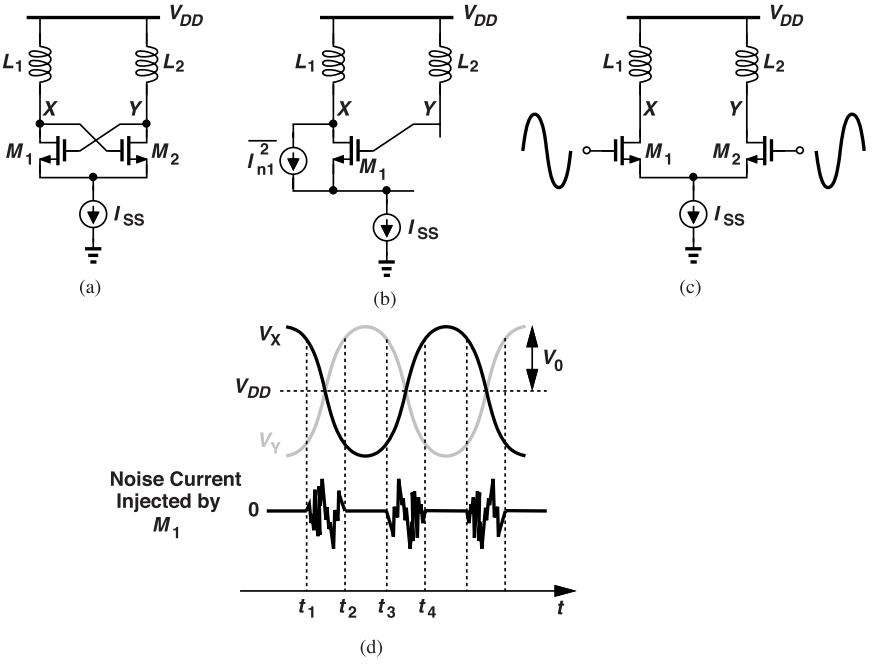


Figure 5.18 (a) Simple LC oscillator, (b) noise current of one transistor, (c) view of transistors as a differential pair, and (d) noise current injected by transistors as a function of time.

gate voltages are at the extremes, $V_{DD} \pm V_0$, M_1 contributes no noise because it is either off or degenerated by I_{SS}. Thus, it is only when M_1 and M_2 are simultaneously on, from t_1 to t_2 or from t_3 to t_4, that they inject noise into the tanks. The highest noise injection occurs around the crossing points of V_X and V_Y. We say the transistors exhibit cyclostationary noise, i.e., periodically-switched noise.

An interesting question that we face is how to measure the spectrum of periodically-switched noise. Recall from Chapter 2 that, to construct the spectrum, we pass the noise through a 1-Hz filter and measure the average output power. Since the output of the filter has a 1-Hz bandwidth and since a signal with a 1-Hz bandwidth reveals its dynamics (changes significantly) for a time scale on the order of 1 second, we surmise that the averaging should be performed for several seconds. In practice, we choose a filter with a bandwidth of, say, 1 kHz, perform the averaging for some milliseconds, and divide the resulting power reading by 1000 to obtain the power in 1 Hz. The key point here is that the spectrum is measured over a time scale much greater than the switching period in Fig. 5.18(d); we do not attempt to measure the spectrum only from t_1 to t_2.

For further study of the noise of the transistors, we should learn how to manipulate cyclostationary noise. The following example is our first step.

Example 5.4

The ambient temperature of a resistor R_1 is periodically and instantaneously switched between ≈ 0 K and T_1 at a frequency of f_0. If the switching duty cycle is 50%, determine the spectrum of the noise produced by R_1. Assume the resistor tracks the ambient temperature at all times.

Solution

The resistor generates negligible noise for half of the time and a two-sided noise spectral density of $2kT_1R_1$ for the other half. From the time-domain waveform in Fig. 5.19(a), we note that the result, $V_n(t)$, is equal to the product of a continuous noise waveform, $V_1(t)$, and a square wave toggling between 0 and 1, $P(t)$. The spectra of $V_1(t)$ and $P(t)$ are therefore convolved [Fig. 5.19(b)]. As the white spectrum is convolved with

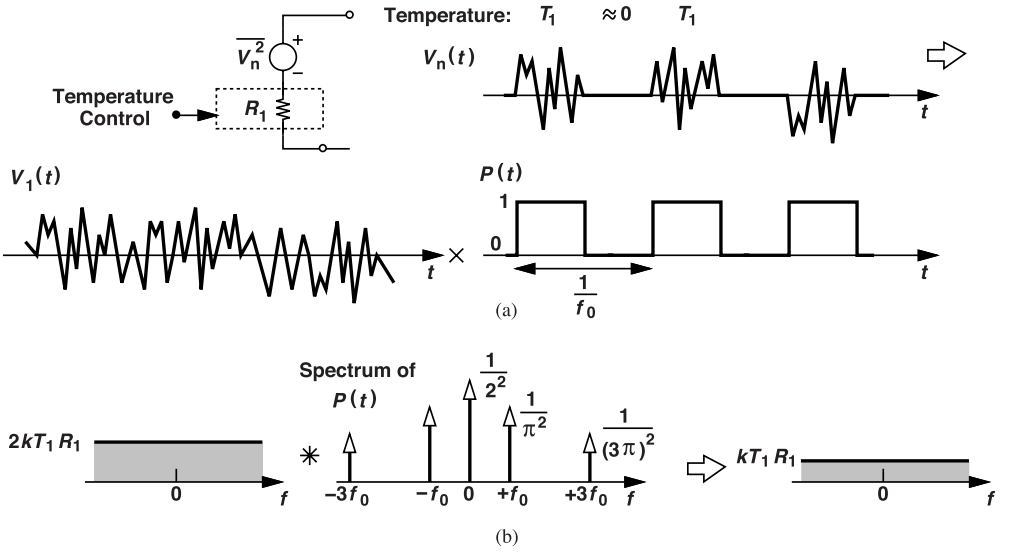


Figure 5.19 (a) Periodically-switched thermal noise, and (b) resulting output spectrum.

each impulse in the spectrum of $P(t)$, it is shifted to the right or to the left and added to other shifted replicas, while still remaining white. The key observation here is that periodically-switched white noise is white.

Let us now calculate the overall noise spectral density. Since the switching action removes the noise for half of the time, we surmise that the output power is half of the unswitched noise power. That is, the two-sided output spectral density should be equal to kT_1R_1. This can also be proven by carrying out the convolution in Fig. 5.19(b). We first note the Fourier transform of $P(t)$ contains an impulse at $f = 0$ with an area of $1/2$ (the dc value), and impulses at the odd harmonics. The first harmonic consists of two impulses, each with an area of $1/\pi$. But we must not confuse the Fourier transform with the spectrum (the power spectral density). What area do we attribute to the impulses in the spectrum of $P(t)$? For example, should we assume an area of $1/\pi$ for the first harmonic? Let us ask, what would we do if $P(t)$ were just a constant (dc) value, a? Since $V_1(t)$ in Fig. 5.19(a) would be multiplied (scaled) by a, we must convolve the power spectral density of V_1 with $a^2\delta(f)$. Thus, the nth harmonic of $P(t)$ must be represented by an area of $1/(n\pi)^2$ rather than by $1/(n\pi)$ if the spectrum of $P(t)$ is of interest.

The two-sided output spectrum then emerges as

$$S_{out}(f) = 2kT_1R_1\left[\frac{1}{4} + \frac{2}{\pi^2} + \frac{2}{(3\pi)^2} + \frac{2}{(5\pi)^2} + \cdots\right], \tag{5.13}$$

where the term $1/4$ accounts for the dc value of $P(t)$ and the factor of 2 in the numerators is due to the impulses at $\pm nf_0$. It follows that

$$S_{out}(f) = 2kT_1R_1\left[\frac{1}{4} + \frac{2}{\pi^2}\left(1 + \frac{1}{3^2} + \frac{1}{5^2} + \cdots\right)\right]. \tag{5.14}$$

The reader is encouraged to prove that $1 + 3^{-2} + 5^{-2} + \cdots = \pi^2/8$. We therefore have

$$S_{out}(f) = kT_1R_1, \tag{5.15}$$

as expected. Note that this is the two-sided spectrum.

A more general result is that, if white noise is periodically turned on for t_1 seconds over a period of t_0, the resulting spectrum is white and its spectral density is scaled by a factor of t_1/t_0, which is the switching duty cycle.

In addition to cyclostationary noise, we are also interested in periodically-switched resistances. The following example proves useful.

Example 5.5

Plot as a function of time the negative resistance presented by M_1 and M_2 in Fig. 5.18(a).

Solution

The negative resistance is equal to $-2/g_m$ at the crossing points of V_X and V_Y and nearly minus infinity at the peaks. The result is shown in Fig. 5.20. The reader is encouraged to explain why the resistance does not go to *plus* infinity at the peaks.

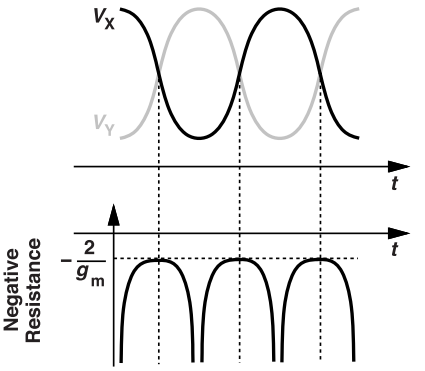


Figure 5.20 Variation of the cross-coupled pair's negative resistance with time.

The resistance seen at the drains of the cross-coupled pair varies periodically with time. Can we write a Fourier expansion for it?

5.2.3 Noise Injected by Cross-Coupled Pair

Norton Noise Equivalent We now formulate the noise produced by M_1 and M_2 in the oscillator shown in Fig. 5.21. To determine how long the transistors remain on simultaneously, we make two observations: (1) Fig. 5.21 suggests that one transistor turns off when the differential voltage, V_{XY}, reaches $\sqrt{2}(V_{GS} - V_{TH})_{eq}$, where $(V_{GS} - V_{TH})_{eq}$ denotes the overdrive in equilibrium, i.e., when $V_{XY} = 0$;[5] (2) V_{XY} can be approximated as $V_1 \sin \omega_0 t$, where $V_1 = (4/\pi)R_p I_{SS}$ (Chapter 1).[6] For V_{XY} to reach $\sqrt{2}(V_{GS} - V_{TH})_{eq}$, we must have in Fig. 5.21:

$$V_1 \sin \omega_0 \Delta T = \sqrt{2}(V_{GS} - V_{TH})_{eq} \tag{5.16}$$

and hence

$$\Delta T = \frac{1}{\omega_0} \sin^{-1} \frac{\sqrt{2}(V_{GS} - V_{TH})_{eq}}{V_1}. \tag{5.17}$$

[5] The value $\sqrt{2}(V_{GS} - V_{TH})_{eq}$ is derived for square-law characteristics but also holds reasonably well for short-channel devices.

[6] We denote the peak differential swing by V_1 to avoid confusion with V_0 in Fig. 5.13(a).

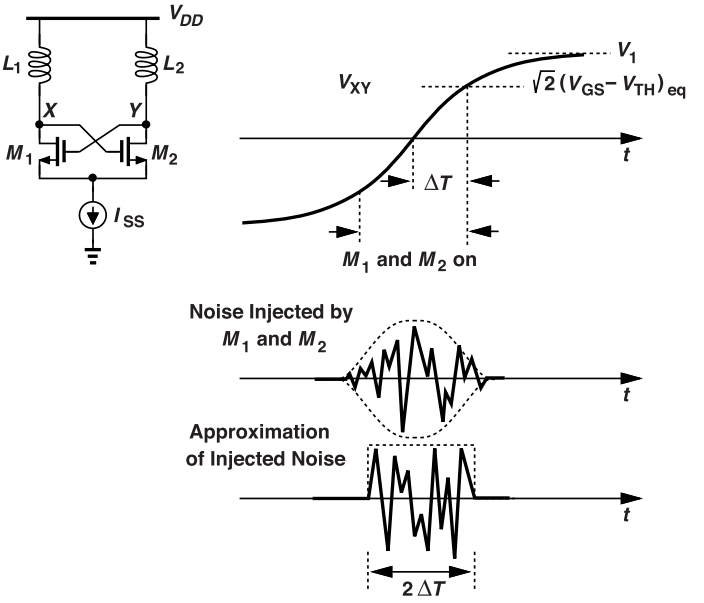


Figure 5.21 LC oscillator and its differential output voltage along with approximation of injected noise.

We now approximate $\sin^{-1}\alpha$ with α, in essence assuming that V_{XY} in Fig. 5.21 linearly goes from 0 to $\sqrt{2}(V_{GS}-V_{TH})_{eq}$ with a slope of $V_1\omega_0 = 2\pi V_1/T_0$. This approximation underestimates ΔT. We have

$$\Delta T \approx \frac{\sqrt{2}(V_{GS}-V_{TH})_{eq}}{2\pi V_1} T_0. \tag{5.18}$$

Note that M_1 and M_2 are simultaneously on for $4\Delta T$ seconds in every period.

Recall from Fig. 5.18(d) that the noise injected by M_1 and M_2 is "strongest" around $V_{XY} = 0$ and becomes weaker as one transistor approaches the off state. We make another simplifying assumption that the noise injected by M_1 and M_2 has a *constant* strength for $2\Delta T$ seconds around each zero crossing (Fig. 5.21). We hope that underestimating ΔT and overestimating the injected noise strength partially cancel each other, yielding a relatively accurate result.

In order to calculate the net white noise injected into the tank, we must multiply the spectral density of the noise current produced by M_1 and M_2 at $V_{XY} = 0$ by the switching duty cycle, $4\Delta T/T_0$. To this end, let us construct a Norton noise equivalent for the cross-coupled pair. As illustrated in Fig. 5.22(a), we seek the short-circuit noise current, I_{Nort}. Recognizing that M_1 and M_2 are now in a diode-connected configuration, we redraw the circuit as shown in Fig. 5.22(b) and note that half of I_{n1} flows through M_1 and the other half through M_2 (why?); the same holds for I_{n2}. That is,

$$I_{\text{Nort}} = \frac{-I_{n1}+I_{n2}}{2}. \tag{5.19}$$

With $\overline{I_{n1}^2} = \overline{I_{n2}^2} = 4kT\gamma g_m$, the spectrum of I_{Nort} would be equal to $8kT\gamma g_m/4$ if there were no switching. In the presence of switching, we express the one-sided spectrum of the noise injected by M_1 and M_2 as

$$S_{inj}(f) = \frac{8kT\gamma g_m}{4} \cdot \frac{4\Delta T}{T_0} \tag{5.20}$$

$$= 8kT\gamma \frac{I_{SS}}{(V_{GS}-V_{TH})_{eq}} \frac{\Delta T}{T_0}, \tag{5.21}$$

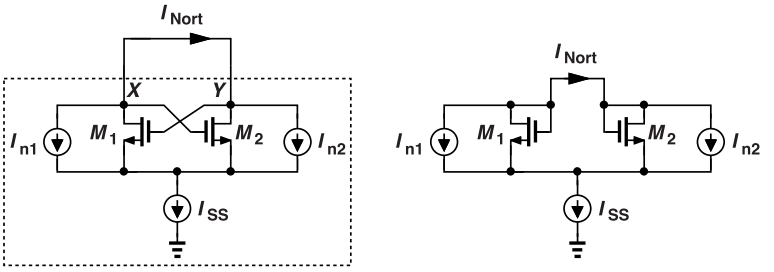


Figure 5.22 (a) Arrangement for computing the Norton noise current, and (b) simplified model.

where g_m is assumed equal to $I_{SS}/(V_{GS} - V_{TH})_{eq}$ at $V_{XY} = 0$. From Eq. (5.18), we have

$$S_{inj}(f) = 8kT\gamma I_{SS}\frac{\sqrt{2}}{2\pi V_1}, \tag{5.22}$$

which, with $V_1 = (4/\pi)R_p I_{SS}$, reduces to

$$S_{inj}(f) = \frac{\sqrt{2}kT\gamma}{R_p}. \tag{5.23}$$

This is the total one-sided spectrum of the noise current injected by M_1 and M_2. Interestingly, it depends only on R_p. We further study these results in Problems 5.11-5.14.

Conversion of Noise Current to Voltage Let us summarize our findings thus far. As illustrated in Fig. 5.23(a), we have modeled the cross-coupled pair's noise contribution by $I_{\rm Nort}$. The transistors also present

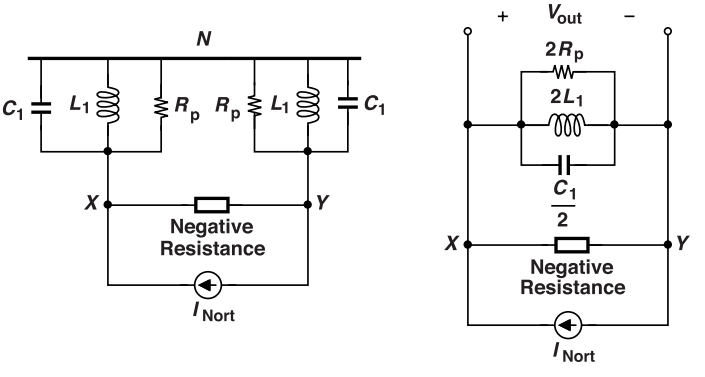
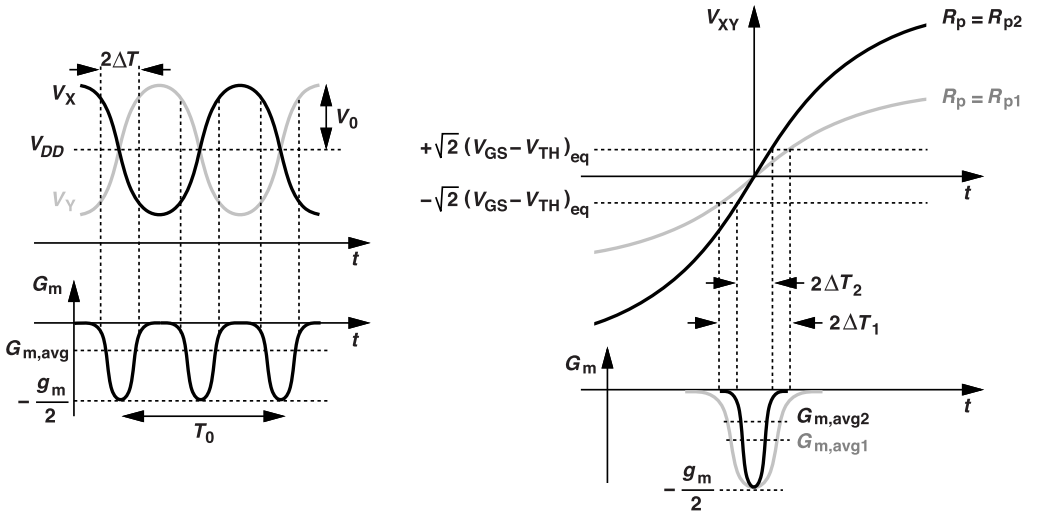


Figure 5.23 (a) LC oscillator model, and (b) its simplified representation.

a time-varying negative resistance between X and Y. To the first order, node N is ac ground, allowing us to merge the two tanks as shown in Fig. 5.23(b). We have determined the spectrum of $I_{\rm Nort}$, Eq. (5.23), and must now compute the spectrum of V_{out} (and eventually the phase of V_{out}).

How much of $I_{\rm Nort}$ in Fig. 5.23(b) flows through the inductor and the capacitor? The answer is, all of it. To understand this point, let us make two observations. First, the negative resistance varies periodically and can be expressed by a Fourier series. Since this resistance assumes an infinite magnitude when $|V_X - V_Y|$ reaches its peak, we examine its inverse, G_m, instead. Illustrated in Fig. 5.24(a), this quantity swings between zero and $-g_m/2$, exhibiting an average value, $G_{m,avg}$, and a pulsewidth of about $2\Delta T$ seconds. Second, $G_{m,avg}$,



(a) (b)

Figure 5.24 (a) Equivalent transcoductance of cross-coupled pair, and (b) effect of larger voltage swings.

i.e., the first term in the Fourier expansion, must be equal to $1/(2R_p)$ for the oscillation amplitude to stabilize. If this average has a magnitude greater than $1/(2R_p)$ (which implies $g_m R_p > 1$), then the amplitude grows further, narrowing the G_m pulsewidth, and vice versa. To see how $G_{m,avg}$ tracks $1/(2R_p)$, suppose we artificially increase R_p in an oscillator from R_{p1} to R_{p2}. Then, as depicted in Fig. 5.24(b), the higher slew rate causes the transistors to switch more abruptly and remain on simultaneously for a lesser amount of time, $2\Delta T_2 < 2\Delta T_1$. As a result, $|G_{m,avg}|$ decreases. We conclude that, over long time scales, the average G_m cancels $2R_p$, allowing the entire I_{Nort} to flow through the ideal tank.

Example 5.6

What happens to the behavior of G_m if the width of the cross-coupled transistors is doubled?

Solution

In this case, the output voltage slew rate is the same, but the greater width translates to a lower equilibrium overdrive, $(V_{GS} - V_{TH})_{eq}$. Thus, as shown in Fig. 5.25, two aspects have changed: the drain currents make faster transitions, and the peak G_m is more negative. The average G_m remains unchanged.

Example 5.7

Prove that the average value of G_m is independent of the transconductance of the transistors.

Solution

In a manner similar to the noise waveforms in Fig. 5.21, we approximate the G_m behavior as shown in Fig. 5.26. The average value is given by

$$G_{m,avg} = \frac{-g_m}{2} \cdot \frac{4\Delta T}{T_0}, \tag{5.24}$$

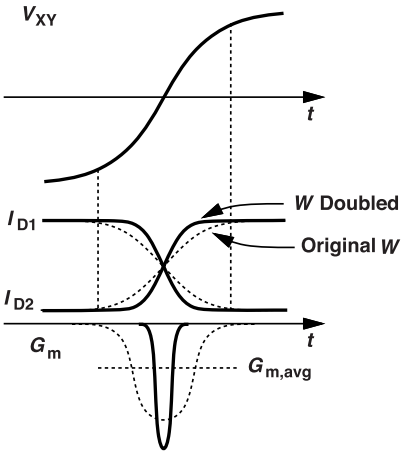


Figure 5.25 Effect of doubling the width of the cross-coupled transistors.

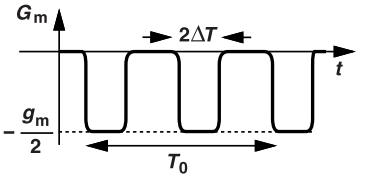


Figure 5.26 Approximation of G_m by a rectangular wave.

which, from Eq. (5.18), reduces to

$$G_{m,avg} = \begin{matrix} -I_{SS} & 4\sqrt{2}(V_{GS}-V_{TH})_{eq} \\ 2(V_{GS}-V_{TH})_{eq} & 2\pi V_1 \end{matrix} \tag{5.25}$$

$$= -\frac{\sqrt{2}}{2}\frac{1}{2R_p}. \tag{5.26}$$

This value is independent of the transistor parameters, but it differs from $1/(2R_p)$ by a factor of $\sqrt{2}/2$ due to the approximate waveform of G_m in Fig. 5.26.

We should remark that the tracking between $G_{m,avg}$ and $1/(2R_p)$ is less accurate if the second harmonic of G_m is significant. But we proceed with the assumption that $G_{m,avg} = 1/(2R_p)$.

5.2.4 Phase Noise Calculation

Our analysis of the effect of white noise has yielded the following: the cross-coupled transistors in Fig. 5.21(a) inject a noise current into the tank whose one-sided spectrum is given by $\sqrt{2}kT\gamma/R_p$. To this we must add the noise contributed by $2R_p$, $4kT/(2R_p)$. The result can then be modeled by $\overline{I_n^2}$ in Fig. 5.13(b).

To create correspondence between Figs. 5.13(b) and 5.23(b), we denote $2L_1$ by L_a and $C_1/2$ by C_a, recognizing that $L_aC_a = 1/\omega_0^2$ and $Q = 2R_p/(L_a\omega_0)$. That is,

$$C_a^2 = \frac{1}{L_a^2\omega_0^4} \tag{5.27}$$

$$= \frac{1}{(4R_p^2/Q^2)\omega_0^2}. \tag{5.28}$$

We return to Eq. (5.12) for the phase noise spectrum and substitute $\sqrt{2}kT\gamma/R_p + 4kT/(2R_p)$ for $\overline{I_n^2}$, the above value for C_a^2, and $(4/\pi)R_pI_{SS}$ for V_0:

$$S_{\phi n}(f) = \left(\frac{\sqrt{2}kT\gamma}{R_p} + \frac{2kT}{R_p}\right)\frac{(4R_p^2/Q^2)\omega_0^2}{4\omega^2}\frac{1}{\frac{4^2}{\pi^2}R_p^2I_{SS}^2} \tag{5.29}$$

$$= \left(\frac{\sqrt{2}}{2}\gamma + 1\right)\frac{\pi^2}{2R_p}\frac{kT}{I_{SS}^2}\left(\frac{f_0}{2Qf}\right)^2. \tag{5.30}$$

As mentioned before, we prefer to use f rather than Δf in this equation because $S_{\phi n}$ is centered about the zero frequency. This result offers insights into the phase noise behavior of cross-coupled LC oscillators. First, the factor $[f_0/(2Qf)]^2$ was predicted in Chapter 3 by our general transformation of delay-line phase noise to oscillator phase noise. With the limited Q of inductors in typical CMOS processes, this factor is not much under the designer's control. Second, this equation normalizes the phase noise to a first harmonic amplitude of $V_0 = (4/\pi)R_pI_{SS}$, whereas the actual drain currents are not exactly square waves and hence produce a smaller voltage swing. If the approximation $V_0 = (4/\pi)R_pI_{SS}$ holds, then $S_{\phi n}(f)$ can be rewritten as

$$S_{\phi n}(f) = \left(\frac{\sqrt{2}}{2}\gamma + 1\right)\frac{2\pi kT}{I_{SS}V_0}\left(\frac{f_0}{2Qf}\right)^2, \tag{5.31}$$

revealing a simple trade-off with I_{SS}—and with no other circuit parameter! The analysis in [12] arrives at the same result but with a factor of $\gamma + 1$ rather than $(\sqrt{2}/2)\gamma + 1$. Third, the phase noise is independent of the transconductance of M_1 and M_2. Fourth, this result is derived with the assumption that M_1 and M_2 do not enter the triode region. Problems 5.16 and 5.17 examine dependencies of the phase noise upon the inductance and the tail current.

Example 5.8

A student predicts that doubling I_{SS} in Fig. 5.21 doubles the output swing but also increases the noise power injected by the transistors. The student then concludes that the dependence of the phase noise on the tail current should be weaker than I_{SS}^2. Explain the flaw in this reasoning.

Solution

Doubling the voltage swings reduces ΔT in Fig. 5.21. As seen from the cancellation of $(V_{GS} - V_{TH})_{eq}$ in Eq. (5.25), the larger noise currents and the shorter time during which they are injected cancel each other, leading to (5.26), which is independent of the g_m. Of course, the assumption is that the transistors do not enter the triode region even with the larger swings. We conclude that doubling I_{SS} lowers the phase noise by 6 dB, a point of contrast to linear scaling.

If we assume V_0 is a certain fraction of the supply voltage, $V_0 = \alpha V_{DD}$, then the trade-off between the phase noise and the power consumption can be written as

$$S_{\phi n}(f)\cdot P = \frac{(\sqrt{2}/2)\gamma + 1}{\alpha}2\pi kT\left(\frac{f_0}{2Qf}\right)^2, \tag{5.32}$$

where $P = I_{SS}V_{DD}$. The figure of merit for oscillators (Chapter 2) normalizes this quantity to f_0^2/f^2 and inverts the result:

$$\text{FOM} = \left(\frac{(\sqrt{2}/2)\gamma + 1}{\alpha}\cdot\frac{2\pi kT}{4Q^2}\right)^{-1}. \tag{5.33}$$

For example, if $\gamma = 1$, $\alpha = 0.5$, and $Q = 8$, then $10 \log \text{FOM} = 215$ dB at $T = 300$ K; since the FOMs reported in the literature express P in milliwatts, our value should be reduced by 30 dB, yielding FOM$= 185$ dB.

Example 5.9

The Q of the tank in a cross-coupled oscillator is doubled. If the transistors do not enter the triode region, explain how the phase noise changes.

Solution

Equation (5.30) suggests that, since both R_p and the Q are doubled, S_ϕ drops by a factor of 8 ($= 9$ dB). This can be explained intuitively as follows. First, consider the noise contributed by R_p. With the Q doubled, the tank resistance injects half as much noise current as before, but the voltage swing is also doubled. Thus, if we return to Eq. (5.11) and write $S_{\phi n} = S_{nQ}/V_0^2$, we observe that the numerator is halved and the denominator is quadrupled. The noise contributed by the transistors experiences a similar scaling factor (why?).

The cross-coupled oscillator of Fig. 5.21 entails two other sources that contribute phase noise, namely, the flicker noise of M_1 and M_2 and the thermal noise of the tail current source. We deal with the latter in the next section and with the former in Section 5.4.

5.3 Tail Noise

The tail current source in a cross-coupled oscillator generates both flicker and white noise, potentially contributing phase noise to the output waveform. Drawing the circuit as shown in Fig. 5.27(a), where I_{nT} models the noise of I_{SS}, and assuming that M_1 and M_2 are noiseless, we wish to follow I_{nT} through the oscillator and

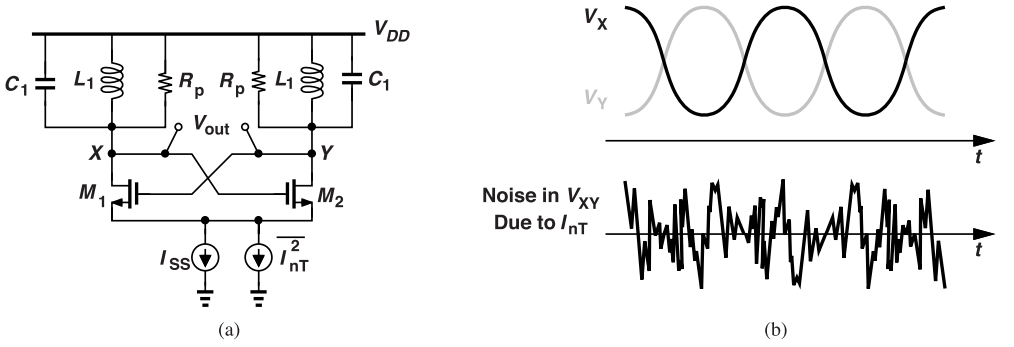


Figure 5.27 LC oscillator with tail noise current.

see whether and how it affects the output phase. It is helpful to view M_1 and M_2 as a differential pair. Also, recall that (1) the differential output voltage is approximately given by $V_{out} = V_{XY} = (4/\pi)R_p I_{SS} \cos \omega_0 t$, and (2) if noise is added to V_{out}, only its quadrature component, in the form of $n(t) \sin \omega_0 t$, contributes phase noise. Note that I_{nT} simply adds to I_{SS}, yielding a tail current $I_{SS} + i_{nT}$ in the time domain. This means that, in the V_{out} expression, we must replace I_{SS} with $I_{SS} + i_{nT}$, obtaining $V_{out} = (4/\pi)R_p(I_{SS} + i_{nT}) \cos \omega_0 t$.

5.3.1 Tail Thermal Noise

Let us first examine the effect of I_{nT} qualitatively. We make two observations. Around the zero crossings of V_{XY}, I_{nT} produces common-mode noise at the output, and hence no noise in V_{XY} [Fig. 5.27(b)]. Also, when V_X and V_Y reach their peak values, I_{nT} does not generate phase noise because the slope of V_{XY} is zero (Section 5.2.1). Interestingly, the longer M_1 and M_2 remain near equilibrium, the less phase noise is generated by I_{nT} because it acts as a CM disturbance for a greater fraction of the period. From this perspective, the

cross-coupled pair should switch more "softly" [12] if we wish to reduce the phase noise due to I_{nT}. At the other extreme, if these transistors turn on and off abruptly, we can view them as switches controlled by complementary clocks and draw the circuit as shown in Fig. 5.28. We recognize that the switching action of M_1 (or M_2) is equivalent to multiplying, in the time domain, the tail current by a square wave toggling between 0 and 1.

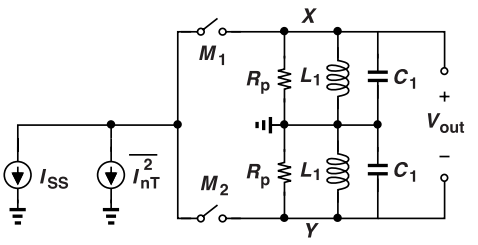


Figure 5.28 Representation of M_1 and M_2 by switches acting on the tail current noise.

For our quantitative analysis, we ask which noise frequency components in the tail should be considered. Let us study those centered (a) around ω_0, and (b) around $2\omega_0$. (We consider noise around the zero frequency in Section 5.3.2.) Since the total tail current is equal to $I_{SS} + i_{nT}$, the output voltage is given by $V_{XY} = (4/\pi)R_p(I_{SS} + i_{nT})\cos\omega_0 t$. We conclude that I_{nT} experiences a "gain" equal to $(4/\pi)R_p$ as it travels through the cross-coupled pair, is translated in frequency, and becomes a voltage quantity. We say M_1 and M_2 act as a mixer.

Example 5.10

Equation $V_{XY} = (4/\pi)R_p(I_{SS} + i_{nT})\cos\omega_0 t$ implies that i_{nT} causes only amplitude modulation. Is this true?

Solution

If, after the mixing action of the cross-coupled pair, i_{nT} has an envelope that is in phase with the output voltage, then it does not cause phase noise. For example, with the desired output given by $(4/\pi)R_p I_{SS}\cos\omega_0 t$, the low-frequency content of I_{nT} is upconverted to $(4/\pi)R_p i_{nT}\cos\omega_0 t$, producing no phase noise [Fig. 5.29(a)]. This observation suggests that, for phase noise to appear, the envelope of the translated tail noise must not go to zero at the output zero crossings. Shown in Fig. 5.29(b) is an example, where the noise injected into the tank has an envelope that is 90° out of phase with respect to the output oscillatory waveform. This was also predicted by Eq. (5.11) and Fig. 5.16(b). We must therefore seek those tail noise components that translate to quadrature noise as they are mixed with V_{XY}.

We now study the effect of band-pass tail noise around ω_0 in Fig. 5.27. Representing this noise by $i_{nT}(t) = n_I(t)\cos\omega_0 t - n_Q(t)\sin\omega_0 t$ (Section 5.2.1), we invoke the simple model in Fig. 5.28 and note that M_1 and M_2 mix i_{nT} with ω_0. The result thus lands at $2\omega_0$ (and at zero), experiencing heavy attenuation by the tanks as it converts to voltage (Fig. 5.30). That is, i_{nT} produces no output component at ω_0 in this case. The effect of this noise is therefore negligible.

The band-pass tail noise around $2\omega_0$, on the other hand, proves serious. Multiplying $i_{nT}(t) = n_I(t)\cos 2\omega_0 t - n_Q(t)\sin 2\omega_0 t$ by $\cos\omega_0 t$ gives a component of the form $n_Q(t)\sin\omega_0 t$, whose envelope is in quadrature with the output waveform, thereby producing phase noise [Fig. 5.29(b)]. To analyze this phenomenon, we return to Fig. 5.28 and examine only the single-ended output at X for now. The noise current carried by M_1 to node X is equal to i_{nT} multiplied by a square wave toggling between 0 and 1. Considering the first harmonic of the square wave, we write this current as $i_{nT}(t)(2/\pi)\cos\omega_0 t$, and subject it to the impedance of the tank near the resonance frequency, $-j/(2C_1\Delta\omega)$ (Section 5.2.1). The resulting voltage at X is added to the single-ended oscillatory waveform at this node and translates to phase noise.

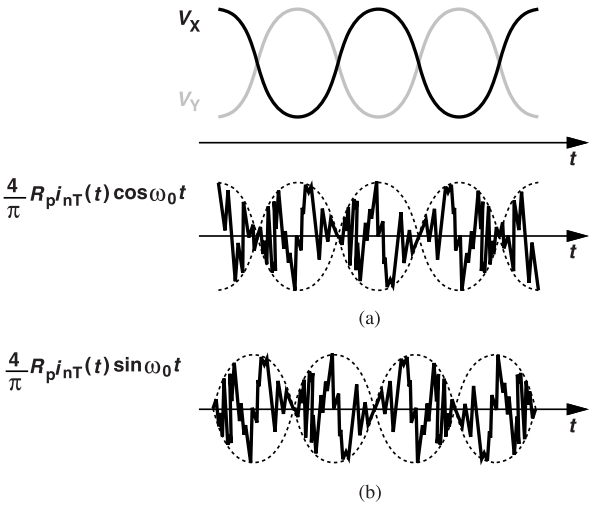


Figure 5.29 Transformation of tail noise current to output noise voltage having an envelope that is (a) in-phase with V_{XY}, or (b) in quadrature with V_{XY}.

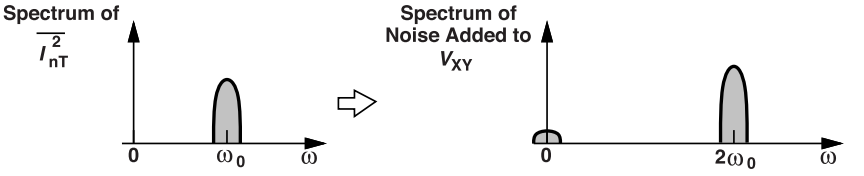


Figure 5.30 Translation of tail thermal noise in the frequency domain after mixing.

Let us express the current carried to X by M_1 as $(2/\pi)[n_I(t)\cos 2\omega_0 t - n_Q(t)\sin 2\omega_0 t]\cos\omega_0 t$, obtaining $(1/\pi)n_Q(t)\sin\omega_0 t$ as the critical component. The one-sided spectral density of $n_Q(t)$ is the same as that of $i_n(t)$ (Section 5.2.1) and denoted by $\overline{I_{nT}^2}$. From Eq. (5.12), the phase noise in V_X is equal to:

$$S_{\phi n}(\omega) = \frac{1}{\pi^2}\overline{I_{nT}^2}\frac{1}{(2C_1\omega)^2}\frac{1}{V_{sing}^2}, \tag{5.34}$$

where V_{sing} is the single-ended peak oscillation swing and equal to $(2/\pi)R_p I_{SS}$.[7] Since $4C_1^2\omega^2 = 4Q^2\omega^2/(R_p^2\omega_0^2)$, we have

$$S_{\phi n}(f) = \frac{\overline{I_{nT}^2}}{4I_{SS}^2}\left(\frac{f_0}{2Qf}\right)^2. \tag{5.35}$$

This is the phase noise observed in each single-ended output. Note that this result is the upper bound because our derivation assumes the transistors switch abruptly. The analysis in [12] arrives at the same result. The following example addresses the phase noise in the differential output.

[7] As explained earlier, the noise traveling through M_1 and M_2 sees the ideal tank because R_p is canceled by the negative resistance.

Example 5.11

While calculating the phase noise in the differential output, a student writes $V_{out} = V_X - V_Y = (2/\pi)R_pI_{SS}\cos(\omega_0 t + \phi_{n,X}) + (2/\pi)R_pI_{SS}\cos(\omega_0 t + \phi_{n,Y})$, where $\phi_{n,X}$ and $\phi_{n,Y}$ denote the phase noise in the single-ended outputs. Since $V_{out} \approx (4/\pi)R_pI_{SS}\cos\omega_0 t - (\phi_{n,X} + \phi_{n,Y})(2/\pi)R_pI_{SS}\sin\omega_0 t$, the student reasons that the normalized phase noise in the differential output, $(\phi_{n,X} + \phi_{n,Y})/2$, is lower by a factor of $\sqrt{2}$ and yields a spectrum given by $(S_{\phi n,X} + S_{\phi n,Y})/4 = S_{\phi n,X}/2$. Is this correct?

Solution

No, it is not. The student has assumed $\phi_{n,X}$ and $\phi_{n,Y}$ are uncorrelated, but these arise from the same noise source, i_{nT}. In fact, the currents injected by M_1 and M_2 in Fig. 5.28 into their respective tanks can be respectively expressed as $(2/\pi)i_n(t)\cos\omega_0 t$ and $-(2/\pi)i_n(t)\cos\omega_0 t$, leading to phase noise quantities that are opposite and equal. Thus, $(\phi_{n,X} + \phi_{n,Y})/2 = (2\phi_{n,X})/2 = \phi_{n,X}$ and has the same spectral density as $\phi_{n,X}$ and $\phi_{n,Y}$. In other words, Eq. (5.35) applies to the differential output as well.

In Fig. 5.28, i_{nT} is also mixed with the third harmonic of the oscillation waveform. The reader can show that the resulting phase noise is uncorrelated with respect to that expressed by Eq. (5.35) and, in the case of abrupt switching, raises this value by one-ninth (≈ 0.46 dB).

Is the phase noise contribution of the tail current source comparable to that of the cross-coupled pair? To answer this question, we express the tail noise as $\overline{I_{nT}^2} = 4kT\gamma g_{mT}$, where $g_{mT} = 2I_{SS}/(V_{GS} - V_{TH})$ denotes the transconductance of the transistor acting as the current source. If $V_{GS} - V_{TH} = V_{DS}$, we have $\overline{I_{nT}^2} = 4kT\gamma(2I_{SS})/V_{DS}$, which, upon substitution in (5.35) and (5.31), yields

$$S_{\phi n}(f) = \frac{2kT}{I_{SS}}\left\{\frac{\pi[(\sqrt{2}/2)\gamma + 1]}{V_0} + \frac{\gamma}{V_{DS}}\right\}\left(\frac{f_0}{2Qf}\right)^2, \tag{5.36}$$

where $V_0 = (4/\pi)R_pI_{SS}$ is the peak differential output swing. Since the available V_{DS} for the tail current source is typically several times less than V_0, we conclude that the γ/V_{DS} term can be comparable with $\pi[(\sqrt{2}/2)\gamma + 1]/V_0$. We envision that tying a capacitance from the tail node to ground can suppress this noise, but such a remedy presents other difficulties (Section 5.4). In Problem 5.20, we investigate another method of reducing the tail noise at $2\omega_0$.

5.3.2 Tail Flicker Noise

The tail current source in an LC oscillator typically suffers from high flicker noise. This can be seen from the current-mirror implementation depicted in Fig. 5.31(a), where $(W/L)_a$ is usually chosen 5 to 10 times

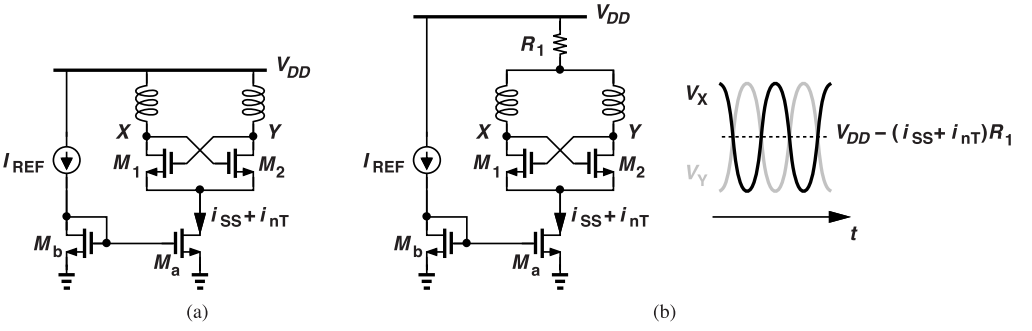


Figure 5.31 (a) Tail current implementation illustrating the problem of flicker noise, and (b) arrangement exhibiting greater conversion of flicker noise to phase noise.

$(W/L)_b$ so as to save power. Unfortunately, such a choice also amplifies the noise of M_b and I_{REF}, which itself is copied from a bandgap reference and possibly contains high flicker noise.

Does the tail flicker noise translate to phase noise at the output? Only in some cases. Let us begin with the following perspective. Recall that $V_{XY} = (4/\pi)R_p I_{SS}\cos\omega_0 t$ in the absence of noise: if we change I_{SS}, the output amplitude changes. Since the tail flicker noise is added to and causes slow fluctuations in the bias current, I_{SS}, we can express the output differential voltage as $(4/\pi)R_p(I_{SS} + i_{nT})\cos\omega_0 t$, where i_{nT} represents the time-domain noise waveform.[8] Thus, i_{nT} slowly modulates the output amplitude and has no effect on the phase in this case. The AM component is removed by the limiting action of the stages following the oscillator.

The flicker noise of the tail current source does produce phase noise in two cases: (a) if it modulates the output common-mode (CM) level, or (b) if the capacitances at X and Y do not exhibit odd-symmetric nonlinearity, causing AM/PM conversion. We study these phenomena here.

CM Level Modulation Recall from Chapter 1 that the voltage-dependent drain-bulk capacitance at the drain nodes in Fig. 5.31(a) allows changes in V_{DD} to modulate the frequency. The same effect occurs if the tail current affects the CM level at X and Y: the noise in I_{SS} perturbs V_X and V_Y together, modulating the capacitance at these nodes.

In the oscillator of Fig. 5.31(a), the output CM level remains close to V_{DD}, experiencing only a small drop due to the series (low-frequency) resistance of the inductors, r_s. That is, $V_{CM} = V_{DD} - (I_{SS}/2)r_s$, and flicker noise in I_{SS} only slightly modulates the voltage-dependent capacitance at the drain nodes. Now, suppose we wish to shift the CM level down for proper interface with the next stage and hence insert a resistor in series with V_{DD} [Fig. 5.31(b)]. In this case, i_{nT} modulates the CM level and the capacitances at X and Y by $i_{nT}R_1$.

AM/PM Conversion Consider the LC oscillator shown in Fig. 5.32(a), noting that the transistors' drain-

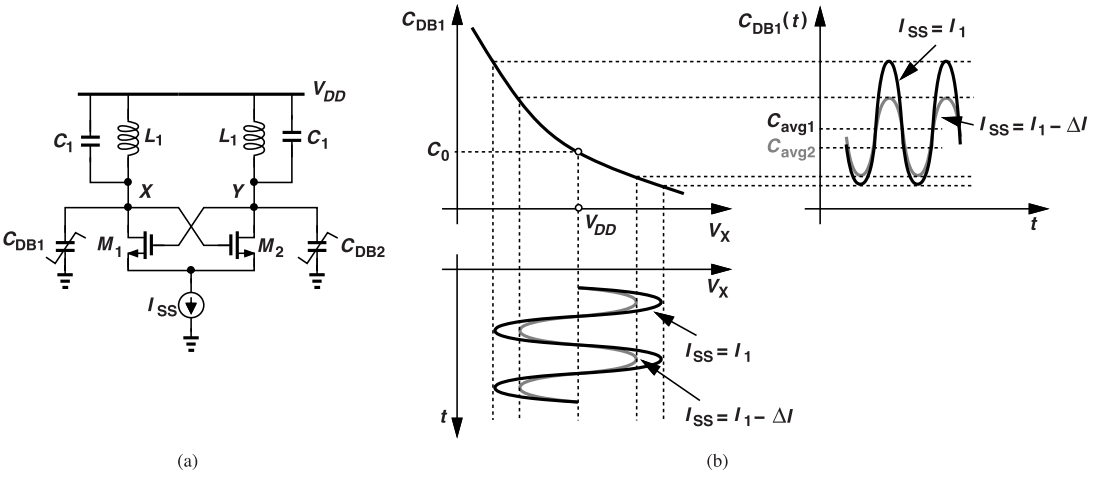


Figure 5.32 (a) LC oscillator including voltage-dependent load capacitances, and (b) C-V characteristic showing how average capacitance changes with I_{SS}.

bulk junction capacitances, C_{DB1} and C_{DB2}, are voltage-dependent. As explained earlier, the low-frequency noise in I_{SS} slowly modulates the output amplitude. In this section, we examine how AM can create PM.

In order to study this phenomenon, we estimate the oscillation frequency for two cases, namely, (1) the tail current is equal to I_1, and (2) the tail current is equal to $I_1 - \Delta I$. If ω_0 remains the same, then the circuit does not exhibit AM/PM conversion. In Fig. 5.32(b), we plot V_X versus time, C_{DB1} versus V_X, and C_{DB1} versus time, noting that a tail current of I_1 yields a peak single-ended swing of $(2/\pi)R_p I_1$ around V_{DD}. The

[8] Since I_{nT} has a low-pass spectrum, it cannot be expressed in terms of quadrature components, n_I and n_Q.

drain junction capacitance decreases as V_X rises and vice versa, and displays a periodic behavior with time. We estimate the oscillation frequency as $\omega_0 = 1/\sqrt{L_1(C_1 + C_{DB,avg})}$, where $C_{DB,avg}$ denotes the average value (the first term in the Fourier series) of $C_{DB1}(t)$.

Now, we change the tail current to $I_1 - \Delta I$ and hence the swing to $(2/\pi)R_p(I_1 - \Delta I)$ [the gray waveforms in Fig. 5.32(b)]. If the average value of $C_{DB1}(t)$ does not change, neither does ω_0. This is possible if the C_{DB1}-V_X characteristic is an odd-symmetric function around the point (V_{DD}, C_0).

The problem of AM/PM conversion becomes more serious when varactors are added to the oscillator. Since neither C_{DB1} nor the varactors' capacitance has an odd-symmetric behavior, the flicker noise in I_{SS} can cause substantial phase noise (Section 5.5.2). We also expect greater upconversion for larger varactor capacitances; i.e., AM/PM conversion is more pronounced for larger K_{VCO} values.

5.4 Effect of Tail Capacitance

Recall from our study of differential ring oscillators in Chapter 4 that the tail capacitance in a differential pair introduces a tail current at the second harmonic of the oscillation frquency. Upon flowing through the transistors, this current leads to asymmetric rise and fall times at the drain nodes, allowing the flicker noise of the two transistors to be upconverted. These observations apply to the cross-coupled oscillator as well. We now study this phenomenon in greater detail. The relationship between asymmetric transitions and flicker noise upconversion is described in Chapter 6.

Consider the circuit depicted in Fig. 5.33(a), where C_T represents (a) the source-bulk capacitances of M_1 and M_2 and the drain capacitance of I_{SS}, and (b) any additional capacitance tied to P to shunt the noise of I_{SS} (at $2\omega_0$) to ground. If the transistors do not enter the triode region, V_P briefly falls at the zero crossings of V_{XY} [Fig. 5.33(b)]. Thus, C_T draws a current that exhibits a "doublet" (the derivative of an impulse) around each zero crossing, i.e., twice per period. This current, $C_T dV_P/dt$, splits equally between M_1 and M_2 when $V_X = V_Y$. We recognize that, if I_{D1} is rising, then $C_T dV_P/dt$ *sharpens* the transition, and vice versa. As a result, the rise and fall times of the drain currents and voltages become asymmetric, leading to the upconversion of the cross-coupled pair's flicker noise. (It can be shown that for perfectly square-law transistors operating in the saturation region, their flicker noise is not upconverted to phase noise [10].)

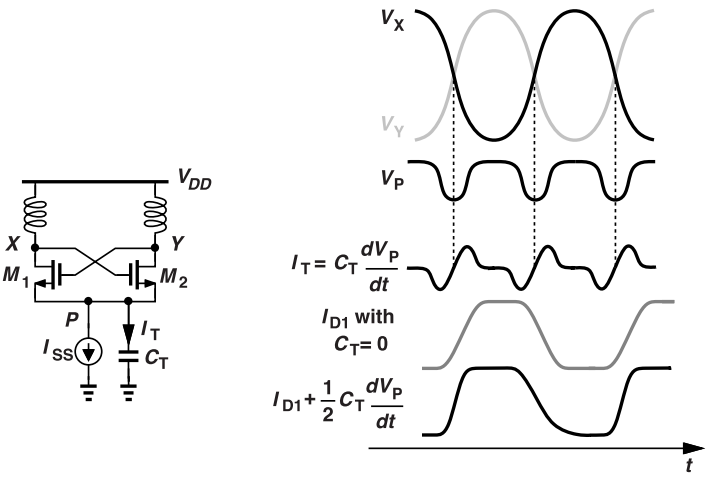


Figure 5.33 Asymmetry in drain current transitions due to tail capacitance.

The effect of the tail capacitance in Fig. 5.33(a) is even more pronounced if M_1 and M_2 enter the triode region. From an intuitive perspective, we can say that, for most of the oscillation cycle, M_1 or M_2 acts as a resistor, tying one end of the tank to C_T. The loss associated with this resistor thus degrades the Q.

In order to minimize the second-harmonic current drawn by the tail in Fig. 5.33(a), we can modify the circuit as shown in Fig. 5.34 [9], where the tank consisting of L_a and C_a resonates at $2\omega_0$, thus acting as a high impedance. Capacitor C_T is still added to supress the noise current generated by I_{SS}. The drawback of this approach is the greater layout complexity.

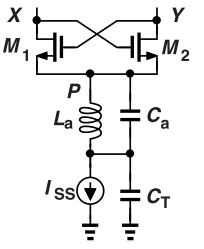


Figure 5.34 Use of a resonant tank to isolate the tail node from the tail capacitance.

5.5 Step-by-Step Design

5.5.1 Preliminary Thoughts

Let us begin with the cross-coupled oscillator shown in Fig. 5.35(a) (Chapter 1). We must select $(W/L)_{1,2}$ and the value of L_1, C_1, and I_{SS} so as to meet certain performance specifications. The question is, how?

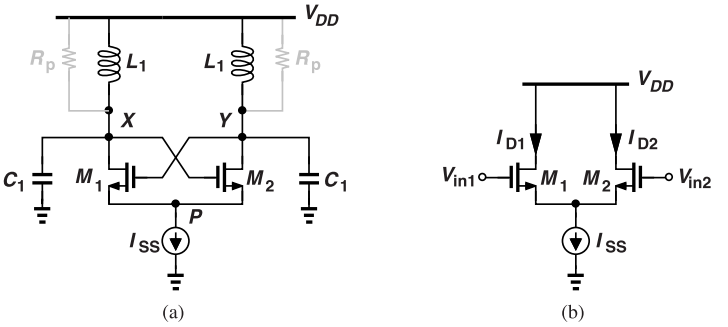


Figure 5.35 (a) Basic LC oscillator, and (b) simplified model for studying drain currents.

In general, it is helpful to assume a power budget, P, and an output voltage swing, V_0, before we embark upon the design. For lowest phase noise, the swing is chosen as large as possible without driving M_1 and M_2 into the triode region, e.g., the peak-to-peak single-ended swing in Fig. 5.35(a) can reach V_{TH} (why?). It follows that $I_{SS} = V_{DD}/P$ is known, and so is R_p because the peak differential voltage swing is equal to $(4/\pi)I_{SS}R_p$. The design space thus reduces to $(W/L)_{1,2}$, L_1, and C_1 for a known oscillation frequency of ω_0.

In the second step, we examine our library of characterized inductors and find the *smallest* one that exhibits at ω_0 the R_p found above. That is, this inductor has the highest Q $[= R_p/(L_1\omega_0)]$ at the frequency of interest. The reader recognizes that the design cannot proceed without a reasonable inductor model, e.g., at least that in Example 5.3.

In the third step, we choose $(W/L)_{1,2}$ so as to allow relatively complete switching of the tail current with a voltage swing of $(4/\pi)I_{SS}R_p$. Recall from Example 1.10 that an optimistic condition for complete switching is $g_{m1,2}R_p = \sqrt{2}\pi/4 \approx 1.11$. But in practice, we construct the simple circuit shown in Fig. 5.35(b), apply $V_{in1} - V_{in2} = (4/\pi)I_{SS}R_p$, and make $(W/L)_{1,2}$ large enough to obtain $I_{D1} \approx I_{SS}$. This condition assumes

complete switching only near the peaks of V_{XY}; we generally prefer to steer the current abruptly and hence select wider transistors.

In the fourth step, we allow C_1 and the capacitances of M_1 and M_2 to resonate with L_1 at ω_0. Note that C_1 would eventually include the inductor parasitics and the input capacitance of the next stage. In the fifth step, we compute the phase noise and decide whether a greater power budget is necessary. If so, we iterate from the first step.

Example 5.12

Determine the capacitances contributed by M_1 and M_2 to nodes X and Y in Fig. 5.35(a). Assume a small swing at node P.

Solution

As illustrated in Fig. 5.36, each transistor presents its C_{GS} from X or Y to the tail node, its C_{DB} from an

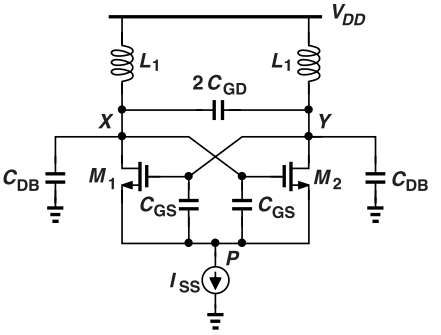


Figure 5.36 LC oscillator including transistor capacitances.

output node to ground, and its C_{GD} between X and Y. The small swing at P means that the C_{GS}'s can be considered tied to ground. The differential swings at X and Y transform $2C_{GD}$ to single-ended capacitors equal to $4C_{GD}$. Thus, the capacitance seen at X or Y to ac ground is equal to $C_{GS} + C_{DB} + 4C_{GD}$.

5.5.2 Design Example

In order to illustrate the foregoing procedure, we design a 5-GHz LC oscillator with a power budget of 1 mW and a peak-to-peak single-ended (or peak differential) swing of 0.5 V. As with the designs in Chapter 3, we assume a worst-case V_{DD} value of 0.95 V with $T = 75\ ^\circ$C and operation in the slow-slow corner of the process.

With $P = 1\ \text{mW} = V_{DD}I_{SS}$, we obtain $I_{SS} \approx 1$ mA. The required voltage swing translates to $(4/\pi)I_{SS}R_p = 0.5$ V and hence $R_p = 393\ \Omega$. We must find the smallest inductor that provides such a value at 5 GHz. For a typical Q of 8, we arrive at $L_1 = R_p/(Q\omega_0) = 1.56$ nH, and for resonance at 5 GHz, the total single-ended capacitance is 660 fF. As a first guess, we allocate 60 fF for the transistor capacitances.

To steer the tail current with a differential peak swing of 0.5 V, simulations necessitate $(W/L)_{1,2} \approx 2\ \mu\text{m}/40$ nm. But we choose $(W/L)_{1,2} \approx 5\ \mu\text{m}/40$ nm for sharper switching. Shown in Fig. 5.37(a) is the design.[9] Note that, at the peaks of the output swing, the drain voltage of one transistor is 0.5 V below its gate voltage, possibly driving the device into the "soft" triode region.

The simulated output waveforms are plotted in Fig. 5.37(b), revealing a peak-to-peak single-ended swing of about 425 mV at X and at Y, and $f_0 = 5.2$ GHz. Why is the swing less than the desired value? Let us

[9]The parasitic capacitance of the inductors is absorbed by the 600-fF capacitor.

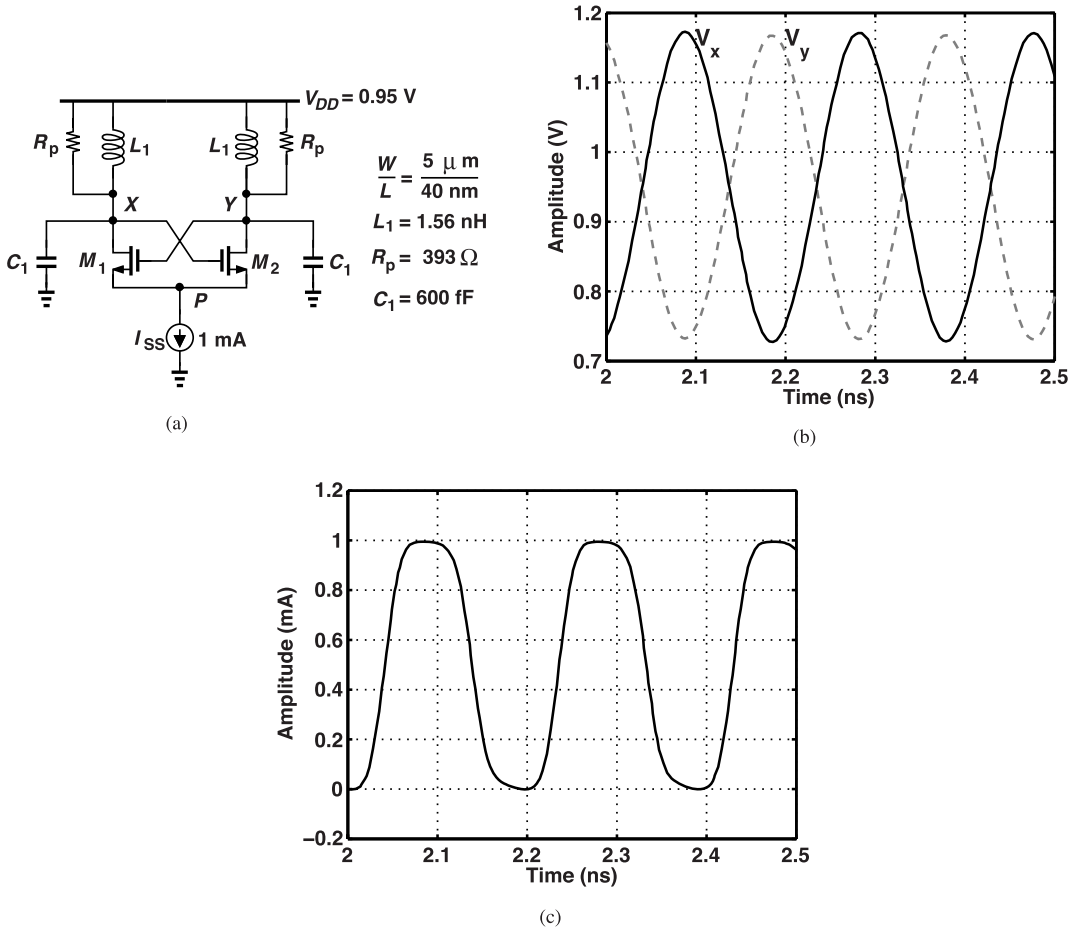


Figure 5.37 (a) LC oscillator design example, (b) output wavforms, and (c) drain current waveform of one transistor.

examine the drain current waveform of one transistor, shown in Fig. 5.37(c). We observe complete switching, but with gradual edges. The first harmonic peak amplitude is therefore smaller than $(2/\pi)I_{SS}$.

How "good" is this design? To answer this question, we examine its supply sensitivity and its phase noise. The former is about 4 MHz/V, much less than that of the differential rings studied in Chapter 3. This is because the nonlinear drain junction capacitance of M_1 and M_2 is overwhelmed by the large, constant single-ended capacitance of the tank, C_1. This benefit would vanish if we had, instead, tied a differential capacitance of 300 fF between X and Y.

Phase Noise Let us use Eq. (5.30) to compute the phase noise of our oscillator. We have $T = 348$ K, $R_p = 393\ \Omega$, $I_{SS} = 1$ mA, $\gamma = 1$, $f_0 = 5.2$ GHz, and $Q = 8$. The phase noise at 1-MHz offset is thus equal to -109.6 dBc/Hz. Figure 5.38 plots the simulation results, revealing remarkable agreement with our analytical prediction.

Another important observation in Fig. 5.38 is that the slope is equal to -20 dB/dec even for offsets as low as 10 kHz, implying little flicker noise contribution. As explained previously and formulated in the next chapter, the flicker noise upconversion depends on the symmetry of the rising and falling transitions at X (and Y) in Fig. 5.37(a), and hence on the second harmonic content of these voltages. According to simulations, the second harmonic in V_X or V_Y is about 50 dB below the fundamental in the present design.

It is instructive to explore the oscillator design space by varying the device parameters and observing the

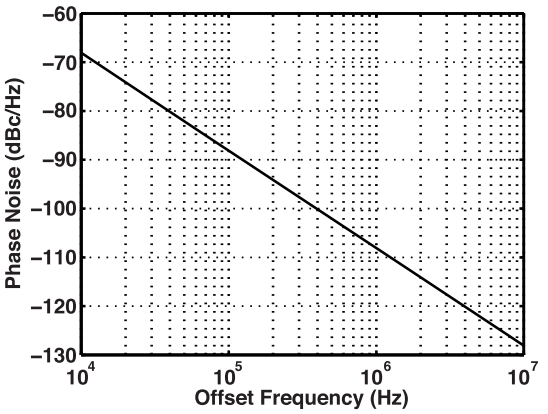
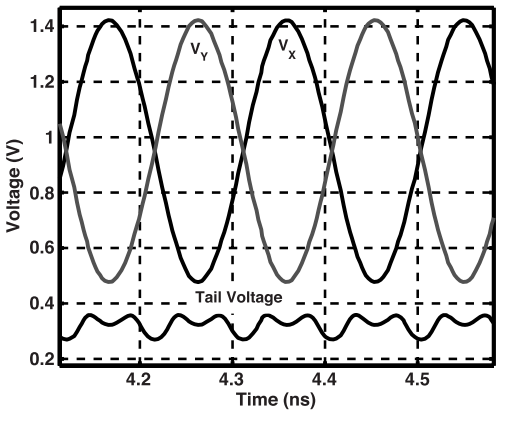


Figure 5.38 Phase noise of design in Fig. 5.37(a).

results. For example, if we increase the widths of M_1 and M_2 in Fig. 5.37(a), we expect ΔT in Fig. 5.24(a) to decrease, leading to a drain current waveform with sharper edges and hence a fundamental amplitude closer to $V_0 = (4/\pi)R_p I_{SS}$. With $W = 10$ μm, the output amplitude increases from 425 $\mathrm{mV_p}$ to 440 $\mathrm{mV_p}$, still short of the theoretical 500 mV, and the phase noise is about the same.

What happens if we double I_{SS} in Fig. 5.37(a) with other parameters unchanged? We expect a twofold increase in V_0 and, from Eq. (5.30), a fourfold (6-dB) decrease in $S_{\phi n}$. Figure 5.39 plots the output and tail node waveforms along with the phase noise for this case. Interestingly, the swing more than doubles, an



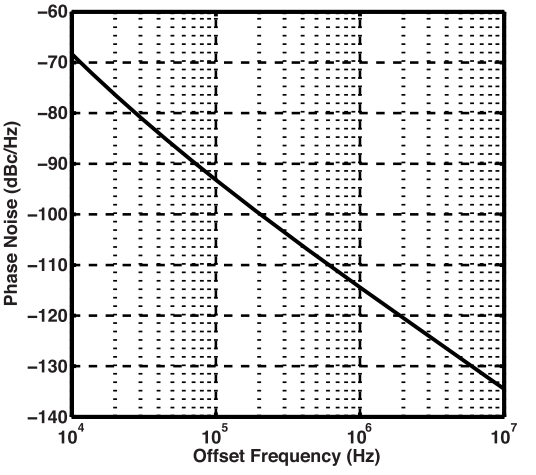


Figure 5.39 Output waveforms and phase noise of design in Fig. 5.37(a) if I_{SS} is doubled.

effect studied in Problem 5.21. We observe that the phase noise at offsets greater than 100 kHz falls by 6 dB. However, the phase noise at 10-kHz offset does not change and exhibits a slope of 30 dB/dec, thus revealing flicker noise contribution. This is because a single-ended peak-to-peak swing of 0.9 V drives the transistors into the deep triode region, raising their flicker noise. The tail voltage has a minimum of about 250 mV, leaving a reasonable headroom for the tail current source.

Example 5.13

Does flicker noise upconversion increase or decrease if the transistor widths in Fig. 5.37(a) are changed from 5 μm to 10 μm while $I_{SS} = 2$ mA? Assume V_0 is relatively constant. Note that M_1 and M_2 enter the triode region.

Solution

The wider transistors exhibit a lower on-resistance, sustaining a *smaller* V_{DS} for the same current. The devices therefore go farther into the triode region and raise the flicker noise upconversion—even though their channel areas are doubled. According to simulations, the phase noise at 10-kHz offset rises by 4 dB in this case.

From the foregoing experiments and our theory of linear scaling (Chapters 2 and 3), we can envision two different scenarios for trading power for phase noise: (1) we double only I_{SS} in Fig. 5.37(a), thereby reducing thermal-noise-induced phase noise by 6 dB while observing greater flicker noise upconversion, or (2) we double I_{SS}, W, and the tank capacitance and halve the inductance (linear scaling), obtaining 3-dB lower phase noise at all frequency offsets. The choice between these two cases depends on the application, but the former is often preferred as it can lead to a smaller *integrated* phase noise. This is seen by noting that the area under the phase noise plot from 10 kHz to 100 kHz (i.e., across 90 kHz) is typically much less than the area from 100 kHz to 10 MHz (i.e., across 9.9 MHz).

Is it possible to alleviate the problem of flicker noise upconversion? That is, can the circuit of Fig. 5.37(a) accommodate larger voltage swings without driving the transistors into the deep triode region? We recognize that the gate and drain voltages change by equal and *opposite* amounts, constraining the peak-to-peak single-ended swing to approximately V_{TH} (why?) if the triode region must be avoided. But, the gate voltage of, say, M_2, need not change by as much as the drain voltage of M_1. Specifically, suppose we insert a capacitive divider between the drain of one transistor and the gate of the other (Fig. 5.40) so that the peak voltage swings

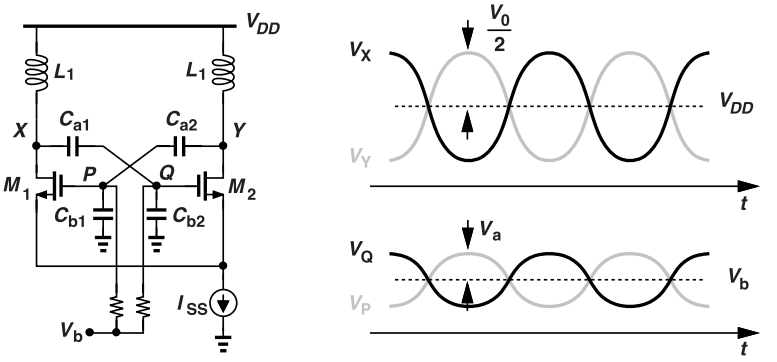


Figure 5.40 Use of capacitive divider to avoid the triode region.

at the gates are reduced from V_0 to $V_a = [C_{a1}/(C_{a1} + C_{b1})]V_0$. For now, we choose the bias voltage V_b equal to V_{DD}. Since the peak of V_P must be at most V_{TH} volts higher than the valley of V_X to ensure M_1 remains in saturation, we have $V_{DD} - V_0/2 - (V_b + V_a) = -V_{TH}$. That is, the allowable peak-to-peak single-ended swing at the drains rises to $2V_{TH} - 2V_1$ (why?) if $V_b = V_{DD}$. Of course, the capacitive attenuation must not drop the loop gain so much as to disallow complete steering of I_{SS}. For this reason, the attenuation factor is typically limited to about 2.

The circuit of Fig. 5.40 merits two remarks. First, in practice, the grounded capacitors, C_{b1} and C_{b2}, are unnecessary; C_{a1} and C_{a2} can be chosen small enough to form an attenuator along with the gate capacitance of M_1 and M_2. Second, it is tempting to choose $V_b < V_{DD}$ to allow even greater voltage swings, but lower values of V_b would leave little voltage headroom for I_{SS}. Third, it can be proved that the capacitive attenuation proportionally raises the MOS noise contribution [10]. For this reason, we do not pursue this topology.

Effect of Tail Noise Let us implement I_{SS} in Fig. 5.37(a) with an NMOS transistor, as shown in Fig. 5.41(a). We allow a V_{DS} ($= V_{GS} - V_{TH}$) of 400 mV for M_3 to minimize its noise current. With $V_0 = 425$ mV

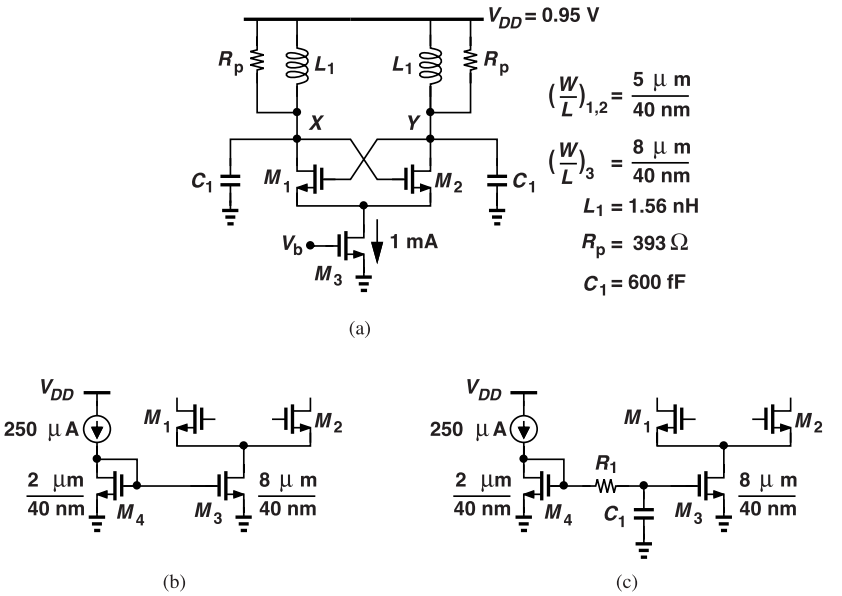


Figure 5.41 (a) LC oscillator with tail transistor, (b) current mirror arrangement defining the tail current, and (c) use of a low-pass filter to reduce noise contribution of reference branch.

and $\gamma = 1$, we obtain the two terms in Eq. (5.36) as $\pi[(\sqrt{2}/2)\gamma + 1]/V_0 = 12.6\,\mathrm{V}^{-1}$ and $\gamma/V_{DS3} = 2.5\,\mathrm{V}^{-1}$. The phase noise thus rises by $10\log(15.1/12.6) = 0.8$ dB due to the tail noise at $2\omega_0$. Upon simulation, however, we see a smaller change in the phase noise. This is because the cross-coupled transistors do not switch abruptly, converting some of the tail noise to a common-mode disturbance.

For robust operation, the tail transistor in Fig. 5.41(a) must be biased by means of a current mirror. Shown in Fig. 5.41(b), such an arrangement allocates a smaller current to the diode-connected device so as to conserve power. In this example, the fourfold scaling factor between M_3 and M_4 raises the total tail noise by a factor of five (why?). In other words, the term γ/V_{DS} in Eq. (5.36) is multiplied by this factor, possibly contributing significantly to the output phase noise—by as much as 3 dB in the presence of abrupt switching. Simulations reveal a 3-dB rise for offsets greater than 100 kHz. At 10-kHz offset, however, the phase noise is about -57 dBc/Hz; this 11-dB degradation arises from the *flicker* noise of M_4. Some of this degradation is due to AM/PM conversion arising from the drain-bulk capacitance of M_1 and M_2.[10]

The thermal noise contributed by M_4 in Fig. 5.41(b) can be suppressed by interposing a low-pass filter between this device and M_3. Depicted in Fig. 5.41(c), this approach chooses a corner frequency much less than $2\omega_0$, allowing reasonable values for R_1 and C_1, e.g., $R_1 = 2\,\mathrm{k}\Omega$ and $C_1 = 0.5$ pF. For the flicker noise of M_4, on the other hand, R_1C_1 must be impractically large. We return to this issue in the next section.

5.5.3 Frequency Tuning

The overall tuning range of LC oscillators must accommodate the following components: (1) PVT variations, which amount to a few percent—considerably less than those of ring oscillators, and (2) the required operating frequency range, for example, from 5.1 GHz to 5.9 GHz for WiFi. We typically design the oscillator for the highest necessary frequency and add sufficient tuning capacitance to reach the lower end.

[10] For a more accurate study of flicker noise, we must include a series resistance for L_1 (as in Fig. 5.12).

Given by $\omega_0 = 1/\sqrt{L_1 C_1}$, the oscillation frequency of the LC oscillator shown in Fig. 5.37(a) can be tuned by varying L_1 or C_1. Since tuning inductors—especially in integrated circuits—is difficult, we study varactor tuning. Recall from Chapter 1 that a MOS varactor is formed by placing an NMOS transistor within an n-well and modeled by a hyperbolic tangent. For continuous, fine tuning, therefore, we can employ MOS varactors, and for discrete, coarse tuning, switched capacitors. The reader is encouraged to review the tuning considerations in Section 3.5.

The choice of varactor capacitance is governed by two opposing issues. (1) We wish to achieve a reasonable continuous tuning range so as to avoid a dead zone in the discrete characteristics. For example, if a total range of 15% is required and the continuous range is around 1%, then we have about 20 to 30 discrete steps, a somewhat reasonable design. (2) We also wish to maintain a moderate K_{VCO}, on the order of a few percent of ω_0 per volt, so as to minimize the effect of noise on the control line. As explained in Section 5.3.2, a high K_{VCO} also exacerbates AM/PM conversion within the oscillator, thus raising the phase noise.

Another consideration relates to the dimensions of the varactors. For a given capacitance, WL is known, providing flexibility in the choice of W and L. However, the value of L entails a trade-off between the Q and the capacitance of the varactor. The following example elaborates on this point.

Example 5.14

Two varactors are laid out as shown in Figs. 5.42(a) and (b), both with a nominal channel area of $2W_1L_1$.

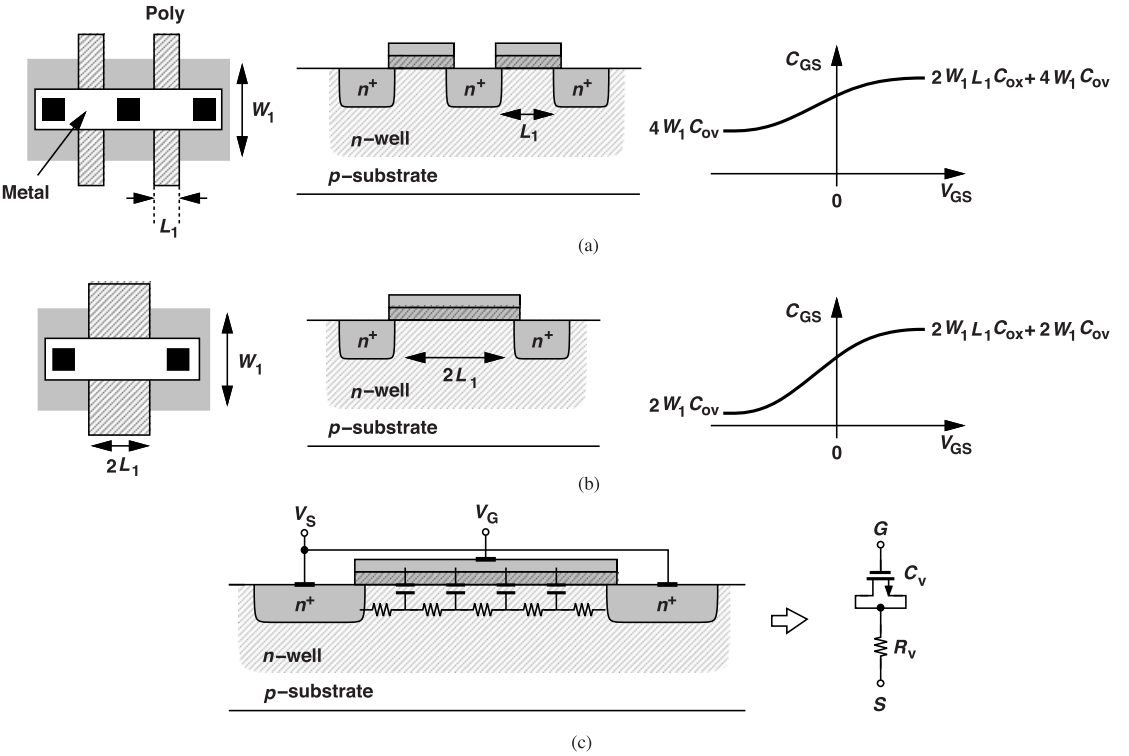


Figure 5.42 (a) MOS varactor, (b) varactor with twice the length, and (c) illustration of effect of channel resistance.

(Assume, for the sake of simplicity, that the drawn and effective lengths are the same.) If L_1 is the minimum allowed by the process, explain which structure exhibits a higher Q and which one a wider tuning range.

Solution

In the structure of Fig. 5.42(a), the overlap capacitance between the gate and the source and drain, $4W_1C_{ov}$,[11] constitutes a large fraction of the total varactor capacitance because L_1 is short. The maximum-to-minimum capacitance ratio is equal to $L_1C_{ox}/(2C_{ov}) + 1$ in this case. In the second structure, on the other hand, the overlap capacitance is equal to $2W_1C_{ov}$, yielding a ratio of $L_1C_{ox}/C_{ov}+1$. That is, the tuning range provided by the latter is greater. Unfortunately, the longer device suffers from a greater distributed resistance between the source and the drain and hence a lower Q. This effect is represented by a lumped model [Fig. 5.42(c)], yielding a Q of $1/(R_vC_v\omega_0)$. For frequencies up to about 10 GHz, a length of 100 to 150 nm can be used for the varactor with little concern about its Q. This is because the varactor is but a small fraction of the total tank capacitance. The trade-off between the capacitance range and the Q manifests itself at frequencies in the range of tens of gigahertz.

Example 5.15

A varactor having a quality factor of Q_v is attached to a tank with a quality factor equal to Q_{tank}. Estimate the overall Q.

Solution

We begin with the arrangement shown in Fig. 5.43(a), where $R_v = 1/(Q_vC_v\omega_0)$. To convert the series resistance of the varactor to a parallel resistance (Chapter 1), we multiply R_v by Q_v^2, obtaining $(R_vC_v^2\omega_0^2)^{-1}$

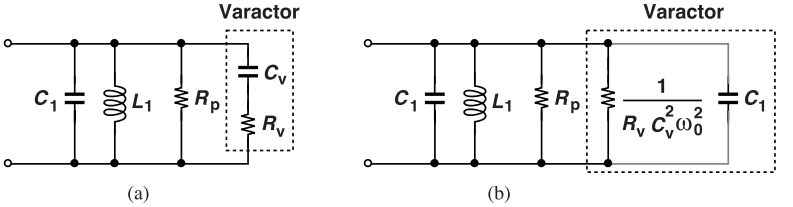


Figure 5.43 (a) Tank including a varactor, and (b) equivalent parallel network.

[Fig. 5.43(b)]. Since this resistance appears in parallel with R_p and since the overall Q can be written as this parallel combination divided by $L_1\omega_0$, we have

$$\frac{1}{Q} = \frac{L_1\omega_0}{R_p \| \dfrac{1}{R_vC_v^2\omega_0^2}} \tag{5.37}$$

$$= L_1\omega_0\left(\frac{1}{R_p} + R_vC_v^2\omega_0^2\right) \tag{5.38}$$

$$= \frac{L_1\omega_0}{R_p} + R_vC_v\omega_0\frac{C_v}{C_1+C_v}. \tag{5.39}$$

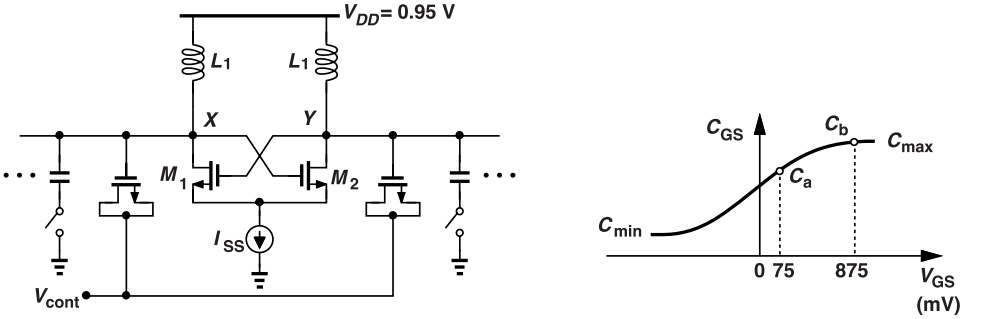
It follows that

$$\frac{1}{Q} = \frac{1}{Q_{tank}} + \frac{1}{Q_v}\frac{C_v}{C_1+C_v}. \tag{5.40}$$

[11] C_{ov} denotes the overlap capacitance per unit width.

For example, if the constant tank capacitance, C_1, is ten times C_v, then the varactor Q is equivalently multiplied by 11, far exceeding the inductor Q.

Let us modify the LC oscillator of Fig. 5.37(a) for operation from 5.1 GHz to 5.9 GHz. To reach $f_0 = 5.9$ GHz with $L_1 = 1.56$ nH, we require a total capacitance of 466 fF. We thus reduce C_1 from 600 fF to about 450 fF, leaving some margin. Next, we estimate that the 800-MHz tuning range can be covered by eight discrete steps, each having a continuous range of about 200 MHz. This partitioning between discrete and continuous control should be revisited if K_{VCO} turns out to be excessively large. Figure 5.44(a) shows the conceptual design.



(a) (b)

Figure 5.44 (a) VCO with discrete tuning, and (b) varactor tuning range.

Continuous Tuning In order to tune f_0 by $\Delta f = 200$ MHz, the varactor capacitance must vary by approximately $2C_1\Delta f/f_0 = 2(200\,\text{MHz}/5.9\,\text{GHz}) \times 450\,\text{fF} = 31$ fF (why?). In Fig. 5.44, V_{cont} is limited to about 75 mV to 875 mV due to the stage preceding it (Chapter 7). We also note that the common-mode level at X and Y is approximately equal to V_{DD}, concluding that the average voltage across the varactors can vary from $0.95\,\text{V} - 75\,\text{mV} = 0.875$ V to $0.95\,\text{V} - 0.875\,\text{V} = 75$ mV. As shown in Fig. 5.44(b), therefore, the varactor capacitance varies across only a limited range, and we require that $C_b - C_a \approx 32$ fF. That is, the varactor gate area must be large enough to provide such a capacitance range for such a voltage range. We estimate that, if the varactor's C_{max}/C_{min} is around 3 and if $C_{max} \approx 80$ fF, then $C_b - C_a$ may reach the necessary value.

Figure 5.45(a) depicts the VCO design.[12] According to simulations, with $(W/L)_{var} = 30\ \mu\text{m}/240$ nm, f_0 varies from 5.7 GHz to 5.9 GHz as V_{cont} goes from 875 mV to 75 mV. The worst-case phase noise occurs for $V_{cont} = 875$ mV and is plotted in Fig. 5.45(b), revealing a value of -54 dBc/Hz at 10-kHz offset. Recall that this value is around -57 dBc/Hz for the circuit of Fig. 5.41(a), which operates at 5.2 GHz. Taking into account $20\log(5.7\,\text{GHz}/5.2\,\text{GHz}) = 0.8$ dB, we observe that the circuit still incurs an additional 2.2 dB degradation after the varactors are added, an effect arising from the AM/PM conversion of the tail flicker noise.

Noise Reduction Our design efforts in Section 5.5.2 and in this section have encountered two phenomena that contribute phase noise: (1) the thermal noise of M_4 in Fig. 5.45(a) at $2f_0$ raises the profile by 3 dB, and (2) the flicker noise of M_4 leads to 11 dB of degradation at 10-kHz offset. We wish to suppress these effects and return to the original phase noise profile shown in Fig. 5.38.

Suppose we interpose a low-pass filter between M_4 and M_3 in Fig. 5.45(a) to attenuate the flicker noise of M_4 at 10 kHz. Choosing a corner frequency of, say, 1 kHz, we recognize the very large resistor(s) and capacitor(s) necessary in such a filter. Instead, we seek a method of directly reducing the flicker noise of M_4 by enlarging its channel area. The challenge is to maintain proper current copying from M_4 to M_3. For

[12] For simplicity, we hereafter do not show the tank R_p in the circuit diagram.

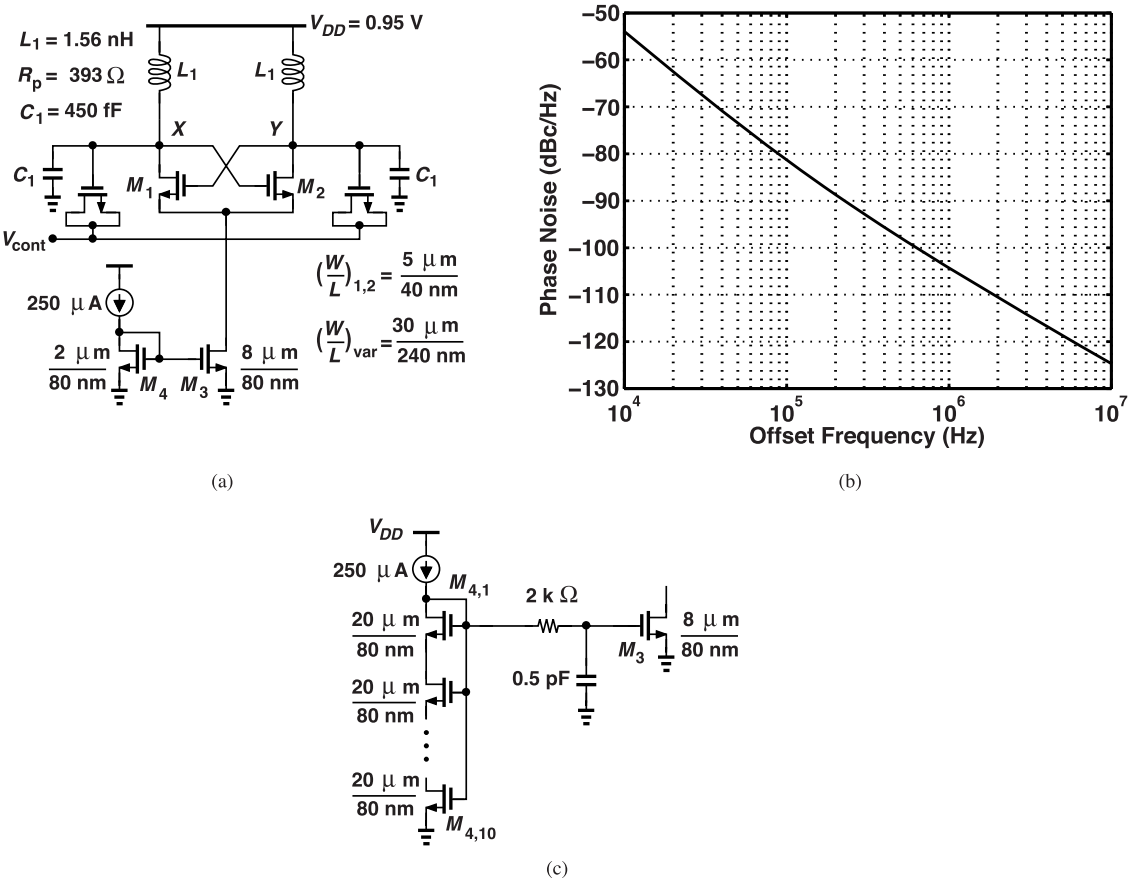


Figure 5.45 (a) LC VCO, (b) simulated phase noise, and (c) reduction of diode-connected transistor's flicker noise.

example, if we simply increase W_4 and L_4 by a factor of 10, then $I_{D3} \neq 4I_{D4}$ because the effective length of M_4 is not the same as 10 times the effective length of M_3. This issue can be resolved as illustrated in Fig. 5.45(c), where M_4 is realized as 10 devices in series, each with $W/L = 20\ \mu\text{m}/80$ nm. Since the effective length of M_4 is now exactly 10 times that of M_3 and since $W_4 = 2.5W_3$, we have $(W/L)_3 = (10/2.5)(W/L)_4 = 4(W/L)_4$ and hence $I_{D3} = 4I_{D4}$. The 100-fold increase in the channel area of M_4 diminishes its flicker noise, and the RC filter reduces its thermal noise. Note that the 10 series devices can be viewed as one diode-connected transistor, i.e., they do *not* act as a cascode topology.

Discrete Tuning In a manner similar to that in Chapter 3, we employ switched capacitors to cover the range from 5.1 GHz to 5.9 GHz.[13] First, we determine the total capacitance, ΔC, that must be added to reach the lower end. Denoting the total capacitance at X in Fig. 5.45(a) by C_X, we seek $C_X + \Delta C$ such that f_0 drops to 5.1 GHz. That is, $(C_X + \Delta C)/C_X = (5.9/5.1)^2$ and hence $\Delta C/C_X = 0.34$. Since $(2\pi\sqrt{L_1 C_X})^{-1} = 5.9$ GHz, we have $\Delta C = 160$ fF. Figure 5.46(a) depicts the result, where C_{e1} and C_{e2} are equal to ΔC.

Example 5.16

In Fig. 5.46(a), M_{S1} and M_{S2} exhibit a channel resistance equal to R_{sw} when they are on. Determine the Q of the circuit in this case if the varactors are lossless. How should R_{sw} be chosen?

[13] The switched capacitors actually need cover only 5.1 GHz to 5.7 GHz because the varactors extend f_0 to 5.9 GHz, but we continue with our conservative design.

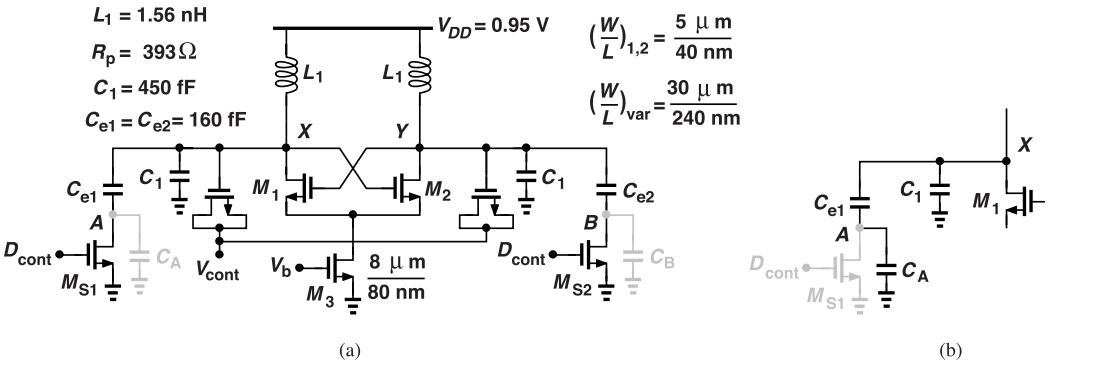


Figure 5.46 (a) VCO of Fig. 5.45(a) with discrete tuning added, and (b) effect of switch parasitic capacitance.

Solution

This situation is similar to that in Example 5.15, yielding

$$\frac{1}{Q} = \frac{1}{Q_{tank}} + \frac{1}{Q_e}\frac{C_e}{C_e + C_1 + C_v}, \tag{5.41}$$

where $Q_{tank} = R_p/(L_1\omega_0)$, $Q_e = 1/(R_{sw}C_e\omega_0)$ is the Q of the switched branches, and $C_e = C_{e1} = C_{e2}$.

In our design, $C_e/(C_e+C_1+C_v) \approx 0.25$, i.e., the Q of the switched capacitors is equivalently quadrupled. For negligible effect on the phase noise, the second term in Eq. (5.41) must be 20 times less than the first. With $Q_{tank} = 8$, we require $Q_e \approx 8 \times 20/4 \approx 40$ and $R_{sw} = 4.9\ \Omega$ at 5.1 GHz. Such a small R_{sw} translates to very wide transistors and considerable parasitic capacitances, C_A and C_B, at nodes A and B, respectively.

The foregoing example reveals that, when M_{S1} and M_{S2} in Fig. 5.46(a) turn off, C_{e1} and C_{e2} do not exit the circuit completely because they now appear in series with their parasitics, C_A and C_B, respectively [Fig. 5.46(b)]. In other words, the discrete tuning range is reduced.

Since an R_{sw} of 4.9 Ω is not practical in this environment, we must relax the factor of 20 assumed for the ratio of the two terms on the right-hand side of (5.41). For example, if $Q_e = 10$, then $R_{sw} = 20\ \Omega$ but the overall Q drops by 20%, degrading the phase noise by about 2 dB. A switch resistance of 20 Ω is still difficult to obtain with moderate transistor dimensions.

In order to remedy the situation, we add a third switch between the bottom plates of the capacitors, as shown in Fig. 5.47(a). Considering the differential waveforms at X and Y and hence at A and B, we note that

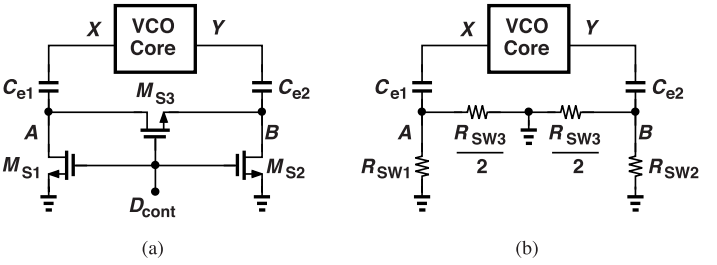


Figure 5.47 Use of differential switch to reduce series resistance.

the on-resistance of M_{S3}, R_{sw3}, can be decomposed into two resistances of value $R_{sw3}/2$, which equivalently

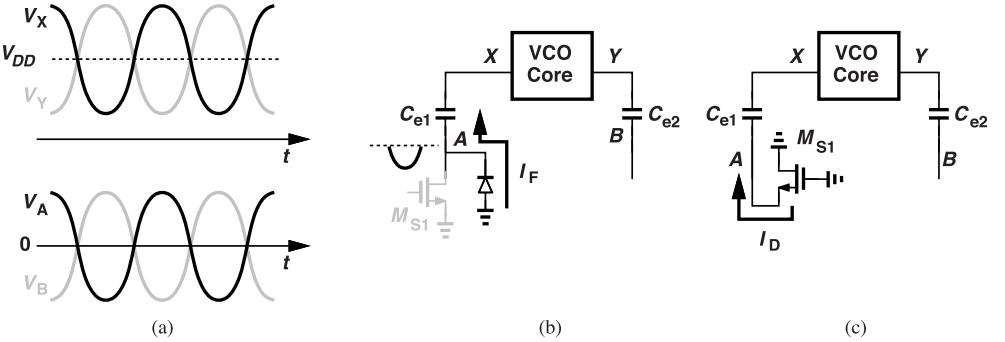
tie A and B to ground [Fig. 5.47(b)]. That is, for a given width, M_{S3} exhibits half the resistance compared to M_{S1} and M_{S2}. In our design example, therefore, we choose $R_{sw3} \approx 40\ \Omega$ and the minimum width for M_{S1} and M_{S2}. The latter two switches are still necessary to keep the dc level at zero.[14]

Example 5.17

If the three switches in Fig. 5.47 are off, plot the waveforms at A and B and explain what happens if the peak voltage swing at X and Y are large.

Solution

With the switches off, the voltage swings at X and Y couple to A and B, respectively, experiencing a slight attenuation due to the parasitic capacitances from A and B to ground. The CM level of V_X and V_Y is close to V_{DD} whereas that of V_A and V_B is around zero. This is because the leakage currents of M_{S1}-M_{S3} discharge these nodes to ground. We thus obtain the waveforms shown in Fig. 5.48(a), recognizing that V_A or V_B travels



(a) (b) (c)

Figure 5.48 (a) Oscillator output and bottom-plate voltage waveforms, and (b) turn-on of the junctions, and (c) turn-on of the MOS devices.

below zero for half of the cycle. As a result, the drain-bulk junction of M_{S1} (or M_{S2}) and the drain-bulk or source-bulk junction of M_{S3} are forward-biased for this duration [Fig. 5.48(c)]. If the peak negative swings at A and B exceed approximately 0.7 V, these diodes turn on heavily, acting as a small incremental resistance to ground.

Even with a peak swing of a few hundred millivolts at A and B, M_{S1} and M_{S2} partially turn on when V_A and V_B are negative enough [Fig. 5.48(c)]. The flicker noise of these transistors then produces significant phase noise because M_{S1} and M_{S2} have small dimensions and they operate in the saturation region (why?).

In order to avoid the issues described in the above example, we modify the design as shown in Fig. 5.49. Here, A and B are tied to V_{DD} through large resistors when M_{S1}-M_{S3} are off. These resistors do not degrade the Q of the oscillator, but their noise modulates the junction capacitances of M_{S1}-M_{S3} and hence the tank resonance frequency.

Upon simulation of the above circuit, we find that f_0 reaches only 5.5 GHz when M_{S1}-M_{S3} are off. This is because the parasitic capacitances of these switches load the tank through C_{e1} and C_{e2}. We therefore reduce, in Fig. 5.46(a), the capacitors to 400 fF.

Figure 5.50(a) depicts the VCO design developed thus far, and Fig. 5.50(b) its tuning characteristics along with the phase noise at 10-kHz and 10-MHz offsets at the four boundaries. The 5-dB degradation at 5.71 GHz with respect to the phase noise of the untuned oscillator in Fig. 5.37(a) arises from the AM/PM conversion of the tail transistor flicker noise.

[14] One can argue that the junction leakage of the main switch brings these nodes to zero, obviating the need for the other two switches.

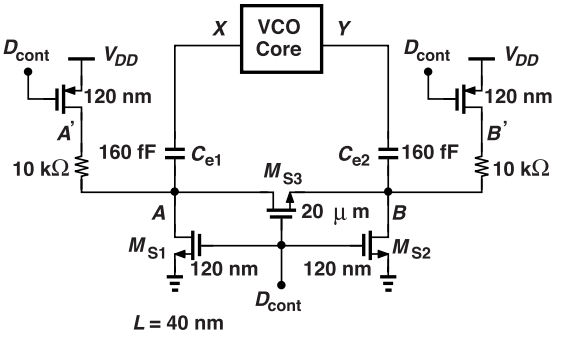


Figure 5.49 Additional pull-up resistors to avoid turn-on of the NMOS switches.

In the last step of our design effort, we decompose the discrete tuning devices in Fig. 5.50(a) so as to obtain overlapping characteristics spanning the range of 5 GHz to 5.9 GHz. As a first guess, we replace C_{e1} and C_{e2} with eight 20-fF capacitors. In principle, all of the devices in the discrete tuning network must now be scaled by a factor of eight, but M_{S1}, M_{S2}, M_{S4} and M_{S5} are minimum-size transistors and cannot be scaled down. Also, the pull-up resistors need not exceed 10 kΩ. We thus arrive at the network depicted in Fig. 5.51(a), where eight identical sections controlled by D_1-D_8 provide the tuning steps. With D_1-D_8 set to logical ZERO, we have $f_0 \approx 5.9$ GHz. To reduce f_0, we raise D_1 to ONE, and then D_2, while D_1 is high, etc. The digital control inputs therefore follow a "thermometer" code. Figure 5.51(b) illustrates the operation.

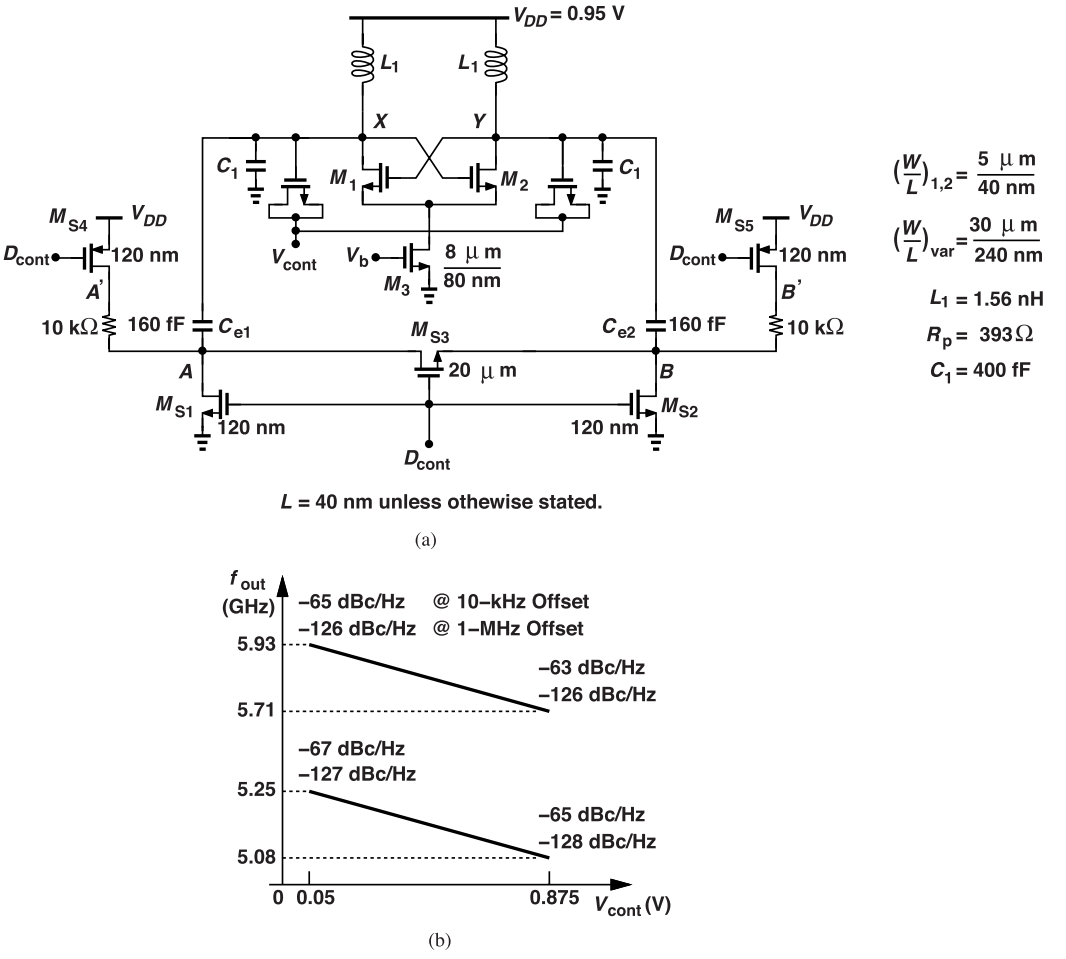
The overall tuning characteristics emerge as shown in Fig. 5.51(c) and merit two remarks. First, the characteristics do not cover equal frequency ranges: the range decreases slightly as more constant capacitors are switched into the tanks (why?). Second, for this reason, the overlap between consecutive characteristics falls from 100 MHz to 80 MHz as D_1-D_8 goes from 0...0 to 1...1. This nonuniformity can be alleviated if the unit capacitors are chosen less than 20 fF for the higher curves and more for the lower.

5.5.4 Summary of Oscillator Design Procedure

The design task carried out in this section follows a systematic procedure:

- Select a power budget, P, and an output voltage swing, V_0; the latter is preferably chosen not to drive the cross-coupled transistors into the triode region. From P and V_0, determine I_{SS} and R_p.
- Find the smallest inductance that yields the necessary value of R_p at f_0. This choice translates to the highest Q for the inductor.
- Choose the necessary width of the cross-coupled transistors to allow complete switching of the tail current with a voltage swing of V_0. In practice, we double or triple this width to ensure abrupt switching, nearly square-wave drain currents, and hence a single-ended peak-to-peak swing close to $(4/\pi)I_{SS}R_p$.
- Add enough capacitance to the tank to obtain the highest desired frequency.
- Using Eq. (5.30), estimate the phase noise; if it does not meet the specification, increase I_{SS} and/or apply linear scaling (or try other oscillator topologies).
- Implement I_{SS} by a current mirror and minimize its noise contribution by proper sizing and filtering.
- Add varactors to obtain a K_{VCO} of about a few percent of ω_0 per volt.
- Add enough constant capacitance to the tanks to obtain the lowest frequency of interest.
- Decompose this capacitance so as to obtain discrete tuning characteristics with adequate overlap.

This procedure provides a first-order design which can then be refined by iteration.



(a)

(b)

Figure 5.50 (a) Final LC VCO design, and (b) its tuning and phase noise performance at 10-kHz and 10-MHz offsets.

References

[1]. S. Jenei, B. K. J. C. Nauwelaers, and S. Decoutere, "Physics-based closed-form inductance expressions for compact modeling of integrated spiral inductors," *IEEE J. of Solid-State Circuits,* vol. 37, pp. 77-80, Jan. 2002.

[2]. S. S. Mohan et al., "Simple accurate expressions for planar spiral inductances," *IEEE J. Solid-State Circuits,* vol. 34, pp. 1419-1424, Oct. 1999.

[3]. A. Niknejad and R. G. Meyer, "Analysis, design, and optimization of spiral inductors and transformers for Si RF ICs," *IEEE J. Solid-State Circuits,* vol. 33, pp. 1470-1481, Oct. 1998.

[4]. M. Kraemer, D. Dragomirescu, and R. Plana, "Accurate electromagnetic simulation and measurement of millimeter-wave inductors in bulk CMOS technology," *IEEE Topical Meeting on Silison Monolithic Integrated Circuits in RF Systems,* Jan. 2010.

[5]. A. Zolfaghari, A. Y. Chan, and B. Razavi, "Stacked inductors and transformers in CMOS technology," *IEEE J. Solid-State Circuits,* vol. 36, pp. 620-628, April 2001.

[6]. B. Razavi, *RF Microelectronics*, Second Edition, Upper Saddle River, NJ: Prentice-Hall, 2012.

[7]. M. Danesh et al., "A Q-factor enhancement technique for MMIC inductors," *Proc. IEEE Radio Frequency Integrated Circuits Symp.*, pp. 217-220, April 1998.

[8]. P. Yue and S. S. Wong, "On-chip spiral inductors with patterned ground shields for Si-based RF ICs," *IEEE J. Solid-State Circuits,* vol. 33, pp. 743-751, May 1998.

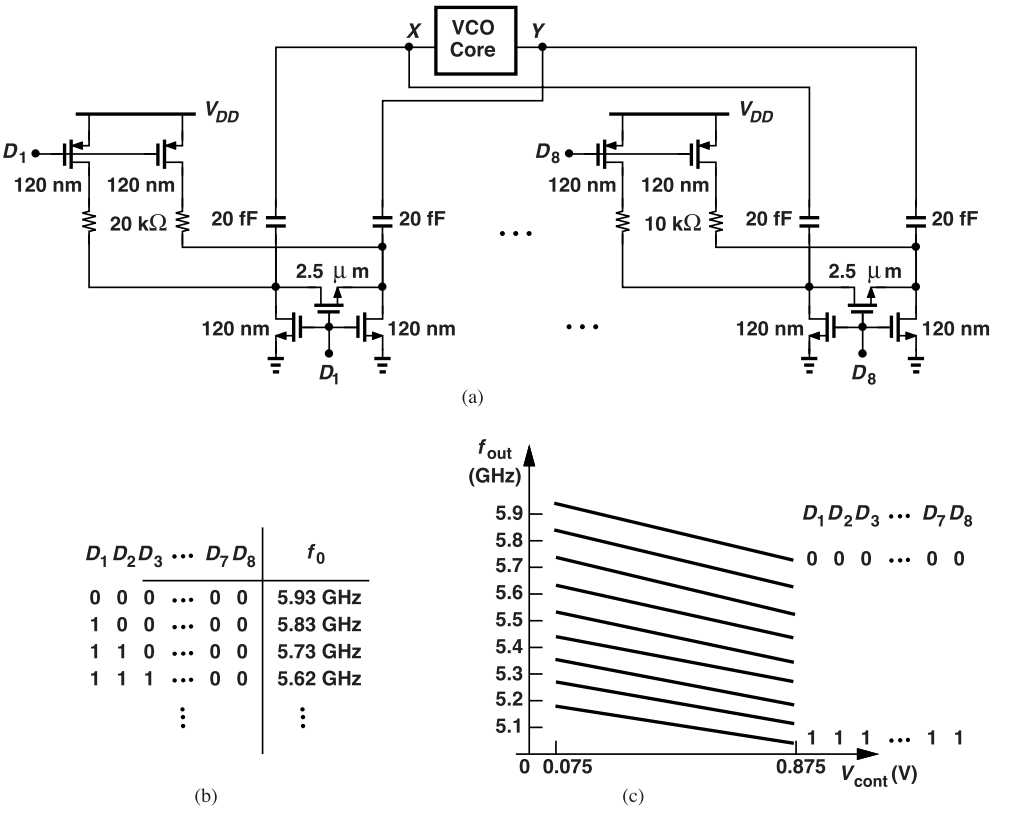


Figure 5.51 (a) VCO of Fig. 5.50(a) with fine discrete tuning, and (b) discrete tuning behavior.

[9]. E. Hegazi, H. Sjoland, and A. A. Abidi, "A filtering technique to lower LC oscillator phase noise," *IEEE J. Solid-State Circuits,* vol. 36, pp. 1921-1930, Dec. 2001.

[10]. P. Andreani and A. Fard, "More on the 1/f phase noise performance of CMOS differential-pair LC -tank oscillators," *IEEE J. Solid-State Circuits,* vol. 41, pp. 2703-2712, Dec. 2006.

[11]. L. W. Couch, *Digital and Analog Communication Systems,* Fourth Edition, New York: Macmillan Co., 1993.

[12]. P. Andreani et al, "A study of phase noise in Colpitts and LC-tank CMOS oscillators," *IEEE J. Solid-State Circuits,* vol. 40, pp. 1107-1118, May 2005.

Problems

5.1. An LC oscillator incorporates a tank consisting of L_1, C_1, and R_p in parallel. Suppose the current injected into the tank resembles a square wave. If L_1 has no series resistance, calculate the ratio of the third harmonic amplitude to the first harmonic amplitude in the voltage across the tank.

5.2. Equation (5.12) appears to suggest that the phase noise does not depend on the tank's inductance value. Is this true?

5.3. Two different (ideal) tanks are designed with the same capacitance but different inductances. Carefully construct the plots in Fig. 5.17 for each tank and identify the similarities and the differences.

5.4. We double the capacitance and halve the inductance in an ideal tank. Explain what happens to the plots in Fig. 5.17.

5.5. Equation (5.12) indicates that the phase noise is inversely proportional to $(C_aV_0)^2$. Note that C_aV_0 is the initial charge on C_a. Can we instead express $S_{\phi n}(f)$ in terms of the peak current flowing through the inductor?

5.6. Prove that, in Fig. 5.18(b), I_{n1} does not flow through L_1 if the impedance seen at the tail node is infinite. Assume M_1 is in saturation. (Hint: write a KCL at the tail node and another at the drain of M_1.)

5.7. Repeat the previous problem if M_1 is in the triode region.

5.8. The parasitic capacitance at the tail node in Fig. 5.18(b) allows some of I_{n1} to flow through L_1. Estimate the maximum tolerable capacitance if 10% of I_{n1} should reach the inductor.

5.9. If V_0 and the oscillation frequency in Fig. 5.18(a) are given, what can we change in the circuit to shorten the time during which both M_1 and M_2 are on?

5.10. In the circuit depicted in Fig. 5.52, switch S_1 turns on and off with 50% duty cycle. Determine the spectrum of $V_{n,out}$.

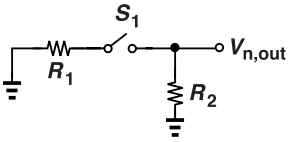


Figure 5.52 Switched network for noise calculation.

5.11. In the circuit of Fig. 5.21, we double I_{SS} and the width of M_1 and M_2. Explain what happens to V_1 and ΔT. Assume the oscillation frequency is relatively constant.

5.12. In the previous problem, what happens to Eqs. (5.21) and (5.23)?

5.13. In the circuit of Fig. 5.21, we halve L_1 and R_p and double the tank capacitance. Explain what happens to V_1 and ΔT. Assume the oscillation frequency is relatively constant.

5.14. In the previous problem, what happens to $S_{inj}(f)$ in Eqs. (5.21) and (5.23)?

5.15. What happens to $G_{m,avg}$ in Fig. 5.24(a) if I_{SS} and $W_{1,2}$ are halved? Assume the oscillation frequency is relatively constant.

5.16. For the LC oscillator of Fig. 5.21, we conduct the following experiment. We double L_1 and L_2 and halve the tank capacitance and I_{SS}. If Q and the oscillation frequency remain constant, R_p is doubled. What happens to the phase noise as predicted by Eq. (5.30)? This case illustrates the trade-off between phase noise and power consumption.

5.17. Repeat the previous problem but with I_{SS} unchanged. Assume that M_1 and M_2 do not enter the triode region despite the larger voltage swings. The lower phase noise simply means that the original design was not optimized—because its voltage swing was not chosen as large as possible.

5.18. We design two simple LC oscillators for frequencies f_0 and $2f_0$. The latter is obtained from the former by halving the tank inductance and capacitance. The other parts of the design remain the same. Using Eq. (5.30), compare the phase noise of the two oscillators if (a) Q is linearly proportional to the frequency (an optimistic assumption), or (b) Q is constant with frequency (a pessimistic assumption).

5.19. In the circuit of Fig. 5.28, the switching action of M_1 and M_2 also mixes I_{nT} with the third harmonic of the oscillation frequency. Examine the noise frequencies in I_{nT} and determine whether any translate to phase noise because of this mixing.

5.20. We wish to reduce the tail noise current at $2\omega_0$ in Fig. 5.27. As shown in Fig. 5.53, we degenerate the current source by a tank resonating at $2\omega_0$. If the tank reduces to a resistance equal to R_T at this frequency, determine the noise current in I_{nT}.

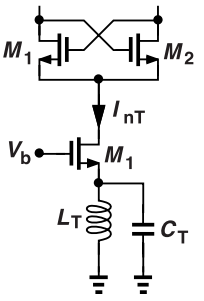


Figure 5.53 Current source degenerated by a tank.

5.21. We observed in Section 5.5.2 that, if the tail current in Fig. 5.37(a) is doubled, the voltage swings more than double (the single-ended swing rises from 425 mV_{pp} to 900 mV_{pp}). Suppose the swing exactly doubles. From Eq. (5.17), determine how ΔT changes. (Hint: the equilibrium overdrive increases by a factor of $\sqrt{2}$.) What does a smaller ΔT mean for the drain current waveform and its first harmonic?

6

Advanced Oscillator Concepts

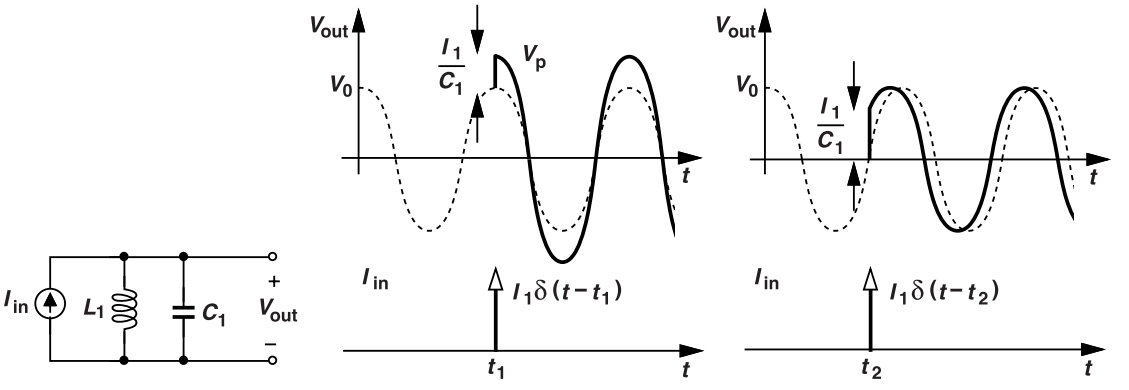
In this chapter, we elevate our study of phase noise and LC oscillators by considering more advanced principles. We first introduce a method pf phase noise analysis that yields some key insights. Next, we analyze advanced LC oscillator topologies that prove useful in practice. Such topologies include complementary cross-coupled oscillators, class-C oscillators, and quadrature oscillators.

6.1 Phase Noise Analysis by Impulse Response

In the previous chapter, we formulated various phase noise mechanisms in the cross-coupled LC oscillator. In some cases, we invoked the principle that, if the rising and falling transitions are symmteric, then the flicker noise of certain transistors does not translate to phase noise. In this section, we follow the analysis in [1] to arrive at this principle.

6.1.1 Phase Impulse Response

Consider an ideal LC tank that, due to an initial condition, produces a sinusoidal output [Fig. 6.1(a)]. During the oscillation, L_1 and C_1 exchange the initial energy, with the former carrying the entire energy at the zero crossings and the latter, at the peaks. Let us assume that the circuit begins with an initial voltage of V_0 across the capacitor. Now, suppose an impulse of current is injected into the oscillating tank at the peak of the output



(a) (b) (c)

Figure 6.1 (a) Ideal tank with a current impulse, (b) effect of impulse injection at peak of waveform, (c) effect of impulse injection at zero crossing of waveform.

voltage [Fig. 6.1(b)], producing a voltage step across C_1. If[1]

$$I_{in}(t) = I_1\delta(t - t_1), \tag{6.1}$$

then the additional energy gives rise to a larger oscillation amplitude:

$$V_p = V_0 + \frac{I_1}{C_1}. \tag{6.2}$$

The key point here is that the injection at the peak does not disturb the *phase* of the oscillation (as shown in the example below).

Next, let us assume the impulse of current is injected at a zero crossing point. A voltage step is again created but leading to a *phase* jump [Fig. 6.1(c)]. Since the voltage jumps from 0 to I_1/C_1, the phase is disturbed by an amount equal to $\sin^{-1}[I_1/(C_1V_0)]$. We therefore conclude that noise creates only amplitude modulation if injected at the peaks and only phase modulation if injected at the zero crossings.

We now plot the change in the output phase as a function of the time at which the current impulse is injected. As explained above, injection at the peaks of V_{out} yields a zero phase change whereas injecting at the zero crossings produces a large change. Figure 6.2 sketches the result, which we can view as the phase impulse response of the oscillator. We derive the shape of this response in Example 6.4.

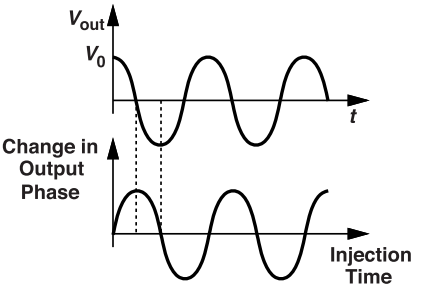


Figure 6.2 Change in output phase for a sinusoidal oscillator.

Example 6.1

Explain how the effect of the current impulse in Fig. 6.1 can be determined analytically. Assume the initial condition on C_1 is created by an impulse at $t = 0$.

Solution

The linearity of the tank allows the use of superposition for the injected currents (the inputs) and the voltage waveforms (the outputs). The output waveform consists of two sinusoidal components, one due to the initial impulse at $t = 0$ (the oscillation waveform) and another due to the impulse at t_1 or t_2. Figure 6.3 illustrates these components for two cases: if injected at t_1, the impulse leads to a sinusoid exactly in phase with the original component, and if injected at t_2, the impulse produces a sinusoid 90° out of phase with respect to the original component. In the former case, the peaks are unaffected, and in the latter, the zero crossings.

Example 6.2

An oscillator produces a square waveform with finite rise and fall times. Sketch the change in the output phase as a function of the time at which a current impulse is injected.

[1] Note that I_1 in this equation is in fact a charge quantity because it denotes the area under the impulse.

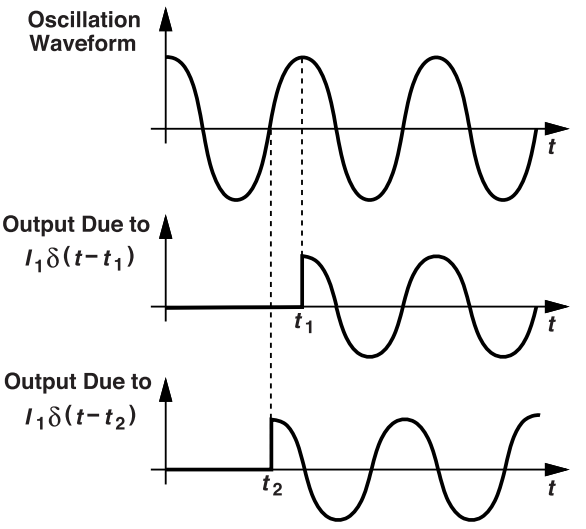


Figure 6.3 Computation of impulse response using superposition.

Solution

As illustrated in Fig. 6.4, the output phase does not change if the impulse "hits" the flat part of V_{out}. It is only

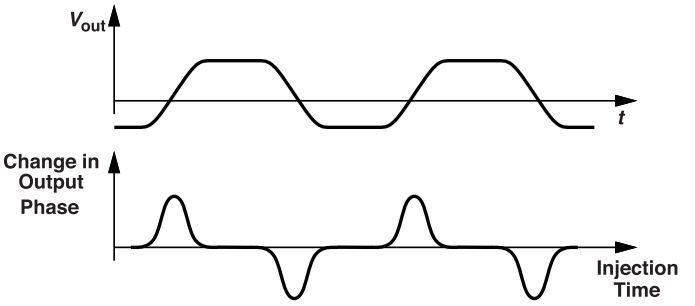


Figure 6.4 Impulse sensitivity for a square-wave oscillator.

during the transitions that a current (or voltage) impulse creates phase change. Note the very different phase impulse responses shown in Figs. 6.2 and 6.4. In both cases, the impulse response suggests a time-variant system because the response depends on the *time* at which the impulse is injected.

The foregoing observations suggest the need for a method of quantifying how and when each source of noise in an oscillator "hits" the output waveform. While the transistors turn on and off, a noise source may only appear near the peaks of the output voltage, contributing negligible phase noise, whereas another may hit the zero crossings, producing substantial phase noise. To this end, we define a linear, *time-variant* system from each noise source to the output phase. The linearity property is justified because the noise levels are very small, and the time variance is necessary to capture the effect of the time at which the noise appears at the output.

For a linear, time-variant system, the convolution property holds but the impulse response varies with time, as exemplified by Figs. 6.2 and 6.4. Thus, the output phase in response to a noise $n(t)$ is given by

$$\phi(t) = h(t,\tau) * n(t), \tag{6.3}$$

where $h(t, \tau)$ is the time-variant impulse response from $n(t)$ to $\phi(t)$. In an oscillator, $h(t, \tau)$ varies *periodically*: as illustrated in Fig. 6.5, a noise impulse injected at $t = t_1$ or at integer multiples of the period thereafter produces the same phase change. Now, the task of output phase noise calculation consists of computing $h(t, \tau)$ for each noise source and convolving it with the noise waveform. The impulse response, $h(t, \tau)$, is called the "impulse sensitivity function" (ISF) in [1].

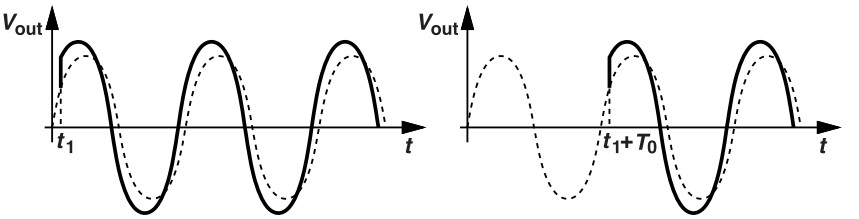


Figure 6.5 Periodic impulse response in an oscillator.

Example 6.3

Explain how the LC tank of Fig. 6.1(a) has a time-variant behavior even though the inductor and the capacitor values remain constant.

Solution

The time variance arises from the finite *initial condition* (e.g., the initial voltage across C_1). With a zero initial condition, the circuit begins with a zero output, exhibiting a time-invariant response to the input. On the other hand, nonzero conditions lead to different responses. This also occurs in a simple, first-order RC circuit.

Example 6.4

Compute the phase impulse response for the lossless LC tank of Fig. 6.1(a), i.e., the shape of the response shown in Fig. 6.2.

Solution

We invoke the superposition perspective of Example 6.1 and wish to calculate the phase change resulting from a current impulse at an arbitrary time t_1 (Fig. 6.6). The overall output voltage can be expressed as

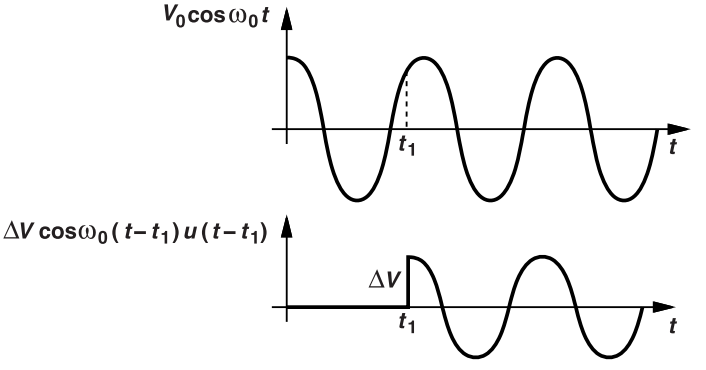


Figure 6.6 Waveforms for computation of phase impulse response of a tank.

$$V_{out}(t) = V_0 \cos\omega_0 t + \Delta V[\cos\omega_0(t - t_1)]u(t - t_1), \tag{6.4}$$

where ΔV is given by the area under the impulse (I_1 in Fig. 6.1) divided by C_1. For $t \geq t_1$, V_{out} is equal to the sum of two sinusoids:

$$V_{out}(t) = V_0 \cos\omega_0 t + \Delta V \cos\omega_0(t - t_1) \quad t \geq t_1 \tag{6.5}$$

which, upon expansion of the second term and regrouping, reduces to

$$V_{out}(t) = (V_0 + \Delta V \cos\omega_0 t_1)\cos\omega_0 t + \Delta V \sin\omega_0 t_1 \sin\omega_0 t \quad t \geq t_1. \tag{6.6}$$

The phase of the output is therefore equal to

$$\phi_{out} = \tan^{-1}\frac{\Delta V \sin\omega_0 t_1}{V_0 + \Delta V \cos\omega_0 t_1} \quad t \geq t_1. \tag{6.7}$$

Interestingly, ϕ_{out} is *not* a linear function of ΔV in general. But, if $\Delta V \ll V_0$, then

$$\phi_{out} \approx \frac{\Delta V}{V_0}\sin\omega_0 t_1 \quad t \geq t_1. \tag{6.8}$$

If normalized to the area under the input impulse (I_1), this result yields the impulse response:

$$h(t, t_1) = \frac{1}{C_1 V_0}\sin\omega_0 t_1 u(t - t_1). \tag{6.9}$$

The response has a quadrature relationship with respect to the oscillation waveform. As expected, $h(t, t_1)$ is zero at $t_1 = 0$ (at the peak of $V_0 \cos\omega_0 t$) and maximum at $t_1 = \pi/(2\omega_0)$ (at the zero crossing of $V_0 \cos\omega_0 t$).

Let us now return to Eq. (6.3) and determine how the convolution is carried out. It is instructive to begin with a linear, time-invariant system. A given input, $x(t)$, can be approximated by a series of time-domain impulses, each carrying the energy of $x(t)$ in a short time span [Fig. 6.7(a)]:

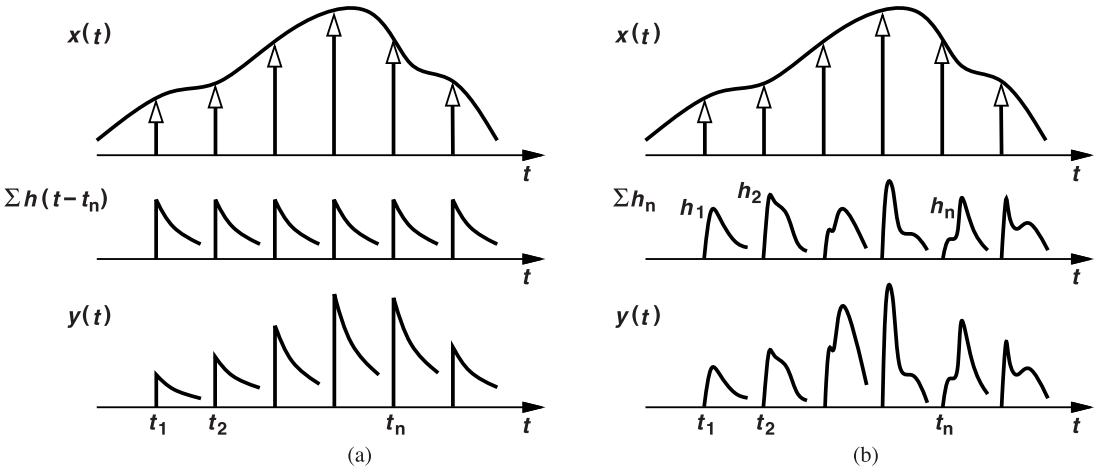


Figure 6.7 Convolution in a (a) time-invariant, and (b) time-variant linear system.

$$x(t) \approx \sum_{n=-\infty}^{+\infty} x(t_n)\delta(t - t_n). \tag{6.10}$$

Each impulse produces the time-invariant impulse response of the system at t_n. Thus, $y(t)$ consists of time-shifted replicas of $h(t)$, each scaled in amplitude according to the corresponding value of $x(t)$:

$$y(t) \approx \sum_{n=-\infty}^{+\infty} x(t_n)h(t - t_n) \tag{6.11}$$

$$= \int_{-\infty}^{+\infty} x(\tau)h(t - \tau)d\tau. \tag{6.12}$$

Now, consider the time-variant system shown in Fig. 6.7(b). In this case, the time-shifted versions of $h(t)$ may be different, and we denote them by $h_1(t)$, $h_2(t)$, ..., $h_n(t)$, with the understanding that $h_j(t)$ is the impulse response in the vicinity of t_j. It follows that

$$y(t) \approx \sum_{n=-\infty}^{+\infty} x(t_n)h_n(t). \tag{6.13}$$

How do we express these impulse responses as a continuous-time function? We simply write them as $h(t, \tau)$, where τ is the specific time shift. For example, $h_1(t) = h(t, 1\text{ ns})$, $h_2(t) = h(t, 2\text{ ns})$, etc. Thus,

$$y(t) = \int_{-\infty}^{+\infty} x(\tau)h(t, \tau)d\tau. \tag{6.14}$$

Example 6.5

A current, $i_n(t)$, having a white spectrum, $S_i(f)$, is injected into the tank of Fig. 6.1(a). Determine the resulting phase noise.

Solution

From (6.9) and (6.14),

$$\phi_n(t) = \int_{-\infty}^{+\infty} i_n(\tau) \frac{1}{C_1V_0} \sin\omega_0\tau u(t - \tau)d\tau \tag{6.15}$$

$$= \frac{1}{C_1V_0} \int_{-\infty}^{t} i_n(\tau) \sin\omega_0\tau d\tau. \tag{6.16}$$

If $i_n(t)$ is white, so is $g(t) = i_n(t) \sin\omega_0 t$ (Problem 6.2), but with *half* the spectral density of $i_n(t)$:

$$S_g(f) = \frac{1}{2} S_i(f). \tag{6.17}$$

Our task therefore reduces to finding the transfer function of a system whose input is $g(t)$ and whose output is $\phi_n(t)$ [Fig. 6.8(a)]. Acting as an integrator, such a system provides a transfer function equal to $H(s) = [1/(C_1V_0)](1/s)$. It follows that

$$S_{\phi n}(f) = |H(j\omega)|^2 S_g(f) \tag{6.18}$$

$$= \frac{1}{C_1^2V_0^2} \frac{1}{(2\pi f)^2} \frac{S_i(f)}{2}. \tag{6.19}$$

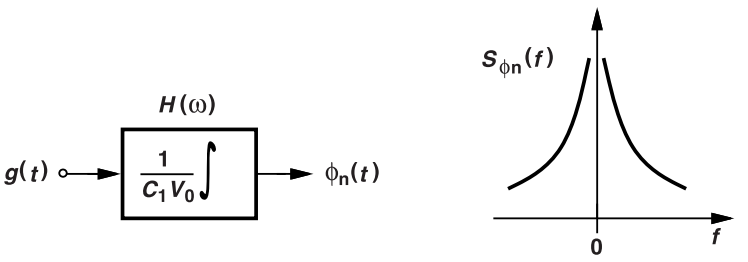


Figure 6.8 (a) Equivalent system for conversion of $g(t)$ to $\phi_n(t)$, (b) resulting phase noise spectrum.

We reconcile this result with Eq. (5.12) in Problem 6.3. As expected, the relative phase noise is inversely proportional to the oscillation peak amplitude, V_0. Depicted in Fig. 6.8(b) is this equation.

The reader may question these results: if the lossless tank with its nonzero initial condition is viewed as an oscillator with infinite Q, why is the phase noise not *zero*? We recognize that, as $Q \to \infty$, the width and bias current of the transistors needed to sustain oscillation become infinitesimally small. The transistors thus inject early zero noise; i.e., if $i_n(t)$ represents transistor noise, $S_i(f)$ approaches zero.

Example 6.6

Which frequency components in $i_n(t)$ in the above example contribute significant phase noise?

Solution

Since $i_n(t)$ is multiplied by $\sin\omega_0 t$, noise components around ω_0 are translated to the vicinity of zero frequency and subsequently appear in (6.19). Thus, for a sinusoidal phase impulse response (ISF), only noise frequencies near ω_0 contribute significant phase noise. In general, we must examine all noise frequency components that can land near zero in $S_{\phi n}(f)$.

In summary, for each noise source, $n(t)$, in an oscillator, we can obtain a phase noise impulse response (an ISF), $h(t,\tau)$, and write $\phi_n(t) = h(t,\tau) * n(t)$ to obtain the corresponding phase noise. As seen above, this time-domain equation must be then manipulated so as to yield the spectrum of ϕ_n.

6.1.2 Effect of Flicker Noise

Owing to its periodic nature, the impulse response of oscillators can be expressed as a Fourier series:

$$h(t,\tau) = [a_0 + a_1\cos(\omega_0 t + \phi_1) + a_2\cos(2\omega_0 t + \phi_2) + \cdots]u(t-\tau), \tag{6.20}$$

where a_0 is the average (or "dc") value of $h(t,\tau)$. In the LC tank studied above, $a_j = 0$ for all $j \neq 1$, but in general this may not be true. In particular, suppose $a_1 \neq 0$ and $a_0 \neq 0$. The former lends itself to the analysis in Example 6.5, translating noise components around ω_0 in $i_n(t)$ to near zero in $S_{\phi n}(f)$. For the latter, on the other hand, the corresponding phase noise in response to an injected noise $i_n(t)$ is equal to:

$$\phi_{n,a0} = \int_{-\infty}^{t} a_0 i_n(\tau)d\tau. \tag{6.21}$$

From Example 6.5, the integration is equivalent to a transfer function of $1/s$, yielding the following spectrum:

$$S_{\phi n,a0}(f) = \frac{a_0^2}{(2\pi f)^2} S_i(f). \tag{6.22}$$

Which components of $S_i(f)$ contribute significant phase noise? Since $S_{\phi n,a0}(f)$ is centered around zero, the *low-frequency* noise in $S_i(f)$, which is simply multiplied by $a_0^2/(2\pi f)^2$, contributes most. The key point here is that, if the "dc" value of $h(t,\tau)$ is nonzero, then the *flicker noise* of the MOS transistors in the oscillator can generate phase noise. Since a_0 relates to the symmetry of $h(t,\tau)$, low upconversion of $1/f$ noise requires a circuit design that exhibits an odd-symmetric $h(t,\tau)$ [1]. This is because an odd periodic function has a zero average. However, the $1/f$ noise of different transistors in the circuit may see different impulse responses, and it may therefore be impossible to minimize the upconversion of *all* $1/f$ noise sources. In general, the phase noise spectrum assumes the shape shown in Fig. 6.9.

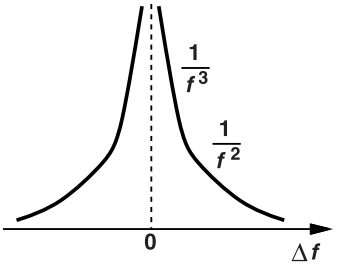


Figure 6.9 Phase noise profile showing regions arising from flicker and white noise.

Let us summarize our thoughts. Suppose for a given noise source in an oscillator, the phase impulse response (the ISF), which is a periodic waveform, is an odd function. Then its Fourier series does not contain a dc term and low-frequency components in this noise source do not become phase noise. On the other hand, if the dc term is nonzero, the large flicker noise of MOSFETs can translate to phase noise.

From the voltage and current waveforms of an oscillator, can we judge whether the ISF is an odd function? In some cases, yes. Specifically, if the rising and falling transitions of voltages or currents are asymmetric, then the phase impulse response also becomes asymmetric. Depicted in Fig. 6.10 is an example illustrating the asymmetric phase jumps due to unequal rise and fall times.

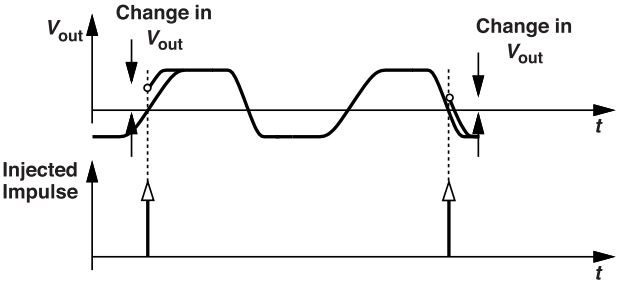


Figure 6.10 Effect of unequal rise and fall times on impulse sensitivity.

This correspondence between rise/fall times and the ISF fails to predict flicker noise upconversion in two cases: (1) in inverter-based ring oscillators, and (2) in the simple cross-coupled LC oscillator if its transistors obey the square law and operate in saturation.

6.1.3 Cyclostationary Noise

We must also incorporate the effect of cyclostationary noise. Such noise can be viewed as stationary noise, $n(t)$, multiplied by a periodic envelope, $e(t)$. Equation (6.14) can thus be written as

$$y(t) = \int_{-\infty}^{+\infty} n(\tau)e(\tau)h(t,\tau)d\tau, \tag{6.23}$$

implying that $e(t)h(t,\tau)$ can be viewed as an "effective" impulse response [1]. In other words, the effect of $n(t)$ on phase noise ultimately depends on the *product* of the cyclostationary noise envelope and $h(t,\tau)$.

This approach to phase noise analysis generally requires that both the noise envelope and the impulse response be determined from multiple simulations for each device. Design optimization may therefore prove a lengthy task.

6.2 Current-Limited versus Voltage-Limited Phase Noise

Our LC oscillator design studies in Chapter 5 have assumed that the bias current is determined by a power budget. Let us explore what happens to the phase noise of the simple cross-coupled topology if the power restriction is lifted but the supply voltage remains constant.

Beginning with a low tail current, we observe the voltage swings are small, the cross-coupled transistors avoid the triode region, and $S_{\phi n} \propto 1/I_{SS}^2$. That is, doubling I_{SS} shifts the entire phase noise profile down by 6 dB. The phase noise is thus "current-limited." This trend continues until the single-ended peak-to-peak swing exceeds V_{TH}, at which point the transistors enter the triode region for part of the period. As observed in Chapter 5, the flicker-noise regime does not shift down anymore.

As we further increase I_{SS}, the voltage swings are ultimately bounded by the supply voltage. Beyond this point, raising the bias current does not lower $S_{\phi n}$. The phase noise is now "voltage-limited." In a more general sense, we say the phase noise is voltage-limited if increasing the voltage swing (e.g., by increasing I_{SS} or R_p) does not lower the phase noise.

6.3 Oscillators with Complementary Cross-Coupled Pairs

Recall from Chapters 1 and 5 that cross-coupled transistors act as a negative resistance in LC oscillators. We wish to explore the use of both NMOS and PMOS cross-coupled pairs for this purpose. To this end, we begin with a "single-pair" LC oscillator, disconnect the common node of the two tanks from V_{DD} and add a PMOS pair as shown in Fig. 6.11(a). Here, the bias current path from V_{DD} to ground is provided by M_3 and M_4 rather than by the tanks. In practice, of course, the two tanks are merged into one, and the inductor is realized as a symmetric spiral. This arrangement is sometimes called a complementary LC oscillator.

From a small-signal perspective, the two pairs operate in parallel, providing a total transconductance of $g_{mN}+g_{mP}$, i.e., a higher loop gain for a given I_{SS}. However, the principal advantage of this topology relates to its greater voltage swings compared to those of its single-pair counterparts. To understand this point, we draw the circuit at one extreme—when M_2 and M_3 are off [Fig. 6.11(b)] and recognize that I_{SS} now passes through two tanks *in series*, thereby generating a differential peak voltage swing equal to $(8/\pi)R_pI_{SS}$. Note that each tank carries a peak current of $+I_{SS}$ in one half cycle and $-I_{SS}$ in the other half. By comparison, the tanks in the LC oscillators of previous chapters carry I_{SS} or zero. Thus, for a given supply current and inductor Q, the complementary topology provides twice as much voltage swing.

Example 6.7

Determine the output common-mode level of the oscillator in Fig. 6.11(a).

Solution

We consider the time at which $V_X = V_Y$ and note that each side carries a current equal to $I_{SS}/2$. In fact, as far as the CM level is concerned, the circuit can be reduced to that shown in Fig. 6.12, revealing that M_3 and

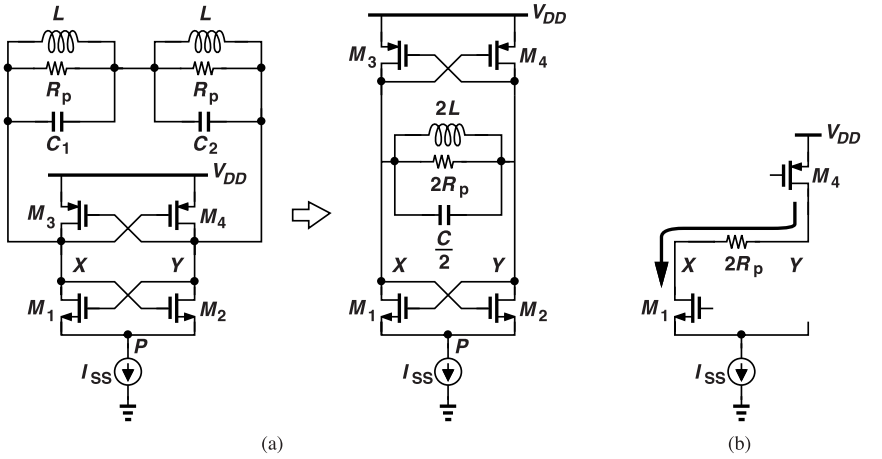


Figure 6.11 (a) Oscillator with complementary cross-coupled pairs, and (b) snapshot of the circuit.

M_4 appear as two diode-connected devices. Alternatively, we can say that the inductor acts as a short circuit

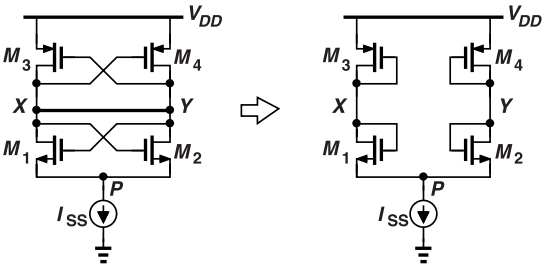


Figure 6.12 Computation of output CM level.

for low-frequency noise arriving from I_{SS}. The CM level is given by

$$V_{CM} = V_{DD} - |V_{GSP,eq}|, \tag{6.24}$$

where $|V_{GSP,eq}|$ denotes the gate-source voltage of the PMOS transistors when their drain currents are equal to $I_{SS}/2$. An interesting point here is that small fluctuations in I_{SS} are multiplied by $1/(g_{m3} + g_{m4})$ as they translate to CM disturbance. Does this effect occur in the topology of Fig. 5.46(a)? (Hint: how much is the output CM level in that circuit?)

The twofold increase in voltage swings afforded by the complementary pairs in Fig. 6.11(a) implies a more relaxed trade-off between power consumption and phase noise. Recall from Example 5.8 that doubling the swing halves the time during which the transistors inject noise into the tank. The same effect occurs here, leading to a 6-dB reduction in phase noise even though, in this case, four transistors produce noise [2]. That is, Eq. (5.30) should be divided by 4. It is important to note that this much advantage accrues only if two conditions are met. First, M_1-M_4 do not enter the triode region. Second, the (parasitic) capacitances seen from X and Y in Fig. 6.11(a) to *ground* are negligible with respect to the "floating" tank capacitances, C_1 and C_2 [2]. Otherwise, the PMOS transistors—with their sources tied to V_{DD} rather than to a current source—inject additional noise, degrading the phase noise [2]. As explained below, the improvement is in fact around 3 dB in low-voltage designs.

6.3.1 Design Issues

The complementary oscillator of Fig. 6.11(a) faces several serious difficulties in today's technology. In this section, we examine the circuit more closely. For simplicity, some diagrams do not show the tank capacitances.

Unlike the topology of Fig. 5.46(a), the complementary arrangement consumes one more V_{GS} in the voltage headroom. As illustrated in Fig. 6.12, the circuit resembles a stack of two diode-connected devices and a current source when V_X and V_Y cross. Thus, $V_{GSN} + |V_{GSP}| + V_{ISS} = V_{DD}$, where V_{ISS} is the voltage required by I_{SS}. We can allocate a smaller V_{ISS} but at the cost of higher tail noise current. For this reason, M_1-M_4 in Fig. 6.11(a) are chosen much wider than M_1-M_2 in Fig. 5.46(a). Also, if available in the process technology, low-threshold devices are used for M_1-M_4.

Example 6.8

The design example described in Section 5.5.2 assumed $I_{SS} = 1$ mA and a peak-to-peak single-ended voltage swing of 0.5 V. How do we modify the circuit for complementary operation?

Solution

If we keep $I_{SS} = 1$ mA and add a PMOS pair, we expect the output swing to double, but with a 0.95-V supply, that is not possible. In reality, M_1-M_4 in Fig. 6.11(a) enter the triode region and the tail transistor's V_{DS} collapses. This means that, for a given R_p, I_{SS} must be *halved* as we change to the complementary topology (if we wish to avoid deep-triode operation) (Fig. 6.13). Also, if M_3 and M_4 enter the triode region, they act as resistors, discharging the single-ended capacitance at their respective drains to V_{DD} and hence lowering the Q.

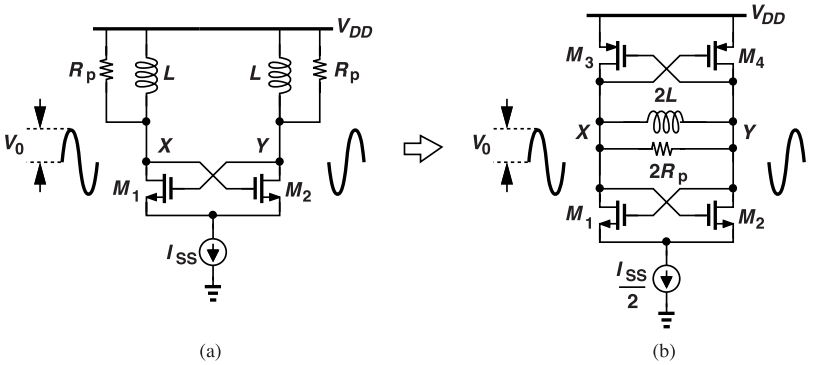


Figure 6.13 (a) Single-pair, and (b) complementary oscillators using scaled tail current sources.

How does the transformation depicted in Fig. 6.13 affect the phase noise? In principle, the phase noise does not change because the 6-dB advantage of the complementary arrangement has been nullified by the twofold reduction in I_{SS} [Eq. (5.30)]. We can now apply linear scaling (Chapter 5) to reduce the phase noise by 3 dB while raising the tail current back to I_{SS}. The key point here is that the complementary oscillator provides a 3-dB advantage if, due to voltage headroom constraints, the output swing must remain constant.

Another important drawback of the complementary oscillator is that its output CM level is modulated by the tail current source. Illustrated in Fig. 6.14(a), this effect can be formulated as $V_{CM} = I_{nT}/(2g_{m3,4})$ (Problem 6.4). Now consider the VCO shown in Fig. 6.14(b), noting that, for a fixed V_{cont}, the average voltage across the varactors changes with the output CM level and hence with the tail noise. In other words, the tail noise modulates the oscillation frequency because the CM disturbance given by $I_{nT}/(2g_{m3,4})$ can be

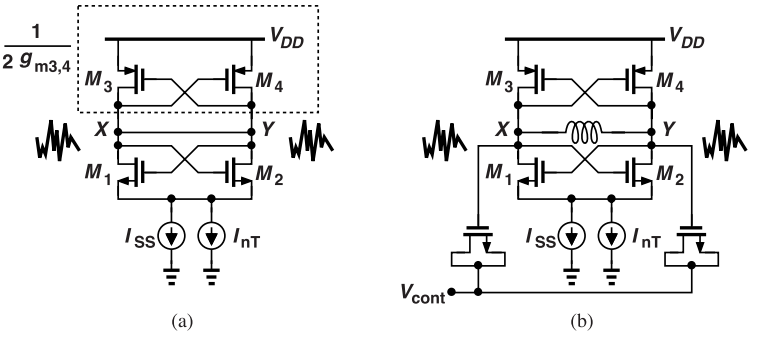


Figure 6.14 (a) Effect of tail noise on output CM level, and (b) resulting phase noise with varactors present.

simply placed in series with V_{cont}. The phase noise thus generated is written as

$$S_{\phi n}(f) = \frac{I_{nT}^2}{4g_{m3,4}^2} \frac{K_{VCO}^2}{(2\pi f)^2}. \tag{6.25}$$

We should remark that this phenomenon entails direct modulation of the varactors and is different from and stronger than AM/PM conversion (Section 5.3.2). In Problem 6.5, we explain why the above effect is less pronounced in the oscillator of Fig. 5.46(a).The wide transistors necessary in this topology also limit its maximum oscillation frequency.

Example 6.9

In a "top-biased" LC oscillator, we place the current source as shown in Fig. 6.15(a) so that the output CM level is around $V_{DD}/2$ and hence more compatible with the next stage. Does the low-frequency noise of I_T modulate the varactors?

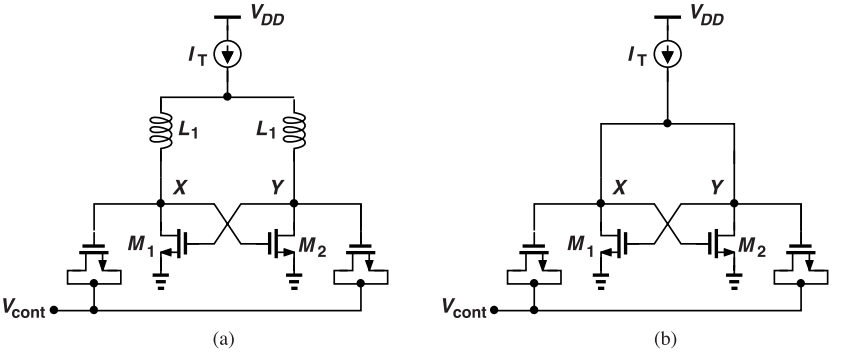


Figure 6.15 (a) Top-biased LC oscillator, and (b) its equivalent for calculating the output CM level.

Solution

Yes, it does. In a manner similar to that in Fig. 6.12, we can simplify the structure as shown in Fig. 6.15(b), concluding that fluctuations in I_T translate to CM modulation.

6.3.2 Design Example

Let us modify the design in Fig. 5.37(a) for complementary operation. Following the observations in the previous section and the transformation depicted in Fig. 6.13, we construct the oscillator shown in Fig. 6.16(a). Wide, low-threshold transistors are chosen to allow a reasonable voltage headroom for the tail current source.

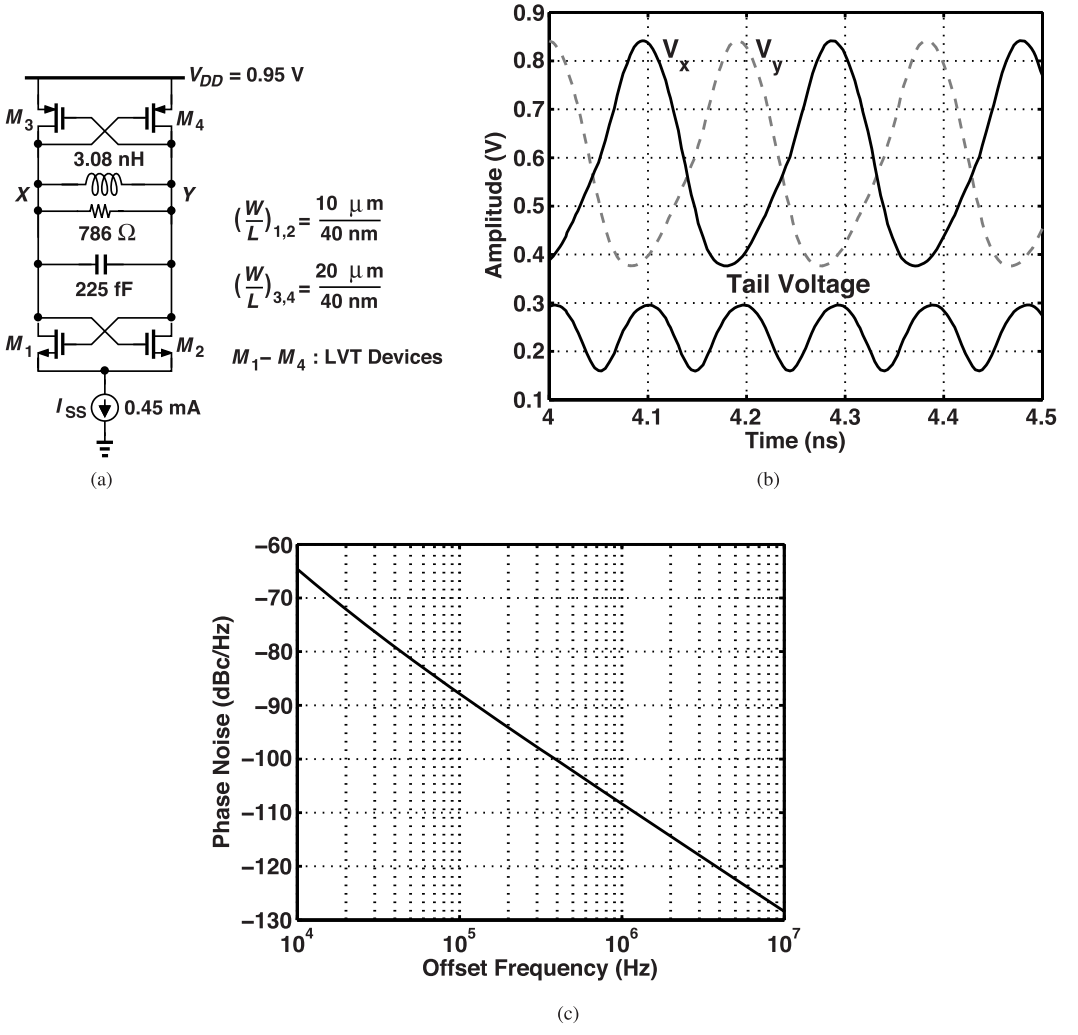


Figure 6.16 (a) Complementary oscillator design, (b) simulated output waveforms, and (c) simulated phase noise.

The constant tank capacitance is reduced to allow oscillation at 5 GHz. Interestingly, the new oscillator has *greater* output swings even though its tail current is reduced to 0.5 mA. This is because the active pull-up provided by the PMOS transistors sharpens the turn-on transition for the NMOS devices, thus raising the first harmonic amplitude. According to simulations, the higher swings drive the transistors further into the triode region and lead to substantial flicker noise upconversion. For this reason, we lower I_{SS} to 450 μA.

Figure 6.16(b) plots the circuit's waveforms, and Fig. 6.16(c) the phase noise. Owing to the asymmetric rise and fall times at X (or at Y), the flicker noise of M_1 and M_2 raises the phase noise at 10-kHz offset to -65 dBc (3 dB higher than that in Fig. 5.38).[2] But, at higher offset frequencies, the phase noise is close

[2]One can adjust W_P/W_N to improve the symmetry, but according to simulations, it does not have much effect.

to the profile in Fig. 5.38. We can, of course, apply linear scaling and lower the entire profile by 3 dB with $I_{SS} \approx 900\ \mu$A.

Let us summarize the pros and cons of the complementary LC oscillator of Fig. 6.16(a) with respect to the single-pair topology of Fig. 5.37(a). If the transistors must avoid the triode region—where their flicker noise contribution rises considerably—then the complementary circuit provides a 3-dB phase noise advantage in the thermal regime. However, it must also deal with both the limited voltage headroom and the modulation of the output CM level by the tail noise current. We can replace the tail current source with a short circuit, but then the bias current varies significantly with PVT. The following example addresses this point.

Example 6.10

In most integrated systems, the LC oscillators are fed by a dedicated supply voltage that is derived from the global supply. Shown in Fig. 6.17 is an example employing a "low-drop-out" (LDO) regulator, which creates

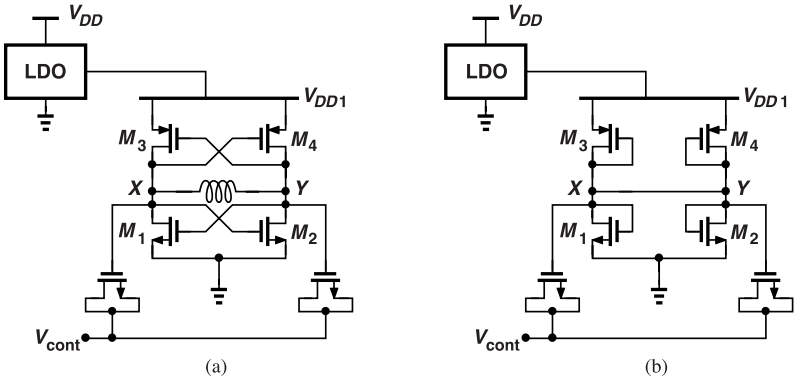


Figure 6.17 (a) Use of an LDO to reduce supply noise, and (b) effect on output CM level and varactors.

a "clean" supply from V_{DD}. (a) Does this approach solve the problem of PVT variations? (b) If the LDO output contains high flicker noise, explain what happens.

Solution

(a) No, it does not. Since typical LDO designs aim to keep their output voltage *constant*, PVT variations still produce large changes in the bias current. Taking a snapshot of the oscillator when $V_X = V_Y$, we construct the equivalent circuit shown in Fig. 6.17(b) and recognize that V_{DD1} is impressed across two diode-connected devices in series, namely, $M_1||M_2$ and $M_3||M_4$. Thus, if V_{DD1} remains constant, the variation of transistor characteristics with process and temperature leads to a substantial change in the bias current and hence the voltage swings.

(b) If V_{DD1} suffers from high flicker noise, then so does the CM level at X and Y. For small signals, we can write a voltage divider expression from Fig. 6.17(b) as $V_{CM} = [(g_{m3}+g_{m4})/(g_{m1}+g_{m2}+g_{m3}+g_{m4})]V_{DD1}$. That is, if the NMOS and PMOS devices have comparable transconductances, roughly half of the flicker noise voltage in V_{DD1} appears at the output and modulates the varactors.

Another issue related to grounded sources is that, if M_1 and M_2 enter the triode region, they discharge their respective drain capacitances, thereby lowering the Q.

It is possible to suppress the effect of low-frequency noise on the CM level seen by the varactors. We recognize that the varactors must connect to the tank at *high* frequencies but must remain isolated from the CM level at *low* frequencies. This can be accomplished by ac coupling. Illustrated in Fig. 6.18(a), the idea is to tie the varactors to the oscillator core through coupling capacitors, C_C, and select the bias, V_b, around $V_{DD}/2$. The bias resistors, R_b, must be large enough not to degrade the tank Q. Now, the low-frequency

noise in I_{SS} modulates the CM level at X and Y but, as seen in the equivalent circuit of Fig. 6.18(b), flows through a high-pass filter consisting of C_C and R_b before reaching the varactors.

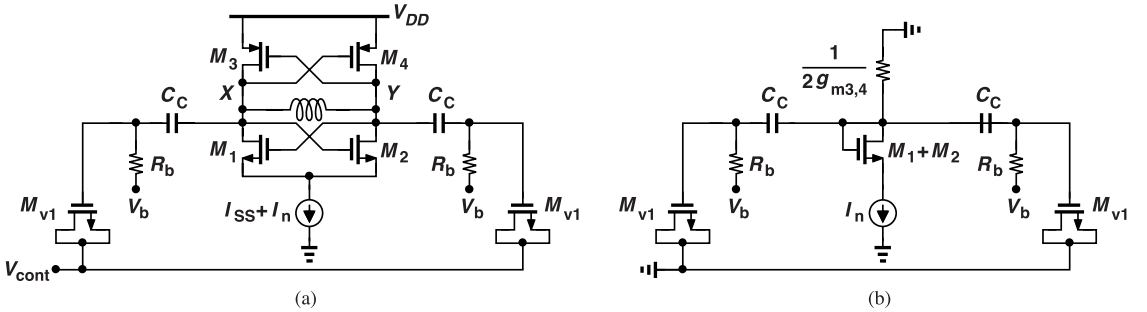


Figure 6.18 (a) Complementary oscillator with ac-coupled varactors, and (b) equivalent circuit for low-frequency tail noise.

Capacitive coupling also proves useful in this and other LC oscillator topologies for suppressing the effect of low-frequency *supply* noise on the varactors. In Fig. 6.18(b), slow fluctuations in V_{DD} do not propagate to the varactors. The principal drawback, however, is that the low-frequency thermal noise of each R_b modulates its corresponding varactor. That is, R_b cannot be excessively small or large.

6.4 Class-C Oscillators

In the differential LC oscillators that we have thus far developed, an implicit assumption has been that the maximum current carried by each transistor is equal to the tail current, I_{SS}. This property naturally accompanies differential operation, but it need not generally hold. In this section, we explore how and why we may depart from this condition, eventually arriving at "class-C oscillators" [3].

Consider the oscillator shown in Fig. 6.19(a), where a large capacitor, C_T, is tied from P to ground so

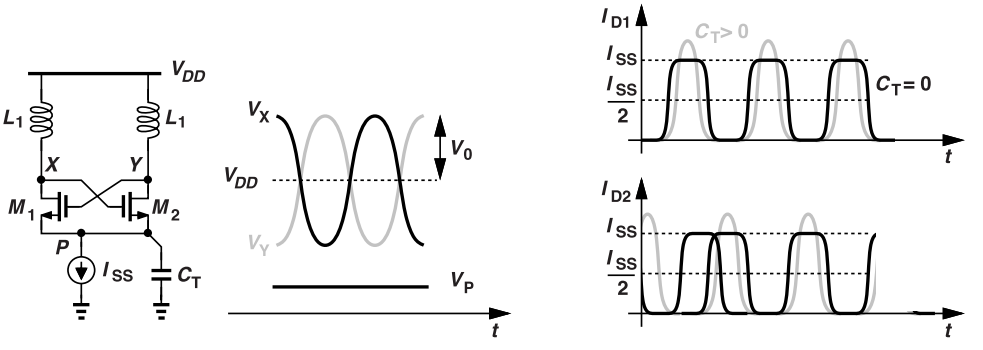


Figure 6.19 (a) LC oscillator with tail capacitance, and (b) effect of capacitance on drain current waveforms.

that V_P does not change with time. We assume that M_1 and M_2 do not enter the triode region. What happens when V_X reaches its peak value, $V_{DD}+V_0$? The drain current of M_2 is now determined by the strength of this transistor and $V_{GS} = V_{DD}+V_0-V_P$ and need not be limited to I_{SS}. For an ideal square-law device, we can write $I_{D2} = (1/2)\mu_n C_{ox}(W/L)(V_{DD}+V_0-V_P-V_{TH})^2$. (The same holds for M_1 when $V_Y = V_{DD}+V_0$.)

We now ask, if the peak currents of M_1 and M_2 in Fig. 6.19(a) exceed I_{SS}, how can the average current drawn from V_{DD} still be equal to the tail current, I_{SS}? This is possible only if each transistor conducts for *less* than half a cycle so as to maintain an average drain current of $I_{SS}/2$. Plotted in Fig. 6.19(b), the drain

currents exhibit class-C operation, i.e., a transistor conduction angle of less than 180°. The key point here is that the taller current pulses provide a first harmonic amplitude greater than $(4/\pi)(I_{SS}/2)$, thereby generating proportionally larger voltage swings. This swing advantage is, in the ideal case, equal to $\pi/2 = 3.9$ dB [3].

The current waveforms shown in Fig. 6.19(b) may also suggest that the transistors inject less noise into the tank and remain simultaneously on for a shorter amount of time. However, it is shown in [3] that the transistors' contribution to the phase noise does not change when C_T is added. In other words, the principal advantage of class-C operation is larger voltage swings and, as a result, lower normalized phase noise—by 3.9 dB in the ideal case. Of course, C_T also removes the high-frequency noise of I_{SS} (Section 5.3).

We should make three remarks here that are explained in [3]. First, the phase noise advantage of class-C operation vanishes if the transistors enter the deep triode region. Second, if C_T in Fig. 6.19(a) is excessively large, then the output waveform experiences amplitude modulation, an effect arising from low-frequency instability. To avoid this issue, C_T should be chosen less than about three to four times the tank capacitance. Third, it is possible to create a bias network, such as that in Fig. 5.40, that ensures "deeper" class-C operation.

6.4.1 Design Example

Let us redesign the basic topology of Fig. 5.37(a) for class-C operation. We add a C_T of 2 pF and also widen the transistors to 20 μm so as to achieve a large peak drain current. Shown in Fig. 6.20(a) is the result along with the drain current waveforms in Fig. 6.20(b) and V_X and V_Y in Fig. 6.20(c). The tail bias remains unchanged. We observe a peak current of 1.75 mA and a peak-to-peak single-ended voltage swing of 620 mV, about 2 dB higher than that of the original circuit.[3] Figure 6.20(d) plots the phase noise, confirming the 2-dB advantage. In the case of high supply noise, we prefer to tie C_T from the tail node to V_{DD} so that supply variations do not modulate the drain currents of M_1 and M_2 (Problem 6.7).

The 2-dB improvement in the performance costs no additional power consumption but requires a 2-pF capacitor. If, for example, we target a phase noise of -113 dBc/Hz at 1-MHz for a particular application, we must resort to linear scaling, which doubles I_{SS} and C_T, halves the inductance, and demands a large area for the tail capacitor.

6.5 Phase Noise Reduction by Frequency Division

Our phase noise studies in Chapter 3 suggest that no phase noise FOM advantage accrues by operating a ring oscillator at a higher frequency and dividing its output down to obtain the desired f_0. In this section, we explore this approach for LC oscillators and show that it does prove beneficial in certain cases.

We begin with the circuit shown in Fig. 6.21(a) and assume it is designed for $\omega_0 = 1/\sqrt{L_1 C_1}$, with C_1 representing the entire single-ended capacitance seen at X or Y. Let us redesign the oscillator for operation at $\omega_1 = 2\omega_0$ while keeping I_{SS} and W/L constant. To this end, we halve both L_1 and C_1. What happens to Q and R_p? Let us consider two extreme scenarios.

In the optimistic scenario, we make the following observations: (1) the simplest physical inductor model includes only a constant series (wire) resistance, R_s, and exhibits a Q equal to $L_1\omega_0/R_s$ at $\omega = \omega_0$; (2) if L_1 is halved, we can assume that R_s is approximately halved; and (3) the Q at $\omega = 2\omega_0$ is therefore twice that at $\omega = \omega_0$. As a result, R_p goes from $L_1\omega_0 Q|_{\omega 0}$ to $(L_1/2)(2\omega_0)Q_{2\omega 0}$, i.e., it is doubled. Interestingly, in this idealized scenario, the output voltage swing is doubled (which may not be possible in reality). Figure 6.21(b) shows the resulting circuit. We should caution the reader that, as explained in Section 5.1, in practice, the tank Q does not double because of the substrate loss, the skin effect, and the declining Q of varactors and switched capacitors.

Let us investigate the thermal phase noise in this scenario. The expression derived in Section 5.2.4,

$$S_{\phi n}(f) = \left(\frac{\sqrt{2}}{2}\gamma + 1\right)\frac{\pi^2 kT}{2R_p I_{SS}^2}\left(\frac{f_0}{2Qf}\right)^2, \tag{6.26}$$

[3] In this case, we assume that the original circuit is designed with $W_{1,2} = 20$ μm (rather than 5 μm) to achieve high currents in class-C operation.

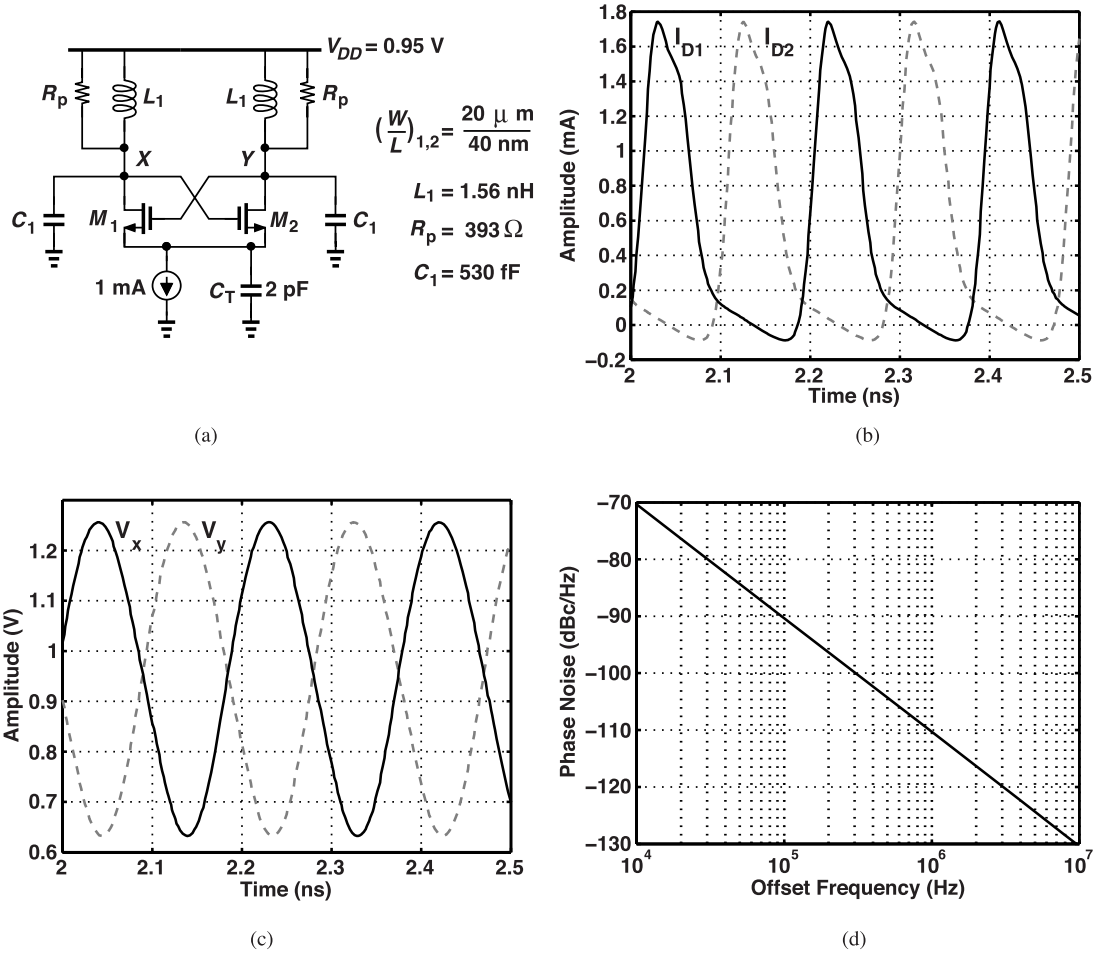


Figure 6.20 (a) Class-C LC oscillator design, (b) drain current waveforms, (c) output voltage waveforms, and (d) phase noise.

can be rewritten as

$$S_{\phi n}(f) = \left(\frac{\sqrt{2}}{2}\gamma + 1\right) \frac{\pi kT}{4I_{SS}^2} \frac{f_0}{Q^3 L_1} \frac{1}{(2f)^2}. \tag{6.27}$$

Thus, if f_0 is doubled, L_1 is halved, and Q is doubled, then $S_{\phi n}(f)$ falls by a factor of 2. Upon dividing the output frequency by a factor of 2 to obtain the desired value, we benefit from another factor of 4 reduction in $S_{\phi n}$ because the waveform changes from $\cos(2\omega_0 t + \phi_n)$ to $\cos(\omega_0 t + \phi_n/2)$. We then conclude that, in the optimistic scenario, the FOM improves by a factor of 8 if (1) $Q \propto \omega$, and (2) the frequency divider phase noise and power consumption are negligible. This behavior stands in sharp contrast to that of ring oscillators.

As a pessimistic scenario, let us now assume that the Q remains *constant* as we go from ω_0 to $2\omega_0$. Thus, R_p also remains constant because $R_p = L_1\omega_0 Q|_{\omega 0} = (L_1/2)(2\omega_0)Q|_{2\omega 0}$. Equation (6.26) then yields a fourfold increase in $S_{\phi n}$, which, after frequency division, returns to its original value. In this case, frequency division offers no phase noise benefit.

Our study can be summarized as follows. The phase noise-power trade-off can be improved by frequency division if two conditions hold: (1) the Q of the tank rises at higher frequencies, and (2) the divider draws

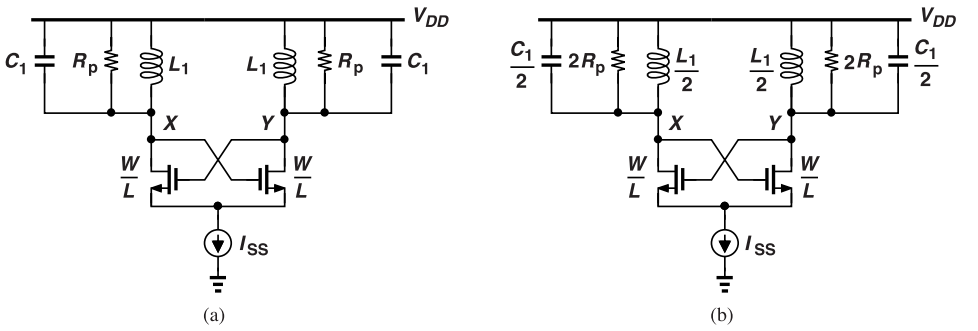


Figure 6.21 (a) LC oscillator, and (b) its redesign for operation at twice the original frequency.

negligible power.[4] Today's technology satisfies these conditions for oscillation frequencies up to a few tens of gigahertz. This method provides two other advantages: the oscillator inductor occupies less area at twice the desired frequency, and the divider can provide quadrature phases (Section 6.6). We should add, however, that the varactor Q falls at high frequencies, posing limitations on this technique.

6.6 Quadrature Generation Techniques

In the study of ring oscillators in Chapter 3, we noted that differential topologies can readily yield quadrature phases for an even number of stages in the loop. In this section, we wish to devise other methods of quadrature generation.

6.6.1 Frequency Division

A common approach to producing quadrature phases at f_0 is to design a differential oscillator running at $2f_0$ and follow it by a master-slave $\div 2$ circuit. Illustrated conceptually in Fig. 6.22(a), such an arrangement generates an in-phase output, I, and a quadrature output, Q.

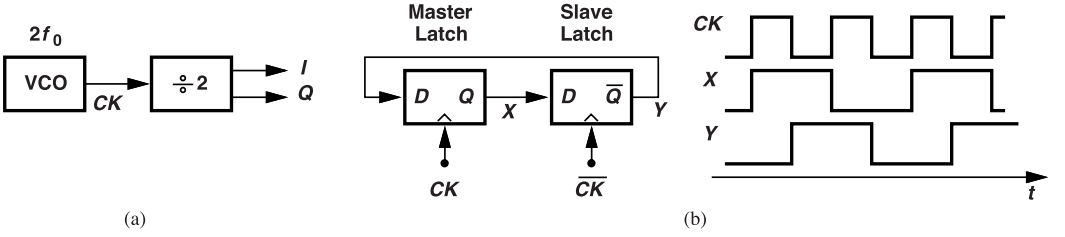


Figure 6.22 (a) Divide-by-2 circuit operating as quadrature generator, and (b) implementation and waveforms.

The $\div 2$ circuit implementation is shown in Fig. 6.22(b): two D latches are placed in a negative-feedback loop (one latch inverts) and clocked by complementary signals, CK and $\overline{CK}$. Recall from basic logic design that a D latch allows its input, D, to go to its output, Q, when its clock is high. For now, we call the two latch outputs X and Y.[5] If, for example, both X and Y begin from zero, then when CK goes high, so does X (why?). Subsequently, when CK falls and $\overline{CK}$ rises, X propagates to the output of the slave latch, forcing Y to go high. On the next rising edge of CK, Y propagates to X and drops X to zero. Similarly, on the next

[4] Dividers' phase noise is typically negligible.

[5] It is an unfortunate coincidence that the letter Q denotes four different quantities in our field: the output of a latch, the quadrature phase of signals, the quality factor of passive devices, and the error function, $Q(x)$.

falling edge of CK, $X = 0$ reaches Y. Since X and Y change only on the rising and falling edges of CK, respectively, they exhibit a frequency of f_0 and a phase difference of 90°. Thus, X and Y correspond to the I and Q outputs in Fig. 6.22(a). Various circuit realizations of the latches are described in Chapter 15.

The use of a $\div 2$ circuit to generate quadrature signals is the most robust and efficient choice, but it requires that the oscillator run at twice the desired frequency and, more importantly, the divider operate reliably at $2f_0$. Low-power $\div 2$ circuits face an upper limit of 15 to 20 GHz in 40-nm technology; for higher speeds, more power-hungry dividers become necessary (Chapter 15).

6.6.2 Quadrature LC Oscillators

In applications where divider-based quadrature generation fails or consumes excessive power, we turn to quadrature LC oscillators. Before studying this approach, we must understand the concept of "injection" or "coupling" into an oscillator.

Consider a simple differential LC oscillator running at f_0. We call f_0 the "natural oscillation frequency." We can inject another periodic signal into the oscillator as shown in Fig. 6.23(a), where M_3 and M_4 convert

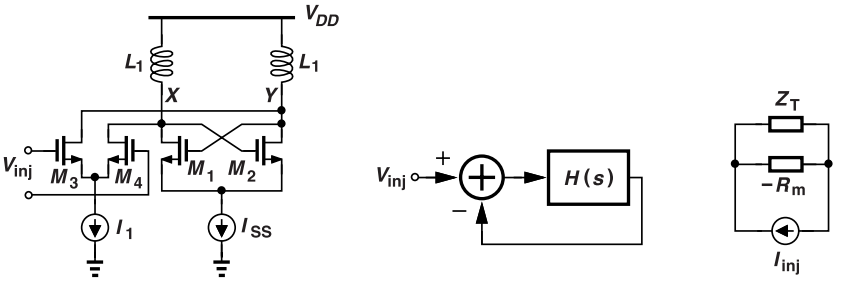


Figure 6.23 (a) Injection into an oscillator, (b) feedback model, and (c) one-port model.

V_{inj} to a differential current. The strength of the injection is determined by the amplitude of V_{inj}, the value of I_1, and possibly the dimensions of M_3 and M_4. For example, with complete current switching in all four transistors, the relative strength, also called the "coupling coefficient," is equal to I_1/I_{SS}.

We can represent the coupling effect by two different models. First, the oscillator can be viewed as a feedback system, with V_{inj} acting as an input [Fig. 6.23(b)]. Second, the oscillator can be reduced to a one-port model (Chapter 1), with the injection represented by a current source [Fig. 6.23(c)]. Here, Z_T represents the tank and $-R_m$ the negative resistance. Both forms help us develop insights.

The reader may wonder how the oscillator behaves upon receiving an external periodic injection. Depending on the frequency and strength of the input and the Q of the tanks, the oscillator can experience "injection locking" or "injection pulling." These interesting effects have been studied extensively [6, 7] but are beyond the scope of this book. The key point for us here is that, if the injection frequency, f_{inj}, is sufficiently close to the natural frequency of the oscillator and if the coupling coefficient is large enough, the oscillator "locks" to the input and runs at f_{inj} rather than at f_0.

Basic Quadrature Generation Let us now couple two nominally-identical oscillators to each other. As illustrated in Fig. 6.24, the output voltage of each is converted to current and injected into the other. We can distinguish between two cases, namely, if the coupling differential pairs invert neither or both outputs [Fig. 6.24(a), "in-phase" coupling], or if only one pair inverts one output [Fig. 6.24(b), "anti-phase" coupling]. Our objective is to determine the oscillation frequencies and the output phases of these two configurations. The use of unilateral injection proves essential here, as explained in the following example.

Example 6.11

Can we replace the (unilateral) coupling differential pairs in Fig. 6.24 with resistors (Fig. 6.25)?

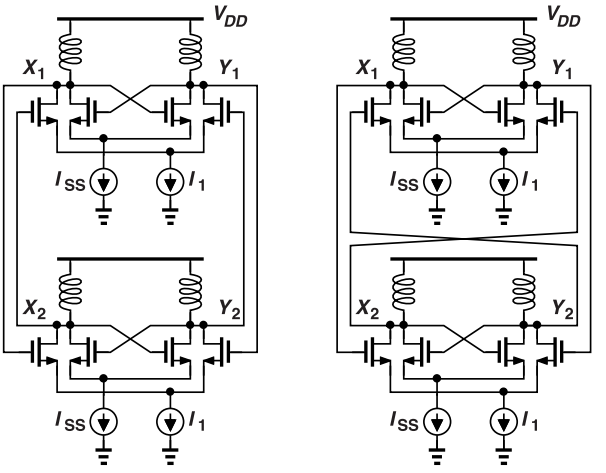


Figure 6.24 (a) In-phase, and (b) anti-phase mutual coupling between two oscillators.

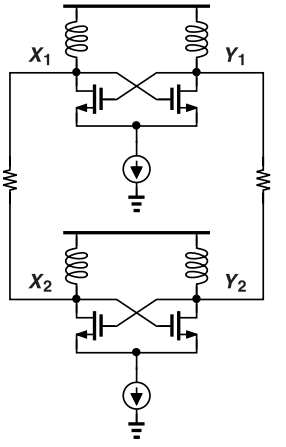


Figure 6.25 Mutual coupling through resistors.

Solution

With resistors, the coupling becomes bilateral (reciprocal), allowing only in-phase coupling. As explained below, quadrature generation requires anti-phase coupling.

We begin our analysis using the feedback model [8]. As shown in Fig. 6.26, the unilateral injection provided by the differential pairs is represented by gain blocks α_1 and α_2, yielding

$$(\alpha_1 A - B)H(s) = B \tag{6.28}$$

$$(\alpha_2 B - A)H(s) = A. \tag{6.29}$$

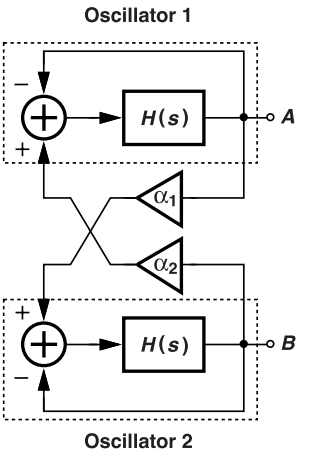


Figure 6.26 Feedback model of two mutually-coupled oscillators.

If the oscillators indeed oscillate, then A and B are nonzero. If, in addition, $\alpha_1 A - B \neq 0$ and $\alpha_2 B - A \neq 0$, we can divide both sides of (6.28) by those of (6.29), obtaining

$$\alpha_1 A^2 = \alpha_2 B^2. \tag{6.30}$$

We observe that $A = \pm B$ if $\alpha_1 = \alpha_2$ (in-phase coupling) and $A = \pm jB$ if $\alpha_1 = -\alpha_2$ (anti-phase coupling). In the former case, the two oscillators operate with a phase difference of 0° or 180° and in the latter, with a phase difference of $\pm 90^\circ$. In Problem 6.12, we use this model to show how two inverter-based rings can operate differentially.

Example 6.12

How do we know that the oscillators in Fig. 6.26 operate at the same frequency?

Solution

Since the forward transfer function, $H(s)$, is assumed the same for both, the two oscillators have equal natural frequencies. In other words, we have assumed sufficiently small mismatches between the two so that they can be locked to each other by injection.

We can repeat the analysis with the aid of the one-port model. Shown in Fig. 6.27, the two oscillators are coupled through dependent current sources $G_{m1}V_A$ and $G_{m2}V_B$, which model the coupling differential pairs

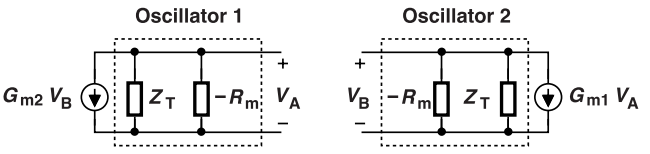


Figure 6.27 One-port model of two mutually-coupled oscillators.

in Fig. 6.24. Writing the parallel combination of Z_T and $-R_m$ as $-Z_T R_m/(Z_T - R_m)$, we have

$$-G_{m1}V_A \frac{-Z_T R_m}{Z_T - R_m} = V_B \tag{6.31}$$

$$-G_{m2}V_B\frac{-Z_T R_m}{Z_T - R_m} = V_A. \tag{6.32}$$

Assuming that we can divide both sides of (6.31) by those of (6.32), we arrive at

$$G_{m1}V_A^2 = G_{m2}V_B^2. \tag{6.33}$$

Thus, $G_{m1} = G_{m2}$ leads to $V_A = \pm V_B$, and $G_{m1} = -G_{m2}$ to $V_A = \pm jV_B$.

Departure from Resonance Let us assume that the oscillators in our study are designed with imaginary (rather than complex) poles, just meeting the startup condition. If the coupling coefficients, α_1 and α_2, in Fig. 6.26 are zero, the two stand-alone oscillators operate at ω_0 such that $H(j\omega_0) = -1$, i.e., $|H(j\omega_0)| = 1$ and $\angle H(j\omega_0) = -180°$. Similarly, if $G_{m1} = 0$ and $G_{m2} = 0$ in Fig. 6.27, the tanks reduce to R_p at ω_0 and $R_m = R_p$ for proper startup. With $\alpha_1 = -\alpha_2 \neq 0$, on the other hand, we have $A = \pm jB$ and, from (6.28),

$$H(j\omega_{osc}) = \frac{-1}{1 \pm j\alpha_1}. \tag{6.34}$$

Since $H \neq -1$, ω_{osc} departs from ω_0. In fact, as shown in Fig. 6.28, ω_{osc} adjusts itself so that

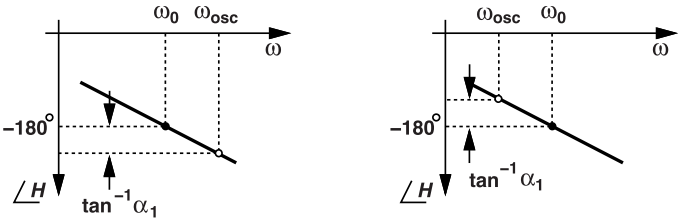


Figure 6.28 (a) Positive, or (b) negative departure from resonance due to mutual coupling.

$$\angle H(j\omega_{osc}) = -180° \pm \tan^{-1}\alpha_1. \tag{6.35}$$

By the same token, substituting $V_A = \pm jV_B$ in (6.31) gives

$$Z_T = \frac{R_m}{1 \pm jG_{m1}R_m} \tag{6.36}$$

and hence

$$\angle Z_T = \pm\tan^{-1}(G_{m1}R_m). \tag{6.37}$$

That is, the tank does not operate at resonance and *must* produce a phase of $\pm\tan^{-1}(G_{m1}R_m)$. Note that $G_{m1}R_m = G_{m1}R_p = \alpha_1$.

Example 6.13

Determine the oscillation frequency of the quadrature topology shown in Fig. 6.29(a). Assume complete current switching for all of the transistors.

Solution

Let us first examine the transistor drain currents closely. As shown in Fig. 6.29(a), V_{X1} and V_{Y1} are differential phasors, and V_{Y2} is perpendicular to them. Noting that I_{D1} and I_{D3} are in phase with V_{Y1} and V_{Y2},

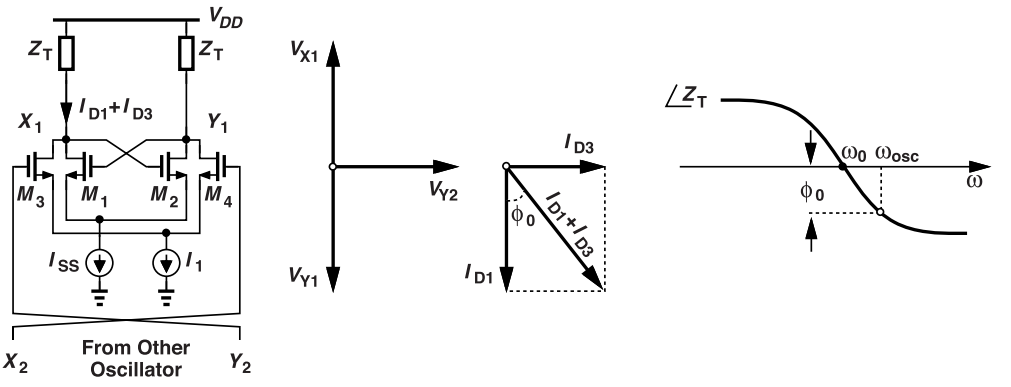


Figure 6.29 (a) Mutually-coupled oscillators with their voltage and current phasors, and (b) departure from resonance to create the additional phase shift ϕ_0.

respectively, we obtain the phasor diagram shown for these currents and their sum. Upon flowing through Z_T, $I_{D1}+I_{D3}$ must generate $-V_{X1}$; i.e., the tank must *rotate* $I_{D1}+I_{D3}$ by an amount ϕ_0 as it converts this current to voltage. This is possible only away from resonance [Fig. 6.29(b)]. Equating the phase generated by the tank to $\phi_0 = \tan^{-1}(I_{D3}/I_{D1})$, we have

$$\frac{\pi}{2} - \tan^{-1}\frac{L_1\omega_{osc}}{R_p(1-L_1C_1\omega_{osc}^2)} = -\tan^{-1}\frac{I_{D3}}{I_{D1}}. \tag{6.38}$$

The reader recognizes $I_{D3}/I_{D1} = I_1/I_{SS}$ as the coupling coefficient, α. Since $Q = R_p/(L_1\omega_{osc})$ and $\tan^{-1}a - \tan^{-1}b = \tan^{-1}[(a-b)/(1+ab)]$, we rewrite (6.38) as

$$\frac{\pi}{2} = \tan^{-1}\frac{1}{Q(1-L_1C_1\omega_{osc}^2)} - \tan^{-1}\alpha \tag{6.39}$$

$$= \tan^{-1}\frac{1-\alpha Q(1-L_1C_1\omega_{osc}^2)}{Q(1-L_1C_1\omega_{osc}^2)+\alpha}. \tag{6.40}$$

Setting the denominator on the right-hand side to zero gives

$$\omega_{osc} = \frac{1}{\sqrt{L_1C_1}}\sqrt{\frac{\alpha}{Q}+1}. \tag{6.41}$$

As expected, the departure from $\omega_0 = 1/\sqrt{L_1C_1}$ becomes smaller if α falls or Q rises.

As explained below, α typically does not exceed 0.25, allowing the following approximation:

$$\omega_{osc} \approx \frac{1}{\sqrt{L_1C_1}}\left(1+\frac{\alpha}{2Q}\right) \tag{6.42}$$

$$\approx \omega_0 + \alpha\frac{\omega_0}{2Q}. \tag{6.43}$$

This deviation can also be estimated from Fig. 6.29(b): since $Q = (\omega_0/2)d\phi/d\omega$, for a phase change of $d\phi = \tan^{-1}\alpha \approx \alpha$, we require a frequency departure of $d\omega = [\omega_0/(2Q)]\alpha$.

Design Issues The quadrature oscillator of Fig. 6.24(b) entails a number of drawbacks with respect to simple, differential LC topologies. First, it requires two symmetric inductors, occupying a larger area and complicating the routing of the signals. Second, it exhibits a higher phase noise. Let us delve into the latter point.

Equation (6.43) implies that, in Fig. 6.29(a),

$$\omega_{osc} \approx \omega_0 + \frac{\omega_0}{2Q}\frac{I_1}{I_{SS}}. \tag{6.44}$$

Thus, the noise in both I_1 and I_{SS} modulates the coupling coefficient and hence the oscillation frequency. Denoting these noise components by I_{n1} and I_{nSS}, respectively, we have

$$\omega_{osc} \approx \omega_0 + \frac{\omega_0}{2Q}\frac{I_1 + I_{n1}}{I_{SS} + I_{nSS}} \tag{6.45}$$

$$\approx \omega_0 + \frac{\omega_0}{2Q}\frac{I_1}{I_{SS}}\left(1 + \frac{I_{n1}}{I_1} - \frac{I_{nSS}}{I_{SS}}\right). \tag{6.46}$$

The frequency experiences random noise equal to $[\omega_0/(2Q)](I_1/I_{SS})(I_{n1}/I_1 - I_{nSS}/I_{SS})$. We must multiply the spectrum of this quantity by $|1/s|^2 = 1/\omega^2$ to obtain the phase noise. Since I_{n1} and I_{nSS} are uncorrelated:

$$S_{\phi n}(\omega) = 2\frac{I_1^2}{I_{SS}^2}\left(\frac{\overline{I_{n1}^2}}{I_1^2} + \frac{\overline{I_{nSS}^2}}{I_{SS}^2}\right)\left(\frac{\omega_0}{2Q\omega}\right)^2, \tag{6.47}$$

where the factor of 2 accounts for tail noise in both the top and the bottom oscillators. The key point is that, in contrast to the behavior of simple differential LC oscillators, the tail flicker noise always produces phase noise here.

From Eq. (6.47), we conclude that $I_1^2/I_{SS}^2 = \alpha^2$ must be minimized. How small can the coupling coefficient be? In practice, the two oscillators in Fig. 6.24(b) incur small mismatches, possibly failing to synchronize (to "injection lock") to each other if α is not large enough. In such a case, the oscillators "pull" each other, producing a heavily-corrupted output. If the resonance frequencies of the oscillators are denoted by ω_1 and ω_2, then the minimum α that guarantees synchronization is given by [9]:

$$\alpha = 2Q\frac{\omega_1 - \omega_2}{\omega_1 + \omega_2}. \tag{6.48}$$

We typically choose α in the range of 0.2 to 0.25. Even if mutually locked, the mismatched oscillators still exhibit phase and amplitude imbalances [10, 11].

The third drawback of quadrature LC oscillators arises from the possibility of operation at one of *two* frequencies, which we call "modes." As illustrated in Fig. 6.28, $\angle H(j\omega_{osc})$ has two solutions. To appreciate this point, we return to Fig. 6.29(a) and recognize the assumption that V_{Y2} *lags* V_{X1} by 90°, which ultimately means that the tank carrying $I_{D1} + I_{D3}$ rotates the current phasor by ϕ_0 *clockwise*. That is, ω_{osc} lies above ω_0. But what happens if V_{Y2} *leads* V_{X1}? As shown in Fig. 6.30(a), the resultant current, $I_{D1} + I_{D3}$, must now rotate counterclockwise by ϕ_0 to yield V_{X1}, requiring that ω_{osc} be *less* than ω_0 [Fig. 6.30(b)]. In this case, the right hand side of (6.38) changes to $+\tan^{-1}(I_{D3}/I_{D1})$, and (6.42) is revised to

$$\omega_{osc} \approx \frac{1}{\sqrt{L_1C_1}}\left(1 - \frac{\alpha}{2Q}\right). \tag{6.49}$$

The two oscillation frequencies have indeed been observed in practice [4, 5]. The difficulty is that, if the oscillator starts at one value, the tuning range must suffice to bring it to the other value. For example, with $\alpha = 0.2$ and $Q = 8$, about 2.5% of the tuning range is "wasted" for this purpose. The work in [5] shows that the shift to one mode or the other depends on the relative losses of the tank's inductance and capacitance and also on the imaginary and real parts of the coupling coefficient (which in reality is complex rather than real).

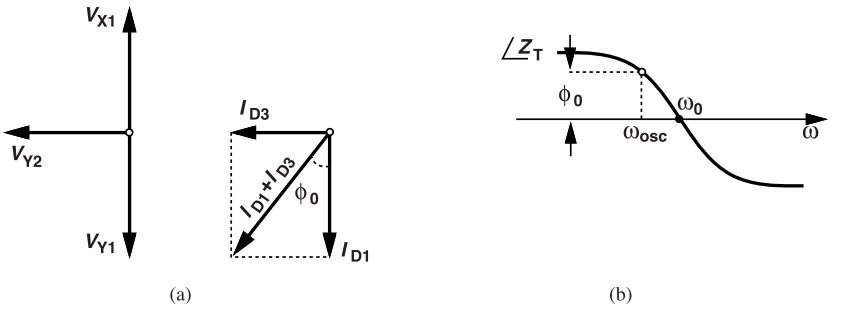


Figure 6.30 (a) Case of quadrature oscillator locked with V_{Y2} leading V_{X1}, and (b) resulting departure from resonance.

Design Example Let us develop a quadrature oscillator using the LC topology shown in Fig. 5.37(a). Depicted in Fig. 6.31(a), the arrangement employs a coupling coefficient of 0.2, realized by scaling the coupling differential pairs and their tail currents with respect to the cross-coupled cores. Figure 6.31(b) plots the four single-ended waveforms, and Fig. 6.31(c) the phase noise in one differential output.

The results in Fig. 6.31 provide a great deal of insight. First, the oscillation amplitude has decreased with respect to that in Fig. 5.37(b), even though the peak current flowing through the tank has increased. This is because the tank impedance decreases away from ω_0. Second, the phase noise has dramatically climbed at low offset frequencies and by 2 dB at 10-MHz offset. This degradation proves even more serious if we recall that the overall power consumption is more than doubled. The 2-dB phase noise increase at 10-MHz offset in Fig. 6.31(c) with respect to the original differential oscillator, taken together with the power penalty, translates to an FOM degradation of 5.4 dB in the thermal regime.

As predicted by (6.47), the phase noise at low offsets is dominated by the flicker noise of the tail current sources. The following example elaborates on this effect.

Example 6.14

Determine the relative significance of the terms $\overline{I_{n1}^2}/I_1^2$ and $\overline{I_{nSS}^2}/I_{SS}^2$ in (6.47) for the circuit of Fig. 6.31(a). Consider only flicker noise.

Solution

Let us express the gate-referred flicker noise of M_a as $[K/(WLC_{ox})](1/f)$ and hence

$$I_{nSS}^2 = g_{ma}^2 \frac{K}{WLC_{ox}} \frac{1}{f}. \tag{6.50}$$

We know that I_1 is scaled down by a factor of 5 with respect to I_{SS}, and so is the width of M_b with respect to that of M_a. The gate-referred flicker noise of M_b is therefore 5 times that of M_a: $[5K/(WLC_{ox})](1/f)$. This noise is multiplied by the square of the g_m of M_b to generate $\overline{I_{n1}^2}$:

$$\overline{I_{n1}^2} = g_{mb}^2 \frac{5K}{WLC_{ox}} \frac{1}{f}. \tag{6.51}$$

Since g_{mb} is $1/5$ of that M_a, we have $\overline{I_{n1}^2} = (1/5)\overline{I_{nSS}^2}$. We conclude that the term $\overline{I_{n1}^2}/I_1^2$ is 5 times the term $\overline{I_{nSS}^2}/I_{SS}^2$. This agrees with simulations, pointing to M_b in Fig. 6.31(a) as the most severe source of phase noise at low offsets.

Oscillation Near Resonance Our investigation in Example 6.11 suggests that coupling of two oscillators by means of bilateral devices such as linear resistors or capacitors does not yield quadrature phases. We

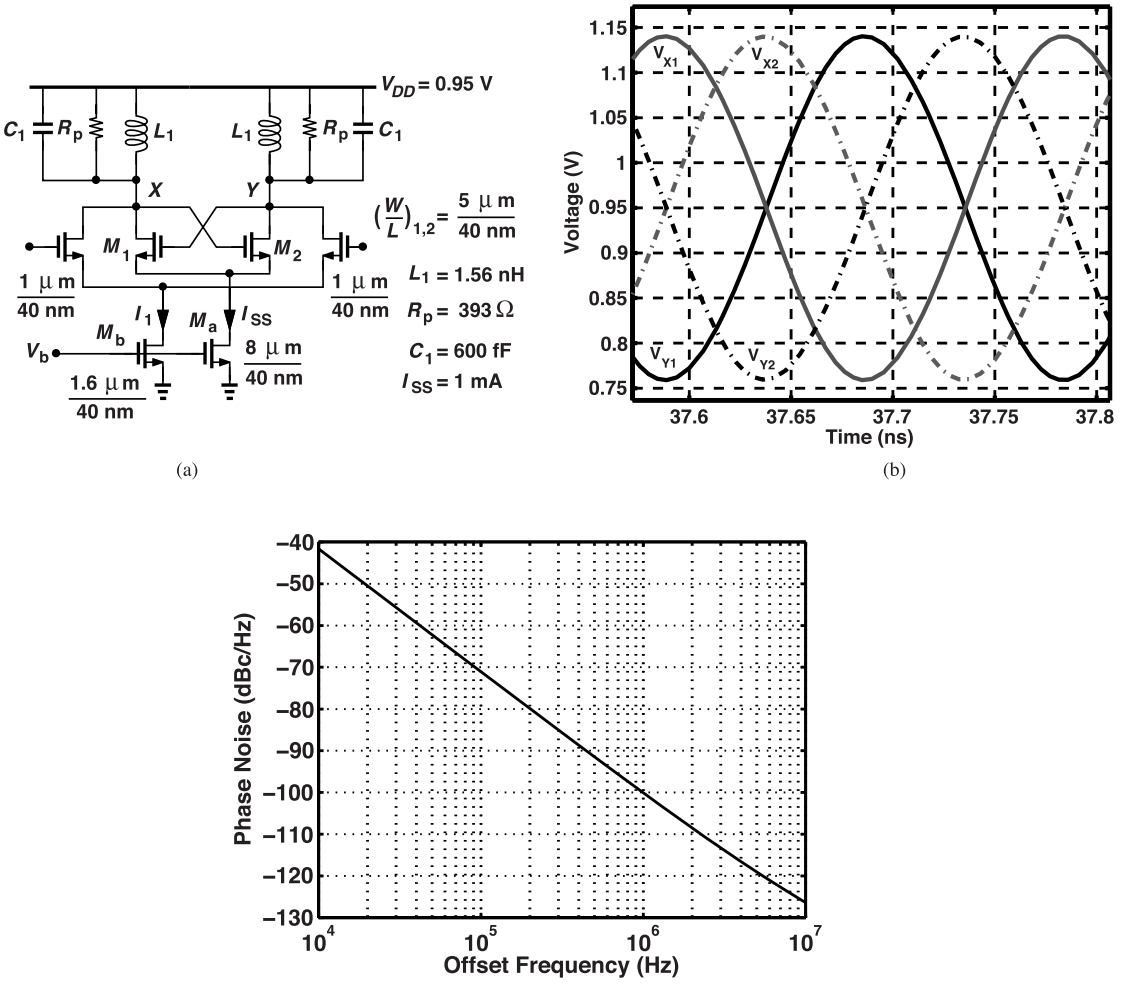


Figure 6.31 (a) Quadrature oscillator design example, (b) its output waveforms, and (c) its phase noise.

also note from Eq. (6.43) that unilateral, anti-phase coupling shifts the oscillation frequency away from the tank's resonance frequency in proportion to the coupling factor, which itself depends on the ratio of the tail currents. Thus, the flicker noise of the tail currents modulates the frequency. A number of techniques have been introduced to minimize the departure from resonance and the flicker noise upconversion.

It is possible for the coupling circuit to introduce phase shift so that the injected current is nearly in phase with the core oscillator current. Illustrated in Fig. 6.32 is a method where the transconductance of the coupling pair is given by

$$\frac{I_{out}}{V_{in}}(s) = \frac{g_m(R_1C_1s+1)}{R_1C_1s + g_mR_1/2 + 1}. \tag{6.52}$$

If we select $1/(2\pi R_1C_1)$ to be around the oscillation frequency, then the coupling pair provides a phase shift of about 45°. Of course, this also means that the coupling strength is less than g_m, requiring wider transistors

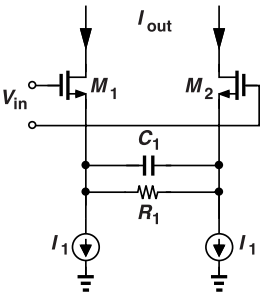
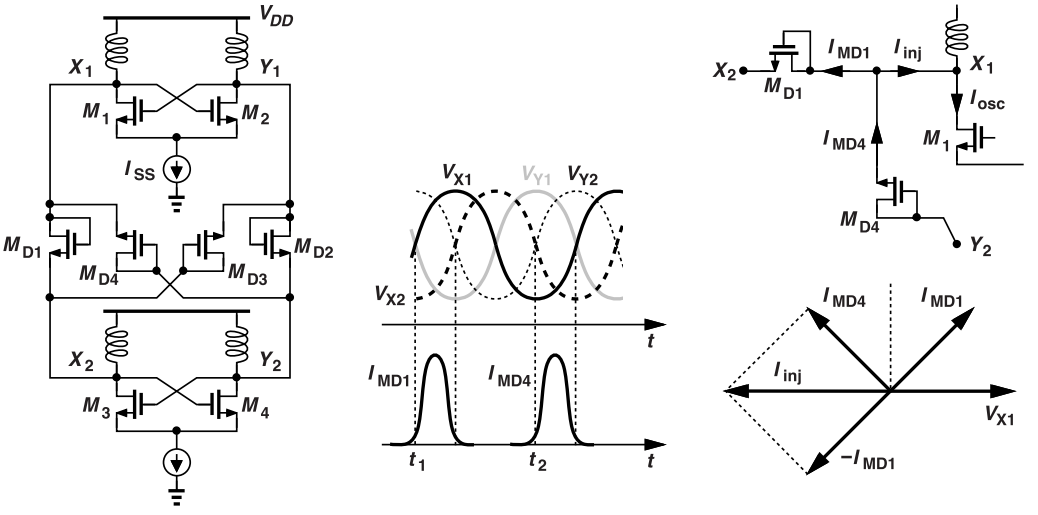


Figure 6.32 Use of capacitive and resistive degeneration to create phase shift .

and higher bias currents.[6]

An interesting quadrature oscillator, proposed in [12], realizes unilateral coupling through diode-connected MOSFETs. Shown in Fig. 6.33(a), the circuit incorporates both in-phase and anti-phase diode coupling



(a) (b) (c)

Figure 6.33 (a) Oscillators coupled through diode-connected devices, (b) voltage and current waveforms, and (c) illustration of currents flowing into one output node.

between the two oscillators. Diode-connected device M_{D1} turns on when $V_{X1} - V_{X2} > V_{TH}$. Similarly, the other diodes conduct when their V_{GS} exceeds V_{TH}. We examine the currents injected by M_{D1} and M_{D4} into X_1. As illustrated in Fig. 6.33(b), M_{D1} turns on around $t = t_1$, when $V_{X1} - V_{X2}$ is sufficiently positive, injecting a current whose peak is approximately 45° behind the rising edge of V_{X1}. Similarly, M_{D4} turns on around $t = t_2$, with a peak current that is 135° behind the falling edge of V_{X1}.

We now construct the phasor diagram shown in Fig. 6.33(c), recognizing that the net injected current, $I_{inj} = I_{MD4} - I_{MD1}$ is approximately *aligned* with V_{X1}. Since the drain current of M_1, I_{D1}, is aligned with V_{Y1}, we observe that I_{inj} and I_{D1} are also aligned. In other words, the tank tied to X_1 need not create a phase shift between its current and voltage, allowing the circuit to oscillate at the resonance frequency. This

[6]These points hold for small-signal operation. If the differential pair experiences complete switching, the phase shift is less than 45° and the strength is given by the total tail current.

means that the oscillation frequency is not modulated by the tail current. We thus expect that the flicker noise upconversion formulated by (6.47) is reduced here.

The principal difficulty with the circuit of Fig. 6.33(a) is that its assumes M_{D1}-M_{D4} inject sufficiently large currents; i.e., the first harmonic of I_{inj} must be about 20% of that produced by M_1. This in turn requires that $V_{X1} - V_{X2}$, for example, reach a peak well above V_{TH} of M_{D1}, a serious issue in low-voltage designs, especially because V_{TH} increases due to body effect.

As an example, let us redesign the circuit of Fig. 5.37(a) with diode coupling. For M_{D1}-M_{D4}, we choose $W/L = 8\ \mu\text{m}/40$ nm, and we raise the tail currents to 1.5 mA so as to provide large voltage swings across the diode-connected devices.[7] Figure 6.34 plots the phase noise, revealing considerable improvement at low

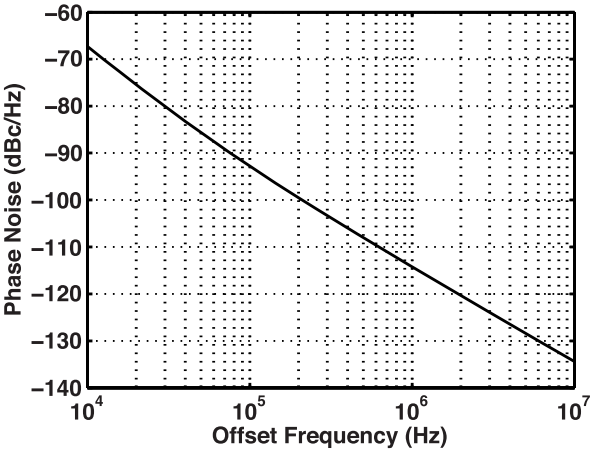


Figure 6.34 Phase noise of the quadrature oscillator shown in Fig. 6.33(a).

offsets with respect to the results in Fig. 6.31(c). At 10-MHz offset, the phase noise (-134 dBc/Hz) is 6 dB lower than that of the original differential oscillator (Fig. 5.38) because both the power and the voltage swings have been increased. However, at 100-kHz offset, the circuit does not benefit from power scaling, still suffering from the flicker noise of the tail currents. In summary, diode coupling proves more efficient than differential-pair coupling but still does not achieve the FOM of a single differential oscillator.

References

[1]. A. Hajimiri, and T. H. Lee, "A general theory of phase noise in electrical oscillators," *IEEE J. Solid-State Circuits,* vol. 33, pp. 179-194, Feb. 1998.

[2]. P. Andreani and A. Fard, "More on the 1/f phase noise performance of CMOS differential-pair LC-tank oscillators," *IEEE J. Solid-State Circuits,* vol. 41, pp. 2703-2712, Dec. 2006.

[3]. A. Mazzanti and P. Andreani, "Class-C harmonic CMOS VCOs, with a general result on phase noise," *IEEE J. Solid-State Circuits,* vol. 43, no. 12, pp. 2716-2729, Dec. 2008.

[4]. S. Li, I. Kipnis, and M. Ismail, "A 10-GHz CMOS quadrature LC-VCO for multirate optical applications," *IEEE J. Solid-State Circuits,* vol. 38, pp. 1626-1634, Oct. 2003.

[5]. I. R. Chamas and S. Raman, "A comprehensive analysis of quadrature signal synthesis in cross-coupled RF VCOs," *IEEE Trans. Circuits and Systems, Part I*, vol.54, pp. 689-704, April 2007.

[6]. R. Adler, "A study of locking phenomena in oscillators," *Proc. IEEE*, vol. 61, No. 10, pp. 1380-1385, Oct. 1973.

[7]. B. Razavi, "A study of injection locking and pulling in oscillators," *IEEE J. Solid-State Circuits,* vol. 39, pp. 1415-1424, Sep. 2004.

[8]. T. P. Liu, "A 6.5-GHz monolithic CMOS voltage-controlled oscillator," *ISSCC Dig. Tech. Papers,* pp. 404-405, Feb. 1999.

[7] But large swings drive the cross-coupled transistors deeper into the triode region, raising their flicker noise.

[9]. B. Razavi, "Mutual injection pulling between oscillators," *Proc. CICC,* pp. 675-678, Sept. 2006.
[10]. L. Romano et al., "Phase noise and accuracy in quadrature oscillators," *Proc. ISCAS,* pp. 161-164, May 2004.
[11]. A Mazzanti, F. Svelto, and P. Andreani, "On the amplitude and phase errors of quadrature LC-tank CMOS oscillators," *IEEE J. Solid-State Circuits,* vol. 41, pp. 1305-1313, June 2006.
[12]. X. Yi et al., "A 57.9-to-68.3GHz 24.6mW frequency synthesizer with in-phase injection-coupled QVCO in 65nm CMOS," *ISSCC Dig. Tech. Papers,* pp. 354-355, Feb. 2013.

Problems

6.1. Repeat Example 6.2 for an oscillator that generates a triangular output waveform.

6.2. Suppose $i_n(t)$ represents the time-domain behavior of a white noise source, whose two-sided spectrum is given by $\overline{I_n^2} = \eta/2$. (For example, $\overline{I_n^2} = 2kT/R$ for a resistor of value R.) Prove that the spectrum of $i_n(t)\sin\omega_0 t$ is also white and equal to $\eta/4$.

6.3. We have obtained two seemingly different expressions for the phase noise of a simple LC oscillator under white noise injection. Equation (5.12) reads

$$S_{\phi n}(f) = I_n^2 \frac{1}{4C_a^2(2\pi f)^2} \cdot \frac{1}{V_0^2} \tag{6.53}$$

whereas (6.19) reads

$$S_{\phi n}(f) = S_i(f) \frac{1}{2C_1^2(2\pi f)^2} \frac{1}{V_0^2}. \tag{6.54}$$

To see how the factor-of-2 discrepancy arises, assume $\overline{I_n^2} = \eta$ in the former (a one-sided spectrum) and $S_i(f) = \eta/2$ in the latter (a two-sided spectrum). Note that the spectrum of $\overline{I_n^2}$ *must* be one-sided (why?) and the spectrum of $g(t)$ in Fig. 6.8 *must* be two-sided because its noise components below and above $f = 0$ translate to corresponding components in $S_{\phi n}(f)$ after multiplication by $|H(j\omega)|^2$.

6.4. If the tail current in Fig. 6.12 contains noise of the form I_{nT}, determine the noise in the output common-mode level. Note that I_{nT} is split equally between the two sides and flows through an average impedance of $1/g_{m3}$ or $1/g_{m4}$ as it translates to CM noise.

6.5. We wish to compare the output CM noise of the oscillator shown in Fig. 6.35 to that of Fig. 6.14(b). Here, R_s

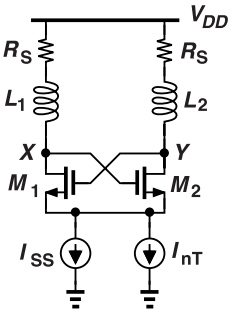


Figure 6.35 Effect of tail noise in the presence of inductors' series resistance.

represents the low-frequency metal resistance incurred by each inductor. Using the approach outlined in the previous problem, find the CM noise in terms of I_{nT}. (Generally, R_s is much less than $1/g_m$ of PMOS devices.)

6.6. Prove that supply noise in Fig. 6.14(b) directly modulates the output CM level and is therefore indistinguishable from noise in V_{cont}. (Hint: return to Fig. 6.12.)

6.7. Suppose the oscillator of Fig. 6.20(a) suffers from low-frequency supply noise. To determine the effect on the drain currents of M_1 and M_2, we construct the model shown in Fig. 6.36. Here, the two transistors are merged and the tank impedance is neglected. Compute the small-signal current in M_1 and M_2 due to $V_{DD,n}$.

6.8. Suppose the Q of an LC tank can be expressed as $Q = \alpha f^m$, where α and m are constant, and $m \approx 0.5$. Repeat the calculations in Section 6.5 for the phase noise.

6.9. If the clock waveform in Fig. 6.22(b) does not have a 50% duty cycle, what happens to V_X and V_Y?

6.10. If α_1 and α_2 in Eq. (6.30) have a small mismatch, $\alpha_1 = \alpha_2 + \Delta\alpha$, explain what happens to A and B for in-phase or anti-phase coupling.

6.11. In Fig. 6.26, the input of α_1 is mistakenly connected to the input of the top $H(s)$ block. Solve the circuit and find a relationship between A and B.

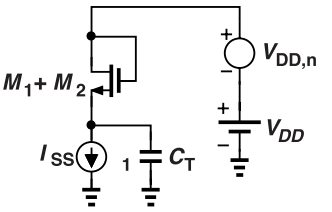


Figure 6.36 Simplified oscillator model for calculation of supply rejection.

6.12. Prove that the coupled ring oscillators shown in Fig. 6.37 satisfy the model in Fig. 6.26 with $\alpha_1 = \alpha_2$, and hence operate differentially.

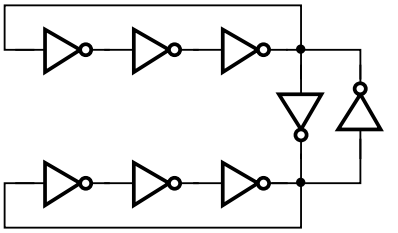


Figure 6.37 Ring oscillators with mutual injection.

6.13. For the quadrature oscillator shown in Fig. 6.24(b), sketch the waveforms at the sources of the transistors.

6.14. With the aid of Eq. (6.35) determine the change in the oscillation frequency of the circuit in Fig. 6.24(b) if the coupling coefficient goes from 0.2 to 0.3.

6.15. A student proposes from Eq. (6.41) that the oscillation frequency can be varied by adjusting the Q. Explain the drawback(s) of this approach.

6.16. Explain why Eqs. (6.35) and (6.41) are different even though they both apply to quadrature LC oscillators.

6.17. If I_1 and I_{SS} in Fig. 6.24(b) are realized by NMOS devices having $g_{m1} = 2I_1/(V_{GS} - V_{TH})$ and $g_{mSS} = 2I_{SS}/(V_{GS} - V_{TH})$, determine the phase noise in Eq. (6.47). Assume $I_1 = 0.2I_{SS}$.

6.18. We double the widths and lengths of the NMOS current sources in the previous problem. Explain how the phase noise changes.

6.19. Does the circuit of Fig. 6.33(a) provide quadrature phases if the gates of M_{D1}-M_{D4} are tied to V_{DD} rather than to their drains?

6.20. Does the circuit of Fig. 6.33(a) exhibit two different ocillation modes [like the topology in Fig. 6.24(b)]?

7

Basic PLL Architectures

Having analyzed and designed various types of VCOs in the previous chapters, we are now ready to study phase-locked loops. Most oscillators must be "phase-locked" so as to minimize their frequency drift and phase noise. Dating back to the 1930s, the concept of phase-locking proves essential in many applications, as seen in this and subsequent chapters. In its simplest form, a PLL consists of a VCO, a "phase detector" (PD), and a low-pass filter. We thus first introduce the operation of PDs.

7.1 Phase Detectors

From basic electronics, we know that an op amp measures the difference between two voltages, ideally exhibiting a linear input-output characteristic. For PLL design, we need a circuit that measures the difference between two *phases*. Illustrated conceptually in Fig. 7.1(a), such a function should produce an output—

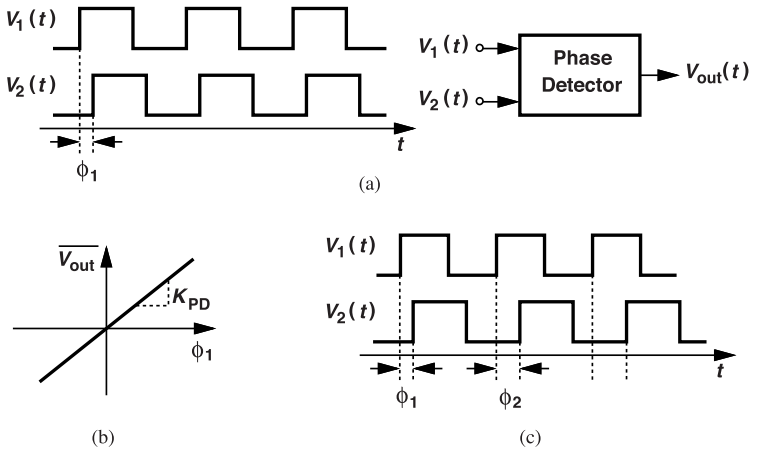


Figure 7.1 (a) PD sensing two periodic signals, (b) PD characteristic, and (c) phase difference variation for two signals of unequal frequencies.

perhaps a voltage—that linearly varies with the phase "error," ϕ_1. Practical PD implementations generate a periodic output whose *average* is proportional to ϕ_1, exhibiting the characteristic shown in Fig. 7.1(b).

The PD behavior plotted in Fig. 7.1(b) merits some remarks. First, the slope or "gain" of the circuit is denoted by K_{PD} and expressed in volts per radian. For example, a PD may produce an output change of 100 mV for a 1-radian change in the input phase difference. Second, this characteristic is meaningful only when $V_1(t)$ and $V_2(t)$ have equal frequencies. Otherwise, their phase difference changes with time [Fig. 7.1(c)].

A PD can be realized by means of an exclusive OR (XOR) gate. We note from Fig. 7.2(a) that the XOR

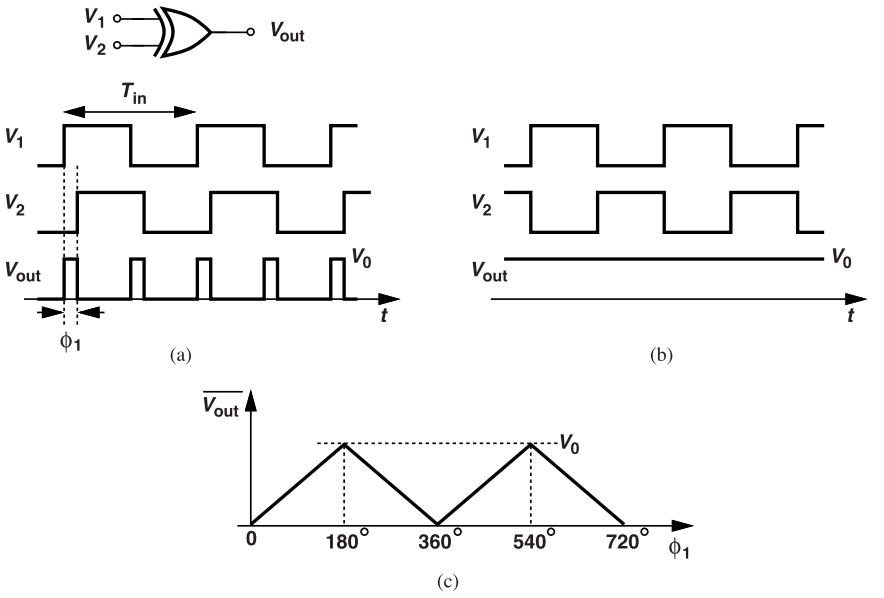


Figure 7.2 (a) XOR gate operating as a PD, (b) input and output waveforms for a phase difference of 180°, and (c) input-output characteristic.

generates a ONE if $V_1(t)$ and $V_2(t)$ are unequal, thereby creating a series of pulses, at twice the input frequency, whose width represents the phase difference between the two inputs. The area under these pulses varies with ϕ_1, leading to an output average proportional to the phase error. To determine the gain, we recognize that for a phase error of ϕ_1 radians and hence $[\phi_1/(2\pi)]T_{in}$ seconds, the average output is equal to $2[\phi_1/(2\pi)]T_{in}V_0/T_{in} = (V_0/\pi)\phi_1$. It follows that $K_{PD} = V_0/\pi$.

The waveforms in Fig. 7.2(a) suggest the PD characteristic begins at zero for $\phi_1 = 0$ and linearly rises as ϕ_1 increases. But we also observe from Fig. 7.2(b) that the average reaches a peak of V_0 at $\phi_1 = 180°$. Beyond this point, the characteristic falls linearly, returning to zero at $\phi_1 = 360°$ and repeating its behavior thereafter. Figure 7.2(c) shows the result.

Example 7.1

An analog multiplier (mixer) can serve as a phase detector. Determine the characteristic of such a device.

Solution

Shown in Fig. 7.3(a), the mixer produces an output equal to $V_{out} = kV_1(t)V_2(t)$, where k is a proportionality

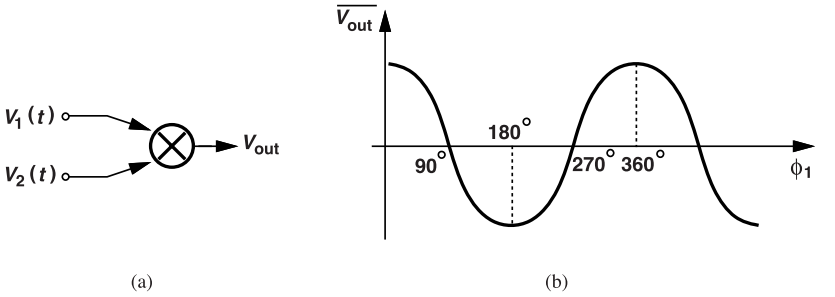


Figure 7.3 (a) Mixer acting as a PD, and (b) its input-output characteristic.

factor. In this case, it is simpler to represent the inputs as sinusoids, $V_1(t) = V_0 \cos \omega_1 t$, $V_2(t) = V_0 \cos(\omega_1 t + \phi_1)$, and hence

$$V_{out}(t) = \frac{kV_0^2}{2} \cos(2\omega_1 t + \phi_1) + \frac{kV_0^2}{2} \cos \phi_1. \tag{7.1}$$

The first term on the right hand side has a zero average, leading to

$$\overline{V_{out}(t)} = \frac{kV_0^2}{2} \cos \phi_1. \tag{7.2}$$

Plotted in Fig. 7.3(b), this characteristic has a *zero* gain at $\phi_1 = 0$; i.e., the PD loses its ability to measure the phase difference in this vicinity. The PD gain reaches a maximum at $\phi_1 = 90°, 270°$, etc.

7.2 Phase Control by Feedback

Suppose an oscillator having a nominal frequency of f_0 provides a clock, CK_{osc}, to a system, e.g., a microprocessor. What happens if the ambient temperature slowly fluctuates? The oscillator output frequency does, too, causing random phase accumulations. This can be seen by plotting this output against an ideal clock [Fig. 7.4(a)]. We observe that the edges of CK_{osc} slowly drift away from ideal points in time. A similar situation

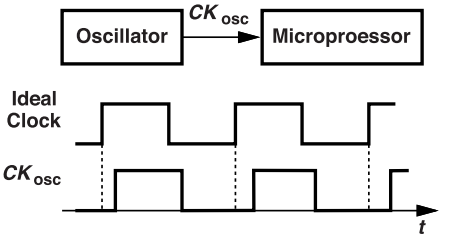


Figure 7.4 Oscillator providing a clock to a microprocessor.

arises if the oscillator phase noise is considered. The key point here is that the phase error incurred by the oscillator can grow so large as to cause failure in the processor. In order to correct this error, we must control the oscillator phase.

We wish to align the output phase of a VCO with that of a reference clock, assuming that the rising edges of the VCO have a skew of Δt seconds with respect to V_{CK} [Fig. 7.5(a)]. To change the VCO phase, we *must* change the frequency and allow the integration $\phi = \int (\omega_0 + K_{VCO} V_{cont}) dt$ to take place. For example, suppose as shown in Fig. 7.5(b), the VCO frequency is stepped to a higher value at $t = t_1$. The circuit then accumulates phase faster, gradually decreasing the phase error. At $t = t_2$, the phase error between V_{CK} and V_{VCO} drops to zero and, if V_{cont} returns to its original value, the two signals remain aligned. Interestingly, the alignment can also be accomplished by stepping the VCO frequency to a *lower* value for a certain time interval. Notably, phase alignment can be achieved only by a (temporary) frequency change.

The foregoing experiment suggests that the output phase of a VCO can be aligned with the phase of a reference if (1) the frequency of the VCO is changed momentarily, (2) a means of comparing the two phases, i.e., a phase detector, is used to determine when the VCO and reference signals are aligned. The task of aligning the output phase of the VCO with the phase of the reference is called "phase locking." In the study of PLLs, which operate with phase quantities, we often draw upon analogies with the more familiar voltage-mode circuits, as illustrated below.

In order to control the output phase of an oscillator, we can employ negative feedback. In a manner similar to voltage-domain feedback [Fig. 7.6(a)], we compare the output (phase) with that of a reference and use

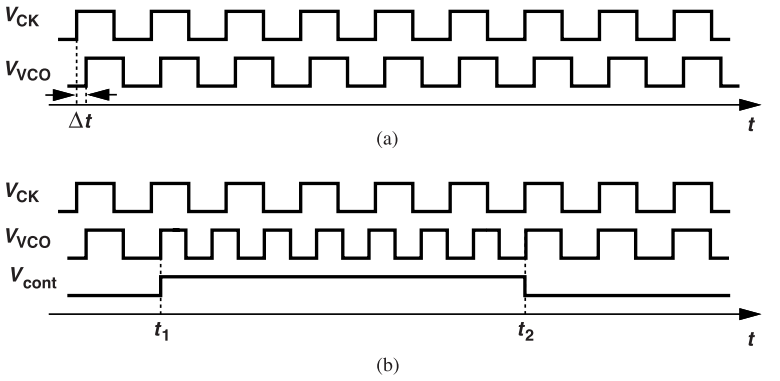


Figure 7.5 (a) Skew between a clock and a VCO output, (b) change of VCO frequency to eliminate the skew.

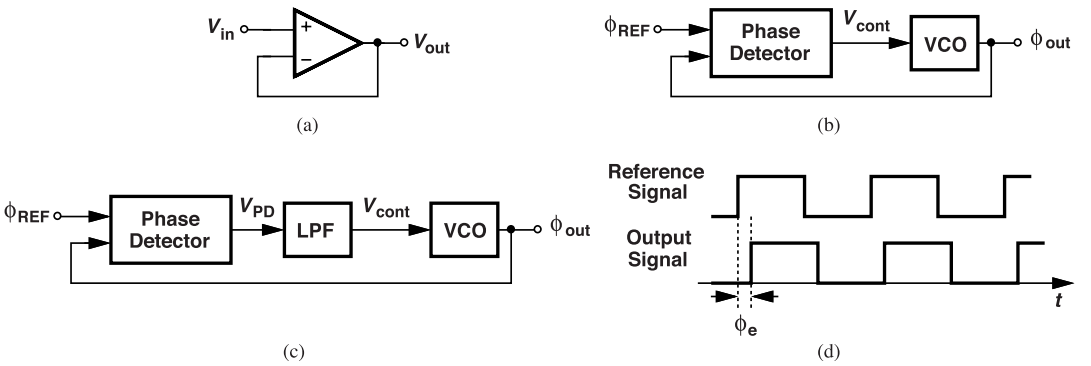


Figure 7.6 (a) Unity-gain voltage follower, (b) phase-control feedback loop, (c) PLL including an LPF, and (d) its input and output waveforms.

the resulting error to adjust the VCO's frequency and phase [Fig. 7.6(b)]. If the loop gain is large enough, $V_{out} \approx V_{REF}$ in Fig. 7.6(a) and $\phi_{out} \approx \phi_{REF}$ in Fig. 7.6(b). If the VCO phase drifts, the negative-feedback loop provides correction so that ϕ_{out} "tracks" ϕ_{REF}. As explained in Chapter 1, the only path controlling ϕ_{out} is through V_{cont}, which directly adjusts the output *frequency*, f_{out}; for a change in f_{out} to translate to a change in ϕ_{out}, we must wait.

Recall from Section 7.1 that the PD output contains repetitive pulses, which, in Fig. 7.6(b), inevitably disturb the VCO. To resolve this issue, we interpose a low-pass filter (LPF) (also called the "loop filter") between the PD and the VCO [Fig. 7.6(c)], suppressing the high-frequency components of the PD output. This negative-feedback circuit is called a "phase-locked loop" (PLL). We say the loop is "locked" when $\phi_{in} - \phi_{out}$ does not change with time (and is preferably small). Figure 7.6(d) illustrates the input and output waveforms in the locked state. We shall study this circuit extensively, but let us first make some preliminary remarks.

- The term "phased-locked loop" is somewhat strange; after all, the circuit in Fig. 7.6(a) is not called a voltage-locked loop! However, the feedback systems in Figs. 7.6(a) and 7.6(c) bear some interesting differences that justify this naming convention. We elaborate on this point in Section 7.3.
- As we travel around the loop in Fig. 7.6(c), the signal of interest changes its dimension: the PD measures phase quantities and generates a voltage (or current), which, after low-pass filtering, reaches the VCO; the VCO delivers a periodic waveform whose phase is fed back. Of course, one can also say that within the op amp of Fig. 7.6(a), the input voltage difference is converted to a current, passed through a load impedance to become a voltage (e.g., as in a cascode stage), and possibly travels through another gain

stage that performs similar conversions (e.g., as in a two-stage op amp).

- It appears that our PLL design effort has thus far yielded a circuit whose output is very similar to its input ($\phi_{out} \approx \phi_{REF}$). Then, why not replace the PLL with a piece of wire?! We will appreciate the role of phase-locking as we delve into the operation of the circuit.

7.3 Analysis of Simple PLL

7.3.1 Static Behavior

We wish to analyze the PLL of Fig. 7.6(c) and understand its properties and limitations. To enhance our intuition, we replace each building block with a simple circuit realization, as depicted in Fig. 7.7(a). Our objective is to calculate the quantities shown around the loop in the steady state.

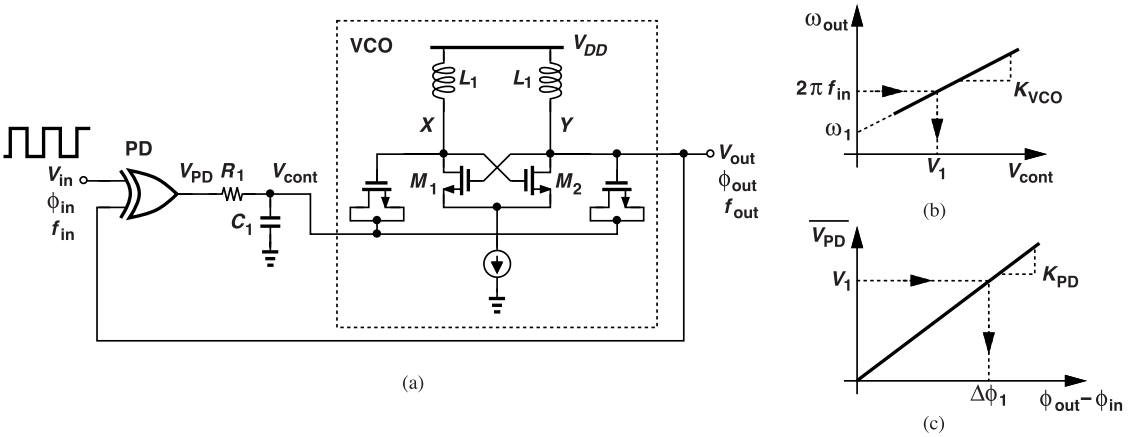


Figure 7.7 (a) Simple PLL implementation, (b) VCO characteristic, and (c) PD characteristic.

A key point that facilitates our task is offered by the phase-locking property; since $\phi_{out} - \phi_{in} = \text{constant}$ in the locked state, we have

$$\frac{d\phi_{out}}{dt} = \frac{d\phi_{in}}{dt}. \tag{7.3}$$

Recall from Chapter 1 that the instantaneous frequency is equal to the time derivative of the phase. It follows that

$$2\pi f_{out} = 2\pi f_{in}. \tag{7.4}$$

This result agrees with the waveforms shown in Fig. 7.6(d): for the input and output edges to sustain a constant phase difference, the periods must be equal at all times. We can thus assume that the VCO runs at a frequency equal to f_{in} if the loop operates properly, i.e., if it is locked.

We now utilize the VCO and XOR characteristics to compute the quantities of interest. From the plot in Fig. 7.7(b), we note that the VCO requires a certain control voltage, V_1, to provide $f_{out} = f_{in}$. Such a voltage must be generated by the low-pass filter, which computes the average of V_{PD}. Turning to the XOR characteristic shown in Fig. 7.7(c), we note that $\overline{V_{PD}} = V_1$ translates to a certain phase error, $\Delta\phi_1 = V_1/K_{PD}$. In summary, the knowledge that $f_{out} = f_{in}$, along with the VCO and XOR characteristics, leads to a unique set of values around the loop. Called the "static phase offset," $\Delta\phi_1$ proves undesirable in many applications. It can be reduced by increasing K_{PD} (or K_{VCO}).

Example 7.2

Plot the voltage waveforms in Fig. 7.7(a) if $(2\pi R_1 C_1)^{-1} \ll f_{in}$. Assume the XOR output toggles between 0 and V_{DD}.

Solution

As illustrated in Fig. 7.8, V_{in} and V_{out} incur a phase difference, $\Delta\phi_1 = V_1/K_{PD}$. The PD output pulses

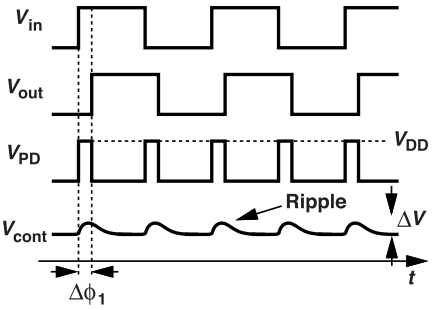


Figure 7.8 Waveforms in PLL of 7.7(a).

arrive at the low-pass filter and cause a response of the form $V_{DD}[1 - \exp(-t/\tau)]$, where $\tau = R_1 C_1$, on the rising edges. Since τ is much greater than the input period, we have $\exp(-t/\tau) \approx 1 + t/\tau$, which simplifies the response to $V_{DD}t/\tau$. Thus, V_{cont} exhibits a peak-to-peak change equal to $\Delta V = V_{DD}[\Delta\phi_1/(2\pi)]T_{in}/\tau$ (why?). Since $\Delta\phi_1 = V_1/K_{PD}$ and $\tau = R_1 C_1$, we have

$$\Delta V = \frac{V_{DD} V_1 T_{in}}{2\pi R_1 C_1 K_{PD}}. \tag{7.5}$$

The periodic disturbance in V_{cont} is called the "ripple."

Example 7.3

In Example 7.2, we increase the input frequency by Δf. Explain what happens to the phase offset and the ripple. Assume the loop remains locked.

Solution

If the input frequency rises to $f_{in} + \Delta f$, Fig. 7.7(b) implies that V_{cont} must increase to $V_1 + 2\pi\Delta f/K_{VCO}$. This translates to a greater phase error in Fig. 7.7(c), given by $V_1/K_{PD} + 2\pi\Delta f/(K_{VCO}K_{PD})$. Similarly, the ripple amplitude rises to $[V_{DD}T_{in}/(2\pi R_1 C_1)](V_1 + 2\pi\Delta f/K_{VCO})/K_{PD}$, where it is assumed that T_{in} changes negligibly.

Example 7.4

Owing to an ambient temperature change, the VCO characteristic in Fig. 7.7(b) shifts up by $2\pi\Delta f$. If K_{VCO} and the input frequency remain unchanged, determine the new values of V_1 and $\Delta\phi_1$. Assume the loop is locked.

Solution

A shift of $2\pi\Delta f$ upward means that the VCO now requires a control voltage equal to $V_1 - 2\pi\Delta f/K_{VCO}$ to operate at $2\pi f_{in}$. Thus, $\Delta\phi_1$ decreases by an amount equal to $2\pi\Delta f/(K_{VCO}K_{PD})$.

7.3.2 Frequency Multiplication

In Section 7.2, we asked how a PLL is different from a piece of wire. In this section, we learn about an important and useful property of PLLs, namely, frequency multiplication. Suppose a precise, stable reference oscillator (e.g., a crystal oscillator) produces an output at 20 MHz and we wish to generate a 4-GHz clock from it. One approach would be to consider the 200th harmonic of the 20-MHz reference, but such a component is extremely small and prone to extreme noise.[1] By contrast, a PLL affords an elegant solution. Illustrated in Fig. 7.9(a), the idea is to employ a loop containing a 4-GHz VCO but insert in the feedback path a

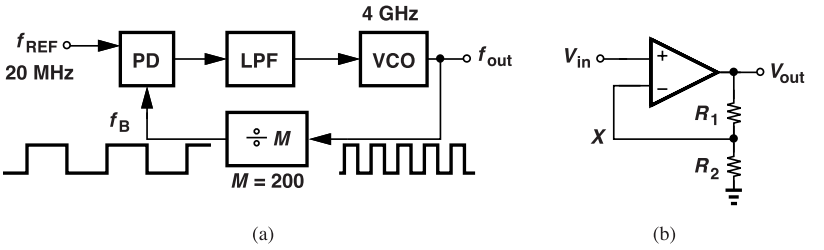


Figure 7.9 (a) Frequency-multiplying PLL, and (b) voltage-domain analogy.

"frequency divider," a counter that generates one output pulse for every $M = 200$ input pulses. If the loop is locked, the PD senses a constant phase difference and hence equal frequencies. That is, $f_B = f_{REF}$, yielding $f_{out} = Mf_B = Mf_{REF}$. The beauty of this arrangement is that it also provides a programmable output frequency: if M goes from 200 to 201, f_{out} changes from 4 GHz to $201 \times 20\ \text{MHz} = 4.02$ GHz.

The notion of dividing the output of a feedback system before returning it to the input is, of course, familiar. In the voltage-mode circuit of Fig. 7.9(b), for example, $V_X = V_{out}R_2/(R_1 + R_2)$ and hence $V_{out} \approx (1 + R_1/R_2)V_{in}$. We can say this topology performs voltage multiplication.

We study the static phase error of this PLL in Problem 7.9. We also analyze frequency-multiplying PLLs later in this chapter.

7.3.3 Dynamic Behavior

Let us qualitatively study the transient behavior of the PLL shown in Fig. 7.10(a). As with more familiar voltage-mode circuits, we must apply a step at the input and observe the output. In this case, however, the input quantity recognized by the PD is the phase or the frequency. We begin our analysis with an input frequency step for it is more intuitive.

As illustrated in Fig. 7.10(b), the input frequency jumps from ω_1 to $\omega_1 + \Delta\omega$ at $t = t_1$. Since the low-pass filter initially keeps V_{cont} at V_1, the VCO still runs at ω_1, the input and output frequencies sustain a difference of $\Delta\omega$, and their phase difference grows. The increasingly wider pulses generated by the PD gradually raise V_{cont} and hence ω_{out}. This continues until ω_{out} becomes equal to $\omega_1 + \Delta\omega$, at which point the loop is locked again. Note that the PD output pulses are now wider so as to yield a greater control voltage.

The settling behavior of V_{cont} in Fig. 7.10(b) may not be as smooth as illustrated; V_{cont} can exhibit "ringing" if the loop is underdamped (not stable enough). In analogy with feedback amplifiers, we must formulate the closed-loop response in terms of such parameters as the open-loop poles and zeros and the phase margin. We also recognize that the loop is unlocked from t_1 to t_2 because $\phi_{out} - \phi_{in}$ changes with time during this period.

Next, we study the response of the PLL to an input phase step. Depicted in Fig. 7.11, such a step simply means that the next input edge occurs earlier or later than expected. Neither V_{cont} nor the output phase can change instantaneously; thus, the phase step immediately translates to phase error, leading to wider PD pulses. The gradual change in V_{cont} and hence ω_{out} allows the VCO to accumulate additional phase and eventually catch up with the input. In contrast to the transient behavior in Fig. 7.10, here the loop quantities V_{cont} and $\phi_{out} - \phi_{in}$ are the same at t_2 as at t_1.

[1] It also requires a very sharp filter to separate it from the other harmonics.

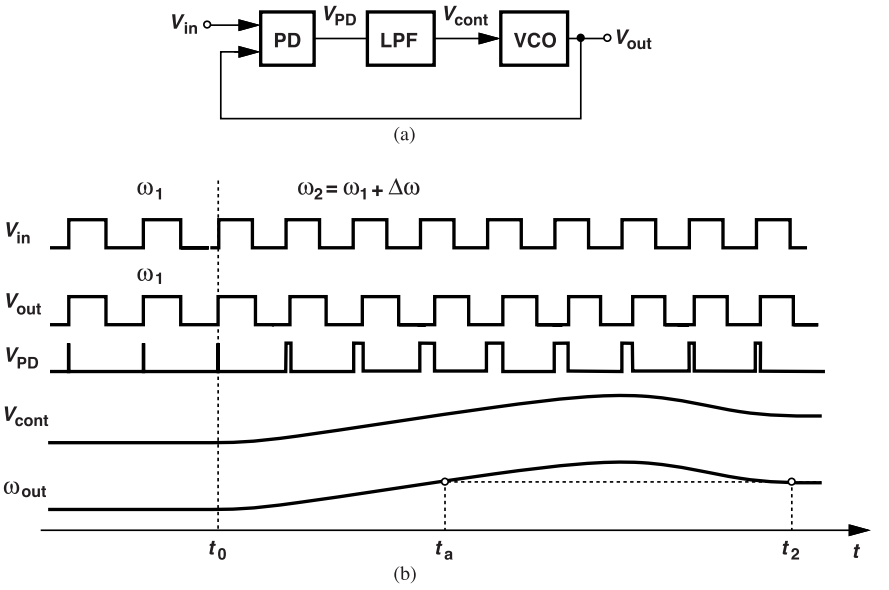


Figure 7.10 (a) PLL sensing an input frequency step, and (b) its waveforms.

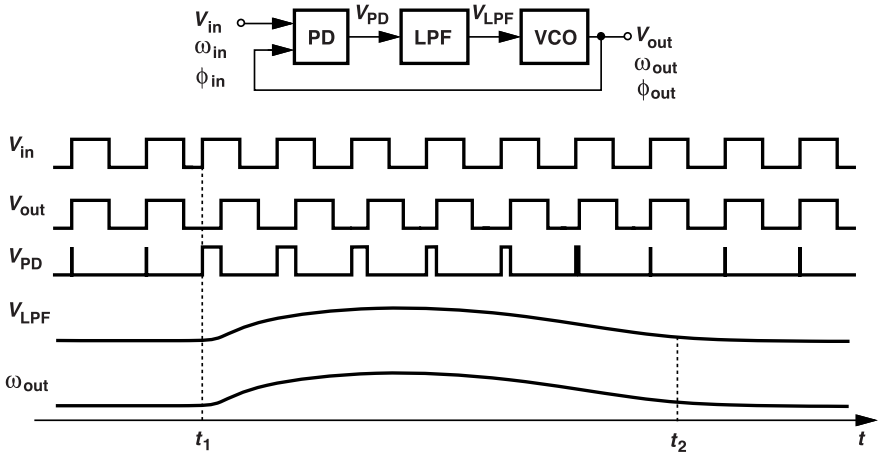


Figure 7.11 PLL sensing an input phase step and its waveforms.

Let us make some interesting observations. First, the amplifier circuit of Fig. 7.6(a) can sense only a *voltage* step at its input, whereas the PLL of Fig. 7.10(a) can receive a frequency step or a phase step. As exemplified by the V_{cont} plots in Figs. 7.10(b) and 7.11, we expect different expressions for the two different step responses. Second, the most convenient point in a PLL for monitoring the transient behavior is the VCO control line; other quantities, such as ϕ_{out} or $\phi_{out} - \phi_{in}$, do not lend themselves to intuitive interpretation. Third, Fig. 7.10(a) suggests that the loop does not settle at $t = t_a$, even though ω_{out} has the correct value (the same value as at $t = t_2$). This is because the phase error (the width of the PD output pulses) has not reached the proper value yet. In other words, the loop settles only when both $\Delta\phi_1$ and ω_{out}, which are related by an integral, settle. We say both "phase acquisition" and "frequency acquisition" must be completed.

Example 7.5

Without deriving its transfer function, prove that the PLL of Fig. 7.10(a) behaves as a low-pass system.

Solution

As a simple text, we first apply a frequency step at the input. In response to a step of height h, the output of a low-pass system begins from an initial condition and, eventually, reaches a value equal to a scalar times h. For example, the unity-gain buffer in Fig. 7.6(a) generates an output approximately equal to $V_0[1 - \exp(-t/\tau)]$ in response to an input step of height V_0. We observe from Fig. 7.10(b) that ω_{out} changes by $\Delta\omega$ after a sufficiently long time. Thus, the transfer function from the input frequency to the output frequency represents a low-pass system, as visualized Fig. 7.12(a).

We expect the same to hold for the input and output *phase* quantities as well. Let us redraw the system as shown in Fig. 7.12(b), where cascading an integrator and a differentiator does not alter the transfer function. Since all of the functions in this cascade are linear and time-invariant, we can swap the LPF with the integrator [Fig. 7.12(c)], noting that the phase is equal to the time integral of the frequency. We thus conclude that ϕ_{out} and ϕ_{in} are related by the same low-pass function. We formulate this property of PLLs in Section 7.3.4.

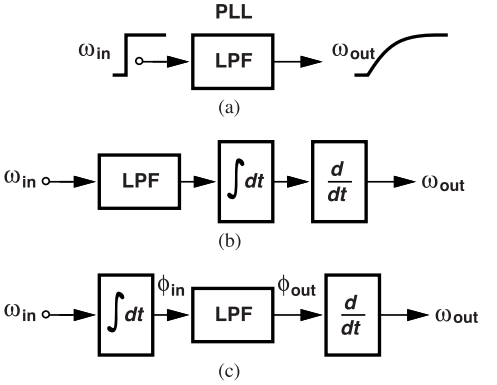


Figure 7.12 (a) Response of a PLL to a frequency step, (b) addition of an integrator and a differentiator, and (c) reordered cascade to show that the PLL acts as a low-pass filter for phase quantities.

Example 7.6

Explain from Fig. 7.11 why a PLL's loop gain for phase quantities is infinite.

Solution

To check for infinite loop gain, we can apply a step at the input and determine whether the feedback error eventually falls to zero. In the amplifier of Fig. 7.6(a), for example, a step of $V_0u(t)$ produces a final output value equal to $V_0A_0/(A_0+1)$, where A_0 is the op amp gain, and an error, $V_{out}-V_{REF}$, given by $V_0/(A_0+1)$. The error is zero if $A_0 = \infty$. For a PLL, the response in Fig. 7.11 reveals that $\phi_{out} - \phi_{in}$ returns to its original value after the transient subsides; the loop gain is thus infinite. This is expected as the loop includes an ideal integrator (the VCO).

Example 7.7

A student sees a contradiction between the above example's result and our earlier static calculations showing a phase error equal to $\Delta\phi_1 = V_1/K_{PD}$. Why is the "static phase error" not zero even though the loop gain is infinite?

Solution

Consider the voltage-domain feedback system shown in Fig. 7.13(a), where an ideal integrator provides an

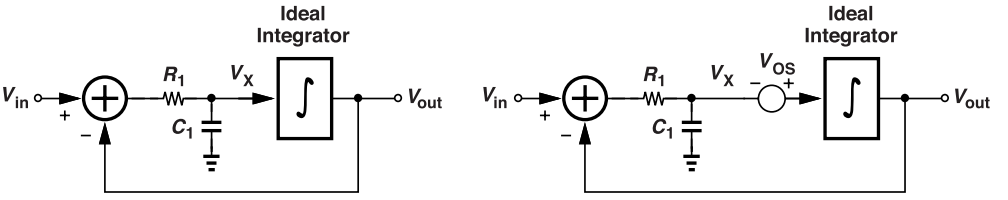


Figure 7.13 (a) A feedback loop containing an integrator, and (b) the effect of an offset voltage.

infinite loop gain. If V_{in} is constant and has a value of $V_0 \neq 0$, so does V_{out}. This can be seen by noting that, if $V_{out} < \infty$, then the integrator's infinite gain requires that $V_X = 0$ and hence $V_{in} - V_{out} = 0$.

Now, let us insert an offset voltage at the input of the integrator [Fig. 7.13(b)]. The integrator input is still zero, but $V_X = -V_{OS}$ and $V_{in} - V_{out} = -V_{OS}$. A PLL would be similar to the system in Fig. 7.13(a) if the VCO could generate the desired frequency for $V_{cont} = 0$. In reality, the VCO may require $V_{cont} > 0$, which corresponds to the offset voltage in Fig. 7.13(b). The loop gain is infinite, but a static feedback error arises from the offset within the loop.

Example 7.8

Explain why a PLL's loop gain is also infinite for frequency quantities.

Solution

We can intuitively see that the PLL has no frequency error after it responds to a frequency step, revealing an infinite loop gain. Let us approach the problem from another viewpoint. If frequency quantities are of interest, we must model the functions within the loop accordingly. Since the phase detector fundamentally measures the phase difference, we can write

$$\phi_{in} - \phi_{out} = \int (\omega_{in} - \omega_{out}) dt \tag{7.6}$$

and hence represent the PD as shown in Fig. 7.14(a) for frequency quantities. For the VCO, on the other hand,

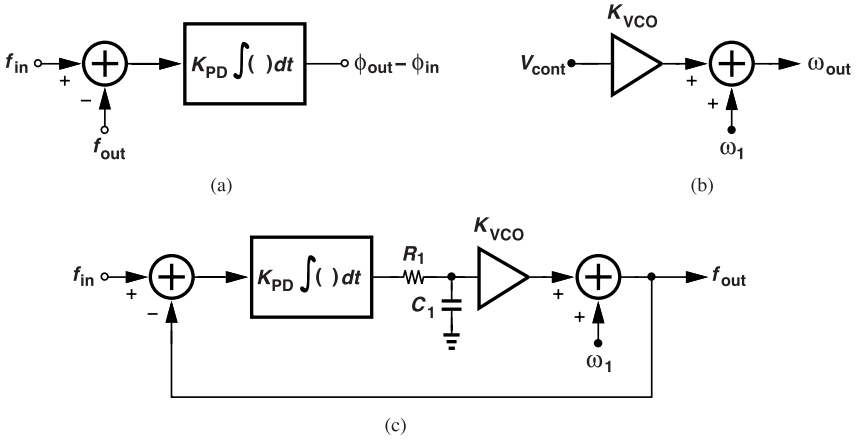


Figure 7.14 (a) PD model for input frequency quantities, (b) VCO model, and (c) closed-loop model.

the output frequency is expressed as $\omega_{out} = \omega_1 + K_{VCO} V_{cont}$, with ω_1 acting as an offset and K_{VCO} as

memoryless gain [Fig. 7.14(b)]. The overall loop is therefore modeled as depicted in Fig. 7.14(c). Interestingly, as with the phase model in Fig. 7.13(b), this system too contains an integrator but within the PD. On the other hand, the offset, ω_1, in Fig. 7.14(c) appears *after* the ideal integrator, contributing no static frequency error (why?). Thus, $\omega_{out} = \omega_{in}$.

7.3.4 PLL Transfer Function

With the qualitative studies in the previous section, we are now ready to formulate the dynamic behavior of PLLs. Our objective is to answer the typical questions addressed in more familiar voltage-mode circuits: (1) How do we interpret and determine the transfer function? (2) How do we evaluate the stability of the loop? (3) What filtering properties does the loop exhibit? (4) How does the loop behave in the time domain?

The examples in the previous section give some intuition into the response of PLLs, but we must still ponder the meaning of a PLL's transfer function. To this end, we return to the third question above and employ an analogy with voltage-mode circuits. Suppose a one-pole op amp is placed in unity-gain feedback (Fig. 7.15). We know that slowly-changing inputs travel to the output unattenuated, i.e., the output "tracks"

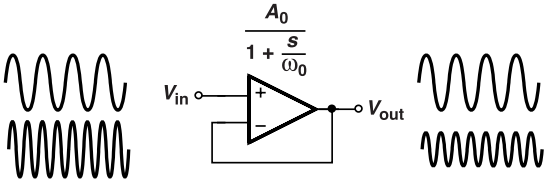


Figure 7.15 Unity-gain voltage follower for studying frequency response.

a slow input faithfully. On the other hand, for fast input changes, e.g., for sinusoids beyond the closed-loop bandwidth, the output cannot keep up, exhibiting smaller voltage swings. The circuit thus acts as a low-pass filter.

To extend these concepts to PLLs, we must explain what we mean when we say the input *phase* changes slowly or fast. In this context, the term phase in fact refers to the excess phase in the input, i.e., any disturbance beyond the linear ωt term. We model such an input waveform as $V_0 \cos[\omega_1 t + \phi_{in}(t)]$, and wish to determine how the PLL output phase behaves if $\phi_{in}(t)$ varies slowly or fast. Figure 7.16 conceptually illustrates two cases of slow and fast phase variations: in the former, the period drifts slowly and in the latter, it changes

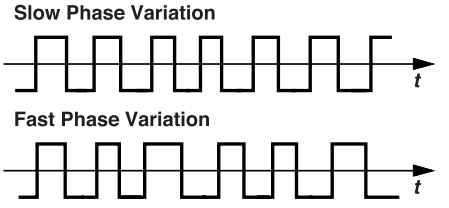


Figure 7.16 Slow and fast phase fluctuations.

rapidly. Our quantitative analysis of the PLL will reveal how these fluctuations propagate to the output phase. From Example 7.5, we predict that the PLL passes only slow phase variations.

Example 7.9

How do we express a sinusoidal waveform that has slow phase fluctuations?

Solution

We write $V = V_0 \cos[\omega_1 t + \phi_{in}(t)]$ and choose $\phi_{in}(t)$ to be a slowly-varying function, e.g., $\phi_{in}(t) =$

$\alpha \cos \omega_m t$ where ω_m is a small number. For example, if $\omega_m = 2\pi(1\ \text{kHz})$, the excess phase varies at a rate of 1 kHz.

In order to compute the transfer function, ϕ_{out}/ϕ_{in}, of the PLL shown in Fig. 7.6(c), we replace each building block with a linear representation, bearing in mind that the VCO must be modeled as an integrator because it is its output phase that is sensed by the PD (Chapter 1). For a first-order loop filter, we arrive at the small-signal model depicted in Fig. 7.17, where the open-loop transfer function is equal to

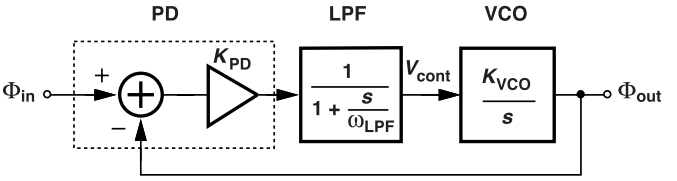


Figure 7.17 Linear model of simple PLL.

$[K_{PD}/(1 + s/\omega_{LPF})]K_{VCO}/s$. We can also view this quantity as the loop gain or the "loop transmission."[2] It follows that the closed-loop transfer function is given by

$$H(s) = \frac{\phi_{out}}{\phi_{in}}(s) \tag{7.7}$$

$$= \frac{K_{PD}K_{VCO}}{\dfrac{s^2}{\omega_{LPF}} + s + K_{PD}K_{VCO}}. \tag{7.8}$$

We prefer, from control theory, to express the denominator in the form $s^2 + 2\zeta\omega_n s + \omega_n^2$, where ζ and ω_n are the "damping factor" and the "natural frequency," respectively. Thus,

$$H(s) = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2}, \tag{7.9}$$

where

$$\omega_n = \sqrt{K_{PD}K_{VCO}\omega_{LPF}} \tag{7.10}$$

$$\zeta = \frac{1}{2}\sqrt{\frac{\omega_{LPF}}{K_{PD}K_{VCO}}}. \tag{7.11}$$

Of particular interest to us is ζ as it reveals how stable the loop is. For a critically-damped response, $\zeta = \sqrt{2}/2$; we typically choose ζ in the range of $\sqrt{2}/2$ and 1. The topology of Fig. 7.17 is called a "type-I" PLL because its open-loop transfer function contains one ideal integrator (one pole at the origin).

Let us examine Eq. (7.9) closely. With a first-order LPF and the VCO (acting as an integrator), we expect a second-order transfer function. We also confirm a low-pass response: for slow input phase fluctuations, s is small and $H(s) \approx 1$. The closed-loop -3-dB bandwidth (sometimes called the "loop bandwidth") can be obtained by setting $|H(s = j\omega)|$ equal to $\sqrt{2}/2$ and computing ω. But what does ω represent here? In this context, ω is not the input or the VCO frequency, but the rate at which the input excess phase, $\phi_{in}(t)$, varies. In the voltage-mode circuit of Fig. 7.15, we would apply an input such as $V_{in} = V_0 \cos \omega_{in} t$ and observe a drop in the output voltage amplitude as ω_{in} exceeds the -3-dB bandwidth. In the PLL environment, on the other hand, we apply $V_{in} = V_0 \cos[\omega_0 t + \phi_{in}(t)]$, where $\phi_{in}(t) = \alpha \cos \omega_m t$. If ω_m is small, then $\phi_{in}(t)$

[2] In a strict sense, the "loop gain" refers to the low-frequency value and the "loop transmission" to the frequency-dependent function.

varies slowly, $|H| \approx 1$, and the PLL output emerges as $V_{out} \approx V_1 \cos(\omega_0 t + \alpha \cos \omega_m t)$, exhibiting the same phase excursions as the input. But if ω_m goes beyond the loop bandwidth, then the output phase fluctuations are smaller.

It is instructive to explain intuitively the dependence of ζ in (7.11) upon ω_{LPF} and $K_{PD}K_{VCO}$. To this end, we construct the Bode plots of the open-loop transfer function, $[K_{PD}/(1 + s/\omega_{LPF})]K_{VCO}/s$ (Fig. 7.18). In contrast to more familiar op amp designs, the PLL contains a pole at the *origin*, displaying an

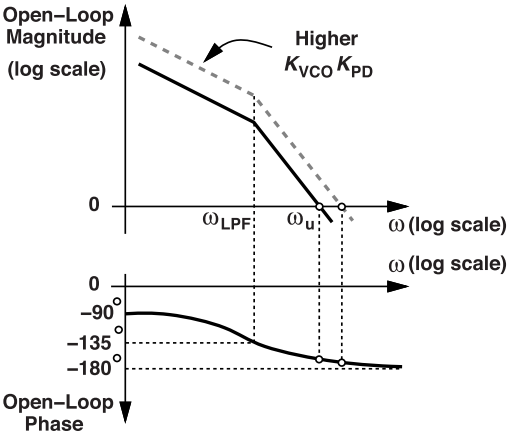


Figure 7.18 Bode plots of simple PLL's open-loop response.

open-loop gain that begins from infinity and falls at -20 dB/dec up to $\omega = \omega_{LPF}$. Beyond this frequency, the slope changes to -40 dB/dec. Also, the phase begins at -90° due to the pole at the origin, reaches -135° at $\omega = \omega_{LPF}$, and asymptotically approaches -180°.

The dependence of ζ upon $K_{PD}K_{VCO}$ is straightforward. In Fig. 7.18, increasing $K_{PD}K_{VCO}$ simply shifts the gain plot up, raising ω_u, but does not alter the phase plot (why not?). Thus, the phase margin decreases. This trend is similar to that in feedback amplifiers.

Recall that the low-pass filter suppresses the high-frequency components generated by the PD. Equation (7.11) suggests that ζ falls if ω_{LPF} is reduced; that is, if we lower ω_{LPF} so as to attenuate the PD injections, we inevitably degrade the stability. The Bode plots in Fig. 7.18 do not readily show this trend. However, we can return to the open-loop transfer function, $[K_{PD}/(1 + s/\omega_{LPF})](K_{VCO}/s)$, and calculate the unity-gain bandwidth, ω_u, by equating its magnitude squared to 1:

$$\frac{K_{PD}^2 K_{VCO}^2}{(1 + \dfrac{\omega_u^2}{\omega_{LPF}^2})\omega_u^2} = 1. \tag{7.12}$$

That is,

$$\frac{\omega_u}{\omega_{LPF}} = \sqrt{\frac{K_{PD}^2 K_{VCO}^2}{\omega_u^2} - 1}. \tag{7.13}$$

The open-loop phase at ω_u is equal to

$$\angle H(j\omega_u) = -90^\circ - \tan^{-1} \frac{\omega_u}{\omega_{LPF}} = -90^\circ - \tan^{-1} \sqrt{\frac{K_{PD}^2 K_{VCO}^2}{\omega_u^2} - 1}. \tag{7.14}$$

We note from Fig. 7.18 that, if ω_{LPF} is reduced, so is ω_u, leading to a larger argument for $\tan^{-1}$ and hence a smaller phase margin.

Example 7.10

Derive the closed-loop transfer function for the frequency-multiplying PLL of Fig. 7.9(a).

Solution

The open-loop forward transfer function is still given by $[K_{PD}/(1+s/\omega_{LPF})](K_{VCO}/s)$, but the feedback factor has fallen from 1 to $1/M$. Thus, we can simply replace K_{VCO} in the denominator of (7.8) with K_{VCO}/M:

$$H(s) = \frac{K_{PD}K_{VCO}}{\dfrac{s^2}{\omega_{LPF}} + s + K_{PD}K_{VCO}/M}. \tag{7.15}$$

It follows that

$$\omega_n = \sqrt{\frac{K_{PD}K_{VCO}}{M}\omega_{LPF}} \tag{7.16}$$

$$\zeta = \frac{1}{2}\sqrt{\frac{\omega_{LPF}}{K_{PD}K_{VCO}/M}}. \tag{7.17}$$

Equation (7.17) implies that the loop becomes more stable as M increases. Indeed, if we plot the magnitude and phase of the loop transmission (Fig. 7.19), we observe that the frequency divider weakens the feedback

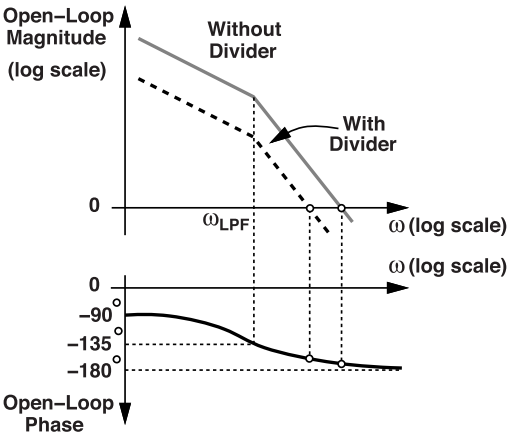


Figure 7.19 Magnitude and phase plots of loop transmission illustrating the effect of feedback divider on frequency response.

and increases the phase margin. These trends are a property of type-I PLLs (and op amps) and stand in contrast to the behavior of type-II loops (Section 7.5).

Example 7.11

A type-I PLL incorporates a voltage-controlled inverter-based ring oscillator. Explain qualitatively what type of transfer function the VCO supply noise experiences as it reaches the output phase.

Solution

As a quick test, we insert a dc voltage in series with the supply line of the VCO (Fig. 7.20) and seek its effect

on ϕ_{out}. This voltage tends to change f_{out}. For the loop to remain locked, f_{out} must still be equal to f_{in} even

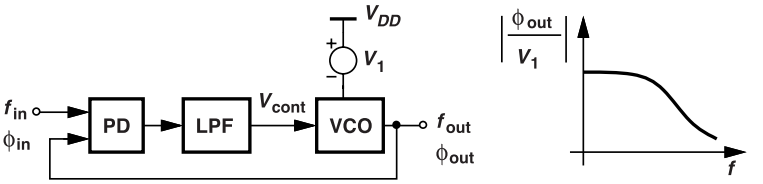


Figure 7.20 PLL experiencing supply noise.

though the VCO supply voltage has changed. This is possible only if V_{cont} changes so as to counteract the effect of V_1. For V_{cont} to change, the phase error must reach a different value. Thus, V_1 translates to a change in ϕ_{out}.

Now, if V_1 varies slowly with time (like low-frequency supply noise), so does ϕ_{out}. Fast variations of V_1, on the other hand, lead to a lesser change in ϕ_{out} (why?). The transfer function from V_1 to ϕ_{out} therefore has a low-pass response. The qualitative method outlined here proves valuable in analyzing the effect of various noise sources in PLLs.

7.3.5 Drawbacks of Simple PLL

Stability Issues The dependencies of ζ in Eq. (7.11) reveal two trade-offs in type-I PLLs. First, if we reduce ω_{LPF} to suppress the disturbance on the control voltage, the loop becomes less stable. Second, if we increase K_{PD} to minimize the static phase error, the loop becomes less stable. Both trends severely limit the performance.

Lock Acquisition The PLL of Fig. 7.6(c) suffers from another serious trade-off as well, namely, that between ω_{LPF} and the "acquisition range." An important question that we have (deliberately) avoided is, when the PLL is turned on, does it necessarily lock? For example, if the VCO initially operates far from the input frequency, can the loop properly tune the VCO and, eventually, lock it to the input? The short answer is, no, it cannot. The maximum initial value of $|\omega_{out} - \omega_{in}|$ for which the loop locks is called the "acquisition range."[3] For a more complete answer, the reader is encouraged to follow the explanation below.

Suppose, upon power-up, the VCO operates at $\omega_{out} \neq \omega_{in}$. For the loop to lock, a proper dc value must be established at the control input of the VCO. Let us open the loop as shown in Fig. 7.21(a) and examine the spectra at different points. Here, ω_{out} denotes a constant, arbitrary frequency. For simplicity, we assume the

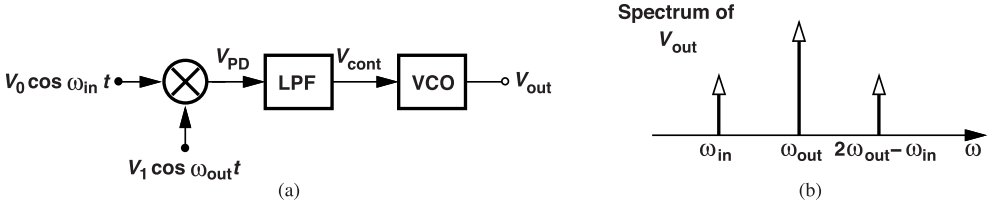


Figure 7.21 (a) Open-loop PLL for capture study, and (b) VCO output spectrum.

PD is realized as a mixer (Example 7.1). The PD output contains a component of the form $kV_0V_1\cos(\omega_{in} - \omega_{out})t$, where k denotes the gain of the mixer. Partially attenuated by the LPF, this component modulates the VCO, generating two sidebands at $\omega_{out} \pm (\omega_{in} - \omega_{out})$ [Fig. 7.21(b)]. In this situation, the dc value of V_{cont}

[3] Traditional PLL design literature distinguishes among acquisition range, lock range, pull-in range, tracking range, etc.

is zero. Now, if the loop is closed, the output component at ω_{in} mixes with the input (at ω_{in}), generating a dc value that drives the VCO frequency toward ω_{in}.

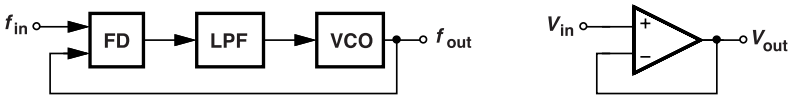
This qualitative study suggests that, if the initial difference between ω_{in} and ω_{out} is large, then the PD output given by $kV_0V_1\cos(\omega_{in} - \omega_{out})t$ experiences significant attenuation in the LPF, failing to modulate the VCO with sufficient strength. As a result, the VCO output sideband at ω_{in} and hence the dc component in the PD output are too small to initiate lock acquisition. The key point here is that, if a small ω_{LPF} is chosen to suppress the disturbance on V_{cont}, then the acquisition range is proportionally reduced. In practice, the acquisition range would be so narrow that the loop would readily fail with PVT variations.

In summary, the type-I PLL suffers from three trade-offs: between ζ and ω_{LPF}, between ζ and K_{PD}, and between the acquisition range and ω_{LPF}. We first devise a method of avoiding the limited acquisition range and then address the trade-offs incurred by ζ.

7.4 Phase/Frequency Detector

We wish to develop a PLL that locks regardless of the initial value of the output frequency. Such a PLL would have an acquisition range equal to the VCO tuning range, with no limitations imposed by the loop bandwidth.

The simple PLL of Fig. 7.6(c) suffers from a narrow acquisition range partly because the phase detector fundamentally measures the phase difference rather than the frequency difference, failing to generate a meaningful dc output when $|f_{out} - f_{in}|$ is large. We therefore envision that a "frequency detector" (FD)—a circuit whose average output represents the input frequency error—could resolve this issue. In fact, we might be tempted to replace the PD with an FD [Fig. 7.22(a)]. We call this structure a "frequency-locked loop" (FLL) and surmise that negative feedback forces f_{out} to become equal to f_{in}. But is that true? In analogy with the voltage-mode circuit of Fig. 7.22(b), where $V_{out} \neq V_{in}$ due to the op amp's finite gain and offset, we predict



(a) (b)

Figure 7.22 (a) Frequency-locked loop, and (b) voltage-domain analogy.

two difficulties: (1) with a finite loop gain, $f_{out} - f_{in} \neq 0$, and (2) in the presence of offsets (asymmetries within the FD), $f_{out} - f_{in} \neq 0$ even if the loop gain is infinite. While $V_{out} - V_{in} \neq 0$ is acceptable in Fig. 7.22(b), $f_{out} - f_{in} \neq 0$ is not in Fig. 7.22(a). For example, if a microprocessor's clock frequency is slightly different from its input data rate, the flipflops sampling the data incur errors.

The foregoing thoughts suggest that robust PLL operation requires both phase and frequency detection: when the frequency error is large, the PD output does not contain useful information, but the FD output does, pushing the VCO frequency toward the desired value. As $|f_{out} - f_{in}|$ becomes small enough, the PD begins to generate a meaningful dc value, eventually guaranteeing phase-lock and hence $\omega_{out} = \omega_{in}$.

Illustrated in Fig. 7.23 is a realization with two intertwined loops. Loop 1 is the PLL studied in the previous sections and Loop 2 is an FLL, both sharing the VCO. Upon power-up or a large transient, Loop 2 comes into action, with the FD measuring $f_{out} - f_{in}$ and adjusting the VCO with negative feedback. The output frequency therefore moves toward f_{in} until $|f_{out} - f_{in}|$ is within the acquisition range of Loop 1. Now, this loop takes over and gradually adjusts V_{cont}. After phase-locking, we have $f_{out} = f_{in}$, $V_2 \approx 0$, and $V_{cont} \approx V_1$. In principle, the frequency-acquiring loop is no longer necessary and can be disabled—until a transient causes the PLL to lose lock.

The architecture of Fig. 7.23 raises two questions. (1) Is it possible for the two loops to "fight" each other, prohibiting convergence and phase-locking? Yes, it is possible. We would choose a much narrower bandwidth for Loop 2 to ensure it has slow dynamics and steadily reaches its final state before Loop 1 begins to control the VCO. (2) Is it possible to merge the PD and the FD? Yes, we elaborate on this idea below.

Using sequential logic, we can construct a circuit that operates as an FD when the input frequencies are not equal and as a PD when they are. Depicted in Fig. 7.24, such a phase/frequency detector (PFD) receives

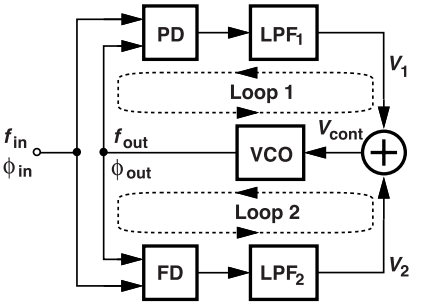


Figure 7.23 PLL with frequency acquisition aid.

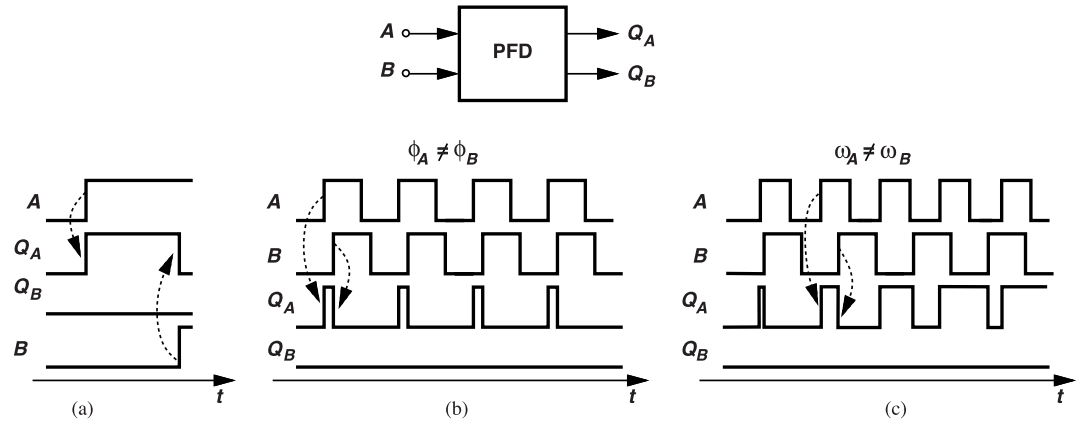


Figure 7.24 (a) Conceptual PFD operation, (b) case of input phase difference, and (c) case of input frequency difference.

two inputs and generates two outputs. The operation is based on two principles [Fig. 7.24(a)]: (1) if Q_A and Q_B are low and a rising edge occurs on A, then Q_A goes high while Q_B remains low, and (2) if Q_A is high and a rising edge occurs on B, then Q_A is reset. The circuit is symmetric with respect to A and B.

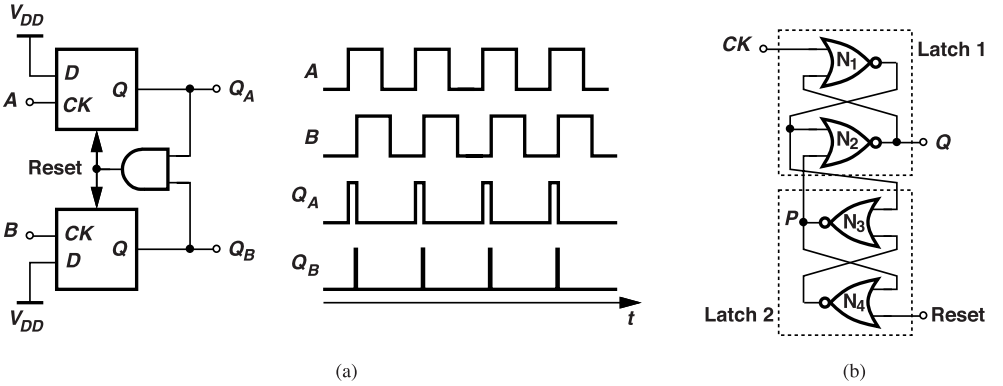
Let us see how these principles lead to phase and frequency detection. Shown in Fig. 7.24(b) is the case of equal frequencies, revealing an output pulsewidth equal to the phase difference. (In contrast to the XOR PD, this circuit produces pulses on only the rising edges.) Thus, the average value of Q_A minus that of Q_B is linearly proportional to the phase error. Similarly, for $\omega_A > \omega_B$, Fig. 7.24(c) demonstrates that Q_A exhibits a greater average than Q_B.

We now implement the foregoing PFD at the logic level. Illustrated in Fig. 7.25(a), the arrangement consists of two resettable D flipflops and an AND gate. The flipflops' D inputs are held at a logical ONE while their clock inputs sense A and B. The AND gate causes reset when Q_A and Q_B are both high. If, for example, A goes high, the logical high at the D input of the top FF travels to Q_A. Next, when B and hence Q_B rise, the AND gate generates a ONE, forcing Q_A and Q_B to return to zero. Unlike the idealized situations portrayed in Figs. 7.24(b) and 7.24(c), here Q_A and Q_B are simultaneously high for a brief period; we shall study the consequences of this effect in Chapter 8.

A simple, robust realization of the resettable D flipflops is shown in Fig. 7.25(b). The latches consist of cross-coupled NOR gates, and the output of each latch drives one input of the other.

Example 7.12

If the output Q in Fig. 7.25(b) is high and the input Reset goes high, how many gate delays does it take for Q to fall?



(a) (b)

Figure 7.25 (a) PFD implementation, and (b) example of resettable latch topology.

Solution

A rising edge on Reset propagates through N_4 and N_3 to reach P (Fig. 7.26), and the edge on P takes another gate delay to travel to Q. If the AND gate in Fig. 7.25(a) is realized as a NAND gate and an inverter, the total delay from Q_A or Q_B through the reset loop is about five gate delays, and so is the width of the narrow pulses on Q_B in Fig. 7.25(a). This rule of thumb proves useful in our future studies.

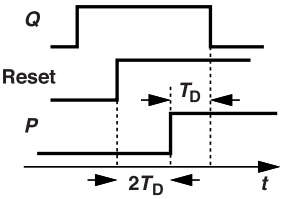


Figure 7.26 Propagation of Reset to Q.

The PFD concept eliminates the trade-off between the acquisition range and the loop bandwidth. As an example, Fig. 7.27 depicts a PLL using a PFD whose outputs are filtered and subtracted to produce V_{cont}.

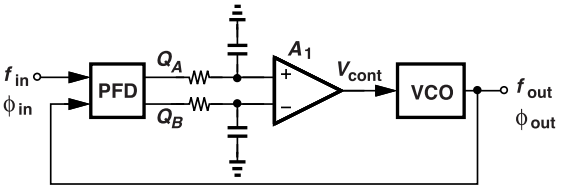


Figure 7.27 PLL employing a PFD.

But the type-I loop still poses trade-offs between stability, the loop bandwidth, and the static phase error. We develop in the next section a solution to deal with these trade-offs.

Example 7.13

Amplifier A_1 in Fig. 7.27 may suffer from high flicker noise. Explain qualitatively what type of transfer function this noise experiences as it translates to output phase noise.

Solution

As in Example 7.11, we insert a dc voltage in series with one of the amplifier inputs (Fig. 7.28) and determine how it affects ϕ_{out}. Such a voltage tends to change V_{cont}, but the PLL ensures that $f_{out} = f_{in}$ and hence

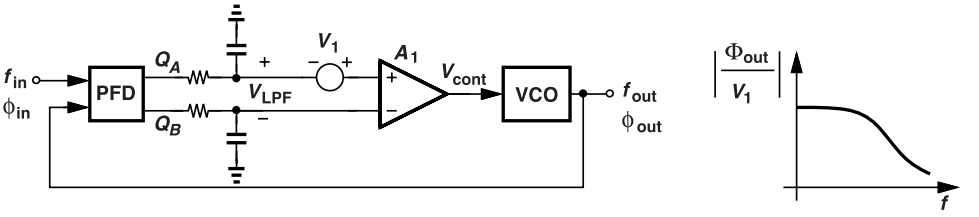


Figure 7.28 Effect of amplifier noise in a PLL.

V_{cont} is constant. Thus, V_{LPF} must contain an additional dc voltage equal to $-V_1$ so that the net difference sensed by A_1 is the same as before V_1 was inserted. For V_{LPF} to change by this amount, the phase error measured by the PFD must change, and so must ϕ_{out}. In summary, $V_1 \neq 0$ translates to a change in the output phase. If V_1 now varies slowly, ϕ_{out} does, too. The transfer function from V_1 to ϕ_{out} therefore has a low-pass shape, making the flicker noise of A_1 critical. In Problem 7.28, we show that the thermal noise of the resistors in Fig. 7.28 has the same effect.

7.5 Charge-Pump PLLs

Most of today's PLLs employ a charge pump (CP) in conjunction with a PFD, exhibiting interesting and useful characteristics. We begin our study with the charge pump itself.

7.5.1 Charge Pumps

In the context of PLL design, a charge pump is a circuit that sources or sinks charge for a controlled amount of time. Shown in Fig. 7.29(a) is a simple realization: if S_1 is on, I_1 charges C_1, and if S_2 is on, I_2 discharges

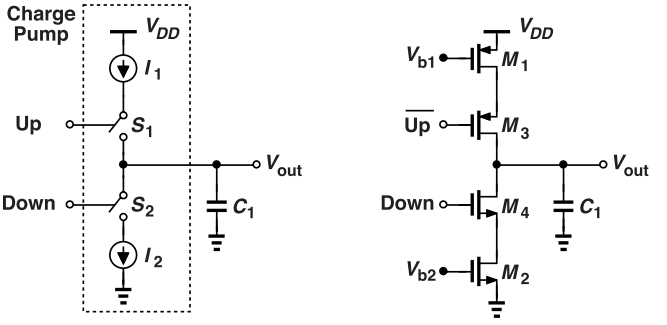


Figure 7.29 (a) Basic charge pump, and (b) transistor-level realization.

it. The controls are called Up and Down, respectively, as they determine whether the output voltage rises or falls. We assume $I_1 = I_2 = I_p$. Depicted in Fig. 7.29(b), the transistor-level implementation uses M_1 and M_2 as current sources and M_3 and M_4 as switches. Note that the gate control of M_3 is called $\overline{\text{Up}}$ to emphasize that, when Up is high, this switch turns on. Since the switches are placed in series with the drains of the current sources, this topology is called a "drain-switched" CP.

7.5.2 PFD/CP/Capacitor Cascade

Consider the cascade shown in Fig. 7.30, assuming A and B have equal frequencies but a finite phase

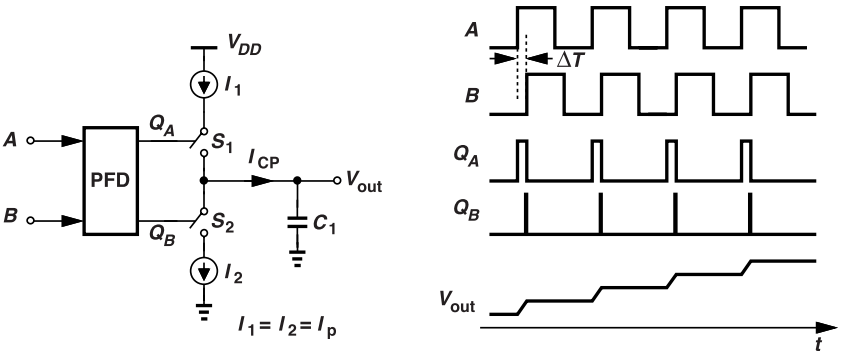


Figure 7.30 Cascade of PFD, charge pump, and capacitor along with the waveforms.

error. Each time a phase comparison is made, Q_A goes high, S_1 turns on, I_1 charges C_1, and V_{out} rises by $\Delta V = (I_p/C_1)\Delta T$. Interestingly, if the input phase difference remains constant, V_{out} goes toward infinity. We intuitively see that this combination provides an infinite "gain" because, for a finite phase error, V_{out} grows indefinitely. From another perspective, for V_{out} to be *finite*, the phase error must be zero so that the CP injects no net charge in every cycle. We return to this important point shortly.

7.5.3 Basic Charge-Pump PLL

Let us now construct a PLL incorporating the PFD/CP/capacitor cascade (Fig. 7.31). If the loop is locked,

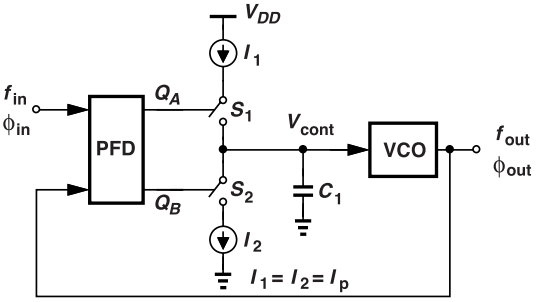


Figure 7.31 PLL using PFD/CP/capacitor cascade.

$f_{out} = f_{in}$, and V_{cont} has a certain finite value. For V_{cont} to remain constant, the charge pump must inject no net charge into C_1, a condition requiring that $\phi_{out} = \phi_{in}$. In other words, regardless of such parameters as the input frequency, K_{VCO}, and I_p, the loop locks with a *zero* "static" phase error. Arising from the integration action of the CP/capacitor combination, this property stands in contrast to the behavior of the type-I PLL.

Example 7.14

Explain qualitatively how the PLL in Fig. 7.31 reacts if the VCO suffers from low-frequency phase noise, i.e., it has slow phase fluctuations.

Solution

As ϕ_{out} drifts away from ϕ_{in} by some amount $\Delta\phi$, the PFD generates a pulse with a width equal to $[\Delta\phi/(2\pi)]T_{in}$ seconds on Q_A or Q_B, turning on one of the CP switches. The corresponding current source then charges or discharges C_1 briefly, changing V_{cont} and f_{out}. As a result, the VCO accumulates phase at a

different rate, cancelling the drift. This process continues for some cycles until ϕ_{out} becomes approximately equal to ϕ_{in}. That is, the loop attempts to correct for the VCO phase noise. We quantify this behavior in Chapter 8.

7.5.4 PFD/CP/Capacitor Transfer Function

In order to formulate the dynamics of the PLL shown in Fig. 7.31, we must develop a transfer function for the PFD/CP/capacitor cascade. The difficulty here is that the input is a phase quantity, and the output a voltage, with a great deal of switching circuitry in between. We begin our analysis in the time domain, compute the impulse response, and take the Laplace transform of the result to obtain the transfer function.

To study the impulse response of the cascade, we must apply a *phase* impulse to the input, i.e., the phase difference between A and B in Fig. 7.30 must jump to infinity and return to zero. Since such a scenario is not intuitive, we apply a phase *step*, determine the output, and then take its derivative to find the impulse response. As illustrated in Fig. 7.32, the B input experiences a phase step equal to $\Delta\phi_1$ at $t = t_1$, causing Q_A to remain high for $[\Delta\phi_1/(2\pi)]T_{in}$ seconds at each phase comparison instant. The control voltage thus rises by $\Delta V = [\Delta\phi_1/(2\pi)]T_{in}(I_p/C_1)$ each time.

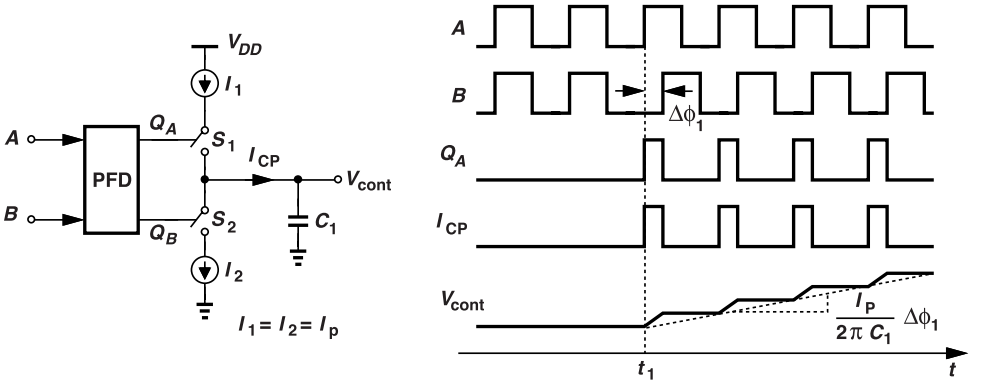


Figure 7.32 PFD/CP/capacitor cascade sensing a finite input phase error.

The V_{cont} waveform shown in Fig. 7.32 is the true step response of the system, but it also reveals that the cascade is nonlinear! As a quick linearity test, we double $\Delta\phi_1$, observing that the flat sections of V_{cont} double in value, but the ramp sections retain their slope, I_1/C_1. That is, the overall V_{cont} waveform is not multiplied by two. We can nonetheless approximate V_{cont} by a continuous ramp as shown by the dashed line, attributing to it a slope of $\Delta V/T_{in} = [\Delta\phi_1/(2\pi)](I_p/C_1)$. That the step response of the cascade is a linear ramp suggests an integration operation. From this step response, we can determine the impulse response to be $[\Delta\phi_1/(2\pi)](I_p/C_1)u(t)$ (for an input impulse of the form $\Delta\phi_1\delta(t)$]. The transfer function is given by the Laplace transform of the impulse response:

$$\frac{V_{cont}}{\Delta\phi}(s) = \frac{I_p}{2\pi C_1 s}. \tag{7.18}$$

The $1/s$ dependence signifies a pole at the origin and hence an ideal integrator. We can take our approximation one step further, write $V_{cont}/\Delta\phi = [I_P/(2\pi)][1/(C_1 s)]$, and consider $I_p/(2\pi)$ the transfer function of the PFD/CP cascade, i.e.,

$$\frac{I_{CP}}{\Delta\phi}(s) = \frac{I_p}{2\pi}. \tag{7.19}$$

With the aid of Eq. (7.18), we can construct the linear phase model depicted in Fig. 7.33, obtaining the open-loop transfer function as $[I_p/(2\pi C_1 s)](K_{VCO}/s)$. This architecture is called a "type-II" PLL because

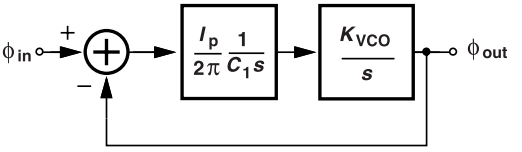


Figure 7.33 Linear model of PLL with PFD/CP/capacitor cascade.

it contains two poles at the origin (two integrators). The closed-loop transfer function emerges as

$$H(s) = \frac{I_p K_{VCO}}{2\pi C_1 s^2 + I_p K_{VCO}}, \tag{7.20}$$

exhibiting two imaginary poles and implying an unstable loop. (Note that $\zeta = 0$ here.) This unfortunate situation is caused by the two ideal integrators.

To stabilize a loop containing two ideal integrators, we have two options: (a) make one of the integrators "lossy" so that its transfer function is of the form $1/(s + \omega_1)$, or (b) add an open-loop zero so that one integrator's transfer function changes to $\alpha + \beta/s = (\alpha s + \beta)/s$. We pursue the latter here. We place a resistor in series with the capacitor (Fig. 7.34). Note that now, the control voltage abruptly jumps by an amount of $I_p R_1$ when the charge pump turns on or off, an effect absent in the loop without a zero. As mentioned earlier,

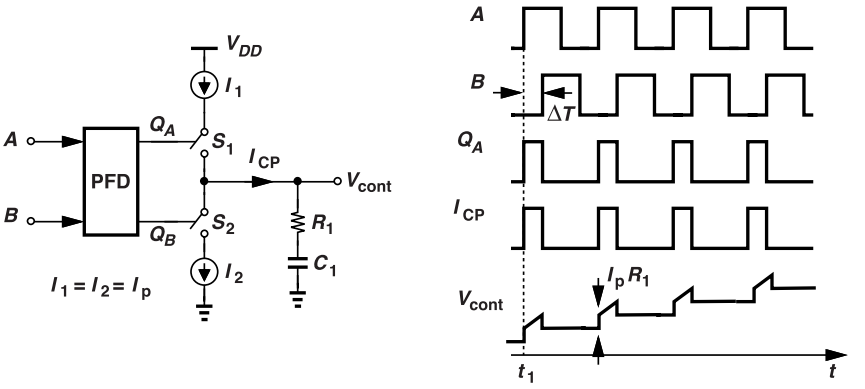


Figure 7.34 Addition of resistor R_1 in series with C_1 to create a zero.

we can attribute a transfer function of $I_p/(2\pi)$ to the PFD/CP cascade, obtaining

$$\frac{V_{cont}}{\Delta\phi}(s) = \frac{I_{CP}}{\Delta\phi} \cdot \frac{V_{cont}}{I_{CP}} \tag{7.21}$$

$$= \frac{I_p}{2\pi}\left(\frac{1}{C_1 s} + R_1\right). \tag{7.22}$$

The new cascade has a zero at $-1/(R_1 C_1)$. Using this arrangement, we construct the basic "charge-pump PLL" (CPPLL) shown in Fig. 7.35 and call the $R_1 C_1$ branch the "loop filter." Note that the static phase error

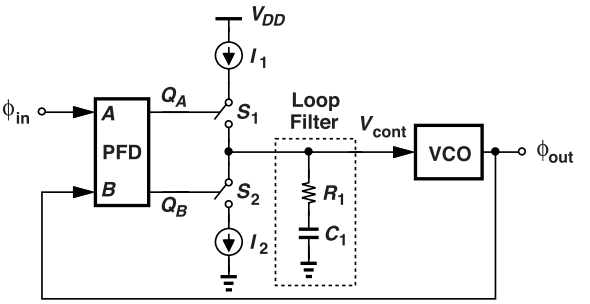


Figure 7.35 Charge-pump PLL.

is still zero due to the integration provided by the CP and C_1. The closed-loop transfer function is given by

$$H(s) = \frac{\frac{I_p K_{VCO}}{2\pi C_1}(R_1 C_1 s + 1)}{s^2 + \frac{I_p}{2\pi} K_{VCO} R_1 s + \frac{I_p K_{VCO}}{2\pi C_1}}. \tag{7.23}$$

It follows that

$$\omega_n = \sqrt{\frac{I_p K_{VCO}}{2\pi C_1}} \tag{7.24}$$

$$\zeta = \frac{R_1}{2}\sqrt{\frac{I_p K_{VCO} C_1}{2\pi}}. \tag{7.25}$$

Let us ponder the results obtained thus far. The closed-loop transfer function in (7.23) contains one zero and two poles, exhibiting a low-pass response. That is, slow input phase fluctuations travel to the output, but fast phase changes are filtered out. Also, (7.25) indicates that ζ *increases* with C_1, a desirable trend in contrast to that in the type-I PLL (Section 7.3.5). It is possible to calculate the -3-dB bandwidth and plot the frequency response of the system, but we defer these to Chapter 8. In summary, the architecture of Fig. 7.35 resolves all of the three drawbacks identified for the type-I PLL in Section 7.3.5.

In Eqs. (7.24) and (7.25), I_p and K_{VCO} appear only as a product, suggesting that if they change in opposite directions, the response remains the same. However, a number of issues arise if K_{VCO} is high or I_p is low (Chapter 8). For a VCO frequency of ω_0 rad/s, we limit K_{VCO} to less than $\omega_0/10$ rad/s/V (Chapters 3 and 5).

Example 7.15

Using Bode plots, examine the stability of the CPPLL for $R_1 = 0$ and $R_1 > 0$.

Solution

With $R_1 = 0$, the system consists of two ideal integrators, displaying the open-loop characteristics shown in Fig. 7.36(a) and a zero phase margin. With $R_1 > 0$, an open-loop zero appears at $-1/(C_1 R_1)$, deflecting both the magnitude and the phase upward [Fig. 7.36(b)]. Thus, the unity-gain frequency, ω_u, can be chosen to obtain a desired phase margin (Chapter 8). Interestingly, the phase margin is equal to 45° if $\omega_u = 1/(R_1 C_1)$, a useful rule of thumb.

Example 7.16

Suppose ω_u in Fig. 7.36(b) is around $1/(R_1 C_1)$. For $\omega < 1/(R_1 C_1)$, we approximate the magnitude plot as

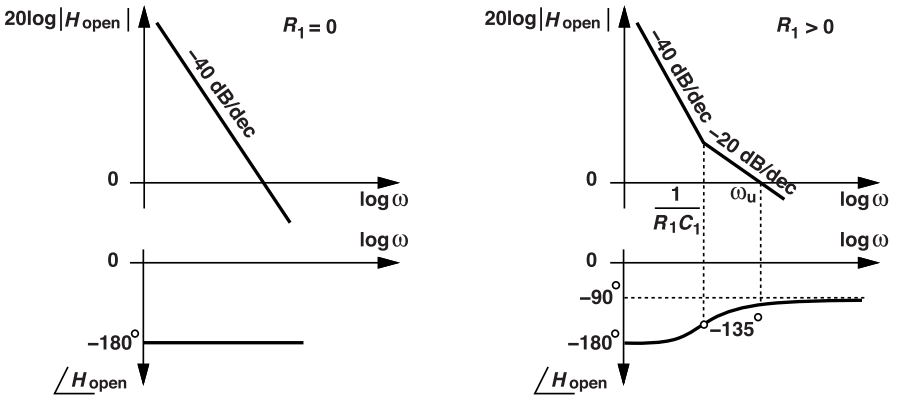


Figure 7.36 Bode plots of charge-pump PLL with (a) $R_1 = 0$, and (b) $R_1 > 0$.

$|[I_p/(2\pi)][1/(C_1 s)](K_{VCO}/s)|$, equate it to unity, and choose $\mathrm{PM} = 45°$. Calculate ζ in this case.

Solution

We first calculate ω_u by writing

$$\frac{I_p}{2\pi}\frac{K_{VCO}}{C_1\omega_u^2} = 1 \tag{7.26}$$

and hence

$$\omega_u = \sqrt{\frac{I_p K_{VCO}}{2\pi C_1}}. \tag{7.27}$$

For $\mathrm{PM} = 45°$, the zero coincides with ω_u, $\omega_u = 1/(R_1 C_1)$, and we obtain $R_1 = \sqrt{2\pi/(I_p K_{VCO} C_1)}$. From Eq. (7.25), $\zeta = 1/2$, an inadequate value for a well-behaved response. In other words, R_1 should be chosen larger, which means the zero frequency must lie below ω_u.

We should remark that the polarity of feedback in Fig. 7.35 does matter; i.e., the VCO output must return to the PFD such that the CP drives V_{cont} in a direction that reduces $|f_{out} - f_{in}|$.

Unlike type-I topologies, the CPPLL becomes *more* stable as K_{VCO} increases [Eq. (7.25)]. This can be seen from the characteristics in Fig. 7.36(b): a higher K_{VCO} shifts the magnitude plot up but does not alter the phase plot; consequently, ω_u and the phase margin increase. The charge pump current has the same effect. This point reveals that the stability degrades if the effective $I_p K_{VCO}$ falls. For example, the presence of a feedback frequency divider reduces the loop gain. In fact, our studies in Section 7.3.2 and Example 7.10 imply that, for the frequency-multiplying CPPLL shown in Fig. 7.37, we have

$$\omega_n = \sqrt{\frac{I_p K_{VCO}}{2\pi C_1 M}} \tag{7.28}$$

$$\zeta = \frac{R_1}{2}\sqrt{\frac{I_p K_{VCO} C_1}{2\pi M}}. \tag{7.29}$$

Thus, R_1, I_p, or C_1 must be increased so as to compensate for the effect of M on ζ.

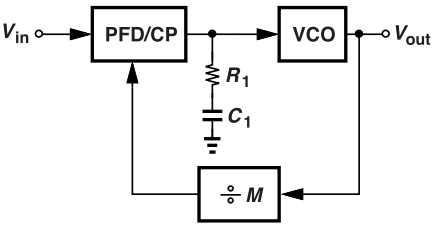


Figure 7.37 Frequency-multiplying charge-pump PLL.

Example 7.17

A 5-GHz WiFi transceiver employs a CPPLL with a 20-MHz reference frequency. If $K_{VCO} = 300$ MHz/V (Chapter 5) and $\zeta = 1$, study the range of values for I_p, C_1, and R_1.

Solution

To multiply the reference to 5 GHz, we have $M = 250$. From Eq. (7.29),

$$\zeta = 1 = \frac{R_1}{2}\sqrt{\frac{I_p(2\pi \times 300 \times 10^6)C_1}{2\pi \times 250}} \tag{7.30}$$

and hence

$$R_1\sqrt{I_pC_1} = 1.83 \times 10^{-3}. \tag{7.31}$$

Let us explore the design space. For example, with $I_p = 100\ \mu$A and $C_1 = 5$ pF, we have $R_1 = 81.8\ \text{k}\Omega$ and $\omega_n = 2\pi(780\ \text{kHz})$. Would such a design be acceptable? We return to this point later in this chapter and in Chapter 9.

The following examples help us develop some intuition.

Example 7.18

A student decides to interpose an amplifier between the loop filter and the VCO in Fig. 7.35 so as to ease the trade-off between I_p and C_1. Explain the issues in such an architecture.

Solution

The amplifier changes the product I_pK_{VCO} to $I_pA_1K_{VCO}$ because we can simply view A_1 and the VCO as a new VCO with a gain equal to A_1K_{VCO}. However, the amplifier may contain poles that degrade the phase margin. Also, the noise of the amplifier translates to output phase noise. To study the noise, we insert, in a manner similar to that in Example 7.13, a constant voltage source at the input of the amplifier [Fig. 7.38(a)] and seek its effect on ϕ_{out}. Since f_{out} does not change after this voltage is added, the filter voltage, V_1, must decrease by V_n so that V_{cont} remains constant. In a CPPLL, the infinite gain provided by the PFD/CP/filter cascade can create any V_1 without changing ϕ_{out}. That is, a constant V_n (a dc offset) has no effect on the output.

Now, we allow V_n to vary with time. Since the PFD/CP/filter cascade can no longer provide an infinite gain, ϕ_{out} changes to some extent in response. That is, the transfer function from V_n to ϕ_{out} has a high-pass behavior, a point of contrast to Example 7.13. The type-II loop proves beneficial in suppressing the flicker noise of A_1.

In reality, for very high-frequency components in V_n, ϕ_{out} is negligibly modulated. This can be seen by noting that the gain of the PFD/CP/filter integrator drops, the feedback action vanishes, and, hence, $\phi_{out} \approx$

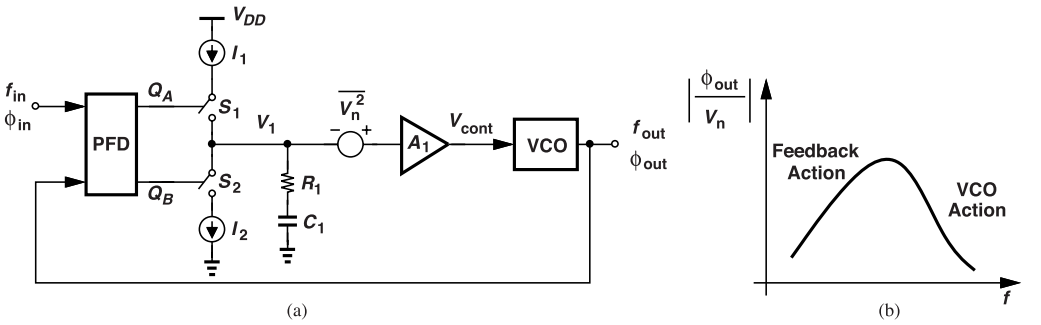


Figure 7.38 (a) Effect of amplifier noise in CPPLL, and (b) noise transfer function.

$A_1 V_n K_{VCO}/s \to 0$. That is, the transfer function in fact has a band-pass shape. Figure 7.38(b) illustrates the result.

Example 7.19

Resistor R_1 in Fig. 7.35 generates thermal noise, disturbing V_{cont}. What type of transfer function do we expect from this noise to the output phase noise.

Solution

Following our approach in Examples 7.13 and 7.18, we insert a voltage source in series with R_1 (Fig. 7.39) and examine its effect on ϕ_{out}. If V_n is constant, it attempts to raise V_{cont}, but the loop reacts so as to keep

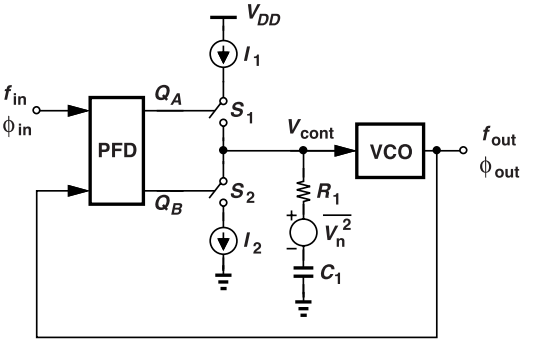


Figure 7.39 Effect of resistor noise in PLL.

f_{out} constant. Thus, the voltage on C_1 decreases by an amount equal to V_n, a change accommodated by the PFD/CP/filter integrator without affecting ϕ_{out}. On the other hand, if V_n varies fast enough, V_{cont} begins to change because the feedback loop cannot correct for it. The transfer function therefore exhibits a high-pass response (in fact, a band-pass response as in Example 7.18).

Example 7.20

A student mistakenly ties the control of the VCO to the capacitor rather than the resistor [Fig. 7.40(a)]. Is such a loop stable?

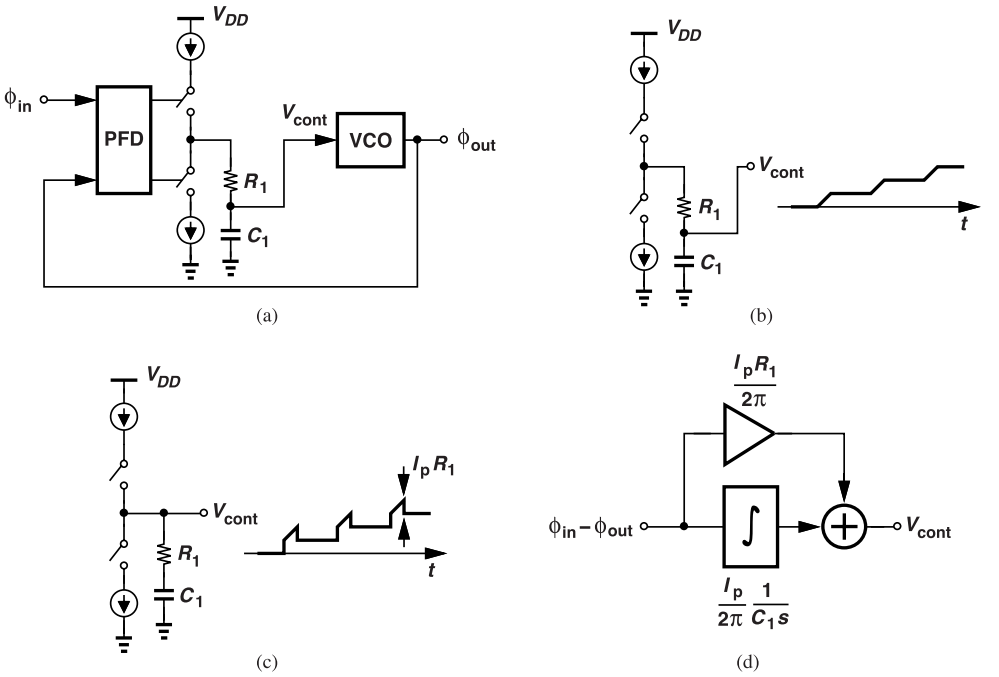


Figure 7.40 (a) CPPLL with VCO control sensed at capacitor terminal, (b) open-loop response, (c) open-loop response showing jumps, and (d) conceptual model.

Solution

No, it is not. Since R_1 appears in series with a current source when S_1 or S_2 is on, it plays no role. The loop thus reduces to the two-integrator arrangement shown in Fig. 7.31.

We can also examine the correct and incorrect arrangements in the time domain. As illustrated in Figs. 7.40(b) and (c), when one switch turns on, the control voltage jumps abruptly in one case but not in the other case. It is indeed this "unintegrated" jump that stabilizes the loop. Going back to the frequency domain, we can visualize the PFD/CP/filter cascade as (1) an ideal integrator, and (2) a feedforward path that carries the signal directly to the VCO [Fig. 7.40(d)]. The feedforward introduces a zero and is necessary for stability.

Example 7.21

A microprocessor must run at a clock frequency of 100 MHz during regular operation and 3 GHz for computationally intensive applications. Explain the effect of changing the PLL feedback divide ratio by a factor of 30.

Solution

From Eqs. (7.28) and (7.29), we observe that both ω_n and ζ change by $\sqrt{30} \approx 5.5$. We must therefore adjust the other loop parameters to compensate for this variation. For a 30-fold increase in M, we can raise the charge pump current by the same factor, thus keeping ζ and ω_n constant. That is, the CP current sources can be programmable and simply scaled by this factor. However, this can lead to a large ripple on the control voltage. We therefore adjust the loop filter as well. In practice, the charge pump is decomposed into parallel identical "slices" that can be turned on or off to establish the desired I_p, and the loop filter is also made programmable.

7.5.5 Phase Margin Calculation

Drawing upon the open-loop transfer function of the CPPLL, we can compute its phase margin. We first determine the frequency at which the magnitude of the loop transmission falls to unity:

$$\left| \frac{I_p}{2\pi} \left(R_1 + \frac{1}{C_1 s} \right) \frac{K_{VCO}}{s} \right|^2_{s=j\omega u} = 1. \tag{7.32}$$

That is,

$$\left(\frac{I_p K_{VCO}}{2\pi} \right)^2 \frac{R_1^2 C_1^2 \omega_u^2 + 1}{C_1^2 \omega_u^4} = 1. \tag{7.33}$$

Utilizing the ζ and ω_n expressions given by (7.24) and (7.25), respectively, we have

$$\omega_u^4 - 4\zeta^2 \omega_n^2 \omega_u^2 - \omega_n^4 = 0. \tag{7.34}$$

It follows that

$$\omega_u^2 = (2\zeta^2 + \sqrt{4\zeta^4 + 1})\omega_n^2. \tag{7.35}$$

The phase of the loop transmission (loop gain) at $\omega = \omega_u$ [Fig. 7.36(b)] is given by the zero's contribution minus 180° (from the two poles at the origin), yielding

$$\text{PM} = \tan^{-1} \frac{\omega_u}{\omega_z} \tag{7.36}$$

$$= \tan^{-1}(2\zeta \sqrt{2\zeta^2 + \sqrt{4\zeta^4 + 1}}), \tag{7.37}$$

because $\omega_n/\omega_z = 2\zeta$. In Problem 7.19, we prove that $\text{PM} \approx \tan^{-1} 4\zeta^2$ for $\zeta > 0.8$, a simple and useful expression to remember. If the loop contains a divider, $\zeta = (R_1/2)\sqrt{I_p K_{VCO} C_1/(2\pi M)}$. For the typical case of $\zeta = 1$, we have $\text{PM} = 76^\circ$ and $\omega_u/\omega_z \approx 4$. That is, the unity-gain bandwidth in Fig. 7.36(b) is four times the zero frequency. In a more aggressive design, we choose $\zeta = \sqrt{2}/2$, $\text{PM} = 65^\circ$, and $\omega_u/\omega_z \approx 2.2$.

Example 7.22

As ζ goes from $\sqrt{2}/2$ to large values, we have from Eq. (7.35)

$$\omega_u \approx 2\zeta\omega_n \tag{7.38}$$

$$\approx \frac{I_p R_1 K_{VCO}}{2\pi}. \tag{7.39}$$

That is, the unity-gain bandwidth is relatively independent of C_1. Explain intuitively why.

Solution

Let us write $R_1^2 C_1^2 \omega_u^2$ as ω_u^2/ω_z^2, noting that this value is well above unity for $\zeta = \sqrt{2}/2$. Thus, in Eq. (7.33), we have $(R_1^2 C_1^2 \omega_u^2 + 1)/(C_1^2 \omega_u^4) \approx (R_1^2 C_1^2 \omega_u^2)/(C_1^2 \omega_u^4) = R_1^2/\omega_u^2$. It follows that $\omega_u \approx I_p K_{VCO} R_1/(2\pi)$. From another standpoint, we can say that, in Eq. (7.32), $R_1 \gg 1/|C_1 s|$ at $s = j\omega_u$ because ω_u lies well above $\omega_z = 1/(R_1 C_1)$. Thus, $[I_p/(2\pi)]R_1 K_{VCO} = \omega_u$. We offer another explanation in Chapter 8.

Example 7.23

Consider the open-loop PLL response plotted in Fig. 7.41. At ω_1, the loop gain is greater than unity and the phase is (evidently) $-180°$. Does this mean that the PLL becomes unstable and oscillates at this frequency?

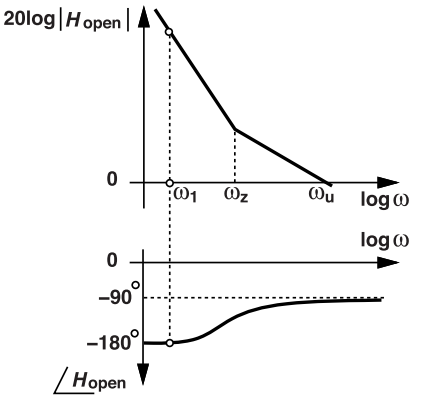


Figure 7.41 Open-loop PLL gain and phase values at ω_1.

Solution

No, it does not. The short explanation is that the zero, ω_z, does not allow $\angle H_{\text{open}}$ to reach exactly $-180°$ even if ω goes to very low values. This can be seen by first expressing the open-loop transfer function (the loop transmission) as $H_{\text{open}}(s) = \alpha(1 + s/\omega_z)/s^2$, where $\alpha = I_p K_{VCO}/(2\pi C_1)$ and $\omega_z = 1/(R_1 C_1)$. We have $\angle H_{\text{open}}(j\omega) = -180° + \tan^{-1}(\omega/\omega_z)$, a value less negative than $-180°$. That is, the zero's role is to ensure the open-loop phase does not reach $-180°$.

A second perspective is to assume that the loop oscillates at some general frequency, $\sigma_a + j\omega_a$, and determine whether σ_a and ω_a exist. For this to occur, we must have $H_{\text{open}}(\sigma_a + j\omega_a) = -1$ and hence

$$\alpha\left(1 + \frac{\sigma_a + j\omega_a}{\omega_z}\right) = -(\sigma_a + j\omega_a)^2. \tag{7.40}$$

Equating the imaginary parts of both sides gives

$$\alpha\frac{\omega_a}{\omega_z} = -2\sigma_a\omega_a, \tag{7.41}$$

which is possible if $\omega_a = 0$ or if $\sigma_a = -\alpha/(2\omega_z)$, neither of which enables oscillation. Again, we recognize that the finite zero frequency, ω_z, produces a finite, negative σ_a, thereby allowing only *decaying* oscillatory transients.

The third approach is based on Nyquist's stability theorem and examines the real and imaginary parts of $H_{\text{open}}(s)$ as s travels through the complex plane. Nyquist's method stands in contrast to Bode plots, which assume that the "test" frequency, s, is equal to $j\omega$ and hence confined to the imaginary axis. Such plots therefore provide little information about the behavior of the loop for *complex* s values, i.e., for $s = \sigma + j\omega$.

To apply Nyquist's approach to a type-II PLL, the reader is referred to [1], but we illustrate the results in Fig. 7.42. We allow s to move on a small circuit near the origin, travel on the $j\omega$ axis, and eventually enter the right half plane (RHP) at a very large radius. Figure 7.42(a) shows the behavior of $H_{\text{open}}(s)$ in the absence of the zero: the H_{open} contour begins at a large radius (point M), crosses the imaginary axis (point N), returns to the real axis, crosses the point $(-1\ ,0)$ for $s = j\sqrt{\alpha}$, and approaches the origin as $s \to +j\infty$. Since the H_{open} contour crosses $(-1, 0)$ twice, we conclude that the closed-loop system exhibits two poles on the $j\omega$ axis. This agrees with our initial results for charge-pump PLLs and the need for stabilization.

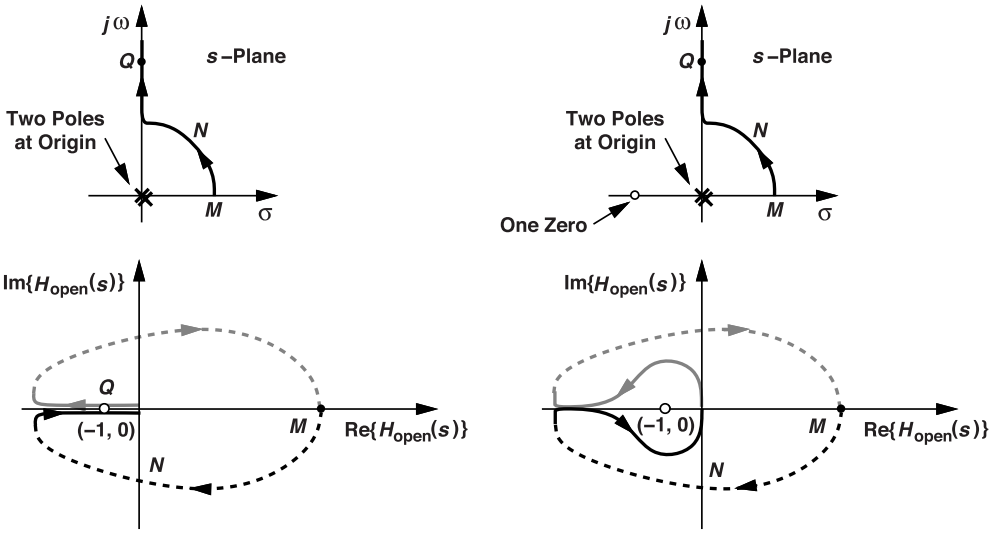


Figure 7.42 Nyquist plots for PLL open-loop transfer function with (a) $R_1 = 0$, and (b) $R_1 > 0$.

In Fig. 7.42(b), we repeat the foregoing analysis but assume the loop contains a zero. We observe that the H_{open} contour does not pass through or encircle the point $(-1, 0)$. The closed-loop system is therefore stable.

7.6 Higher-Order Loops

Owing to various imperfections (Chapter 8), the CPPLL of Fig. 7.35 in reality suffers from considerable ripple on the control voltage, requiring at least one more capacitor in the loop filter. As an example of such imperfections, suppose we employ the CP of Fig. 7.29(b) and precede the PMOS switch with an inverter so as to turn it on when Q_A goes high [Fig. 7.43(a)]. When the loop is locked, the phase error is zero, and Q_A

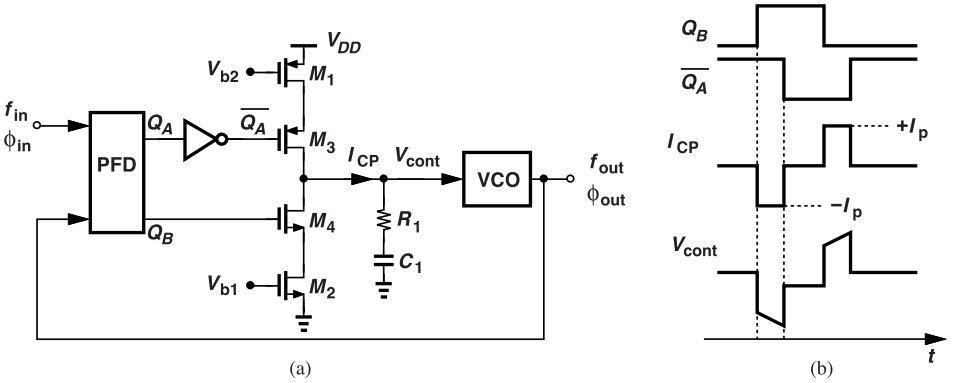


Figure 7.43 (a) PLL with transistor-level CP implementation, and (b) PLL waveforms.

and Q_B rise and fall together. However, M_3 turns on and off later than M_4. That is, the up and down pulses suffer from a skew. Let us study the current injected by the charge pump, I_{CP}, into the loop filter with the aid of the waveforms shown in Fig. 7.43(b). On the rising edge of Q_B, M_4 turns on, drawing a current equal to

I_p while M_3 is off. Thus, V_{cont} experiences a negative jump equal to $-I_pR_1$, and C_1 is slightly discharged before M_3 turns on on the falling edge of Q_A. When both M_3 and M_4 are on, I_{CP} is zero. On the falling edge of Q_B, the reverse occurs: I_{CP} is positive for about one gate delay, causing a positive jump on V_{cont}. Of course, the skew created by the inverter can be partially removed if a complementary pass gate is interposed between the PFD and the gate of M_4. (Why does the pass gate not remove the skew completely?)

How significant are the jumps on V_{cont}? From Example 7.17, we have $I_pR_1 = 100\ \mu\text{A} \times 81.8\ \text{k}\Omega = 8.18$ V!! Limited by V_{DD} and the drain-source voltage drops, the jumps are still enormous, violently modulating the VCO.

Example 7.24

A CPPLL incorporates a divider having a divide ratio of M. Determine the output spectrum in the presence of the up and down skews illustrated in Fig. 7.43(b). Use the narrowband FM approximation (Chapter 2).

Solution

The double pulses on V_{cont} in Fig. 7.43(b) have a width, T_{sk}, of about one gate delay and are separated by an amount, ΔT, of about five gate delays (Example 7.12). Since both of these values are much smaller than the PLL input period, T_{in}, we can approximate the pulses by impulses [Fig. 7.44(a)], each having an area of $I_pR_1T_{sk}$. The Fourier transform of these impulses is given by two impulse trains:

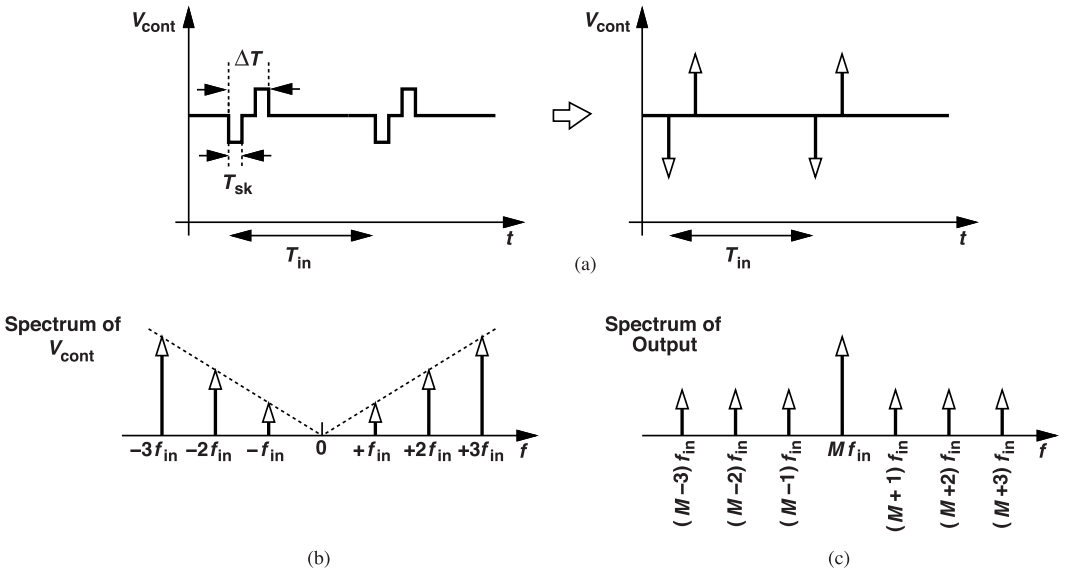


Figure 7.44 (a) Approximation of control voltage ripple by impulses, (b) spectrum of V_{cont}, and (c) VCO output spectrum.

$$V_{cont}(f) = \frac{I_pR_1T_{sk}}{T_{in}} \left[\sum_{k=-\infty}^{+\infty} \delta(f - kf_{in}) - \sum_{k=-\infty}^{+\infty} e^{-j2\pi f\Delta T} \delta(f - kf_{in}) \right], \tag{7.42}$$

where the factor $\exp(-j2\pi f\Delta T)$ signifies the time delay between the impulses. Since ΔT is small, we have $\exp(-j2\pi f\Delta T) \approx 1 - j2\pi f\Delta T$, and

$$V_{cont}(f) = \frac{I_pR_1T_{sk}}{T_{in}} (j2\pi f\Delta T) \sum_{k=-\infty}^{+\infty} \delta(f - kf_{in}). \tag{7.43}$$

The frequency-domain impulses therefore exhibit a height that rises with frequency [Fig. 7.44(b)]. Upon traveling through the VCO, these impulses are multiplied by K_{VCO}/ω and translated to a center frequency of $f_{out} = Mf_{in}$ [Fig. 7.44(c)]:

$$V_{out}(f) = j\frac{K_{VCO}I_pR_1T_{sk}\Delta T}{T_{in}} \sum_{k=-\infty}^{+\infty} \delta(f - Mf_{in} - kf_{in}), \tag{7.44}$$

where $V_{out}(f)$ represents only the sidebands. Note that the magnitude of this spectrum is normalized to the carrier magnitude. These sidebands create difficulties in both RF and timing applications.

Having seen the severity of the ripple, we must now devise a method of reducing it. A simple solution is to tie a capacitor directly from the control line to the ground. Illustrated in Fig. 7.45, the idea is to force the initial charge pump current to flow through C_2, thereby lowering the jump in V_{cont}. In the up/down

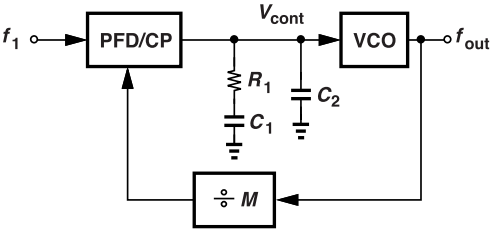


Figure 7.45 Addition of a second capacitor to loop filter.

skew scenario of Fig. 7.43, the peak-to-peak ripple decreases to $(I_p/C_2)T_{sk}$. For example, if $I_p = 100\ \mu$A, $C_2 \approx 0.2C_1 = 1$ pF, and $T_{sk} = 10$ ps, then the ripple amplitude resulting from the skew is about 1 mV.

The addition of C_2 to the filter in Fig. 7.45 raises the loop's order to 3, thereby degrading the phase margin. Nonetheless, a C_2 as large as $0.2C_1$ negligibly affects the loop settling time.

Example 7.25

Estimate the phase margin of the PLL shown in Fig. 7.45.

Solution

Since the loop filter impedance is now given by $[R_1 + 1/(C_1 s)]||[1/(C_2 s)]$, the loop transmission assumes the form:

$$T(s) = \frac{I_p}{2\pi}\frac{R_1C_1s + 1}{R_1C_{eq}s + 1}\frac{1}{(C_1 + C_2)s}\frac{K_{VCO}}{Ms}, \tag{7.45}$$

where $C_{eq} = C_1C_2/(C_1 + C_2)$. We note that $T(s)$ exhibits a third pole, $|\omega_{p3}| = (R_1C_{eq})^{-1}$, located *higher* than the zero (Fig. 7.46).

How should we choose ω_{p3} with respect to ω_u? With $\omega_{p3} = \omega_u$, the phase margin is about 45° (why?). Thus, we generally place ω_{p3} well above ω_u so as to negligibly degrade the phase margin. In such a case, ω_u itself is also negligibly affected by ω_{p3}. We can therefore utilize the ω_u expression in Eq. (7.35), with ζ still defined by (7.25). We have

$$\text{PM} = \tan^{-1}\frac{\omega_u}{\omega_z} - \tan^{-1}\frac{\omega_u}{\omega_{p3}} \tag{7.46}$$

$$= \tan^{-1}(R_1C_1\omega_u) - \tan^{-1}(R_1C_{eq}\omega_u), \tag{7.47}$$

noting that the phase margin degrades by an amount equal to $\tan^{-1}(\omega_u/\omega_{p3})$. For example, $\omega_u/\omega_{p3} = 1/3$ degrades the PM by 18°.

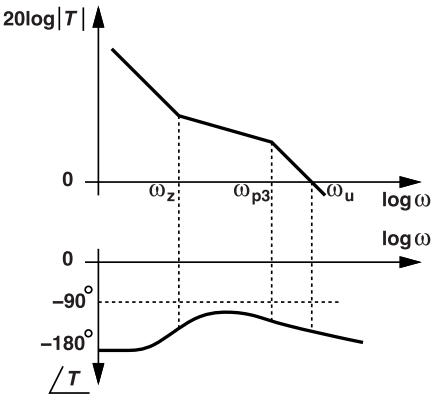


Figure 7.46 Bode plots of third-order PLL.

Let us write $\tan^{-1}(R_1 C_{eq}\omega_u) = \tan^{-1}[R_1 C_1 \omega_u C_2/(C_1 + C_2)] = \tan^{-1}[(\omega_u/\omega_z)C_2/(C_1 + C_2)]$, and select $\omega_u/\omega_z \approx 2.2$ for $\mathrm{PM} = 65°$ without C_2 (Section 7.5.5). Now, if we add $C_2 \approx 0.2C_1$, the phase margin degrades by $\tan^{-1}(2.2/6) = 20°$. In practice, this degradation negligibly alters the settling behavior of the loop.

Example 7.26

An engineer constructs a PLL with $\zeta = 1$ and $C_2 = 0.2C_1$. Does the third pole, ω_{p3}, lie above ω_u in this design?

Solution

In Chapter 8, we show that $\omega_u \approx (R_1 I_p K_{VCO})/(2\pi M)$ if $\omega_{p3} > \omega_u$. Let us proceed with the assumption $\omega_{p3} > \omega_u$ and see whether we obtain a consistent result. Since

$$\zeta^2 = 1 \tag{7.48}$$

$$= \frac{R_1^2}{4}\frac{I_p K_{VCO} C_1}{2\pi M} \tag{7.49}$$

$$= \frac{R_1 C_1 \omega_u}{4}, \tag{7.50}$$

we have $\omega_u = 4/(R_1 C_1)$. On the other hand, if $C_2 \ll C_1$, then $\omega_{p3} \approx 1/(R_1 C_2) \approx 1/(0.2 R_1 C_1)$. It follows that

$$\omega_u \approx 0.8\omega_{p3}. \tag{7.51}$$

Thus, ω_{p3} is slightly higher than ω_u.

Another second-order filter topology is shown in Fig. 7.47. In this case, the voltage across R_1 is allowed to jump, but the subsequent RC stage attenuates the ripple. The filters in Figs. 7.45 and 7.47 provide about the same ripple for a given PM degradation.

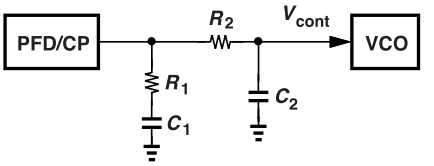


Figure 7.47 Addition of a series resistor in loop filter.

Example 7.27

In Example 7.20, we observed that the initial jump on V_{cont} stabilizes the loop. Examine the situation in the presence of C_2 in Fig. 7.45.

Solution

As shown in Fig. 7.48, suppose S_1 turns on briefly, and the CP current flows into the loop filter. Since $C_1 \gg C_2$, the voltage on C_1 changes negligibly during this time. After S_1 turns off, C_2 shares its charge with C_1, and V_{cont} falls. The pulse thus created in V_{cont} stabilizes the loop.

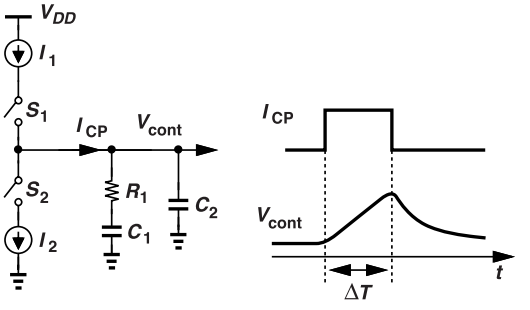


Figure 7.48 Effect of second capacitor on control voltage ripple.

7.7 Basic Charge Pump Topologies

The charge pump circuit introduced in Fig. 7.29(b) is called a "drain-switched" topology because the switches are placed in series with the drains of the current sources. Moving the switches to the source or gate terminals leads to two other configurations.

Shown in Fig. 7.49(a) is a "source-switched" charge pump. Here, M_2 and M_3 operate as current sources, and M_1 and M_4 as switches. Note that the current sources are degenerated by the on-resistance of the switches. For proper definition of the current, V_{b1} and V_{b2} must be established by a current mirror arrangement, as depicted in Fig. 7.49(b). In this arrangement, M_{11} and M_{44} emulate the degenerating behavior of M_1 and M_4, respectively, when the latter are on. For example, M_{11}, M_{22}, M_{33}, and M_{44} can be one-fifth as wide as their corresponding transistors in the main branch (and have the same lengths) so as to provide a scaling factor of 5.

One advantage of the source-switched topology stems from the degeneration provided by the switches. That is, for a given voltage headroom, this arrangement exhibits a higher output resistance than the drain-switched CP does.

Another advantage relates to the clock feedthrough of the switches. First, we consider the drain-switched section shown in Fig. 7.49(c), noting that when Down goes from zero to V_{DD}, C_{GD3} conducts this edge to

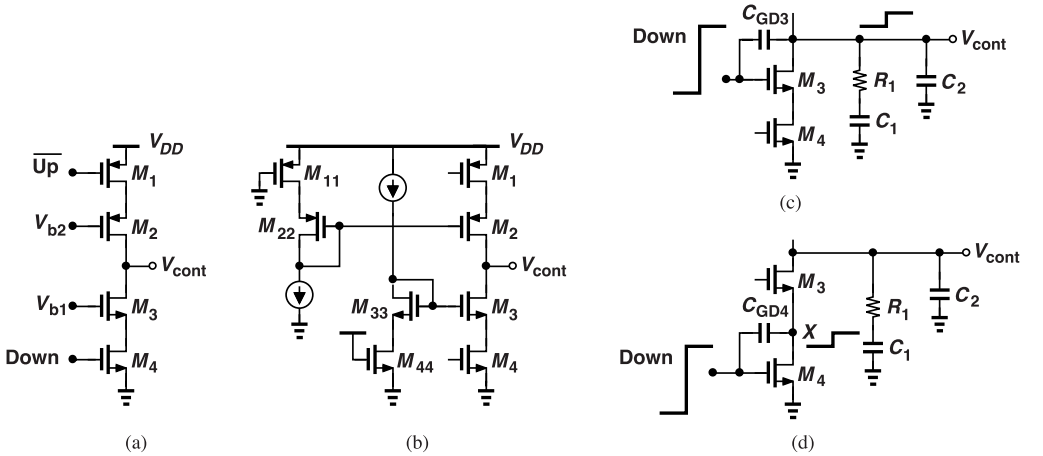


Figure 7.49 (a) Source-switched CP, (b) bias implementation, (c) clock feedthrough in drain-switched CP, and (d) clock feedthrough in source-switched CP.

the output, producing a jump on V_{cont} equal to $V_{DD}C_{GD3}/(C_{GD3}+C_2)$. Second, we examine the source-switched counterpart depicted in Fig. 7.49(d), recognizing that the jump at X can only cause a step in I_{D3} and hence no instantaneous jump in V_{cont}. In other words, M_3 shields, to some extent, the filter from the switch clock feedthrough. The reader is encouraged to repeat this analysis for the falling edge of Up. We examine this effect more closely in Chapter 8.

The drain-switched and source-switched charge pumps suffer from some voltage headroom loss due to the on-resistance of the switches. It is possible to avoid this issue through the use of "gate switching." Illustrated in Fig. 7.50(a), such a topology controls the current sources by connecting their gates to a bias voltage or to

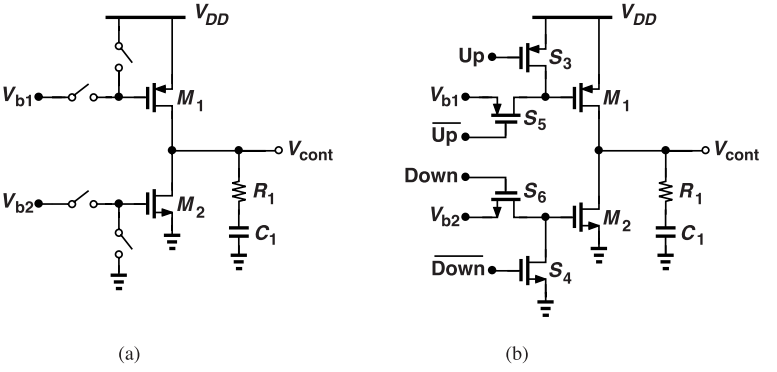


Figure 7.50 (a) Gate-switched CP, and (b) transistor-level realization.

their source terminal. The output can therefore accommodate a voltage range equal to $V_{DD}-|V_{DS1,min}|-V_{DS2,min}=V_{DD}-|V_{GS1}-V_{TH1}|-(V_{GS2}-V_{TH2})$. Also, this structure entails no charge sharing (Chapter 8).

The principal drawback of the gate-switched CP of Fig. 7.50(a) is the skew between the up and down paths. To understand this point, let us implement the switches as shown in Fig. 7.50(b) and observe that M_1 turns on when $\overline{\text{Up}}$ falls to $V_{DD}-|V_{TH1}|-|V_{TH5}|$ whereas M_2 turns on when Down rises to $V_{TH2}+V_{TH6}$. Consequently, from the fast-slow corner (low NMOS threshold, high PMOS threshold) to the slow-fast corner (high NMOS threshold, low PMOS threshold), the turn-on times of M_1 and M_2 experience significant mismatch. Similarly, the turn-off times are skewed as well. As explained in Chapter 8, these skews translate to ripple on the control voltage. We study other charge pump topologies in Chapter 8.

7.8 Settling Time

Many applications require fast-locking PLLs. For example, if the clock frequency of a microprocessor must change to a higher value so as to serve a more demanding application, the PLL must quickly settle to the new frequency. Another reason relates to the *test* cost of chips: a longer test time in the production line translates to a higher cost per chip, demanding that the PLL lock fast enough.

From the closed-loop transfer function given by (7.23), we can estimate the settling time of the PLL in response to a step. Let us approximate the closed-loop behavior by a one-pole response having the same -3-dB bandwidth. As explained in Chapter 8, this bandwidth must be chosen less than $\omega_{in}/10$. Thus, the closed-loop time constant, τ, is at least equal to $10/\omega_{in} = 10T_{in}/(2\pi)$. Allowing four or five time constants for settling, we estimate a settling time on the order of $6T_{in}$ to $8T_{in}$. In practice, however, the choice $\omega_{-3dB} \approx \omega_{in}/10$ leads to excessive ripple in V_{cont}; for the PLL output sidebands to be acceptably small, the bandwidth is less and the settling time reaches about $100T_{in}$.

References

[1]. B. Razavi, *Design of Analog CMOS Integrated Circuits,* Boston: MgGraw-Hill, Second Edition, 2017.

Problems

7.1. Consider the mixer described in Example 7.1. Assume $V_1(t)$ and $V_2(t)$ are square waves toggling between 0 and 1. Plot the output and explain whether the circuit can operate as a phase detector.

7.2. Repeat the previous problem if $V_1(t)$ and $V_2(t)$ toggle between -1 and $+1$.

7.3. If we double R_1 or C_1 in Fig. 7.7(a), does anything change in Figs. 7.7(b) and (c)?

7.4. If we double the PD gain, K_{PD}, in Fig. 7.7(a), how do Figs. 7.7(b) and (c) change?

7.5. If we double the varactor values in Fig. 7.7(a), how do Figs. 7.7(b) and (c) change? Assume ω_1 is relatively constant.

7.6. In Fig. 7.8, the ripple waveform results from applying V_{PD} to a first-order RC filter. Draw carefully the ripple if R_1 in Fig. 7.7(a) is doubled. Is the peak value of the ripple halved? Do not assume τ is much less than the input period.

7.7. Suppose the XOR gate in Fig. 7.7(a) is replaced with the mixer in Example 7.1 (and the mixer inputs are approximated by sinusoids). Construct the plots in Figs. 7.7(b) and (c) for this case and calculate the phase error.

7.8. From Eq. (7.5), plot ΔV versus the static phase offset.

7.9. For the frequency-multiplying PLL in Fig. 7.9(a), construct the plots shown in Figs. 7.7(b) and (c) and calculate the static phase offset. Note that a frequency divider also divides the phase. That is, the VCO output phase is divided by M as it reaches the PD.

7.10. Using the analysis depicted in Fig. 7.13, explain why the phase offset in Fig. 7.7 is not zero.

7.11. Repeat the previous problem if $\omega_1 = 0$ in Fig. 7.7(b).

7.12. The varactor values in Fig. 7.7(a) are doubled. Explain what happens to ω_n and ζ.

7.13. As the supply voltage of an XOR gate fluctuates, so does the circuit's output voltage swing. Explain how supply noise affects the waveforms in Fig. 7.2(a). If the supply changes by ΔV, by how much does th eaverage value of V_{out} change?

7.14. From the previous problem and following the procedure outlined in Example 7.11, explain qualitatively what type of transfer function (low-pass, band-pass, or high-pass) we should expect from the PD supply noise to ϕ_{out} in Fig. 7.7(a).

7.15. Repeat Example 7.13 for the noise of the resistors in the low-pass filter.

7.16. Suppose amplifier A_1 in Fig. 7.38 suffers from supply noise. Assume this noise simply adds a component of the form $V_{n,VDD}$ to the amplifier's output. Do we expect a band-pass transfer function from $V_{n,VDD}$ to ϕ_{out}?

7.17. As mentioned in Section 7.5.4, a feedback loop containing two integrators can be stabilized by making one of the integrators lossy. Suppose the PFD/CP/capacitor cascade in Fig. 7.31 is modified in this manner, its transfer function assuming the form $\alpha/(s+\beta)$. Determine the PLL closed-loop transfer function and find ω_n and ζ.

7.18. Can the lossy integrator in the previous problem be represented by a feedforward model similar to that in Fig. 7.40(c)?

7.19. Prove that, for $4\zeta^2 \gg 1$, the phase margin of a type-II PLL can be expressed as

$$\text{PM} \approx \tan^{-1} 4\zeta^2. \tag{7.52}$$

Note that this is a good approximation for ζ values as low as 0.8.

7.20. From the previous problem, show that

$$\text{PM} \approx \tan^{-1} \frac{R_1^2 I_p C_1 K_{VCO}}{2\pi} \tag{7.53}$$

$$\approx \tan^{-1}(R_1 C_1 \omega_u). \tag{7.54}$$

7.21. A third-order PLL employs loop filter values $R_1 = 10\,\text{k}\Omega$, $C_1 = 20$ pF, and $C_2 = 4$ pF. What happens to the phase margin if R_1 is doubled? Assume ω_u is (a) unchanged, or (b) is halved.

8

PLL Design Considerations

The charge-pump phase-locked loop introduced in Chapter 7 must be designed according to given specifications while dealing with a number of challenges that stem from its imperfections. All of the building blocks within the loop, as well as the input, degrade the performance, but to varying degrees. This chapter first delves into PLL transfer functions, describes these issues, and finally presents circuit techniques to overcome them. An important aspect of our analysis is to determine which nonidealities become critical in a given application.

8.1 More on PLL Transfer Functions

Our study of simple charge-pump PLLs has led to various closed-loop transfer functions from different points to the output. In summary, we have found that (a) the input phase fluctuations experience a second-order low-pass transfer with one zero, and (b) noise voltages in series with the control or supply of the VCO see a band-pass response. In this section, we further explore the PLL transfer functions.

The input/output transfer function expressed by (7.23) can be rewritten as

$$H(s) = \frac{2\zeta\omega_n s + \omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2}. \tag{8.1}$$

The zero is located at $\omega_z = -\omega_n/(2\zeta)$ and the poles at

$$\omega_{p1,2} = (-\zeta \pm \sqrt{\zeta^2 - 1})\omega_n. \tag{8.2}$$

For $\zeta = 1$, we have $\omega_z = -\omega_n/2$ and $\omega_{p1} = \omega_{p2} = -\omega_n$. If, on the other hand, $\zeta^2 \gg 1$, then $\sqrt{\zeta^2 - 1} \approx \zeta[1 - 1/(2\zeta^2)] = \zeta - 1/(2\zeta)$, yielding

$$\omega_{p1} \approx -\frac{\omega_n}{2\zeta} = -\frac{1}{R_1 C_1} \tag{8.3}$$

$$\omega_{p2} \approx -2\zeta\omega_n = -\frac{R_1 I_p K_{VCO}}{2\pi}. \tag{8.4}$$

For a feedback divide ratio of M, we must replace K_{VCO} with K_{VCO}/M in this result. We make two observations: (1) ω_{p1} is now in the vicinity of ω_z, partially cancelling its effect, and (2) $\omega_{p2}/\omega_{p1} = 4\zeta^2 \gg 1$, implying that the transfer function is roughly that of a one-pole system with a corner frequency equal to $\omega_{p2} = -(R_1 I_p K_{VCO})/(2\pi)$. Figure 8.1 plots the magnitude of $H(s)$ for three values of ζ. Note that $|H| > 1$ for some range of jitter frequencies; we say the loop exhibits "jitter peaking," i.e., it amplifies the input phase noise. We calculate the amount of peaking in Chapter 14. The one-pole approximation proves useful for bandwidth calculations, but jitter peaking is always present.

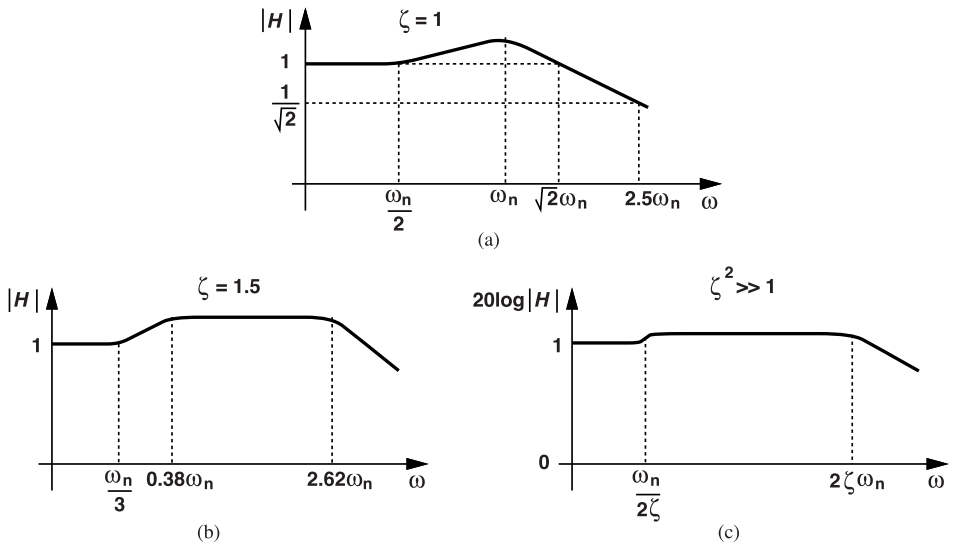


Figure 8.1 Closed-loop PLL response for different ζ values.

The -3-dB bandwidth of the foregoing responses is also of interest. For example, in some cases, we wish to determine how much the input phase noise is filtered by the PLL. Setting $|H(s = j\omega)|$ in Eq. (8.1) to $1/\sqrt{2}$, we obtain

$$\omega_{-3dB}^2 = [1 + 2\zeta^2 + \sqrt{(1 + 2\zeta^2)^2 + 1}]\omega_n^2. \tag{8.5}$$

It follows that

$$\omega_{-3dB} \approx 2.5\omega_n \quad \text{for} \quad \zeta = 1 \tag{8.6}$$

$$\approx 3.3\omega_n \quad \text{for} \quad \zeta = 1.5 \tag{8.7}$$

$$\approx 2\zeta\omega_n \quad \text{for} \quad \zeta^2 \gg 1. \tag{8.8}$$

The last case agrees with the one-pole approximation illustrated in Fig. 8.1(c). As ζ goes from $\sqrt{2}/2$ to large values, ω_{-3dB} varies from $3\zeta\omega_n$ to $2\zeta\omega_n$, i.e., the settling time increases proportionally.

In addition to the closed-loop bandwidth, we are also interested in the open-loop unity-gain bandwidth, ω_u. From (7.35), we have

$$\omega_u^2 = (2\zeta^2 + \sqrt{4\zeta^4 + 1})\omega_n^2 \tag{8.9}$$

and hence

$$\omega_u \approx 2.1\omega_n \quad \text{for} \quad \zeta = 1 \tag{8.10}$$

$$\approx 3\omega_n \quad \text{for} \quad \zeta = 1.5 \tag{8.11}$$

$$\approx 2\zeta\omega_n \quad \text{for} \quad 4\zeta^4 \gg 1. \tag{8.12}$$

These results agree with Example 7.22. We observe that ω_{-3dB} and ω_u are fairly close for $\zeta \geq 1$. For this reason, the term "loop bandwidth" is applied to either one.

Example 8.1

Explain why ω_u and ω_{-3dB} become independent of C_1 for large values of ζ.

Solution

Recall that the open-loop transfer function is given by $[I_p/(2\pi)][R_1 + 1/(C_1 s)](K_{VCO}/s)$. For frequencies well below or well above the zero frequency, this expression reduces to $[I_p/(2\pi)][1/(C_1 s)](K_{VCO}/s)$ and $[I_p/(2\pi)]R_1(K_{VCO}/s)$, respectively (Fig. 8.2). If ζ is large enough and ω_u satisfies the latter case, we can

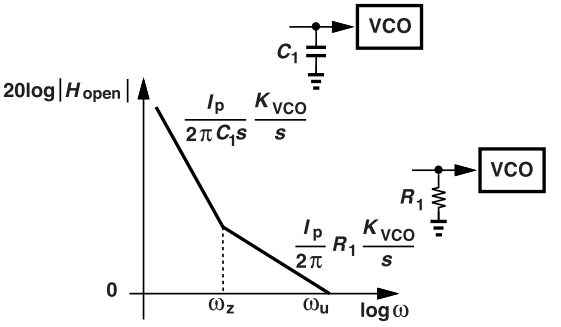


Figure 8.2 Open-loop behavior at different frequencies.

say that C_1 acts as a short circuit at ω_u, no more playing a significant role in the loop bandwidth. We also observe that equating the magnitude of $[I_p/(2\pi)]R_1(K_{VCO}/s)$ to unity yields $\omega_u = I_p R_1 K_{VCO}/(2\pi)$.

The closed-loop bandwidth, too, becomes independent of C_1 because the pole-zero cancellation yields $\omega_{-3dB} \approx \omega_{p2} \approx 2\zeta\omega_n$.

In the study of phase noise in PLLs later in this chapter, we encounter a third quantity that is also loosely called the loop bandwidth. Suppose we model the phase noise produced by the VCO as an additive phase quantity at its output [Fig. 8.3(a)]. We wish to compute $\phi_{out}/\phi_{n,VCO}$ while the input has no phase noise.

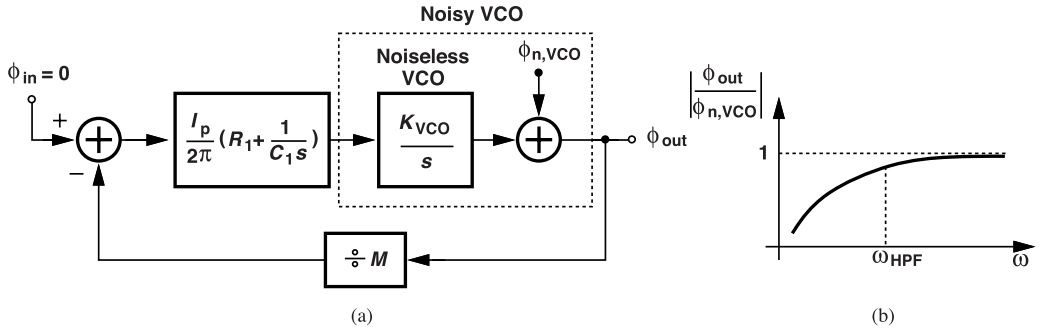


Figure 8.3 (a) Model of noisy VCO in a CPPLL, and (b) the corresponding response.

With $\phi_{in} = 0$, the PFD/CP cascade delivers a current equal to $(0 - \phi_{out}/M)[I_p/(2\pi)]$ to the filter, generating a control voltage given by $(-\phi_{out}/M)[I_p/(2\pi)][R_1 + 1/(C_1 s)]$. The VCO multiplies the result by K_{VCO}/s, yielding

$$-\frac{\phi_{out}}{M}\frac{I_p}{2\pi}\left(R_1 + \frac{1}{C_1 s}\right)\frac{K_{VCO}}{s} + \phi_{n,VCO} = \phi_{out}. \tag{8.13}$$

That is,

$$\frac{\phi_{out}}{\phi_{n,VCO}}(s) = \frac{s^2}{s^2 + 2\zeta\omega_n s + \omega_n^2}. \tag{8.14}$$

This high-pass response represents the transfer function from the VCO phase noise to the output [Fig. 8.3(b)] and agrees with our intuitive study in Example 7.14. In order to minimize the VCO phase noise contribution, we must maximize the loop bandwidth, ω_{HP}.

Example 8.2

Explain intuitively why $\phi_{out}/\phi_{n,VCO}$ has a high-pass behavior.

Solution

Let us redraw the system of Fig. 8.3(a) as shown in Fig. 8.4, where

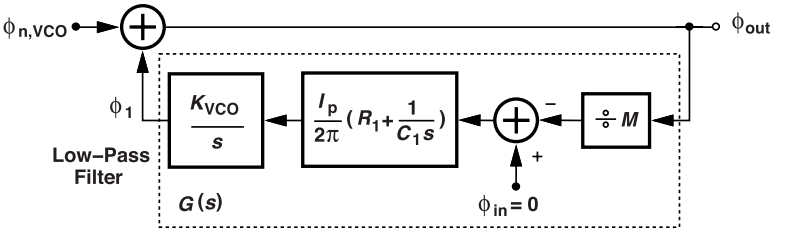


Figure 8.4 Alternative view of CPPLL.

$$G(s) = \frac{1}{M}\frac{I_p}{2\pi}\left(R_1 + \frac{1}{C_1 s}\right)\frac{K_{VCO}}{s} \tag{8.15}$$

denotes the negative of the loop transmission. By virtue of its high loop gain at low frequencies, the system forces ϕ_1 toward $\phi_{n,VCO}$ and hence ϕ_{out} toward zero. That is, slow VCO phase fluctuations are suppressed by the strong feedback present in the loop. For fast phase variations, on the other hand, the loop gain drops and $\phi_{out} \approx \phi_{n,VCO}$. Thus, the VCO phase noise experiences a high-pass response.

To compute the corner frequency of the high-pass response in Fig. 8.3(b), we equate the magnitude of (8.14) to $1/\sqrt{2}$, obtaining

$$\omega_{HP}^2 = [2\zeta^2 - 1 + \sqrt{(2\zeta^2 - 1)^2 + 1}]\omega_n^2. \tag{8.16}$$

The VCO phase noise is suppressed up to $\omega = \omega_{HP}$. We then have

$$\omega_{HP} \approx 1.55\omega_n \quad \text{for} \quad \zeta = 1 \tag{8.17}$$

$$\approx 2.7\omega_n \quad \text{for} \quad \zeta = 1.5 \tag{8.18}$$

$$\approx 2\zeta\omega_n \quad \text{for} \quad 2\zeta^2 \gg 1. \tag{8.19}$$

For low ζ values, the suppression bandwidth is less than the other bandwidths computed earlier.

Figure 8.5 summarizes the loop's behavior for $\zeta = 1$. The open-loop response changes its slope from -40 dB/dec to -20 dB/dec at $\omega_z = 1/(R_1C_1) = \omega_n/(2\zeta) = \omega_n/2$, and the closed-loop response begins to rise at this frequency. The latter reaches its maximum at $\omega_{p1} = \omega_{p2} = \omega_n$ and its -3-dB point at $2.5\omega_n$. The VCO noise transfer function, $|\phi_{out}/\phi_{n,VCO}|$, rises at 40 dB/dec, approaching unity for $\omega > 1.55\omega_n$.

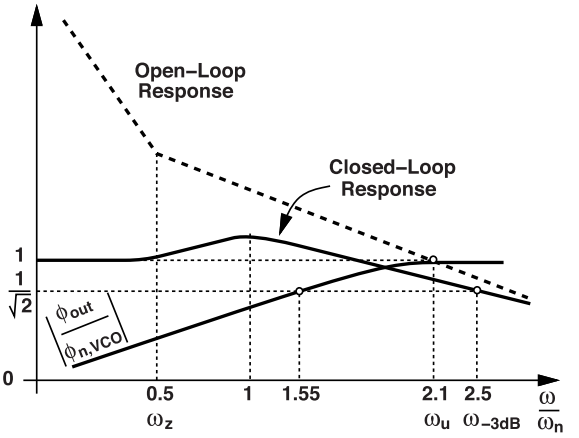


Figure 8.5 Summary of various responses in a CPPLL with $\zeta = 1$.

Example 8.3

Explain why the noise voltage in Example 7.18 sees a band-pass response while the phase noise in Fig. 8.3(a) experiences a high-pass response.

Solution

The output phase of the VCO due to V_n in Fig. 7.38(a) is given by $\phi_{n,VCO} = V_n A_1 K_{VCO}/s$. The two responses in Figs. 8.3(a) and 7.38 differ by a factor of $A_1 K_{VCO}/s$:

$$\frac{\phi_{out}}{\phi_{n,VCO}} = \frac{\phi_{out}}{V_n/(K_{VCO}/s)}. \tag{8.20}$$

If divided by K_{VCO}/s, the band-pass response associated with V_n changes to a high-pass response (why?).

8.1.1 Limitations of Continuous-Time Approximation

Recall from Fig. 7.30 that, to arrive at a linear system, we approximated the step response of the PFD/CP/capacitor cascade by a ramp. This approximation, in fact, also ignores another property of the CP-PLL, namely, its discrete-time (DT) nature. When the loop is locked, the CP turns on briefly at each phase comparison instant, injects some charge onto the capacitor, and then turns off. As a result, the feedback loop remains broken for most of the input period, acting as a discrete-time system. The ramp in Fig. 7.32, on the other hand, assumes that V_{cont} continuously tracks the VCO (and input) phase fluctuations.

What are the consequences of approximating a DT loop by a continuous-time (CT) system? Since the former provides feedback during only a fraction of the input period, we intuitively predict that the latter overestimates the "average" loop gain. In the case of a CPPLL, the CT approximation translates to an optimistic view of stability because it is roughly equivalent to assuming a higher $I_p K_{VCO}$.

In order for Eqs. (7.23)-(7.25) to provide a reasonably accurate estimate of the PLL response, the states within the loop (the VCO phase and frequency) must change *slowly* with time. That is, if the change in these states from one input cycle to the next is small enough, then they can be viewed as varying continuously. To understand this point, let us return to the ramp approximation of the control voltage and consider two cases, assuming that V_{cont} is to reach a certain level, V_1. First, we suppose a certain loop filter capacitance, C_1, and, as shown in Fig. 8.6(a), note that the error resulting from the CT approximation has a peak value V_a.

Second, while keeping the phase error constant, we double this capacitance to slow down the loop dynamics and observe that the peak error, V_b, is halved. In other words, the slower ramp provides a more accurate representation for reaching V_1.

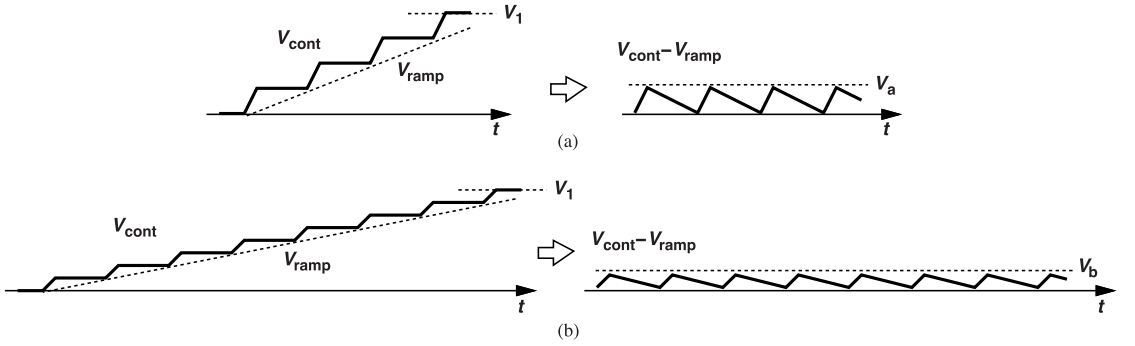


Figure 8.6 (a) Fast, and (b) slow changes in V_{cont} yield different amounts of error in the continuous-time approximation.

A rule of thumb often used to ensure slow changes in the loop is to select the loop bandwidth approximately equal to one-tenth of the input frequency. Indeed, in Example 7.17, we have $2\zeta\omega_n \approx \omega_{in}/12.8$. We utilize this rule in PLL design in Chapter 9.

It is possible to analyze the CPPLL as a linear time-variant system and develop a more accurate view of its stability behavior [1]. This becomes necessary if an aggressive design targets a wide loop bandwidth.

8.2 PFD Issues

The PFD entails three phenomena that affect the performance: random skew between the up and down pulses, random mismatches between the widths of these pulses, and phase noise. The first two arise from random propagation delay mismatches and can be minimized through the use of large transistors. Layout symmetry is, of course, critical as well. In general, these two effects negligibly affect the performance.

The gates within a PFD add phase noise as the edges propagate through them. Consequently, the widths of the up and down pulses are modulated randomly, turning on the two charge pump current sources for different amounts of time. The loop filter therefore receives randomly-modulated up and down charge packets and converts them to random noise in V_{cont}. These effects are studied and formulated in [2].

The PFD phase noise directly rises with the input frequency [2]. From the results in [2], we can estimate a phase noise of around -175 dBc/Hz for an input frequency of 20 MHz. This phase noise rises by 20log of the divide ratio as it appears at the PLL output, manifesting itself in low-noise designs. As a rule of thumb, we rank the PLL building blocks in terms of their phase noise contribution as: the VCO, the CP, the PFD, and the divider.

8.3 Charge Pump Issues

Charge pumps suffer from numerous undesirable effects, contributing both ripple and noise. In this section, we analyze these imperfections.

8.3.1 Up and Down Skew

Recall our study in Chapter 7 of the deterministic skew between the up and down pulses as they arrive at the charge pump. Figure 8.7(a) depicts a first-order correction of this skew, where complementary transistors M_5 and M_6 replicate the inverter delay. Even in this case, the current pulses delivered by the charge pump may not coincide in time [Fig. 8.7(b)]. This is because it is difficult to align the switching times of M_3 and M_4, especially across process and temperature corners. In other words, our ultimate objective is to minimize the

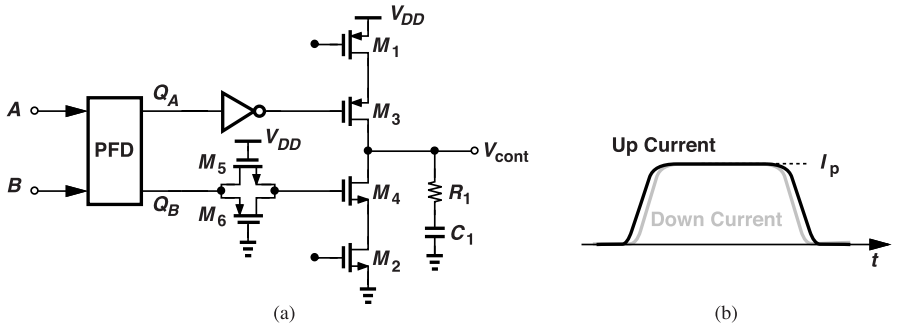


Figure 8.7 Correction of up/down skew by means of a pass gate.

skew between these two currents rather than between the up and down pulses. As explained in Chapter 7, this phenomenon leads to ripple in V_{cont}.

8.3.2 Voltage Compliance and Channel-Length Modulation

The greatest challenge in charge pump design relates to the output voltage "compliance" and channel-length modulation in the current sources. These two effects trade with each other. The former signifies the output voltage range that the charge pump can support while maintaining equal up and down currents. As illustrated in Fig. 8.8(a), we wish to maximize this compliance, $V_{max} - V_{min}$, so as to obtain the widest VCO tuning

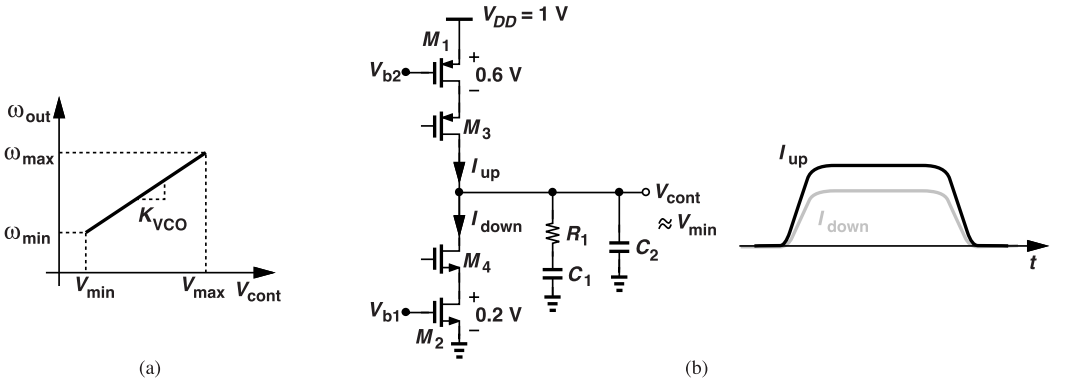


Figure 8.8 (a) VCO characteristic, and (b) mismatch between up and down currents.

range for a given K_{VCO}. Now, we observe that, if $V_{cont} = V_{min}$, the NMOS current source sustains a small V_{DS}, producing a current less than the nominal value, while the PMOS counterpart experiences a large $|V_{DS}|$, generating a larger current [Fig. 8.8(b)]. The reverse occurs if $V_{cont} \approx V_{max}$. The mismatch between the two currents flows through the loop filter, thereby creating ripple.

Example 8.4

How does the PLL behave in the presence of unequal up and down currents in Fig. 8.8(b)?

Solution

Let us begin with the scenario shown in Fig. 8.9(a), where $I_{\rm Up} - I_{\rm Down}$ flows through the loop filter for $T_{res} \approx 5$ gate delays in every input period. The cumulative effect implies that $V_{cont} \rightarrow +\infty$, which cannot happen because the PLL attempts to keep the output frequency constant. Consequently, as illustrated in Fig. 8.9(b), the loop settles with a static phase error, ΔT, such that the smaller current lasts longer. In this example,

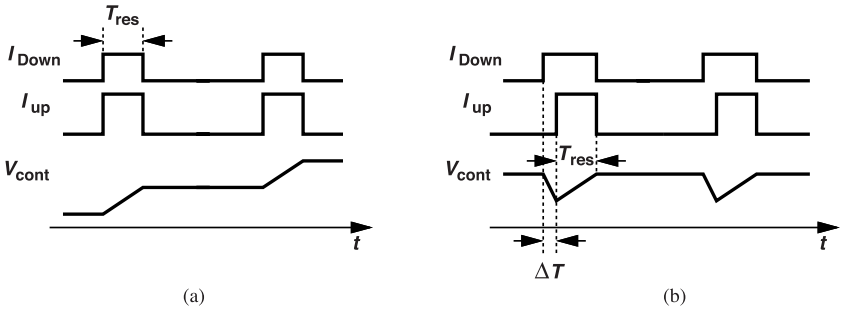


Figure 8.9 (a) Initial guess for lock condition in the presence of CP mismatches, and (b) actual lock situation.

V_{cont} falls at a rate of $I_{\rm Down}/C_2$ for ΔT seconds and rises at a rate of $(I_{\rm Up} - I_{\rm Down})/C_2$ for T_{res} seconds. In the steady state, the fall and the rise must cancel, and hence

$$\Delta T = \frac{I_{\rm Up} - I_{\rm Down}}{I_{\rm Down}} T_{res} \tag{8.21}$$

$$\approx \frac{\Delta I}{I_p} T_{res}, \tag{8.22}$$

where $\Delta I = I_{\rm Up} - I_{\rm Down}$ and I_p is the nominal charge pump current. The peak-to-peak ripple amplitude is equal to $\Delta T I_{\rm Down}/C_2 \approx \Delta I \cdot T_{res}/C_2$. Both the phase error and the ripple are undesirable effects, necessitating that T_{res} be minimized. For example, a T_{res} of about 50 ps and a $\Delta I/I_p$ of 10% translate to $\Delta T \approx 5$ ps.

The effect of CP channel-length modulation on the performance can be measured as follows. We assert both the up and down inputs as shown in Fig. 8.10(a), apply a voltage source to the output, and sweep its

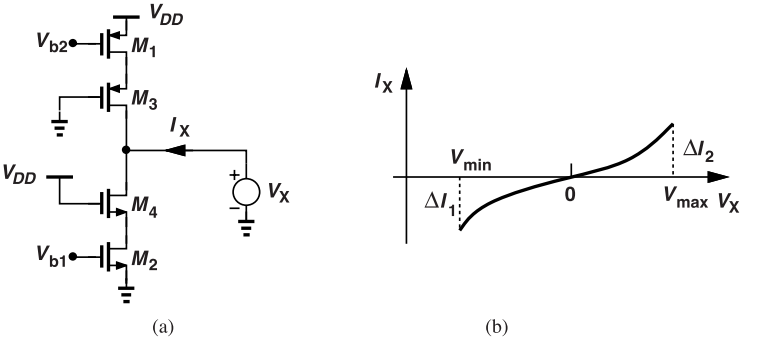


Figure 8.10 (a) Measurement of CP current mismatch, and (b) resulting characteristic.

value across the range of interest, i.e., from V_{min} to V_{max} (the "compliance range"). In the ideal case, V_X need not provide any current, but in reality, I_X is negative for low V_X because $I_{\rm Up} > I_{\rm Down}$, and vise versa [Fig. 8.10(b)]. The greater value of ΔI_1 and ΔI_2 represents the worst-case current mismatch for phase offset and ripple calculations.

In order to deal with channel-length modulation in the charge pump, we can use long (and wide) transistors for the current sources, but the resulting increase in the device capacitances leads to other issues (Section 8.3.4). Other methods of alleviating this problem are described in Section 8.4.

Example 8.5

In order to alleviate channel-length modulation, an integrator can be interposed between the CP and the VCO (Fig. 8.11). Here, R_1 introduces a zero and stabilizes the loop. If the op amp has an open-loop gain of A_1, the

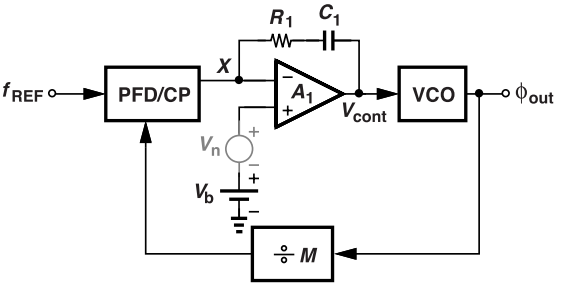


Figure 8.11 CPPLL using an op-amp-based integrator.

CP output voltage range is reduced by the same factor, thereby allowing tighter matching between the up and down currents. That is, V_{cont} can range from V_{min} to V_{max} while V_X changes by only $(V_{max} - V_{min})/A_1$. Explain the issues arising from the integrator.

Solution

The first issue is the noise of the op amp, represented by V_n in Fig. 8.11. Following our analysis methods in Chapter 7, we can qualitatively determine the type of response from V_n to ϕ_{out}. If V_n has a constant value, the loop must keep V_{cont} unchanged, requiring an initial condition on C_1 that cancels the effect of V_n. The PLL creates such a voltage without a phase offset. This can be seen by noting that, in the presence of a constant phase error, the charge pump periodically injects positive (or negative) charge onto C_1, driving V_{cont} toward infinity. We thus conclude that V_n sees a high-pass response. [Since the VCO attenuates the effect of high-frequency noise in V_{cont} (Example 8.3), the transfer function in fact has a band-pass response.]

The other issue relates to the poles contributed by the op amp, which can degrade the PLL phase margin. Also, the op amp must provide an output voltage range from V_{min} to V_{max}. We further study this PLL in Problems 8.9-8.12.

8.3.3 Random Mismatches

The up and down currents also suffer from random mismatches. As an example, let us consider the gate-switched topology of Fig. 7.50 again (Fig. 8.12), observing that M_2 and M_4 exhibit random mismatches, and

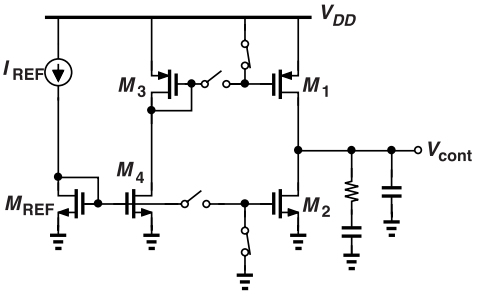


Figure 8.12 Gate-switched CP.

so do M_1 and M_3. The mismatch between $|I_{D1}|$ and I_{D2} thus gives rise to both phase offset and ripple as described in the previous section.

The random mismatch between the up and down currents can be reduced by enlarging the CP transistors, but at the cost of greater capacitances and the difficulties associated with them (Section 8.3.4). In general, the (deterministic) mismatch due to channel-length modulation is more pronounced than random mismatches.

8.3.4 Clock Feedthrough and Charge Injection

The switching action within the charge pump produces ripple at the output even in the absence of the effects described in the previous section. As illustrated in Fig. 8.13, the gate-drain capacitances of M_3 and M_4 respectively couple the transitions at their gates to the output node. Since C_{GD3} and C_{GD4} are generally not equal, a net amount of change appears in V_{cont} given by $V_{DD}(C_{GD3} - C_{GD4})/(C_{GD3} + C_{GD4} + C_2)$.

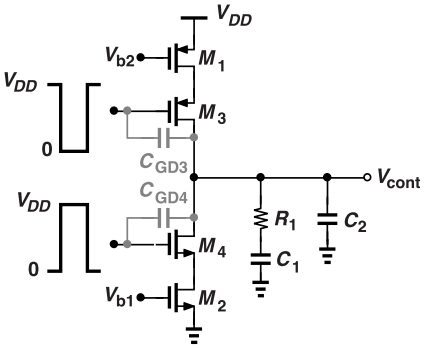


Figure 8.13 Drain-switched CP showing clock feedthrough paths.

In addition, for M_3 and M_4 to turn on or off, they must absorb or release channel charge in the amount of $WLC_{ox}(V_{GS} - V_{TH})$. Since these two devices do not necessarily carry equal charges, a net disturbance arises in V_{cont} on both the rising and the falling edges of the up and down pulses.

The effect of clock feedthrough and charge injection becomes more serious as the CP transistors are enlarged, as required by issues such as channel-length modulation and random mismatches. In such a case, C_2 must be chosen large enough to suppress the ripple, which in turns means that C_1 must also be increased.

Example 8.6

We postulated that the current-source transistors in the source-switched charge pump of Fig. 7.49 shield the output node from the clock feedthrough and charge injection of the switches. Does this mean that no ripple appears in V_{cont}?

Solution

Even though the activities of the switches do not immediately reach the output, their effect still disturbs V_{cont}. To see this point, let us consider the PMOS section, shown in Fig. 8.14(a), when $\overline{\text{Up}}$ goes high. The feedthrough and charge injection of M_1 create an upward jump at X, temporarily raising V_X and $|I_{D3}|$. Now, X discharges slowly through M_3, allowing I_{D3} to charge the loop filter and hence cause ripple [Fig. 8.14(b)].

8.3.5 Other Charge Pump Nonidealities

In this section, we describe two particular phenomena that arise if the charge pump employs short-channel devices. The first relates to the low output impedance of the circuit when both up and down current sources are on [Fig. 8.15(a)], since the loop filter is now shunted by R_{out}, the PFD/CP/filter no longer behaves as an

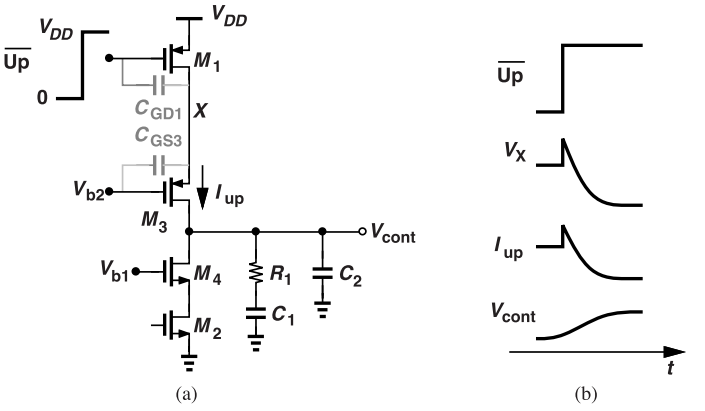


Figure 8.14 Source-switched CP showing clock feedthrough path.

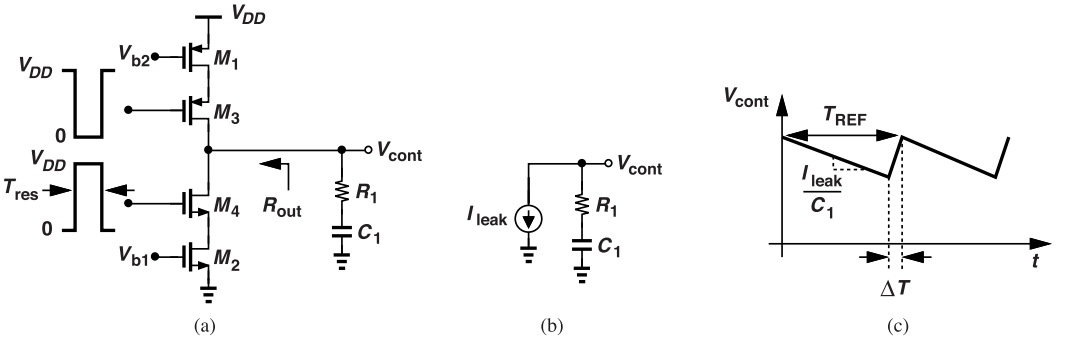


Figure 8.15 (a) CP with finite output resistance, (b) effect of transistors' leakage current, and (c) control voltage waveform due to leakage.

ideal integrator. We may be tempted to write the open-loop PLL transfer function as

$$H(s) = \frac{I_p}{2\pi}\left[R_{out}||\left(R_1 + \frac{1}{C_1 s}\right)\right]\frac{K_{VCO}}{Ms}, \tag{8.23}$$

but R_{out} is present only for T_{res} seconds, where T_{res} denotes the duration of the up and down pulses (about five gate delays). We intuitively see that the effect of R_{out} must be scaled according to T_{REF}/T_{res}:

$$H(s) \approx \frac{I_p}{2\pi}\left[\frac{R_{out}T_{REF}}{T_{res}}||\left(R_1 + \frac{1}{C_1 s}\right)\right]\frac{K_{VCO}}{Ms}. \tag{8.24}$$

Since the equivalent impedance, Z_{eq}, can be expressed as

$$Z_{eq} = \frac{R_{out}T_{REF}}{T_{res}}||\left(R_1 + \frac{1}{C_1 s}\right) \tag{8.25}$$

$$= \frac{(R_1 C_1 s + 1)R_{out}T_{REF}/T_{res}}{(R_1 + R_{out}T_{REF}/T_{res})C_1 s + 1}, \tag{8.26}$$

we recognize that the pole has moved from the origin to approximately $[(R_1 + R_{out}T_{REF}/T_{res})C_1]^{-1}$.

The second imperfection is the leakage of the CP transistors while they are off [Fig. 8.15(b)]. The difference between the up and down leakage currents, I_{leak}, discharges the loop filter for $T_{REF} - T_{res} \approx T_{REF}$

seconds, producing a change of $(I_{leak}/C_1)T_{REF}$ in V_{cont}. In the steady-state operation, this change must be compensated when the charge pump turns on, requiring a phase offset of ΔT [Fig. 8.15(c)] such that $I_p\Delta T/C_1 = (I_{leak}/C_1)T_{REF}$. That is, $\Delta T = (I_{leak}/I_p)T_{REF}$ and the peak-to-peak ripple is equal to $(I_{leak}/C_1)T_{REF}$. Even though I_{leak}/I_p is very small, ΔT can be significant because the leakage lasts for one input period.

The behavior depicted in Figs. 8.15(b) and (c) can be observed in two other cases as well: (1) if the loop filter capacitors have leakage, or (2) if the VCO incorporates large, leaky varactors for frequency control. The former is particularly problematic if the loop filter capacitors are realized by thin-oxide (core) MOSFETs, which exhibit a high gate leakage current. We return to this point in Section 8.7.

8.4 Improved Charge Pumps

A number of charge pump topologies have been introduced to deal with the mismatches between the up and down currents. We describe three here.

Let us return to the gate-switched charge pump shown in Fig. 8.12 and ask how we can suppress the effect of channel-length modulation. We wish to reduce the variation of the up and down currents as V_{cont} approaches 0 or V_{DD}. For example, when V_{cont} is relatively high and I_{D2} greater than $|I_{D1}|$, we can *decrease* the former. This is accomplished as illustrated in Fig. 8.16(a), where M_3 begins to lower the gate voltage of M_2 as V_{cont} rises above V_{TH3}. The same method can be applied to M_1 as well [Fig. 8.16(b)].

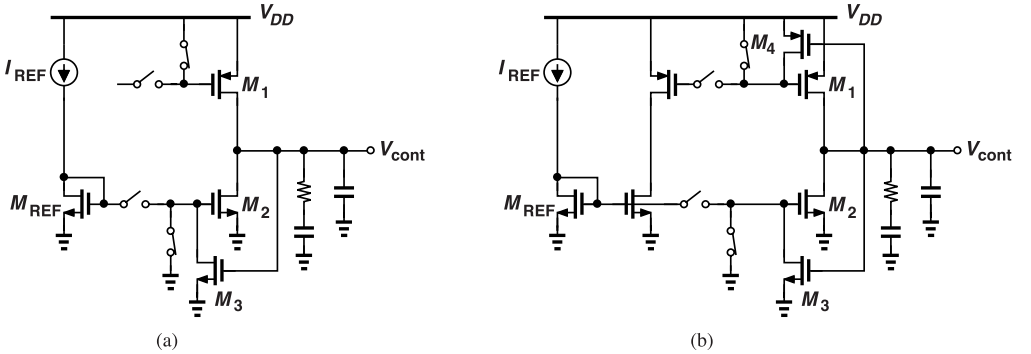


Figure 8.16 Gate-switched CP with feedback (a) around NMOS current source, and (b) around NMOS and PMOS current sources.

Another topology derived from the gate-switched CP is shown in Fig. 8.17 [3]. The circuit incorporates an op amp that senses the difference between V_{cont} and V_X. The op amp adjusts the gate voltage of M_3 and

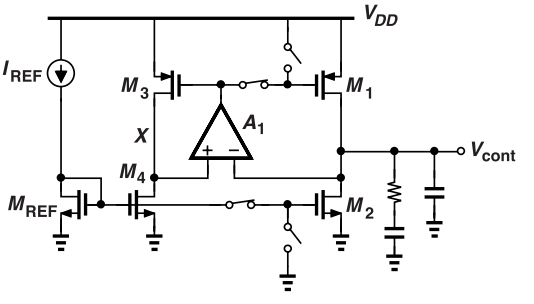


Figure 8.17 Gate-switched CP with a loop suppressing effect of channel-length modulation.

M_1 so as to drive V_X toward V_{cont}. Neglecting random mismatches and assuming for simplicity that M_3 is identical to M_1 and M_4 to M_2, we observe that $I_{D4} = I_{D2}$ because $V_{DS4} = V_{DS2}$. Similarly, $I_{D3} = I_{D1}$.

Also, $|I_{D3}| = I_{D4}$. Thus, $I_{D1} = I_{D2}$ for any value of V_{cont} so long as the loop gain is large enough to force V_X close to V_{cont}. The effect of channel-length modulation is therefore suppressed. In practice, we scale down M_3 and M_4 with respect to the main branch so as to save power.

Example 8.7

The charge pump circuit of Fig. 8.17 contains a negative-feedback loop and a positive-feedback loop. Explain why the circuit operates properly.

Solution

The two feedback loops have equal low-frequency gains if the left branch is simply scaled with respect to the right branch (why?). However, the loop consisting of M_1, M_2, and A_1 turns on only briefly. The positive feedback associated with this loop is thus overwhelmed by the continuous-time negative-feedback loop comprising M_3, M_4, and A_1. Nevertheless, the overall phase margin of the PLL must be evaluated carefully, especially if A_1 employs more than one stage.

As explained in Chapter 7, the gate switching action in Fig. 8.17 causes timing mismatches between the up and down currents. Moreover, the op amp must achieve a wide input common-mode range so as to accommodate the minimum and maximum values of V_{cont}. In Problem 8.14, we investigate the effect of the op amp's noise.

To arrive at a third CP dealing with current mismatches, let us begin with the simple structure in Fig. 8.18 and note that the difference between I_1 and I_2 can be measured when S_1 and S_2 are *off*, for example,

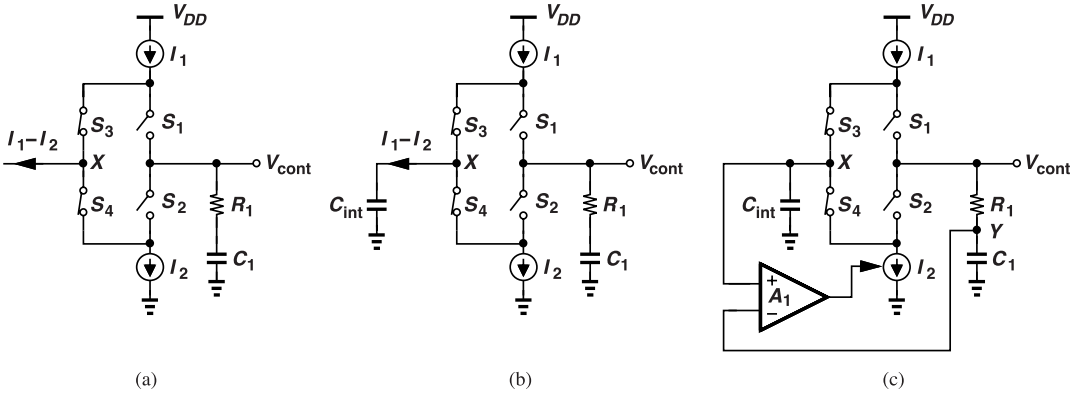


Figure 8.18 (a) Measurement of up and down current mismatch when S_1 and S_2 are off, (b) integration of $I_1 - I_2$ by means of C_{int}, and (c) CP using feedback to eliminate effect of mismatches.

by turning on two other switches, S_3 and S_4, and algebraically adding I_1 and I_2. We must now determine how to *store* this information (as S_3 and S_4 turn off during th next phase comparison) and how to use this information to adjust I_1 and I_2.

It is possible to store the information by injecting $I_1 - I_2$ into a capacitor while S_3 and S_4 are on [Fig. 8.18(b)]. If, for example, $I_1 > I_2$, then $V_X \rightarrow +\infty$, serving as an error that represents both random and deterministic mismatches between I_1 and I_2.

We now utilize V_X to adjust one of the current sources, as shown in Fig. 8.18(c) [4]. Here, the servo loop consisting of A_1 and I_2 forces $V_X - V_Y$ toward zero. This is possible only if $I_1 = I_2$; otherwise, V_X ramps up or down indefinitely. Thus, both random and deterministic mismatches are removed. The circuit senses V_Y, rather than V_{cont}, because the former is quieter.

The CP of Fig. 8.18(c) suppresses the up and down current mismatches, but, like the structure in Fig. 8.17, it also contains both negative- and positive-feedback loops and demands a wide input CM range for the op amp. The effect of the op amp's noise is studied in Problem 8.15.

8.5 PLLs with Discrete VCO Tuning

In our VCO studies in Chapters 3 and 5, we recognized the need for discrete (coarse) VCO tuning. We must now determine how exactly the PLL performs this function. Let us assume the VCO employs switched capacitors for this purpose. Illustrated in Fig. 8.19 is one possible approach. Here, the control voltage is compared to V_{min} and V_{max} by two comparators. The VCO coarse control is provided by a counter, which

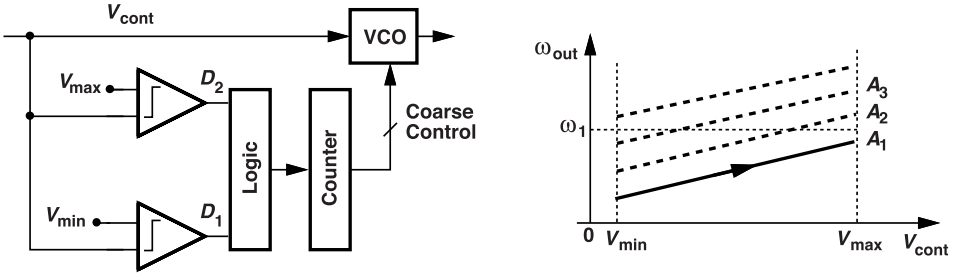


Figure 8.19 Arrangement for discrete VCO tuning.

starts at reset and, for example, switches all of the capacitors into the circuit. Next, the PFD and the charge pump are enabled and the loop attempts to lock, raising V_{cont} along the A_1 tuning curve. If the desired frequency, ω_1, is not reached for $V_{min} < V_{cont} < V_{max}$, the two comparators generate $D_1D_2 = 00$ or $D_1D_2 = 11$, and the counter is incremented by 1 so that one capacitor is switched out of the VCO. Now, the loop attempts to lock again, with the VCO traversing the A_2 curve. This search continues until enough capacitors are switched out and the loop locks with $V_{min} < V_{cont} < V_{max}$. We describe another approach in Chapter 9.

Example 8.8

A student surmises that V_{min} and V_{max} can be chosen close to each other so as to alleviate the CP channel-length modulation issue (Section 8.3.2). Explain whether this is possible.

Solution

For $V_{max} - V_{min}$ to be small, say, around 100 mV, the discrete steps must be sufficiently close to each other to have some overlap and avoid a blind zone. This means that the number of discrete tuning curves must be proportionally larger, leading to greater design complexity.

However, a more serious difficulty is the VCO frequency drift with temperature, e.g., as a user leaves home and steps into a very cold day outside. The drift manifests itself as a vertical shift of the continuous tuning curve (Fig. 8.20). If the new curve does not provide the desired frequency, ω_1, for $V_{min} < V_{cont} < V_{max}$,

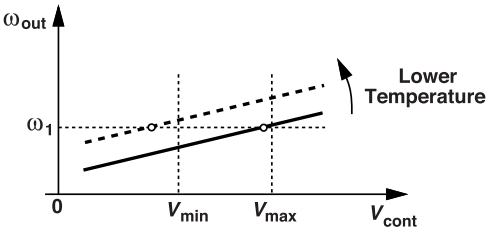


Figure 8.20 Setting V_{min} and V_{max} close to each other.

then the search described above must occur again, interrupting the communication. A similar situation arises if the VCO experiences supply noise.

In practice, it is difficult to avoid this interruption. For example, suppose the loop has locked with V_{cont} just slightly less than V_{max}. A temperature change can thus require $V_{cont} > V_{max}$, causing D_2 in Fig. 8.19 to change and hence the loop to restart the search. In summary, the narrower the $V_{max} - V_{min}$ range is, the higher is the complexity and the more frequently the PLL is interrupted.

8.6 Ripple Reduction by Sampling Filter

An undesirable result of the PFD/CP imperfections is the ripple on the oscillator control voltage. For a given PFD/CP design, the ripple can be reduced only by increasing the loop filter capacitor—unless we devise a new technique.

We wish to isolate the oscillator control from the voltage jumps created by the charge pump. We surmise that V_{cont} need not be connected to the CP output at all times. Consider the arrangement shown in Fig. 8.21(a), where switch S_1 is interposed between the main loop filter and C_2. Suppose S_1 remains off during

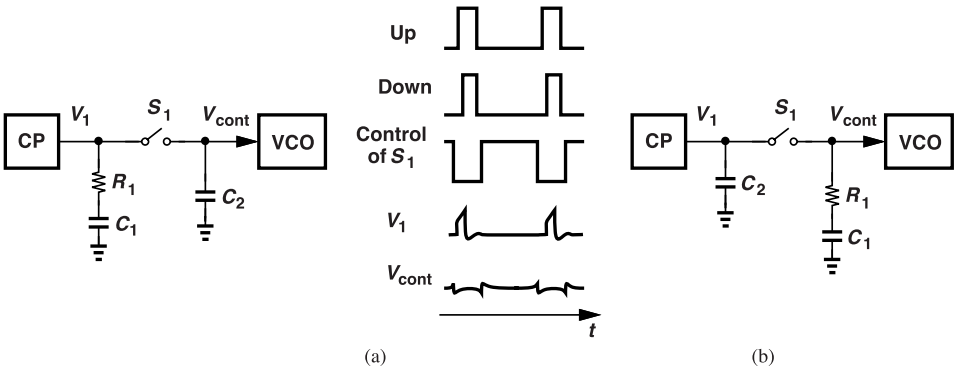


Figure 8.21 (a) First attempt to use sampling between the CP and the VCO, and (b) proper placement of the loop filter components.

PFD/CP activity and turns on only after the CP has turned off. The VCO therefore does not see the "bump" on V_1, exhibiting only a small disturbance due to the clock feedthrough and charge injection of S_1.

Unfortunately, the approach depicted in Fig. 8.21(a) leads to an unstable loop. To understand why, recall from Chapter 7 that R_1 stabilizes the system by causing a jump in the control voltage (in a manner similar to a feedforward path). Here, on the other hand, the initial voltage step across R_1 vanishes before S_1 turns on; i.e., V_{cont} sees no jump. Thus, the instantaneous value of V_{cont} after S_1 turns on is independent of R_1—as if R_1 were zero.

There is still hope. Let us swap the two branches as shown in Fig. 8.21(b). Now, the circuit's behavior is similar to that of a standard topology: the CP briefly delivers charge to C_2 and subsequently turns off. Next, C_2 shares its charge with C_1, causing a (loop-stabilizing) overshoot in V_{cont}. This arrangement is called a "sampling loop filter" [5].

As mentioned above, switch S_1 in Fig. 8.21(b) introduces charge injection and clock feedthrough onto V_{cont}. Some ripple is therefore inevitable.

8.7 Loop Filter Leakage

Some applications require relatively large loop filter capacitors, for example, if the input frequency is low or the loop bandwidth must be small. In such cases, the capacitors can be realized as MOSFETs so as to save chip area, but the gate leakage current proves problematic. Figure 8.22 plots the gate leakage for a 10 μm/0.5 μm

NMOS transistor with a gate dielectric thickness of 20 Å in 45-nm technology [6]. The source, drain, and substrate terminals are grounded.

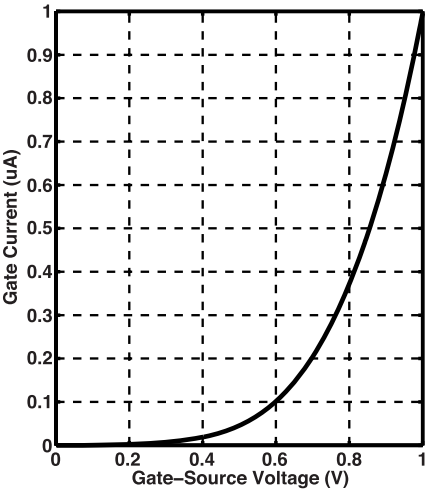


Figure 8.22 Gate leakage of a 10 μm/0.5 μm NMOS in 45-nm technology.

The gate leakage, I_G, discharges the loop filter when the charge pump is off, in the same manner as depicted in Fig. 8.15(b). The potentially large droop in the control voltage introduces significant phase modulation at the VCO output. It can be shown that the peak-to-peak output phase change is given by [6]:

$$\phi_{pp} = \frac{1}{2} K_{VCO} \frac{I_G}{C_2} \left(\frac{T_{REF}}{2} \right)^2, \tag{8.27}$$

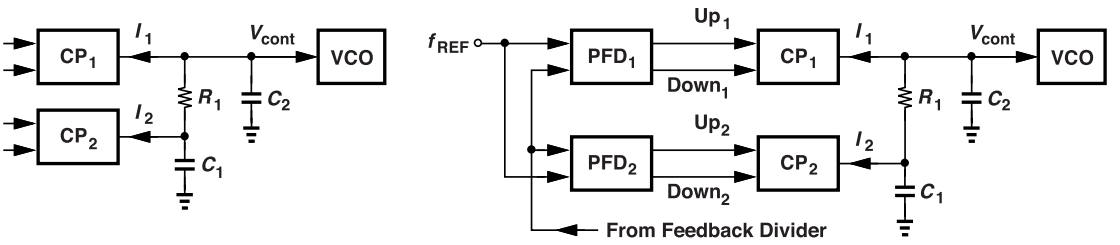
where C_2 is the second, smaller capacitor in the filter and I_G is the total leakage current drawn by C_1 and C_2. Manifesting itself for low reference frequencies, the periodic modulation generates sidebands and deterministic jitter. As illustrated in Fig. 8.22, I_G is a strong function of V_{GS} $(= V_{cont})$, making it difficult to cancel the leakage by means of circuit techniques. If the modulation proves excessive, the capacitors can be implemented by thick-oxide (I/O) transistors or metal-metal capacitors, at the cost of a larger area.

We show in Problem 8.16 that the effect of the loop filter leakage cannot be suppressed by the sampling technique described in the previous section. Digital PLLs, on the other hand, circumvent this issue altogether (Chapter 10).

8.8 Filter Capacitor Reduction

Let us assume $\zeta = 1$. The loop filter capacitor assumes a large value for (a) low reference frequencies, because we choose $2\zeta\omega_n = 2\omega_n = 2\sqrt{I_p K_{VCO}/(2\pi C_1)} \approx 0.1\omega_{REF}$, or (b) a low charge pump current, because $\zeta = (R_1/2)\sqrt{I_p K_{VCO} C_1/(2\pi)} = 1$. We thus seek a technique that equivalently "multiplies" the value of the capacitor.

Depicted in Fig. 8.23(a) is a topology incorporating two charge pumps with nominal currents equal to I_{p1} and I_{p2}, and $I_{p1} = \alpha I_{p2}$. Note that the CP_1 output flows through both R_1 and C_1 whereas the CP_2 output sees only C_1 (if C_2 is relatively small). Suppose the up and down pulses are chosen such that the two charge pumps turn on simultaneously, but their output currents have opposite signs. For example, when the up current in CP_1 is enabled, so is the down current in CP_2. The transfer function from the PLL phase error



(a) (b)

Figure 8.23 (a) Use of two charge pumps to create a larger apparent capacitance, and (b) actual implementation.

to V_{cont} therefore emerges as

$$\frac{V_{cont}}{\Delta\phi}(s) = \frac{I_{p1}}{2\pi}\left(R_1 + \frac{1}{C_1 s}\right) - \frac{I_{p2}}{2\pi}\frac{1}{C_1 s} \tag{8.28}$$

$$= \frac{I_{p1}R_1}{2\pi} + \frac{I_{p1} - I_{p2}}{2\pi}\frac{1}{C_1 s}. \tag{8.29}$$

Since $I_{p1} - I_{p2} = (\alpha - 1)I_{p1}$,

$$\frac{V_{cont}}{\Delta\phi}(s) = \frac{I_{p1}R_1}{2\pi} + \frac{I_{p1}}{2\pi}\frac{1}{\dfrac{C_1}{\alpha - 1}s}, \tag{8.30}$$

revealing that C_1 is equivalently enlarged by a factor of $1/(\alpha - 1)$. For example, $\alpha = 1.2$ yields a fivefold decrease in the area occupied by C_1. It is important to note that this result is *not* equivalent to a mere reduction of the current in a single charge pump; the key point is that the stabilizing jump on R_1 is still equal to $I_{p1}R_1$.

In order to generate proper up and down pulses for the two charge pumps, we employ two PFDs, but with the inputs to one swapped [Fig. 8.23(b)]. In other words, rather than negate the CP_2 output current, we negate the phase error sensed by PFD_2. The reader is encouraged to draw the up and down pulses for a given input phase error and examine the CP currents. This topology proves useful in cases where C_1 tends to be large, but it does suffer from the noise current of two charge pumps.

8.9 Trade-Off Between Bandwidth and Spur Level

We have seen that most of the PFD/CP imperfections introduce ripple on the oscillator control voltage, causing reference sidebands (spurs) and deterministic jitter at the PLL output. If the spur level is unacceptably high, we must inevitably reduce the loop bandwidth, a trade-off that is not readily obvious at this point. In this section, we quantify the trade-off between the unity-gain bandwidth, ω_u, and the spur level.

We begin with a PLL design having $\zeta = 1$ and $C_2 = 0.2C_1$. As seen in Example 7.26, these values yield a third pole, $\omega_{p3} \approx 1/(R_1 C_2)$, that is slightly above ω_u. Now, suppose the output spur level is 20 dB higher than desired. Let us assume that the PFD/CP cascade is already optimized to yield minimum ripple. How do we methodically redesign the loop for 20-dB lower spurs?

We must make a number of observations. First, the ripple on the control can be reduced by only two parameters: (1) we can scale down the CP current, I_p, and the widths of the CP's main transistors proportionally so that the unwanted injections into the loop filter are reduced; since the transistor currents and widths are scaled together, their overdrive voltages and hence the CP output voltage compliance remain unchanged; (2) we can increase the filter's second capacitor, C_2. Second, for the redesigned PLL, we must still have $\omega_{p3} > \omega_u$ and $\zeta \geq 1$.

Let us write

$$\omega_{p3} \approx \frac{1}{R_1 C_2} \tag{8.31}$$

$$\omega_u \approx \frac{R_1 I_p K_{VCO}}{2\pi} \tag{8.32}$$

$$\zeta = \frac{R_1}{2}\sqrt{\frac{I_p K_{VCO} C_1}{2\pi}}. \tag{8.33}$$

Suppose we scale the charge pump down by a factor of 10 for a 20-dB reduction in the output spur level. Then, ω_u falls by the same factor while ω_{p3} remains constant. That is, ω_u is excessively decreased. Instead, we scale the CP down by a factor of $\sqrt{10}$ and increase C_2 by this factor. As a result, both ω_{p3} and ω_u drop by $\sqrt{10}$. We also raise C_1 by the same factor so that ζ does not change. In summary, to lower the spur level by 20 dB, we have redisigned the loop as follows:

$$I_p \rightarrow \frac{I_p}{\sqrt{10}} \tag{8.34}$$

$$C_2 \rightarrow \sqrt{10} C_2 \tag{8.35}$$

$$C_1 \rightarrow \sqrt{10} C_1 \tag{8.36}$$

$$\omega_{p3} \rightarrow \frac{\omega_{p3}}{\sqrt{10}} \tag{8.37}$$

$$\omega_u \rightarrow \frac{\omega_u}{\sqrt{10}} \tag{8.38}$$

$$\zeta = 1. \tag{8.39}$$

The reader can prove that the same objective is met if K_{VCO}, rather than I_p, is scaled down by $\sqrt{10}$. Since we still have $\zeta = 1$ and $C_2 = 0.2C_1$, we conclude that ω_{p3} remains slightly higher than ω_u, keeping the phase margin intact.

8.10 Phase Noise in PLLs

The most important aspect of PLL design relates to phase noise and jitter; these imperfections arise from both the electronic noise and the supply and substrate noise.

Our analysis of various PLL transfer functions in Chapter 7 has given us the foundation for output noise computations. In this section, we examine various noise sources and formulate their contributions at the output. Figure 8.24 highlights the sources of noise: input phase noise, $\phi_{n,in}$; CP noise current, I_n; the loop filter resistor noise, V_n; VCO phase noise, $\phi_{n,VCO}$; divider phase noise, $\phi_{n,div}$; and supply noise, $V_{n,VDD}$. The PFD also contributes some noise (Section 8.2). Among these, $\phi_{n,VCO}$, $\phi_{n,in}$, and $V_{n,VDD}$ prove most serious.

8.10.1 Shaping of Input Phase Noise

According to our derivations in Chapter 7, the input phase noise experiences the low-pass response of the PLL. As illustrated in Fig. 8.25, if $\phi_{n,in}$ has a flat spectrum of S_0—a good approximation for the phase noise of crystal oscillators—then the output spectrum displays an in-band value of $M^2 S_0$ and a roll-off imposed by the PLL's -3-dB bandwidth. We say the input phase noise is "shaped" by the transfer function. To obtain the total output jitter, we must integrate this spectrum. We surmise that the area under the curve is greater than $2M^2 S_0 f_{-3dB}$.

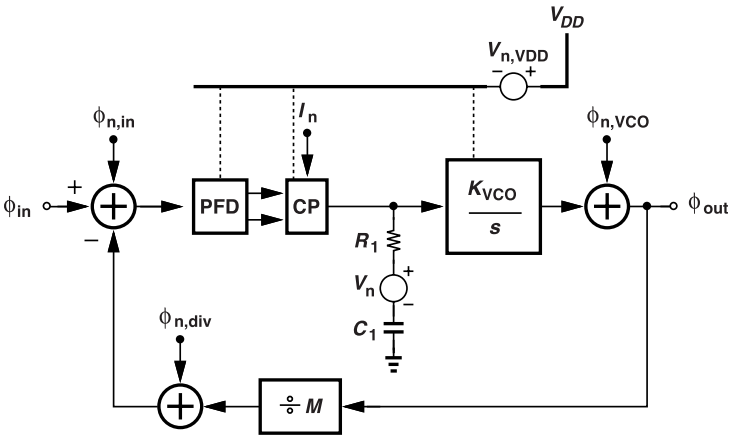


Figure 8.24 Summary of various noise sources in a CPPLL.

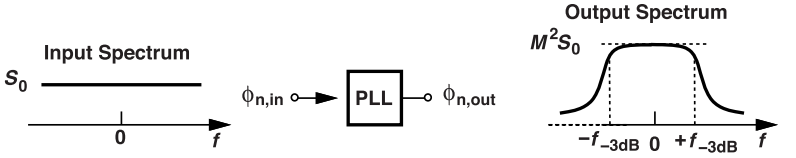


Figure 8.25 Shaping of input phase noise by PLL.

Example 8.9

Why does a PLL multiply the in-band input phase noise spectrum by M^2 but not the input jitter (measured in seconds)?

Solution

For an input of the form $V_1 \cos[\omega_{in}t + \phi_{n,in}(t)]$, the output emerges as $V_2 \cos[M\omega_{in}t + M\phi_{n,in}(t)]$, where $\phi_{n,in}$ represents the in-band phase noise. Thus, the output spectrum is M^2 times that at the input. For jitter (in seconds), on the other hand, we write $\Delta t_{in} = [\phi_{n,in}/(2\pi)]T_{in}$, and $\Delta t_{out} = [M\phi_{n,in}/(2\pi)]T_{in}/M$. These two quantities are equal. Intuitively, we note that, if the input edges drift (slowly) by Δt seconds, so do the output edges because the loop forces the phase error to zero. For example, a slow 1-ps edge displacement at the input produces to a 1-ps change in the output zero crossings as well, but this 1 ps translates to different phase values at the input and at the output because the periods are different.

Recall from Section 8.1 that a ζ value of above 1.5 allows the approximation $\omega_{-3dB} \approx 2\zeta\omega_n = R_1 I_p(K_{VCO}/M)/(2\pi)$ and roughly a one-pole response. We can thus readily integrate the output phase noise spectrum in Fig. 8.25:

$$\phi_{out,rms}^2 \approx \int_{-\infty}^{+\infty} \frac{M^2 S_0}{1 + (\frac{2\pi f}{\omega_{-3dB}})^2} df. \tag{8.40}$$

Since

$$\int \frac{dx}{1+x^2} = \tan^{-1} x, \tag{8.41}$$

we have

$$\phi^2_{out,rms} \approx M^2 S_0(\pi f_{-3dB}) \tag{8.42}$$

$$\approx \frac{M S_0 R_1 I_p K_{VCO}}{4\pi}. \tag{8.43}$$

To find the output rms jitter, we normalize $\phi_{out,rms}$ to 2π and multiply by the output period, T_{out}:

$$\Delta t_{rms} = \frac{1}{2\pi}\sqrt{\pi M^2 S_0 f_{-3dB}}\, T_{out}. \tag{8.44}$$

The jitter can be reduced by decreasing the loop bandwidth.

Example 8.10

A 20-MHz crystal oscillator exhibits a flat phase noise of -150 dBc/Hz. Determine the output jitter of a 6-GHz PLL that uses such a reference. Assume a one-pole response and $f_{-3dB} \approx 2$ MHz.

Solution

In this case, $M = 300$ and the output in-band phase noise is equal to $-150\ \text{dBc/Hz} + 20\log 300 \approx -100$ dBc/Hz. The overall output jitter is computed from (8.44) with $S_0 = -150\ \text{dBc/Hz} = 10^{-15}\ \text{Hz}^{-1}$:

$$\Delta t_{rms} = 0.631\ \text{ps} \tag{8.45}$$

$$= 1.36^\circ. \tag{8.46}$$

Note that this is only the input contribution. Some RF transmitters must incur less than 1° of rms phase noise at their output.

Example 8.11

Analog-to-digital converters typically receive a sampling clock from a PLL. The clock jitter, Δt_{rms}, degrades the signal-to-noise ratio of the sampled signal. As illustrated in Fig. 8.26, this effect appears as nonuniform sampling of the analog signal. For an input of the form $A\cos(2\pi f_{in} t)$, the jitter-induced noise power is equal to $P_{jit} = 2\pi^2 f_{in}^2 A^2 \Delta t_{rms}^2$. How much jitter can be tolerated if a 10-bit 10-GHz ADC must incur no more than 2 dB of SNR penalty? The quantization noise power of such an ADC is given by $P_Q = (2A/2^{10})^2/12$. Assume $f_{in} \approx 5$ GHz.

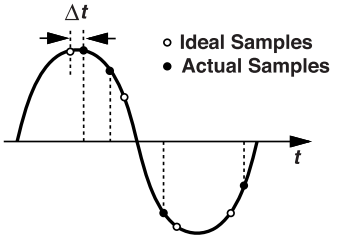


Figure 8.26 Effect of clock jitter on sampling a signal.

Solution

We wish to guarantee that $10\log[P_Q/(P_Q+P_{jit})] = -2$ dB. It follows that $\Delta t_{rms} = 19.4$ fs. This very low value underscores the stringent jitter requirements in high-speed ADC design. In fact, Example 8.10 suggests that even the reference phase noise leads to far more jitter—unless the loop bandwidth is greatly reduced.

8.10.2 Shaping of VCO Phase Noise

Recall from the analysis in Section 8.1 and the behavior illustrated in Fig. 8.3(b) that the VCO phase noise experiences a high-pass response. We must therefore multiply the spectrum of this phase noise by the magnitude squared of the transfer function to obtain the PLL output spectrum. From Eq. (8.14), we have

$$\left|\frac{\phi_{out}}{\phi_{n,VCO}}\right|^2 = \frac{\omega^4}{(\omega_n^2-\omega^2)^2+4\zeta^2\omega_n^2\omega^2}. \tag{8.47}$$

The key observation here is that, to minimize the VCO phase noise, the loop bandwidth must be maximized, a requirement in conflict with that for reducing the effect of the reference phase noise. For slow phase fluctuations, $|\phi_{out}/\phi_{n,VCO}|^2 \approx \omega^4/\omega_n^4$.

Let us look more closely at this "noise-shaping" phenomenon. The free-running VCO phase noise spectrum can be expressed as $S_{\phi,VCO}(\omega) = \alpha/\omega^3 + \beta/\omega^2$, where the two terms denote flicker and white noise contributions, respectively.[1] We find the intersection of the two regimes by writing $\alpha/\omega_c^3 = \beta/\omega_c^2$ and hence $\omega_c = \alpha/\beta$, and call this frequency the flicker noise corner. From Fig. 8.27, we must determine how $S_{\phi,VCO}$ is shaped by the transfer function. We predict that (a) at low ω values, the product of (8.47) and $S_{\phi,VCO}$

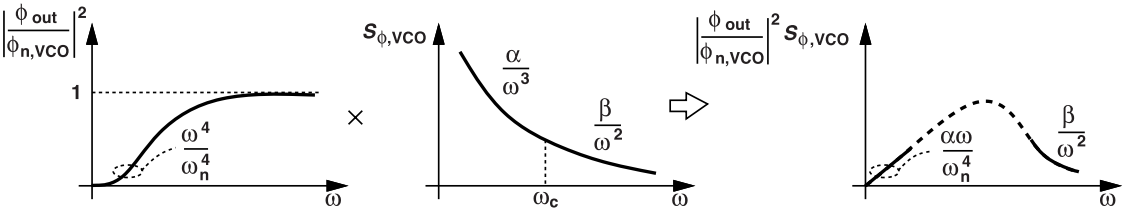


Figure 8.27 Shaping of VCO phase noise by a PLL.

reduces to $(\omega^4/\omega_n^4)(\alpha/\omega^3) = \alpha\omega/\omega_n^4$, rising linearly from zero at $\omega = 0$, and (b) at high ω values, the product is equal to $(\omega^4/\omega^4)\beta/\omega^2 = \beta/\omega^2$, i.e., the same as the shape of the VCO's free-running phase noise. The peaking in the middle depends on the values of ζ, ω_n, α, and β.

In order to gain insight, we consider two special cases, namely, when the loop bandwidth is much less or much greater than the VCO's flicker noise corner frequency. In the former case, only the flicker-noise-induced phase noise is shaped to some extent [Fig. 8.28(a)]. We can write

$$S_{\phi n} \approx \frac{\alpha}{\omega^3}\left|\frac{\phi_{out}}{\phi_{n,VCO}}\right|^2 \tag{8.48}$$

$$\approx \frac{\alpha\omega}{(\omega_n^2-\omega^2)^2+4\zeta^2\omega_n^2\omega^2}. \tag{8.49}$$

Representing how much slow VCO phase fluctuations are suppressed, this spectrum reaches a peak value of $9\alpha/(16\sqrt{3}\omega_n^3) \approx \alpha/(3.1\omega_n^3)$ at $\omega = \omega_n/\sqrt{3}$ if $\zeta = 1$. The free-running VCO, on the other hand, exhibits a

[1] The actual phase noise should be expressed as $\alpha/f^3 + \beta/f^2$. We will account for this change of variables later.

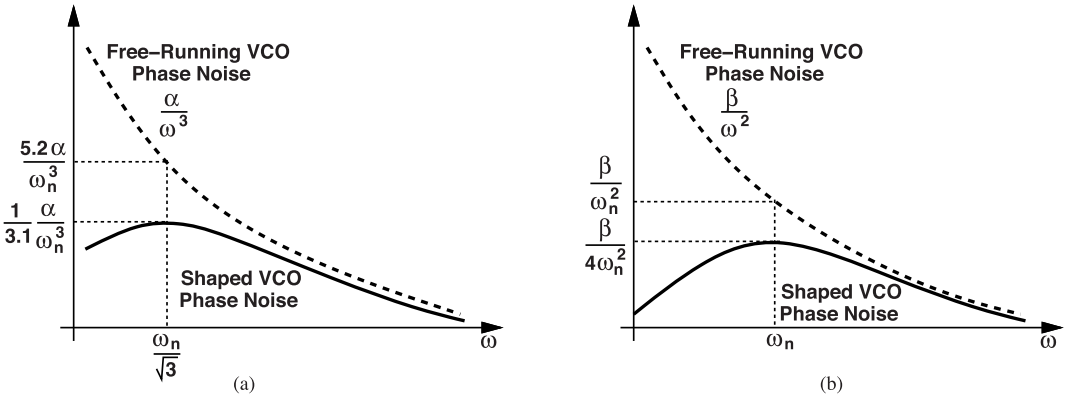


Figure 8.28 VCO phase noise shaping for the case of the loop bandwidth (a) much less, or (b) much greater than the VCO flicker noise corner frequency. It is assumed that $\zeta = 1$.

phase noise of $\alpha/[\omega_n^3/(3\sqrt{3})] \approx 5.2\alpha/\omega_n^3$ at this offset frequency. That is, the phase noise is reduced by a factor of 16, about 12 dB, at $\omega_n/\sqrt{3}$. It is important to note that the peaking frequency, $\omega_n/\sqrt{3}$, is about 40% less than the high-pass corner frequency, ω_n, of the VCO phase noise transfer function (Section 8.1).

In the second special case, the loop bandwidth is much greater than the flicker noise corner, allowing high suppression at low offsets and making the white-noise-induced phase noise, β/ω^2, dominant [Fig. 8.28(b)]. We thus write

$$S_{\phi n} = \frac{\beta\omega^2}{(\omega_n^2 - \omega^2)^2 + 4\zeta^2\omega_n^2\omega^2}. \tag{8.50}$$

This spectrum has a maximum of $\beta/(4\omega_n^2)$ at $\omega = \omega_n$ if $\zeta = 1$. We observe a 6-dB reduction due to phase-locking.

The shaped phase noise peaks in Figs. 8.28(a) and 8.28(b) are inversely proportional to ω_n^3 and ω_n^2, respectively. We say that the loop bandwidth must be maximized so as to suppress the VCO phase noise. However, if $\zeta = (R_1/2)\sqrt{I_p K_{VCO} C_1/(2\pi M)} = 1$, then $\omega_n = \sqrt{I_p K_{VCO}/(2\pi M C_1)}$ can be increased only by reducing C_1 (and increasing R_1 so that ζ does not degrade). Since C_2 must also decrease proportionally, the ripple on the control voltage rises. In other words, the PLL presents a trade-off between the VCO phase noise suppression and the ripple amplitude (deterministic jitter).

Example 8.12

If $\zeta^2 \gg 1$, sketch the PLL output spectrum arising from the VCO phase noise. Neglect the flicker noise contribution.

Solution

Let us return to Section 8.1 and plot the VCO phase noise transfer function from Eq. (8.14) if $\zeta^2 \gg 1$. We observe that the poles are given by (8.3) and (8.4) in this case:

$$\omega_{p1} = -\frac{\omega_n}{2\zeta} \tag{8.51}$$

$$\omega_{p2} = -2\zeta\omega_n. \tag{8.52}$$

The transfer function also has two zeros at the origin, exhibiting the response shown in Fig. 8.29(a). Multiplying the square of this magnitude by β/ω^2 therefore yields a relatively flat spectrum between the two poles [Fig. 8.29(b)]. This is the behavior commonly observed in PLLs.

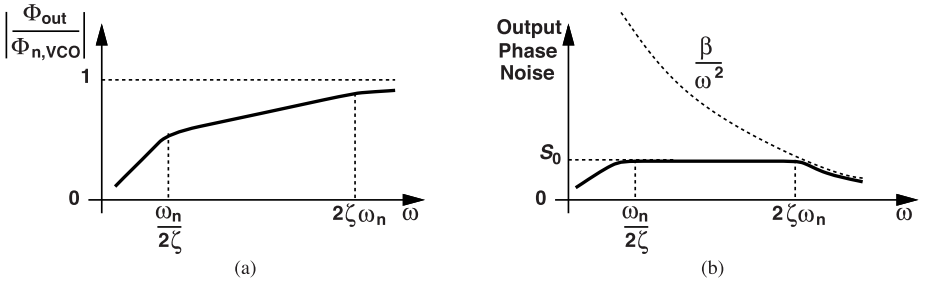


Figure 8.29 (a) VCO noise shaping for the case of $\zeta^2 \gg 1$, and (b) resulting PLL output spectrum.

We can estimate the value of the spectrum, S_0, in the flat region by noting that the free-running and shaped profiles meet at approximately $\omega = 2\zeta\omega_n$. Thus, $S_0 \approx \beta/(2\zeta\omega_n)^2$. However, since the spectral density is defined in terms of f, rather than in terms of ω, we must write $S_0 \approx \beta/(2\zeta f_n)^2$, where $f_n = \omega_n/(2\pi)$. [This point should not be confusing: the unit of phase noise is Hz^{-1} rather than $(\text{rad/s})^{-1}$.]

PLL Jitter Estimate If the VCO phase noise is dominant and $\zeta^2 \gg 1$, the profile in Fig. 8.29(b) implies the two-sided output spectrum shown in Fig. 8.30, where the horizontal scale is in hertz. (In practice, the "notch"

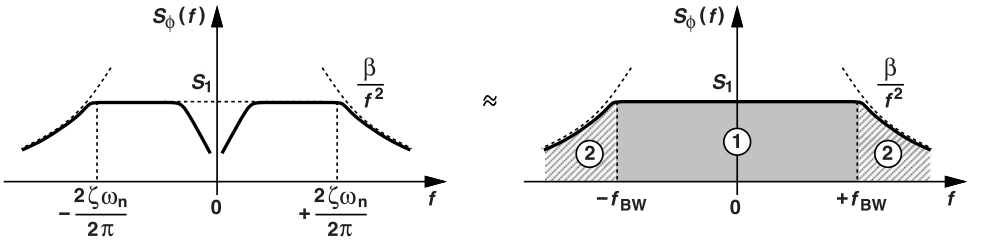


Figure 8.30 Approximation of shaped VCO phase noise.

near $f = 0$ is filled with the noise contributed by the other parts of the PLL.) Recognizing $2\zeta\omega_n/(2\pi)$ as the loop bandwidth, f_{BW}, we approximate the area under this profile from $-f_{BW}$ to $+f_{BW}$ by $A_1 = 2f_{BW}S_1 = 2f_{BW}\beta/(2\zeta f_n)^2 = 2\beta/f_{BW}$. Moreover, the total area from $+f_{BW}$ to $+\infty$ and from $-f_{BW}$ to $-\infty$ is computed as the definite integral of β/f^2 and is also equal to $A_2 = 2\beta/f_{BW}$. It follows that

$$\phi^2_{out,rms} = \frac{4\beta}{f_{BW}} \tag{8.53}$$

$$= 4S_1 f_{BW}. \tag{8.54}$$

This simple result can be remembered as: four times the loop bandwidth times the VCO free-running phase noise at an offset of f_{BW}. The equivalent jitter is given by

$$\Delta t_{rms} = \frac{1}{2\pi}\sqrt{4S_1 f_{BW}}\, T_{out}, \tag{8.55}$$

where T_{out} denotes the output period.

Example 8.13

A PLL suffers from a high ripple on its control voltage. An engineer halves the loop bandwidth to reduce the ripple. What happens to the total integrated phase noise contributed by the VCO? Neglect the flicker noise contribution.

Solution

As illustrated in Fig. 8.31, the in-band phase noise rises from S_1 to $4S_1$ (why?). Equation (8.54) therefore yields a twofold increase in $\phi^2_{out,rms}$ because the bandwidth is halved. The rms jitter climbs by a factor of $\sqrt{2}$.

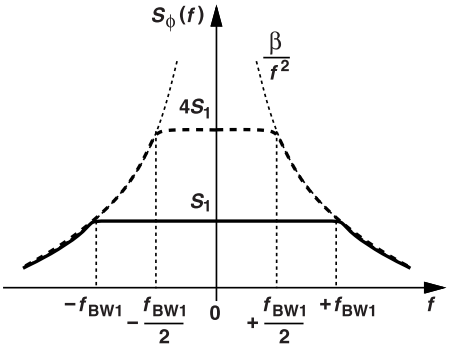


Figure 8.31 Effect of bandwidth reduction on the VCO phase noise.

Phase Noise Optimization In many applications, the VCO contributes higher phase noise than the input does, requiring a wide loop bandwidth. However, in stringent cases, such as in Examples 8.10 and 8.11, the "amplified" input noise also plays a role, thus demanding the optimum choice of the bandwidth. In other words, we must minimize the area under the following output spectrum:

$$S_{\phi,out} = \frac{M^2 S_0 (4\zeta^2 \omega_n^2 \omega^2 + \omega_n^2)}{(\omega_n^2 - \omega^2)^2 + 4\zeta^2 \omega_n^2 \omega^2} + \frac{\beta \omega^2}{(\omega_n^2 - \omega^2)^2 + 4\zeta^2 \omega_n^2 \omega^2}, \tag{8.56}$$

where the flicker noise contribution is neglected, and the two terms on the right-hand side represent the reference and VCO contributions, respectively. This is carried out in [6].

Let us approach the optimization from a different angle. As an approximation, we write the integrated reference and VCO contributions as $\pi M^2 S_0 f_{-3\text{dB}}$ and $4\beta/f_{BW}$, respectively, where S_0 denotes the reference phase noise. We now assume $f_{-3\text{dB}} \approx f_{BW}$ (Fig. 8.32) and minimize the total, $\pi M^2 S_0 f_{BW} + 4\beta/f_{BW}$. It follows that f_{BW} must be chosen equal to $\sqrt{4\beta/(\pi M^2 S_0)}$, yielding *equal* contributions for the reference and the VCO. The total integrated phase noise is then given by $4\sqrt{\beta \pi M^2 S_0}$.

8.10.3 Charge Pump Noise

The two current sources in the charge pump exhibit noise, corrupting the control voltage at each phase comparison instant. Even with perfect PFD/CP operation, the current sources turn on for T_{res} seconds (about five gate delays) (Chapter 7), injecting random noise. Our objective is to compute the PLL output phase noise resulting from this effect. We first take into account only one current source's noise.

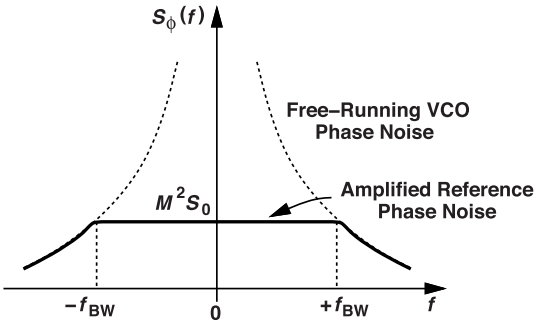


Figure 8.32 Effect of reference phase noise.

Example 8.14

Without writing the transfer function, explain the type of response experienced by the charge pump noise as it reaches the PLL output phase, ϕ_{out}.

Solution

Let us place a constant current source in parallel with one of the CP current sources, noting from Section 8.3.2 that this mismatch leads to a phase offset between the input and the feedback signal. If this current source now varies slowly with time, so do the phase offset and ϕ_{out}. Thus, the CP noise sees a low-pass response.

Effect of White Noise As shown in Fig. 8.33(a), I_1, injects noise for T_{res} seconds every T_{in} seconds. For white noise, we can say that the average injected noise power is equal to $\overline{I_{1,n}^2}\times(T_{res}/T_{in})$, where $\overline{I_{1,n}^2}$ denotes the noise current spectrum associated with I_1 (Chapter 5). This means that we can equivalently model the periodically-switched (cyclostationary) noise by a continuous-time source directly tied to the control voltage that has a spectral density of $\overline{I_{n,eq}^2} = \overline{I_{1,n}^2}\times(T_{res}/T_{in})$ [Fig. 8.33(b)]. We must simply determine the closed-loop transfer function from $I_{n,eq}$ to ϕ_{out}. To this end, we construct the linear model depicted in Fig. 8.33(c), where $\phi_{in} = 0$.

The forward transfer function seen by $I_{n,eq}$ is given by $[R_1 + 1/(C_1 s)]K_{VCO}/s$. It follows that

$$\frac{\phi_{out}}{I_{n,eq}}(s) = \frac{(R_1 + \frac{1}{C_1 s})\frac{K_{VCO}}{s}}{1 + \frac{I_p}{2\pi}(R_1 + \frac{1}{C_1 s})\frac{K_{VCO}}{Ms}} \tag{8.57}$$

$$= \frac{(R_1 C_1 s + 1)(K_{VCO}/C_1)}{s^2 + 2\zeta\omega_n s + \omega_n^2}. \tag{8.58}$$

Note that $\phi_{out}/I_{n,eq}$ has the same zero and poles as ϕ_{out}/ϕ_{in}; the two differ by only a factor of K_{VCO}/C_1. For low-frequency noise, $|s|$ is small and

$$\frac{\phi_{out}}{I_{n,eq}} \approx \frac{K_{VCO}}{C_1\omega_n^2} \tag{8.59}$$

$$\approx \frac{2\pi M}{I_p}. \tag{8.60}$$

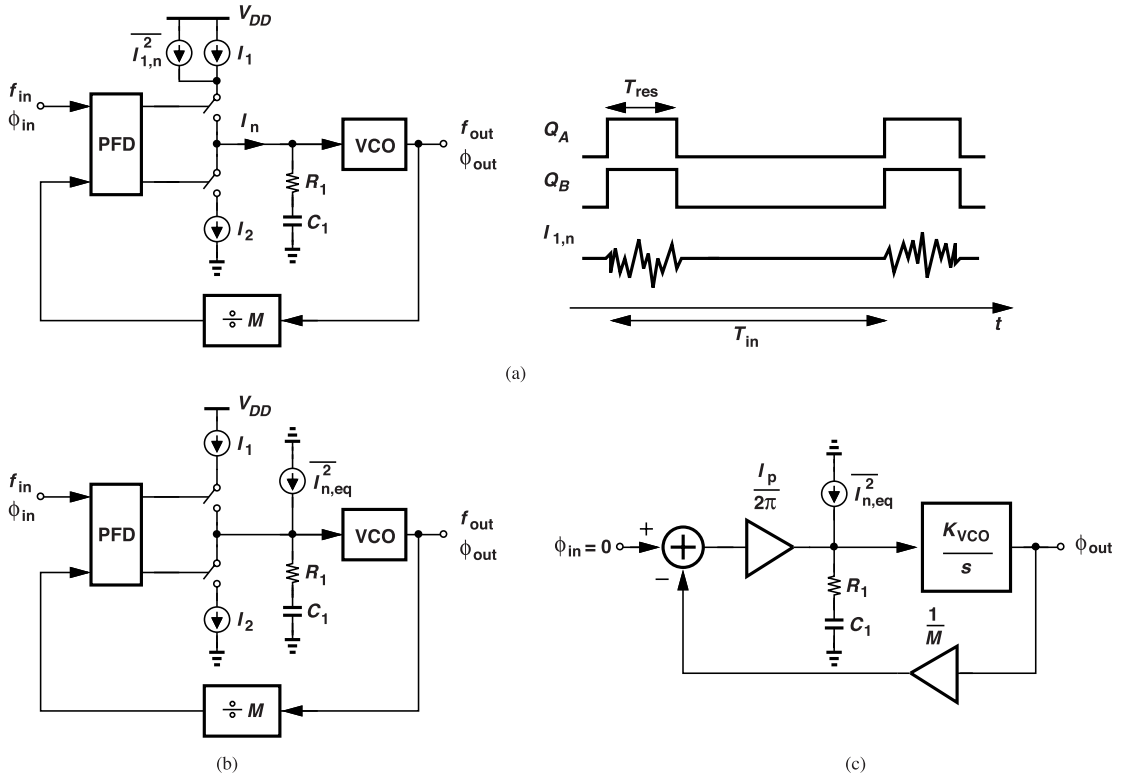


Figure 8.33 (a) Effect of CP noise, (b) equivalent noise current, and (c) the loop model.

The output phase noise within the loop bandwidth due to white noise thus emerges as

$$S_{\phi n,white} = 2I_{1,n}^2 \times \frac{T_{res}}{T_{in}} \times \frac{4\pi^2 M^2}{I_p^2}, \tag{8.61}$$

where the factor of 2 accounts for the contributions of both current sources.[2] At high noise frequencies, $S_{\phi n}$ experiences a low-pass roll-off. Equation (8.61) reveals that T_{res} must be minimized and I_p, maximized.

Example 8.15

A PLL used in a WiFi radio multiplies $f_{REF} = 20$ MHz to $f_{out} = 2400$ MHz. The charge pump has a bias current of 100 μA with an overdrive voltage of 50 mV. If $T_{res} = 50$ ps, determine the output in-band phase noise.

Solution

We have $I_{1,n}^2 = 4kT\gamma g_m = 4kT\gamma[2I_D/(V_{GS} - V_{TH})] = 6.624 \times 10^{-23}$ A^2/Hz if $\gamma = 1$. Also, $T_{in} = 50$ ns and $M = 120$. It follows from (8.61) that

$$S_{\phi n} = 7.53 \times 10^{-12} \text{ Hz}^{-1} = -111 \text{ dBc/Hz}. \tag{8.62}$$

While acceptable in WiFi systems, this amount of phase noise proves objectionable in some applications.

[2] We assume equal noise spectra for the two current sources.

Effect of Flicker Noise To study the effect of CP flicker noise, $I_{nf}(t)$, we recognize from the waveforms in Fig. 8.33(a) that the noise current is multiplied by 1 for T_{res} seconds and by 0 for $T_{in} - T_{res}$ seconds in every cycle [Fig. 8.34(a)]. The CP output noise spectrum therefore appears as shown in Fig. 8.34(b). Owing

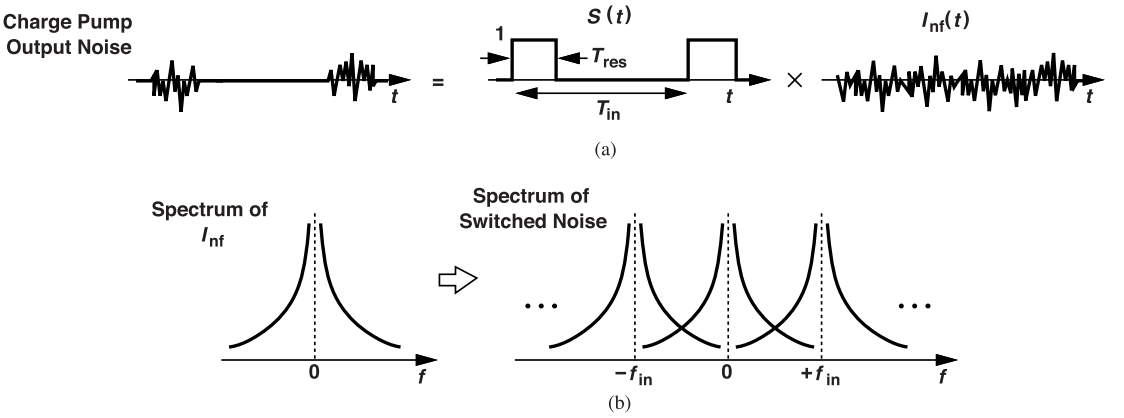


Figure 8.34 (a) Chopping of CP flicker noise, and (b) the resulting spectrum.

to the aliasing of the spectral tails, the calculation is difficult in the general case. However, if the flicker noise corner frequency is less than $f_{in}/2$, we can neglect aliasing and consider the section around $f = 0$ as the main contributor. This section results from the product of $I_{nf}(t)$ and the average value of $S(t)$, which is equal to T_{res}/T_{in}. It is interesting to note that, here, the noise current is weighted by T_{res}/T_{in} whereas, for thermal noise, the current *spectrum* is multiplied by this factor. It follows that the equivalent noise current source in Fig. 8.33(b) is given by $(T_{res}/T_{in})I_{nf}(t)$. Multiplying this current by the transfer function in (8.60) leads to the following output spectrum within the loop bandwidth:

$$S_{\phi n,1/f} = 2\overline{I_{nf}^2}(t) \times \frac{T_{res}^2}{T_{in}^2} \times \frac{4\pi^2 M^2}{I_p^2}, \tag{8.63}$$

where the factor of 2 assumes equal flicker noise powers for the two current sources. Compared to (8.61), this result benefits from the square of T_{res}/T_{in}. This difference arises because white noise is heavily aliased whereas flicker noise is not. The effect of the CP flicker noise is hence negligible, unless we are interested in very low offset frequencies, at which $\overline{I_{nf}^2}$ is large.

8.10.4 Loop Filter Noise

The noise generated by the loop filter's resistor modulates the VCO frequency, as qualitatively analyzed in Example 7.19. To formulate the effect, we draw the PLL as shown in Fig. 8.35(a) and seek the transfer function ϕ_{out}/V_n. Since the feedback signal $\phi_b = \phi_{out}/M$, the PFD/CP cascade generates an output given by $(-\phi_{out}/M)[I_p/(2\pi)]$, yielding a control voltage equal to $(-\phi_{out}/M)[I_p/(2\pi)][R_1 + 1/(C_1 s)]$. Adding V_n to this voltage and applying the result to the VCO, we obtain

$$\left[-\frac{\phi_{out}}{M} \cdot \frac{I_p}{2\pi}\left(R_1 + \frac{1}{C_1 s}\right) + V_n\right]\frac{K_{VCO}}{s} = \phi_{out}. \tag{8.64}$$

That is,

$$\frac{\phi_{out}}{V_n}(s) = \frac{2\pi M K_{VCO} C_1 s}{2\pi M C_1 s^2 + K_{VCO} I_p R_1 C_1 s + K_{VCO} I_p} \tag{8.65}$$

$$= \frac{K_{VCO} s}{s^2 + 2\zeta\omega_n s + \omega_n^2}. \tag{8.66}$$

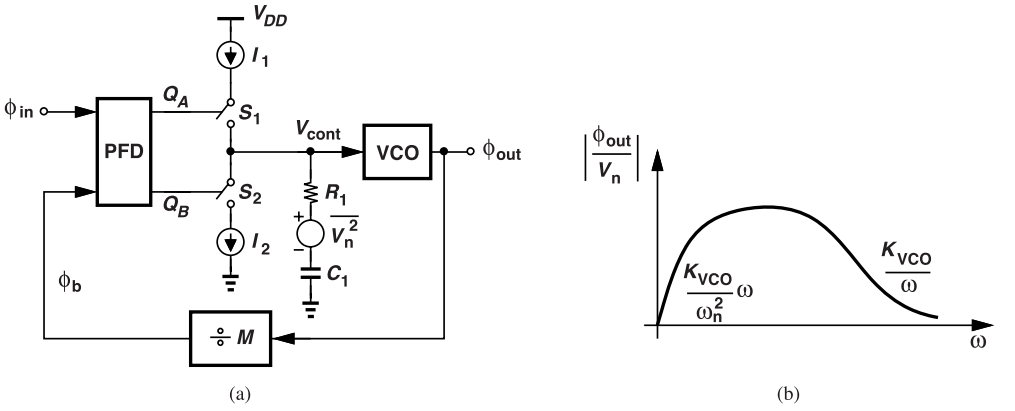


Figure 8.35 (a) Phase noise due to loop filter resistor, and (b) noise transfer function.

This band-pass response reduces to $(K_{VCO}/\omega_n^2)s$ at low frequencies (due to the strong feedback) and to K_{VCO}/s at high frequencies (as a result of the weak feedback) [Fig. 8.35(b)].

The reader can prove that $|\phi_{out}/V_n|^2$ reaches a maximum value of $K_{VCO}^2/(4\zeta^2\omega_n^2)$ at $\omega = \omega_n$. Since $2\zeta\omega_n = I_p K_{VCO} R_1/(2\pi M)$, it follows that the peak output phase noise is given by

$$S_{\phi n,p} = \frac{K_{VCO}^2}{4\zeta^2\omega_n^2}(4kTR_1) \tag{8.67}$$

$$= \frac{16kT\pi^2M^2}{R_1 I_p^2}. \tag{8.68}$$

(This is the one-sided spectrum.) This peak can be lowered by increasing R_1 or I_p.

Example 8.16

For the PLL design in Example 7.17, determine the output peak phase noise resulting from R_1.

Solution

We have $R_1 = 81.8\ \text{k}\Omega$, $K_{VCO} = 2\pi(300\ \text{MHz/V})$, $C_1 = 5$ pF, $M = 250$, and $I_p = 100\ \mu$A. Thus, $\omega_n = 2\pi(780\ \text{kHz})$, and $S_{\phi n,p} = 5 \times 10^{-11}\ \text{Hz}^{-1} = -103$ dBc/Hz at 780-kHz offset. This relatively high value reveals that I_p must be chosen higher.

The phase noise peaking depicted in Fig. 8.35(b) readily manifests itself in practice if R_1 is not sufficiently large. As illustrated in Fig. 8.36, the peak occurs within the loop bandwidth (why?).

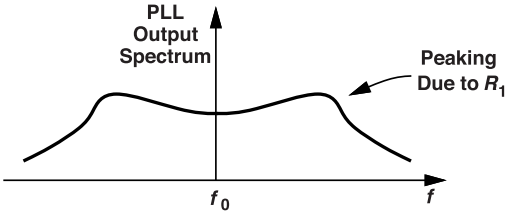


Figure 8.36 Phase noise peaking due to loop filter resistor noise.

8.10.5 Supply Noise

Supply noise translates to phase noise through two principal mechanisms: (1) modulation of the charge pump current, and (2) modulation of the VCO frequency. We study both effects here.

If the supply voltage varies, the up and down currents in the charge pump experience a "common-mode" change and a "differential" change, $I_1 - I_2$ (Fig. 8.37). The former does not translate to phase noise if the up and down pulses have equal widths and no skew (why?). The latter, on the other hand, is similar to CP current mismatch (Section 8.3.2) and CP noise (Section 8.10.3), obeying the same equations.

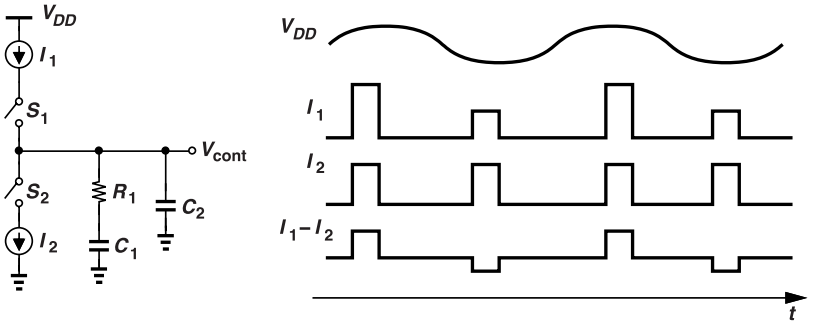


Figure 8.37 Effect of supply noise on CP output.

Example 8.17

A PLL designer shares the supply lines of the charge pump and the feedback divider, with the latter drawing large transient currents. Suppose the resulting supply bounce modulates one of the CP sources in the form of $\Delta I = I_1 \cos \omega_{in} t$, where ω_{in} denotes both the input frequency (and the divider output frequency). Determine the PLL output sidebands.

Solution

Recall from Section 8.1.1 that the PLL bandwidth is typically much less than ω_{in}. We can thus simplify Eq. (8.58) by assuming a large s:

$$\frac{\phi_{out}}{I_{n,eq}}(s) \approx \frac{R_1 K_{VCO}}{s}. \tag{8.69}$$

It follows that

$$\phi_{out} = \frac{R_1 I_1 K_{VCO}}{\omega_{in}} \sin \omega_{in} t, \tag{8.70}$$

yielding a normalized sideband magnitude of $R_1 I_1 K_{VCO}/(2\omega_{in})$. For example, if $I_1 = 0.001 I_p$, the PLL design in Example 7.17 exhibits sidebands -24 dB below the carrier. This calculation shows how serious the CP supply noise is.

The effect of supply noise on the VCO can be modeled as shown in Fig. 8.38(a). We denote the VCO supply sensitivity by $K_{VDD} = \partial \omega_{out}/\partial V_{DD}$ (Chapter 3) and note that the VCO free-running output phase is given by $K_{VDD} V_n/s$. The modulation can be viewed as arising from a noise source in the main control path if we simply refer the noise to V_{cont} [Fig. 8.38(b)]. The reader can prove that this source experiences the same transfer function as that by the noise of R_1 in Fig. 8.35(a). Thus,

$$\frac{\phi_{out}}{V_n}(s) = \frac{K_{VDD} s}{s^2 + 2\zeta\omega_n s + \omega_n^2}. \tag{8.71}$$

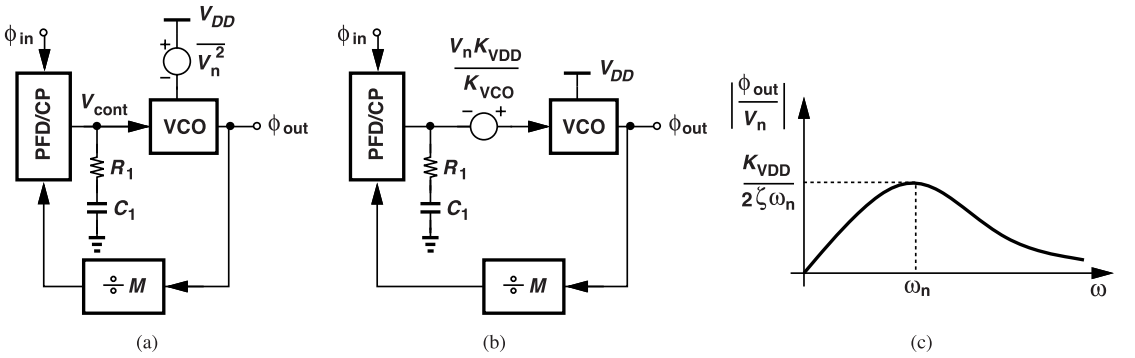


Figure 8.38 (a) Supply noise model, (b) representation of supply noise by a voltage in series with V_{cont}, and (c) noise transfer function.

The band-pass response has a peak value of $K_{VDD}/(2\zeta\omega_n)$ at $\omega = \omega_n$ [Fig. 8.38(c)], suggesting that $2\zeta\omega_n$ must be maximized so as to suppress the supply noise.

Example 8.18

The 5-GHz PLL design in Example 7.17 incorporates a ring VCO with a K_{VDD} of $2\pi(4\ \text{GHz/V})$. If the supply voltage happens to carry a sinusoidal noise component at $\omega = \omega_n$ and modeled as $(0.1\ \text{mV})\cos\omega_n t$, determine the output sidebands and jitter.

Solution

Recall that $\zeta = 1$ and $\omega_n = 2\pi(780\ \text{kHz})$. The output phase at $\omega = \omega_n$ can be expressed as $V_n K_{VDD}/(2\zeta\omega_n)$ and hence

$$\phi_{out}(t) = \frac{K_{VDD}}{2\zeta\omega_n}(0.1\ \text{mV})\cos\omega_n t, \tag{8.72}$$

thus producing a sideband magnitude of $(K_{VDD} \times 0.1\ \text{mV})/(4\zeta\omega_n) = -18$ dBc. The VCO output emerges as $V_0\cos[\omega_0 t + \phi_{out}(t)]$, revealing a peak-to-peak jitter equal to twice the amplitude of $\phi_{out}(t)$, i.e., $[2K_{VDD}/(2\zeta\omega_n)](0.1\ \text{mV}) = 0.51$ radians, which amounts to 16.3 ps,pp. The sidebands and the jitter are unacceptably large, necessitating a lower K_{VDD}. This example illustrates the severity of the VCO supply noise in PLL design.

References

[1]. A. Homayoun and B. Razavi, "On the stability of charge-pump phase-locked loops," *IEEE Trans. Ciruits and Systems - Part I*, vol.63,pp. 626-635, June 2016.

[2]. A. Homayoun and B. Razavi, "Analysis of phase noise in phase/frequency detectors," *IEEE Trans. Circuits and Systems - Part I*, vol. 60, pp. 529-539, March 2013.

[3]. M. Terrovitis et al., "A 3.2 to 4 GHz 0.25 μm CMOS frequency synthesizer for IEEE 802.11a/b/g WLAN," *ISSCC Dig. Tech. Papers,* pp. 98-99, Feb. 2004.

[4]. M. Wakyama, "Low offset and low glitch energy charge pump and method of operating same," US Patent 7057465, April 2005.

[5]. S. E. Meninger and M. H. Perrott, "A 1-MHz bandwidth 3.6-GHz 0.18-μm CMOS fractional-N synthesizer utilizing a hybrid PFD/DAC structure for reduced broadband phase noise," *IEEE J. Solid-State Circuits,* vol. 41, pp. 966-981, April 2006.

[6]. B. Razavi, "The role of PLLs in future wireline transmitters," *IEEE Trans. Circuits and Systems - Part I*, vol. 56, pp. 1786-1793, Aug. 2009.

Problems

8.1. Consider the VCO phase noise response sketched in Fig. 8.3(b). Explain how the plot changes if K_{VCO} is doubled.

8.2. Repeat the previous problem if the charge pump current or the main capacitor in the loop filter is doubled.

8.3. We wish to maximize $omega_{HPF}$ in Fig. 8.3(b) while maintaining $\zeta = 1$. Explain how the loop parameters must be adjusted.

8.4. Consider the three transfer functions plotted in Fig. 8.5. Explain how they change if K_{VCO} is doubled.

8.5. Repeat the previous problem if the loop filter resistor is doubled.

8.6. Consider the drain-switched charge pump shown in Fig. 8.8(b). What happens to the current mismatch if we double the widths and lengths of M_1 and M_2? How does the PLL phase offset change?

8.7. Explain why the plot in Fig. 8.10(b) exhibit an increasing slope (rather than a decreasing slope) near the ends of the voltage range.

8.8. How does the plot in Fig. 8.10 change if the widths of M_3 and M_4 are halved?

8.9. Explain qualitatively whether the noise of R_1 in Fig. 8.11 experiences a low-pass, band-pass, or high-pass response as it translates to ϕ_{out}.

8.10. Derive the open-loop and closed-loop transfer functions of the PLL shown in Fig. 8.11 if $A_1 = \infty$.

8.11. Repeat the previous problem if A_1 is low, say, around 10.

8.12. In the circuit of Fig. 8.11, what is the steady-state voltage at node X (if $V_n = 0$)? Does the charge pump operate properly with such a voltage? How do we fix the circuit?

8.13. Suppose we halve I_{REF} in Fig. 8.12. Considering only channel-length modulation, explain what happens to the current mismatch and the phase offset of a PLL employing this charge pump. (The output resistance of a MOSFET in saturation is inversely proportional to the drain current.)

8.14. Consider the CP shown in Fig. 8.17 and model its noise by a voltage source in series with one of its inputs. First, assume this voltage is *constant* and explain whether the output phase of a PLL emplying this CP would change. Next, determine qualitatively the type of transfer function from this noise to the PLL output phase.

8.15. Repeat the previous problem for the CP in Fig. 8.18(c).

8.16. Suppose capacitor C_1 in Fig. 8.21(b) suffers from leakage. Sketch V_1 and V_{cont} as a function of time when the PLL is locked. Assume $C_2 \ll C_1$.

8.17. We wish to reduce $\phi^2_{out,rms}$ in Eq. (8.43) without changing ζ. Explain how R_1, I_p, and K_{VCO} must be adjusted.

8.18. Noting that, in Eq. (8.44), we have $MT_{out} = T_{REF}$, simplify the expression for Δt_{rms}. If we choose $f_{-3dB} \approx 0.1 f_{REF}$, simplify the equation further.

8.19. Consider the CP noise expressed by (8.61). Suppose we double I_p and the widths of all of the CP transistors. Explain what happens to this phase noise, the loop bandwidth, and the damping factor.

8.20. Equation (8.68) implies that the noise contributed by R_1 is minimized if R_1 is maximized. In the presence of a second capacitor, C_2, in the loop filter, explain what happens to the zero, the unity-gain frequency, and the phase margin of the loop.

9

PLL Design Study

Our analysis of phase-locked loops in previous chapters has prepared us to embark upon their design as well. In this chapter, we deal with the step-by-step design of a "generic" charge-pump PLL in 40-nm CMOS technology. While targeting certain performance specifications, we are primarily interested in how the design proceeds, what aspects of the circuit must be analyzed and simulated, and, most importantly, how the designer thinks. This endeavor also gives us a feel for the performance bounds that we would typically encounter in practice. As we will see, not every design choice leads to success, and we must often iterate on the values or the circuit topologies.

Our design targets are as follows:
Reference Frequency = 20 MHz
Output Frequency = 2.4 GHz
Power Consumption = 2 mW
Deterministic Jitter = 10 ps,pp
Random Jitter = 2 ps,rms
Supply Voltage = 1 V.

We observe that the loop has a divide ratio of $M = 120$. The reader is encouraged to review Chapters 3 and 7 before beginning this design study.

9.1 Design Procedure

A key point in circuit design is to begin with the *simplest* topology and the *smallest* transistors unless our experience suggests otherwise. We therefore start with a standard charge-pump PLL and utilize the 2.5-GHz ring VCO developed in Chapter 3. Whether these choices yield the desired performance needs to be seen.

How do we choose the loop parameters? We seek a method that readily leads to a basic design, anticipating that we must perform refinements later. The two equations derived in Chapter 7 for type-II PLLs provide a starting point:

$$\zeta = \frac{R_1}{2}\sqrt{\frac{I_p K_{VCO} C_1}{2\pi M}} \tag{9.1}$$

$$\omega_n = \sqrt{\frac{I_p K_{VCO}}{2\pi M C_1}}. \tag{9.2}$$

The VCO gain, K_{VCO}, is obtained as explained in Chapter 3, leaving I_p, C_1, and R_1 under our control. We choose $\zeta \approx 1$, but how about ω_n? Recall from Chapter 8 that the discrete-time nature of the loop requires that the bandwidth (BW) be reduced to about one-tenth of the input (reference) frequency. But which bandwidth? More relevant here is the open-loop unity-gain BW, $\omega_u \approx 2.1\omega_n$ for $\zeta = 1$. (Note that the -3-dB closed-loop

BW, $\omega_{-3dB} \approx 2.5\omega_n$, is about the same.) Thus,

$$\omega_u = 2.1\omega_n = \frac{\omega_{REF}}{10}. \tag{9.3}$$

We now have two equations and three unknowns, and hence some freedom in the choice of the loop parameters. The charge pump current, in particular, can assume a wide range, with the lower bound governed by the charge pump noise (Chapter 8) and the size of the loop filter capacitor (because $\zeta \propto \sqrt{I_p C_1}$). Let us select $I_p = 100\ \mu$A. Note that the CP power consumption is typically quite small because, in the locked condition, I_p is drawn from the supply for only a brief amount of time, namely, the PFD reset time.

We have from (9.2) and (9.3):

$$2.1\sqrt{\frac{I_p K_{VCO}}{2\pi M C_1}} = \frac{2\pi \times (20\ \text{MHz})}{10}. \tag{9.4}$$

The discrete tuning plots in Fig. 3.42(c) yield $K_{VCO} \approx 400$ MHz/V. Bearing in mind that K_{VCO} must be expressed in radians/s/V, $K_{VCO} \approx 2\pi \times (400)$ Mrad/s/V, and with $M = 120$, we obtain

$$C_1 = 9.3\ \text{pF}. \tag{9.5}$$

Equating (9.1) to unity gives

$$R_1 = 35.9\ \text{k}\Omega. \tag{9.6}$$

We also select $C_2 = 0.2C_1 = 1.86$ pF. The PLL can now be simulated (Section 9.4). Note that the VCO phase noise suppression bandwidth for $\zeta = 1$ is equal to $\omega_n = 2\pi(0.95\ \text{MHz})$ in this case.

The choice of $C_2 = 0.2C_1 = 1.86$ pF raises one issue: as explained in Chapter 7, $C_1C_2/(C_1+C_2)$ and R_1 form a pole, which in this case resides at 2.8 MHz, just slightly above ω_u. Thus, the phase margin degrades and the PLL input-output response exhibits additional peaking. We return to this point in Section 9.5.

9.2 PFD Design

For the PFD, we return to the NOR-based implementation described in Chapter 7 and choose $W/L = 120\ \text{nm}/40$ nm for all of the transistors. Such small devices exhibit high flicker noise, but still produce negligible phase noise in this design. This can be verified by simulations as explained in [1]. The PFD output pulses have a minimum width of about 100 ps.

9.3 Charge Pump Design

Charge pump design in namometer technologies proves challenging. In this section, we attempt different topologies and learn about these challenges. As practiced in previous chapters, we begin our simulations with a worst-case supply of 1 V -5%, the SS corner, and at $T = 75\ ^\circ$C.

9.3.1 First CP Design

With a worst-case 0.95-V supply, we resort to the gate-switched CP topology (Chapter 7) and size the output transistors so as to accommodate a range of 0.1 V to 0.85 V for the VCO control. Additionally, the transistors must be long enough to provide adequate matching between the up and down currents. For example, suppose the mismatch due to channel-length modulation must remain below 5%. This value is somewhat arbitrary and will be revisited in terms of the ripple that it generates on the VCO control. To appreciate the difficulty in achieving such a small mismatch, let us recall the test setup studied in Chapter 7, repeated in Fig. 9.1(a), recognizing that for $V_X = 0.1$ V or 0.85 V, $|I_X| = |I_{D2} - |I_{D1}||$ must be less than $0.05I_p$, where I_p is the nominal CP current. Thus, as depicted in Fig. 9.1(b), when one device resides at the edge of the triode region

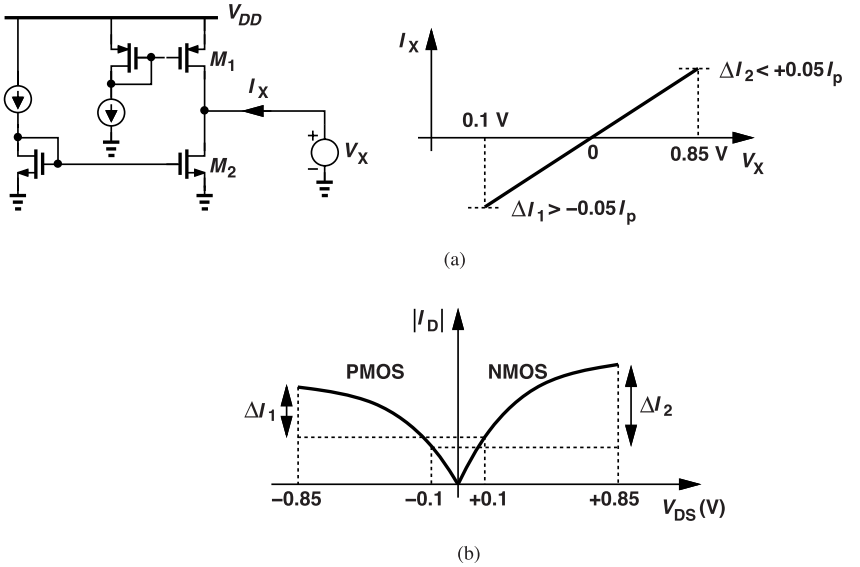


Figure 9.1 (a) Set-up for quantifying mismatch between up and down currents, and (b) illustration of transistor current variations.

($|V_{DS}| = 0.1$ V) and the other sustains a large V_{DS}, their currents must match to within 5%. That is, ΔI_1 and ΔI_2 must be less than $0.05 I_p$.

Using the arrangement in Fig. 9.1(a), we can plot the I_X-V_X characteristic and adjust the transistor dimensions toward the 5% error goal, while keeping I_p around 100 μA. We choose $L = 40$ nm and the transistor widths for an overdrive of about 100 mV. In this case, simulations indicate that I_X varies by ± 75 μA! We then increase L to 800 nm and choose $W_N = 40$ μm and $W_P = 80$ μm [Fig. 9.2(a)], arriving at the simulated I_X-V_X characteristic shown in Fig. 9.2(b). (The reference branches of the current mirrors are unscaled for now.) The variation is still around ± 20 μA, suggesting that the two transistors would need to be excessively

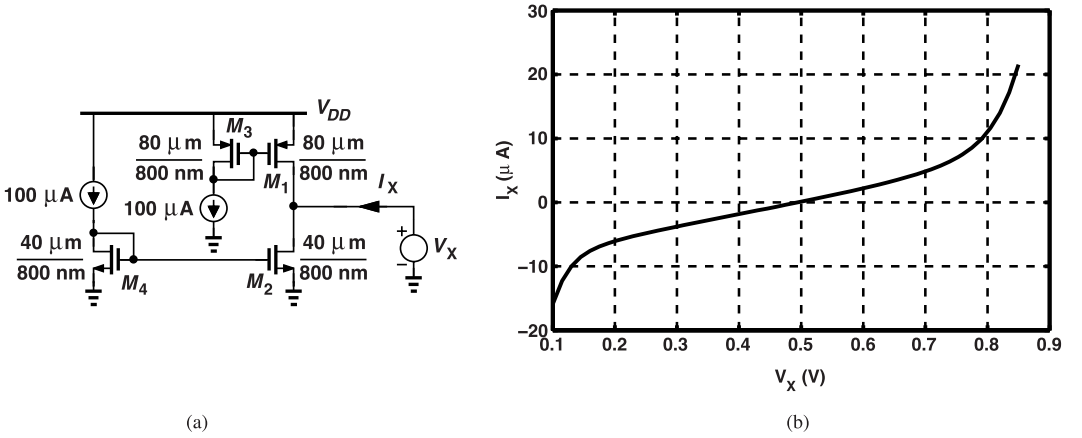


Figure 9.2 (a) Initial design of CP current sources, and (b) net current error.

large for the 5% target. Note that the large gate areas of M_1-M_4 make random mismatches negligible.

Rather than enlarge the transistors in Fig. 9.2(a), we return to the CP circuit techniques in Chapter 7 and explore their utility here. We choose the simple feedback technique illustrated in Fig. 8.16. With the aid of simulations, we reach the design shown in Fig. 9.3(a), and the characteristic depicted in Fig. 9.3(b). Note that M_5 and M_6 provide correction at low and high values of V_X, respectively. If these devices are too strong, then

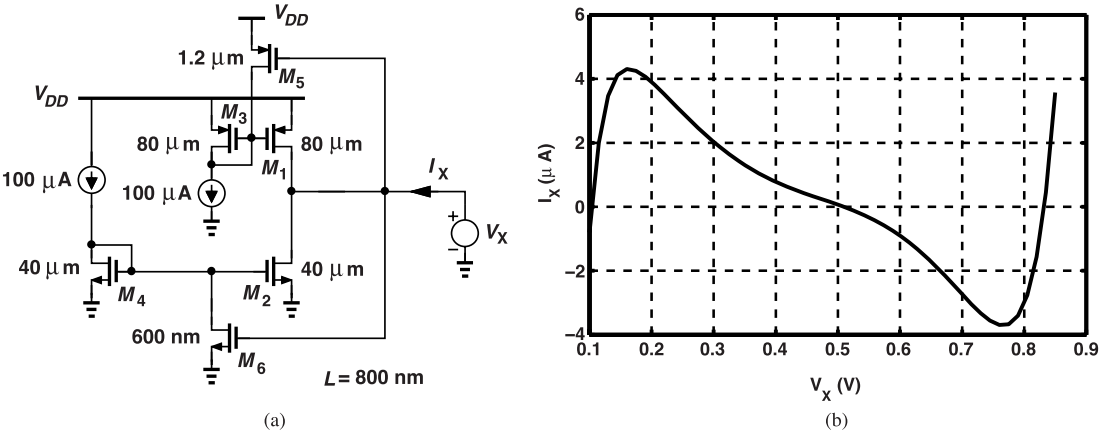


Figure 9.3 (a) Reduction of current variation by feedback, and (b) simulated current error.

the peaks around $V_X = 150$ mV and 750 mV grow. With the present dimensions, the deterministic mismatch between the up and down currents remains less than 4%.

We must now address two issues. First, the unscaled reference branches in Fig. 9.3(a) substantially raise the power consumption and area of the circuit. If we scale down these branches, then M_5 and M_6 must be made weaker, preferably by reducing their widths, but at the cost of higher random mismatches. As a compromise, the reference branches and these two transistor widths can be scaled down by a factor of two,[1] but we proceed without this scaling for now.

Second, we must study the efficacy of the feedback technique at other process and temperature corners. For example, for FF and $0\ ^\circ$C conditions, the correction becomes too strong, leading to an error of about $8\ \mu$A. This is because the circuit has been optimized at SS, $75\ ^\circ$C. We therefore select TT, $25\ ^\circ$C as our starting point and reach the optimum characteristic using $W_5 = 1\ \mu$m and $W_6 = 0.5\ \mu$m. As a result, the maximum current mismatch is equal to $4\ \mu$A in this case and $6\ \mu$A in the FF, $0\ ^\circ$C case.

It is instructive to examine the output noise current of the CP core that we have developed. With $V_X \approx 0.5$ V in Fig. 9.3(a), we can perform a small-signal noise simulation, obtaining the profile shown in Fig. 9.4. Let us make some observations. Dominated by the flicker noise of M_2 and M_4, the noise current is around 40 pA/$\sqrt{\text{Hz}}$ at 10 kHz. It falls to about 7 pA/$\sqrt{\text{Hz}}$ at high frequencies, with nearly equal contributions by the four main transistors. Since the flicker corner frequency is far below 10 MHz, the CP switching action at 20 MHz does not alias this noise. Based on our derivations in Chapter 8, the reader is encouraged to assess the impact of the CP noise on the PLL performance.

Switching Behavior The very long transistors used in the CP core do raise concern regarding the switching speed, especially if the two current sources are controlled through their gates. Considering the conceptual arrangement shown in Fig. 9.5, we wish to estimate the time constant at node X when S_1 turns on. Even if the on-resistance of the switch is negligible, the relatively high small-signal resistance of M_3 and the large gate capacitance of M_1 and M_3 yield a slow response. Specifically, a gate oxide capacitance of 15 fF/μm^2 translates to $2 \times 80 \times 0.8 \times 15\,\text{fF} \times (2/3) = 1.28$ pF for the capacitance at X, where the factor $2/3$ accounts for C_{GS} in the saturation region. Since M_3 is biased in weak inversion,[2] we estimate its transconductance as

[1] The feedback transistors do not act as current mirrors (i.e., do not directly copy currents), but we surmise that choosing $L_5 = L_1$, and $L_6 = L_2$ provides better tracking across corners.

[2] In weak inversion, the transistor drain current is approximately an exponential function of its gate-source voltage, and hence $g_m \approx nI_D/V_T$.

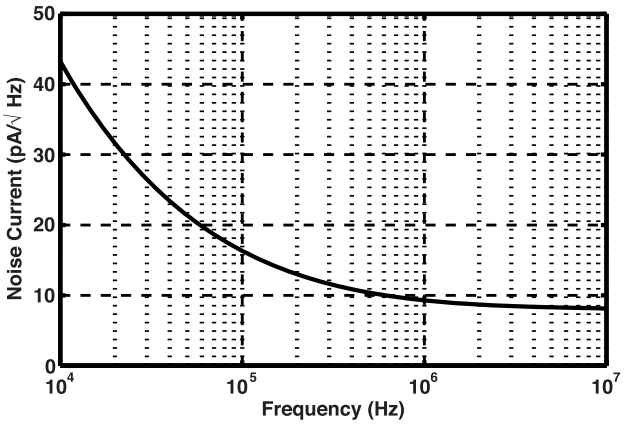


Figure 9.4 Output noise current of the circuit in Fig. 9.3(a).

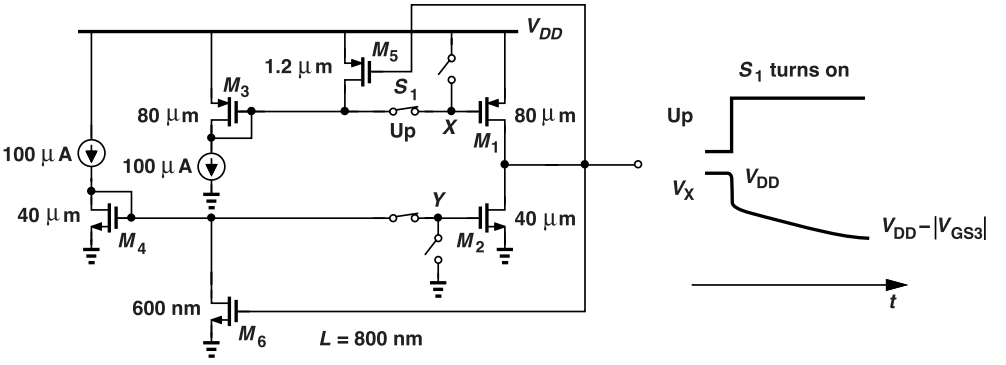


Figure 9.5 Addition of switches to the circuit of Fig. 9.3(a).

nI_D/V_T, where $n \approx 1.5$ and $V_T = kT/q$. It follows that $1/g_{m3} \approx 390\ \Omega$ and hence the time constant at X is around 0.5 ns. Thus, each time S_1 turns, X would require a few nanoseconds to settle.

Example 9.1

A PFD design in 40-nm technology produces up and down pulses with a minimum width of about 100 ps. Explain what happens if such a PFD drives the foregoing charge pump.

Solution

Owing to the CP's long time constant, nodes X and Y do not swing enough to turn on M_1 and M_2, respectively. That is, the actual charge pump current delivered to the loop filter is very small (Fig. 9.6), possibly making the loop unstable.

It is possible to increase the widths of the PFD output pulses by adding a delay in its reset path (why?), but, as explained in Chapter 8, most PLL imperfections intensify as a result.

This study suggests that the gate-switched CP topology cannot provide both a small current mismatch and fast switching. We should remark that the large transistors in this design also introduce considerable clock feedthrough and charge injection.

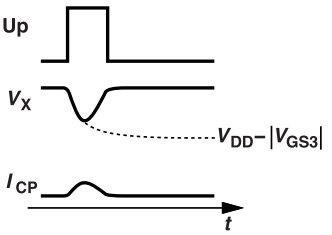


Figure 9.6 Poor turn-on of CP due to excessively narrow up pulse.

9.3.2 Second CP Design

In view of the gate-switched CP issues, we now try source switching. As illustrated in Fig. 9.7(a) for the top half of the charge pump, we insert a wide, short PMOS switch in the source of M_1. With these dimensions, M_7

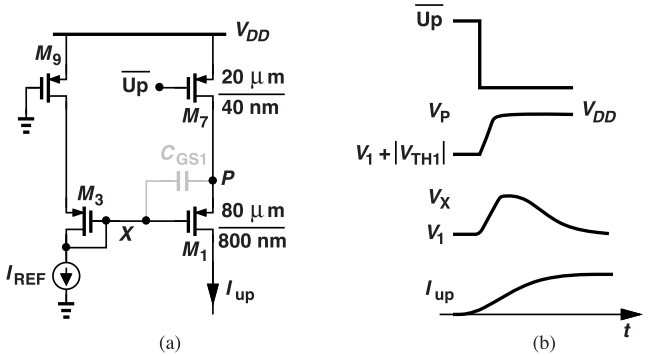


Figure 9.7 (a) PMOS section of a switched-source CP, and (b) its transient waveforms.

sustains a $|V_{DS}|$ of about 5 mV at $I_D = 100\ \mu$A, consuming negligible headroom. Transistor M_9 sustains the same voltage drop, allowing accurate mirroring of I_{REF} onto M_1. The circuit is nonetheless slow. Suppose Up goes down to turn on M_7. Then, V_P rapidly rises from $V_1 + |V_{TH1}|$ to near V_{DD}, where V_1 is the steady-state voltage at X and equal to $V_{DD} - |V_{DS9}| - |V_{GS3}|$ [Fig. 9.7(b)]. Since M_1 is initially off, the change in V_P travels through C_{GS1}, causing a jump in V_X. Now, V_X must return to V_1 with approximately the same time constant as calculated in the previous section for the first CP design, $\tau \approx 0.5$ ns. As a result, I_{up} rises slowly.

Can we tie a capacitor from X in Fig. 9.7 to V_{DD} so as to suppress the jump in V_X? Such a capacitor is effective only if it is much greater than C_{GS1}, e.g., around 100 pF; if less, it also increases the time constant at this node, thereby producing a long tail in I_{up}.

We conclude from this brief study that a source-switched CP is also slow if it must meet our 5% current error criterion. We thus seek other solutions.

9.3.3 Third CP Design

With gate and source switching having provided disappointing results, we consider a drain-switched CP structure and try the design shown in Fig. 9.8(a). (Note that the wide switches consume a negligible voltage headroom.) Suppose $V_{cont} = V_0$. When M_7 turns on, V_P falls from V_{DD} to approximately V_0, and this change couples through C_{GD1} to X, creating a *smaller* disturbance than in Fig. 9.7. This is because, as depicted in Fig. 9.8(b), the change in V_P is attenuated by approximately a factor of $C_{GD1}/(C_{GD1} + C_{GS1} + C_{GS3})$ as it reaches X. However, the clock feedthrough of M_7 and M_8 leads to a large surge in I_{CP} because $C_{GD7} \neq C_{GD8}$. Since it is difficult to match the clock feedthrough and charge injection of M_7 and M_8, a high ripple results. Let us then try one more design.

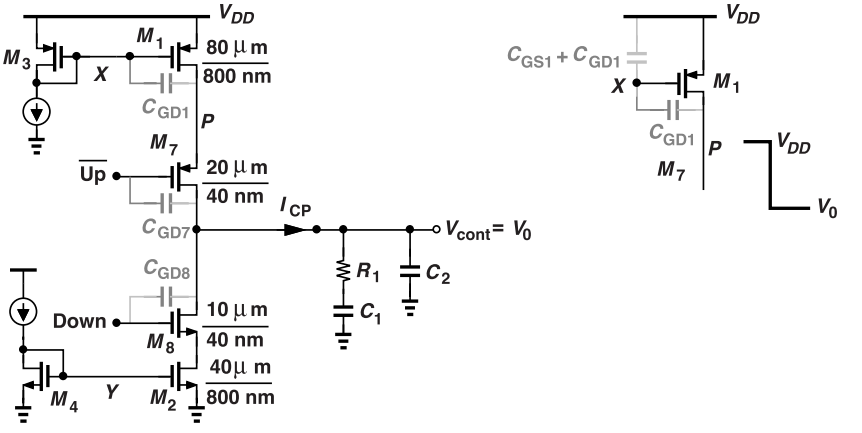


Figure 9.8 (a) Switched-drain CP design, and (b) effect of change in V_P at X.

9.3.4 Fourth CP Design

For fast gate switching, we must employ transistors with a short channel length. Let us investigate such a design with $L = 40$ nm. Figure 9.9(a) shows the new topology and Fig. 9.9(b) its I-V characteristic. Owing to severe channel-length modulation, the current varies by more than 100% as V_X goes from 100 mV to 850 mV. The individual up and down currents are plotted in Fig. 9.9(c).

We know from Chapter 7 that an integrator interposed between the CP and the VCO can suppress the effect of channel-length modulation, thus relaxing the mismatch requirement. Let us then continue with the simple charge pump in Fig. 9.9(a) and utilize an integrator if necessary. Figure 9.10(a) shows the CP with the switches and Fig. 9.10(b) plots the gate voltages of M_1 and M_2 in the SS, 75 °C corner as they turn on and off. We observe a long time constant as V_X falls and V_Y rises. This is primarily because M_5 and M_6 suffer from a small overdrive voltage and hence a high on-resistance. As a result, I_{up} and I_{down} exhibit considerable misalignment and different peak values even if the up and down pulses are aligned [Fig. 9.10(c)].

How can we speed up the waveforms at X and Y? Since these voltages must change in the same direction as $\overline{\text{Up}}$ and Down, respectively, we ask whether the latter signals can "help" the former. To this end, we add a feedforward capacitor from $\overline{\text{Up}}$ to X and another from Down to Y [Fig. 9.11(a)]. The CP currents now behave as shown in Fig. 9.11(c), displaying closer alignment. In practice, we realize the 10-fF and 4-fF capacitors by PMOS and NMOS devices, respectively, to ensure tracking with the switches across PVT corners.

9.3.5 PFD/CP Interface

The CP of Fig. 9.11(a) requires up and down pulses as well as their complements. We configure the PFD/CP interface as depicted in Fig. 9.12, where the pass gate G_4 is added to match the delay of G_3, and G_5 and G_6 serve as buffers. For all transistors, $W = 1$ μm and $L = 40$ nm, except for G_5 and G_6, where $W = 2$ μm. Since the capacitances seen at the inputs of G_3 and G_4 are different (why?), inverters G_1 and G_2 are also inserted as buffers. With these choices, the CP up and down currents remain fairly aligned.

9.4 Behavioral Simulations of PLL

9.4.1 Loop Simplification

The simulation of PLLs typically takes a very long time. Consider the case at hand, noting the vastly different time scales in the design: the loop requires about 100 input cycles ($= 5$ μs) to lock while the oscillator runs at 2.4 GHz (a period of about 417 ps). Transient simulations must therefore proceed with a time step well below the output period and continue for about 5 μs, e.g., for 100,000 points.

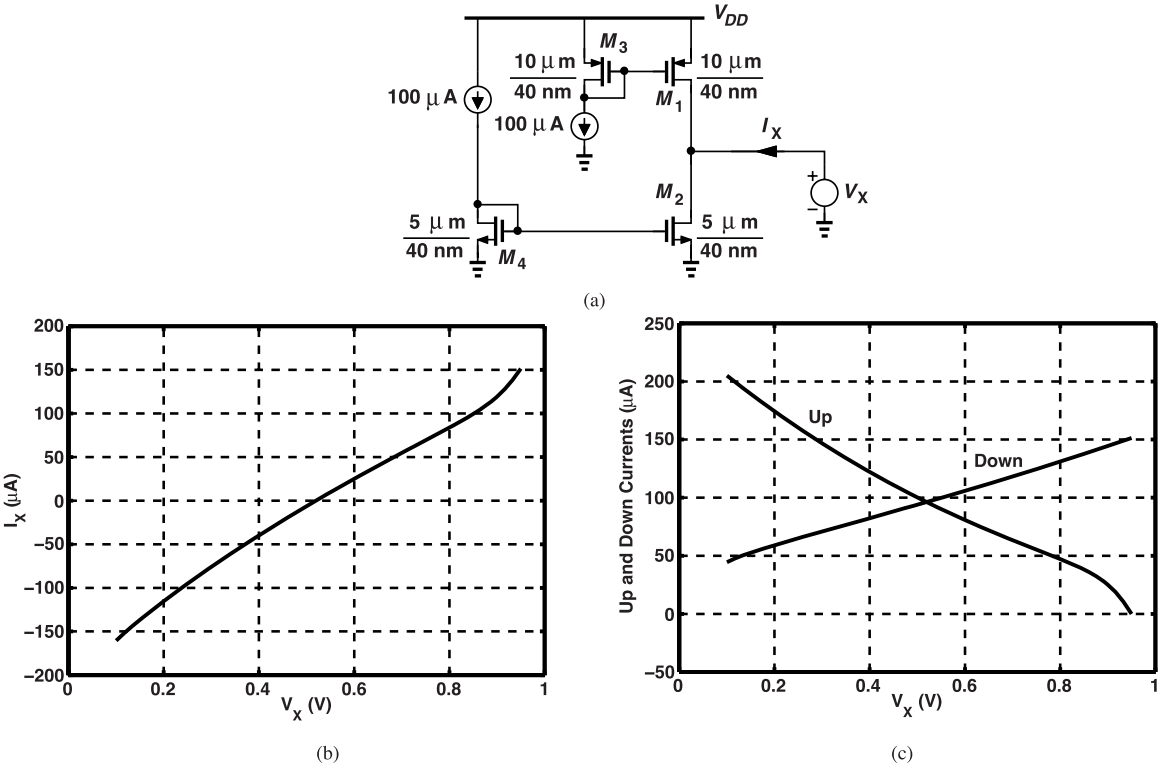


Figure 9.9 (a) CP core using short-channel devices, (b) simulated error current, and (c) simulated up and down currents.

Another equally serious reason for the long simulation time is that the overall loop incorporates a large number of transistors. For example, the PFD/CP cascade developed so far contains about 60 transistors, and the three-stage ring VCO designed in Chapter 3 has 30. Moreover, the feedback divider will also employ many devices.

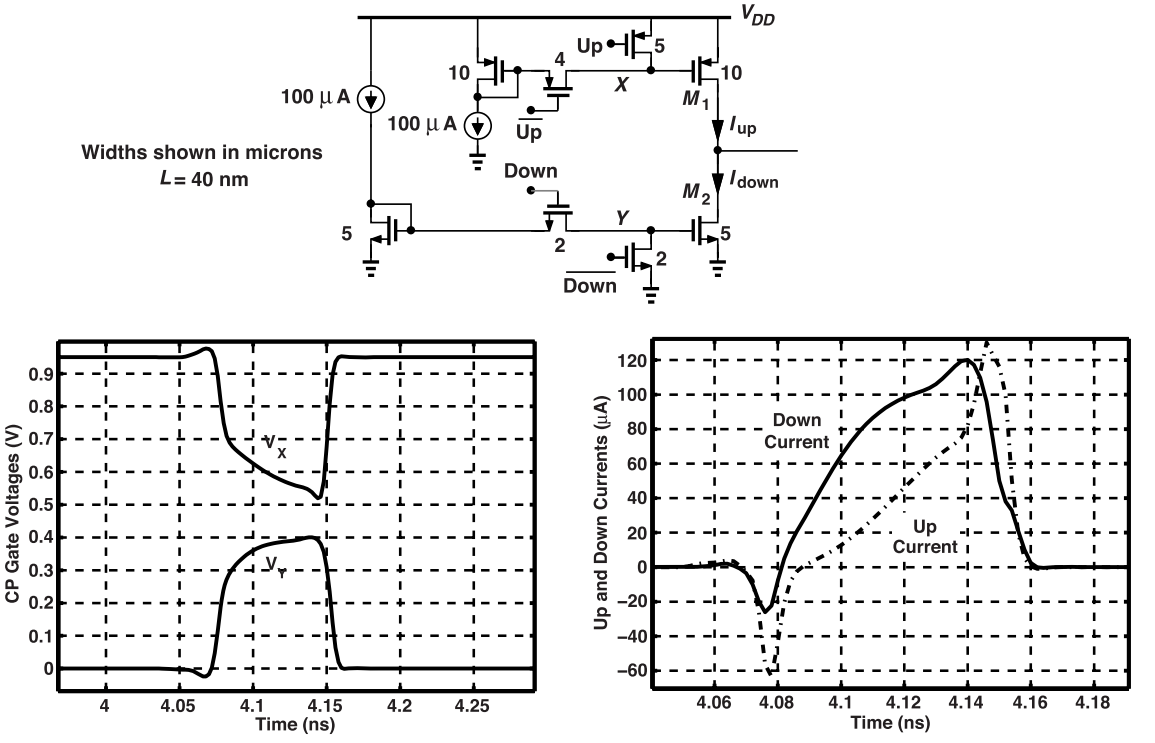
Our general approach to designing PLLs is as follows. We have thus far completed the VCO, the loop filter, and the PFD/CP cascade. Our next task is to study the loop dynamics by means of transient simulations, anticipating tens or hundreds of trials as we optimize the bandwidth and reduce the ripple on the control voltage. For these studies, we replace the VCO and the divider with a simple "behavioral" model so as to reduce the simulation time—from hours to minutes.

Another key point leading to faster simulations is to avoid signal sources that have sharp changes, and more specifically, discontinuous derivatives. For example, if the PLL reference is represented by a periodic pulse [Fig. 9.13(a)], then the simulator is forced to perform detailed calculations around t_1 and t_2, which may not be necessary. We instead model the reference by a sinusoid and convert it to a square wave by an inverter.

These methods afford fast, efficient simulations. After the loop design reaches a satisfactory state, we can perform a few final simulations with all of the building blocks represented by actual transistor-level implementations.

The reader may wonder why we do not use a behavioral model for the PFD/CP chain. This is because, as seen in Chapters 7 and 8, the nonidealities introduced by this circuit lead to ripple and hence a trade-off with the loop bandwidth and VCO phase noise suppression.

Let us consider a behavioral model for the VCO, bearing in mind that sinusoidal sources allow faster



(b) (c)

Figure 9.10 (a) Switched-gate CP using short-channel devices, (b) simulated gate voltages, and (c) simulated up and down currents.

simulations. The VCO output is expressed as

$$V_{VCO}(t) = V_0 + V_0 \cos(\omega_c t + \int K_{VCO} V_{cont} dt), \tag{9.7}$$

and such a model is available in Cadence as a Verilog A function. We must follow this stage with a divider. Rather than implement a stand-alone divider model, we recognize from Fig. 9.14(a) that we simply need an output of the form

$$V_{div}(t) = V_1 + V_1 \cos \frac{\omega_c t + \int K_{VCO} V_{cont} dt}{M}, \tag{9.8}$$

which can be viewed as the output of a VCO having a center frequency of ω_c/M and a gain of K_{VCO}/M. This observation leads us to the simple arrangement shown in Fig. 9.14(b), where the divider is also modeled by a VCO.

Returning to our objectives, we note that our analysis is presently concerned with the loop dynamics and the effect of the ripple. Can we eliminate the VCO model in Fig. 9.14(b)? The incentive would be to avoid high-frequency signal excursions, which slow down the simulation. In other words, can the simplified architecture shown in Fig. 9.14(c) provide all of the information that we seek? Two points help us appreciate the beauty of this approach: (1) the loop dynamics do not depend on the VCO block in Fig. 9.14(b) (why?). (2) The sidebands due to the ripple can be observed at the divider output. The VCO output would have sidebands that are $20 \log M$ dB higher in magnitude (Chapter 2). Note that this loop contains only 20-MHz waveforms and lends itself to much faster simulations.

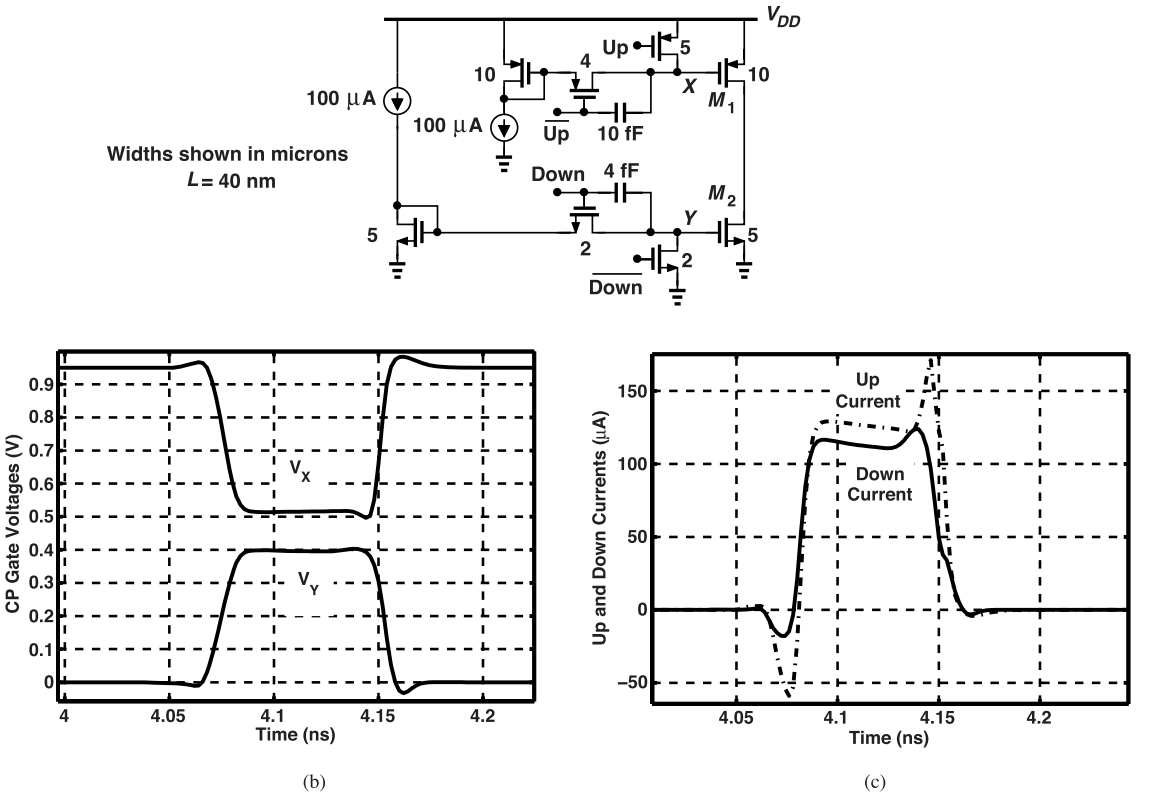


Figure 9.11 (a) CP using capacitive clock feedforward to improve settling, (b) gate voltage waveforms, and (c) up and down currents.

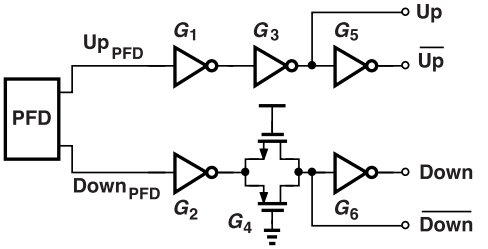
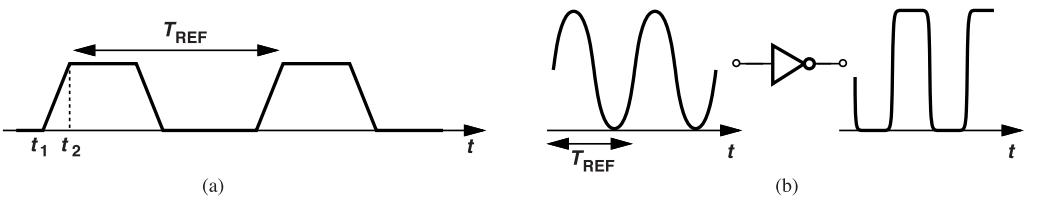


Figure 9.12 Interface between PFD and CP.

We should point out that the reference spurs due to the ripple lie at $f = 0$ and $f = 2f_{REF}$ at the divider output. The latter creates two difficulties. First, it can be confused with the second harmonic. Fortunately, the sinusoidal Verilog A divider model generates no such harmonic. Second, the component at $2f_{REF}$ does *not* manifest itself as jitter because it simply repeats at twice the main component at $f = f_{REF}$, i.e., f_{REF} and $2f_{REF}$ form a periodic waveform. Thus, the jitter (in seconds) must be indirectly obtained from the relative magnitude of the component at $2f_{REF}$ in the frequency domain.

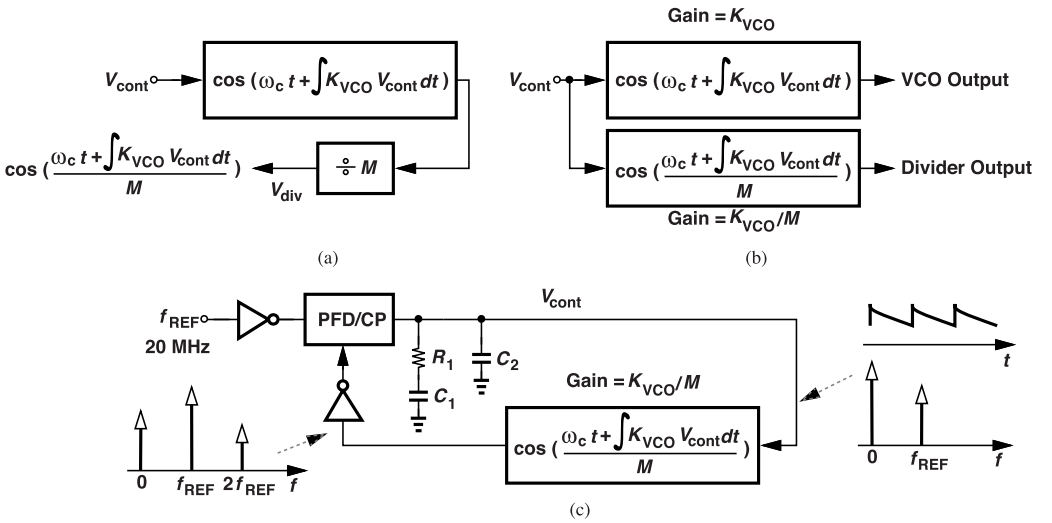
9.4.2 Loop Dynamics

The compact PLL model of Fig. 9.14(c) can be readily simulated in the time domain. From Section 9.1, we choose $R_1 = 35.9\ \text{k}\Omega$, $C_1 = 9.3$ pF, and $C_2 = 1.86$ pF. Figure 9.15(a) plots V_{cont} as a function of time, indicating a lock time of about 4 μs, which is shorter than our rule of thumb of 100 input cycles (5 μs). Of



(a) (b)

Figure 9.13 (a) Using a square wave, or (b) a sine wave followed by an inverter.



(a) (b) (c)

Figure 9.14 (a) VCO/divider chain, (b) alternative view, and (c) behavioral model of the actual loop.

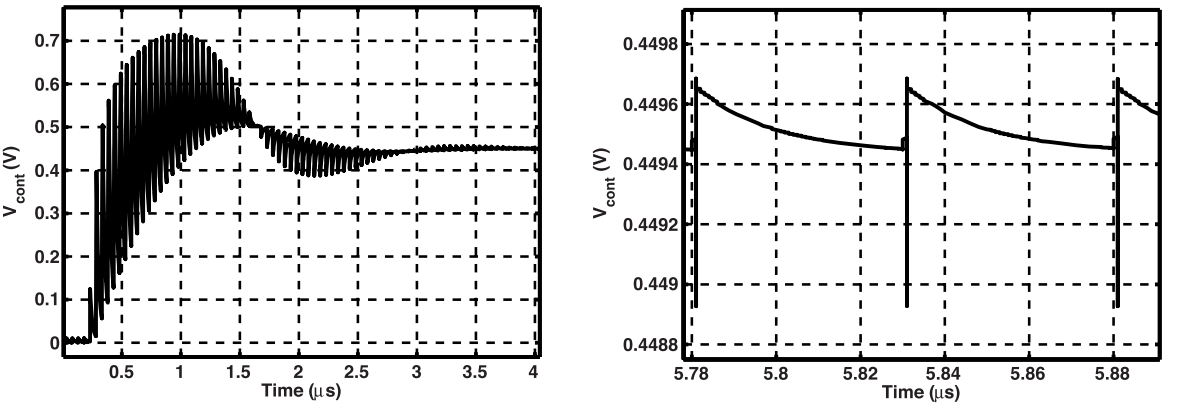
course, we must also examine the effect of the ripple and determine whether it is acceptable. The settling behavior in Fig. 9.15(a) does not contain much ringing, suggesting that $\zeta = 1$ is approximately satisfied. The static phase error in this case is a few picoseconds. A close-up of the ripple is shown in Fig. 9.15(b).

9.4.3 Effect of Ripple

It is instructive to examine the ripple waveform in Fig. 9.15(b) closely. The initial glitch arises from the turn-on and turn-off of the charge pump and lasts about 5 gate delays. But, as illustrated in Fig. 9.16, the subsequent discharge brings V_{cont} $(= V_{C2})$ *below* V_{C1}. Why does this happen? The discharge in fact arises from the net leakage current of the charge pump when its devices are off, a noticeable effect here because of the transistors' short channel lengths. In response to this leakage, the loop develops a phase offset so that one CP current source turns on earlier, thereby replenishing the charge that the loop filter has lost in the previous reference cycle. We then face two issues: (1) due to the phase offset, the initial glitch is wider than 5 gate delays, carrying more energy, and (2) the discharge waveform of V_{cont} leads to additional VCO phase drift.

Let us quantify the effect of the above two imperfections at the VCO output. For the initial glitch, the reader can show that narrow periodic rectangular pulses of width ΔT and height V_0 appearing on V_{cont} produce sidebands with a normalized magnitude equal to $V_0 \Delta T K_{VCO}/(2\pi)$, or, more generally, equal to the area under the pulse times $K_{VCO}/(2\pi)$. In the ripple waveform of Fig. 9.15(a), the glitch area is about 5×10^{-14} V·s, yielding a sideband magnitude of -94 dBc at the VCO output. The initial glitch is thus negligible.

For the discharging sections of V_{cont}, we approximate them by a sawtooth waveform and note that the Fourier series of a sawtooth waveform of height V_1 contains a first harmonic with a peak amplitude of V_1/π. In Fig. 9.15(b), $V_1 \approx 0.2$ mV, and the corresponding sideband magnitude is equal to $(V_1/\pi)K_{VCO}/(2\omega_{REF}) =$



(a) (b)

Figure 9.15 (a) Simulated transient behavior of the PLL, and (b) its close-up view.

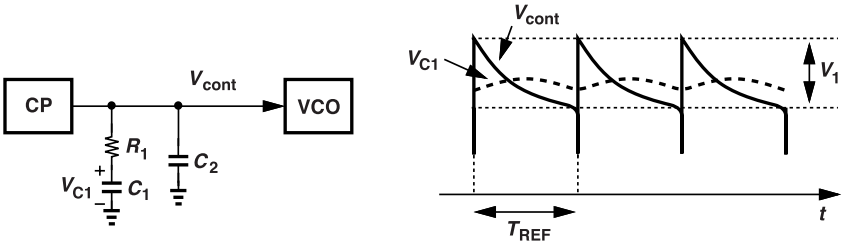


Figure 9.16 PLL control voltage and capacitor voltage waveforms.

-64 dBc. In the worst case, namely, with $V_{cont} \approx 0.85$ V, the sawtooth ripple amplitude increases to 1 mV_{pp} because the down current source experiences a greater V_{DS} and a higher leakage. The resulting sidebands are around -50 dBc and, from simulations, the phase offset is about 350 ps.

Example 9.2

Determine the deterministic jitter resulting from the sawtooth ripple.

Solution

As illustrated in Fig. 9.17, we approximate the ripple by a linear ramp that lasts T_{REF} seconds. Here, $\alpha =$

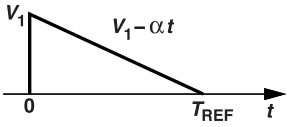


Figure 9.17 Approximation of ripple.

V_1/T_{REF}. The VCO phase departure in one input period is given by

$$\Delta\phi = \int_0^{T_{REF}} K_{VCO} V_{cont} dt \tag{9.9}$$

$$= \int_0^{T_{REF}} K_{VCO}(V_1 - \alpha t) dt \tag{9.10}$$

$$= \frac{V_1 K_{VCO} T_{REF}}{2} \text{ rad.} \tag{9.11}$$

With $V_1 = 1$ mV (in the worst case of $V_{cont} = 0.85$ V), $K_{VCO} = 400$ MHz/V, and $T_{VCO} = 417$ ps, the peak-to-peak jitter is equal to $[V_1 K_{VCO} T_{REF}/(4\pi)]T_{VCO} = 4.2$ ps at the VCO output. This value satisfies our design requirement.

We conclude from the foregoing calculations that the simple charge pump shown in Fig. 9.11(a) suffices for our design specifications. In more demanding applications, additional CP techniques may be necessary (Chapter 8).

9.5 Simulation of the PLL Transfer Function

As we have seen in previous chapters, the bandwidth of PLLs plays a critical role in the performance. We therefore wish to assess, by simulations, the validity of our simple bandwidth calculations in Section 9.1. We consider three approaches here.

9.5.1 One-Pole Approximation

The first approach approximates the settling behavior of the PLL by a one-pole response, and hence derives the closed-loop time constant, τ. The closed-loop bandwidth is then equal to $1/(2\pi\tau)$ hertz. As shown in Fig. 9.18, suppose V_{cont} begins from V_0 and settles to V_1. The time necessary for V_{cont} to reach halfway,

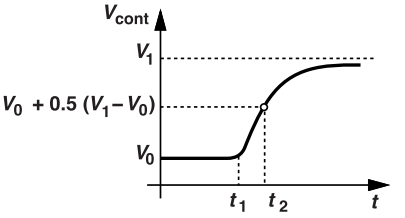


Figure 9.18 Simple method of calculating loop time constant.

$V_0 + 0.5(V_1 - V_0)$, is given by $t_2 - t_1 = 0.69\tau$ (why?) if K_{VCO} remains constant across this range. Note that the loop is locked before t_1, and then we apply a stimulus that requires V_{cont} to change. For example, the input frequency is raised by a small amount at t_1.

To check the validity of this approximation, we simulate the PLL with different values of V_0 and determine whether the τ values thus obtained are the same. Simulations indicate that this is not the case for the PLL under study in this chapter. This is because a ζ of unity leads to *coincident* poles with $\omega_{p1} = \omega_{p2} = 2\omega_z$, disallowing the one-pole approximation.

9.5.2 Use of Input FM Source

The second approach is more rigorous and computes the actual closed-loop transfer function from the input to the output, a technique also applicable to laboratory measurements. Recall that the transfer function signifies how much of the input phase fluctuation appears at the output. We must then devise a method of modulating the input phase. A simple technique is to modulate the supply voltage of the input inverter in Fig. 9.14(c) and hence its output phase. For example, we can place a 10-mV 1-MHz sinusoidal voltage source in series with this inverter's supply to create input phase modulation at this rate. Alternatively, we can select, in Cadence, a sinusoidal voltage source for the reference input and specify how its *frequency* is modulated. Such a source

provides an output expressed as

$$V_{FM}(t) = V_a + V_a \cos[\omega_{REF} t + K \int \cos \omega_m t dt], \tag{9.12}$$

where K and ω_m are the FM parameters. Since we are interested in input *phase* fluctuations, we rewrite $V_{FM}(t)$ as

$$V_{FM}(t) = V_a + V_a \cos\left(\omega_{REF} t + \frac{K}{\omega_m} \sin \omega_m t\right), \tag{9.13}$$

recognizing that the input peak-to-peak phase variation is equal to $2K/\omega_m$ radians. For example, if ω_m is chosen equal to the closed-loop -3-dB bandwidth, then the output phase variation, i.e., the output (deterministic) jitter, is $(\sqrt{2}/2)(2K/\omega_m)$. We thus perform a transient simulation with a frequency-modulated reference having a certain ω_m, wait until the loop locks, and then divide the output jitter by the input jitter to obtain the magnitude of the transfer function at $\omega = \omega_m$. Since our calculations deal with jitter (in seconds), the "output" can be simply V_{div} in Fig. 9.14(c).[3] Multiple simulations with different values of ω_m yield the overall transfer function, predicting the bandwidth and even jitter peaking.

As an example, Fig. 9.19 plots the output eye diagram of our PLL in response to a peak-to-peak input jitter of 31 ps at $\omega_m = 2\pi(1\ \text{MHz})$. (The transient simulation is long enough to provide 1 μs of the steady state

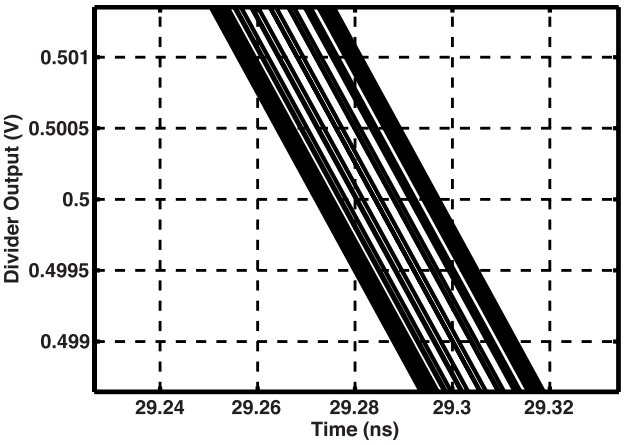


Figure 9.19 PLL output eye diagram for 31-ps,pp input phase modulation at 1 MHz.

after the loop locks.) Note that this jitter is measured at the divider output in Fig. 9.14(c). From the output jitter of 25 ps,pp, we conclude that the PLL attenuates the input by about 1.9 dB at this jitter frequency.

Based on this approach, we construct the input-output response of the PLL as shown in Fig. 9.20. We observe that, even though the zero is located at $\omega_z = (35.9\ \text{k}\Omega \times 9.3\ \text{pF})^{-1} = 2\pi(477\ \text{kHz})$, the response begins to exhibit peaking from 100 kHz, reaching a maximum of 3 dB at 500 kHz. As mentioned in Section 9.1, this peaking partially occurs because the pole formed by C_2 and R_1 is not far from ω_u in this case. The -3-dB bandwidth is around 1.2 MHz, about half of the theoretical value, $\omega_{-3dB} = 2.5\omega_n = 2\pi(2.4\ \text{MHz})$. This is a serious issue and is resolved by the third approach described below.

This study merits a few remarks. First, for input phase modulation at 100 kHz, the simulation must run for at least 15 μs, demonstrating the advantage of the simplified loop design described in Section 9.4.1. Second, to obtain more accurate results, we can examine the *spectra* of the input and output signals rather than their eye diagrams. As illustrated in Fig. 9.21, both spectra exhibit sidebands due to phase modulation. If Δ_1 and Δ_2 are expressed in decibels, then $\Delta_2 + 20 \log 120 - \Delta_1$ yields the magnitude of the PLL response (from

[3] Unlike the jitter due to the ripple on the control voltage, this jitter is observable at the divider output.

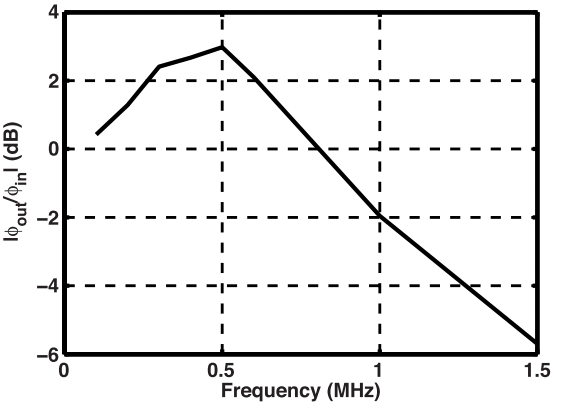


Figure 9.20 Response of PLL to input phase modulation.

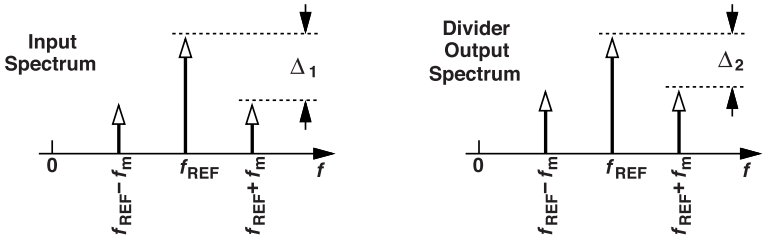


Figure 9.21 Input and divider output spectra in response to phase modulation.

ϕ_{REF} to ϕ_{out}) at $f = f_m$. (Recall that our divide ratio is equal to 120.)

9.5.3 Use of Random Phase Modulation

The third approach obtains the entire transfer function in one simulation and proves the most efficient and accurate. Let us insert a voltage source having a white noise spectrum in series with the supply voltage of the input inverter in Fig. 9.14(c), as conceptually depicted in Fig. 9.22(a). We observed in Chapter 2 that this noise directly modulates the delay of the inverter and hence its output phase:

$$V_{inv}(t) = V_a + V_a \cos(\omega_{REF} t + K_{DD} V_n), \tag{9.14}$$

where K_{DD} denotes the inverter's supply sensitivity and is expressed in rad/V. The phase-modulated input then travels through the PLL, experiencing its transfer function and revealing the -3-dB bandwidth and any peaking behavior.

This approach can be realized in Cadence by performing "periodic steady state" (pss) and "periodic noise" (pnoise) analyses. For the former, we must specify a "stabilization" time at least equal to the PLL's settling time to ensure that the loop has reached the steady state. Even then, however, pss has difficulty with convergence, primarily because the cross-coupled NOR gates within the PFD cause state uncertainties for Cadence. This issue is resolved by attaching large (1-MΩ) resistors from all of the NOR output nodes to the ground [Fig. 9.22(b)].

The white-noise phase modulation of the 20-MHz reference is implemented as shown in Fig. 9.22(c). The RC network generates band-limited thermal noise and a voltage-dependent voltage source amplifies the noise voltage by a factor of 1000 so that the phase modulation far exceeds the intrinsic PLL phase noise.

Figure 9.23 plots the PLL output phase noise for $C_2 = 1.86$ pF and $C_2 = 0.93$ pF. Since the PLL's internal noise sources contribute much less phase noise, these profiles accurately reflect the input-output transfer

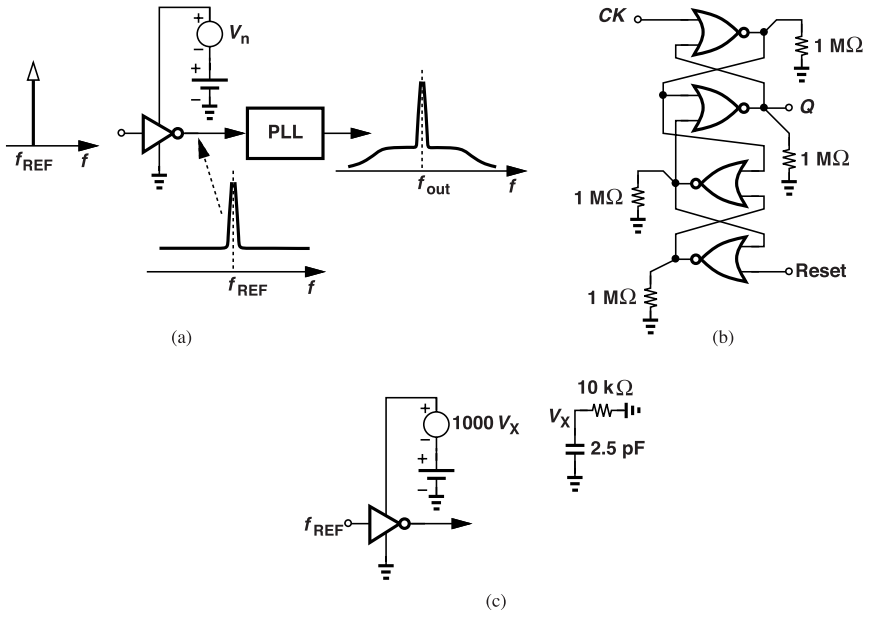


Figure 9.22 (a) Random phase modulation of PLL input, (b) addition of resistors to PFD for convergence of simulations, and (c) themal noise generator for phase modulation.

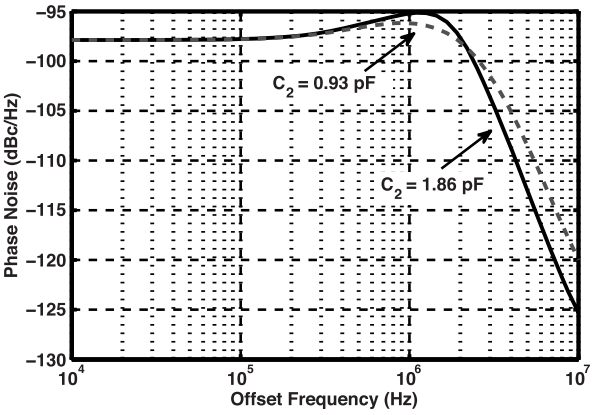


Figure 9.23 Simulated input-output response of the PLL for two different values of C_2.

function. In the case of $C_2 = 1.86$ pF, the peaking is about 2.7 dB, close to that predicted by the approach described in the previous section, with a -3-dB bandwidth of 2.5 MHz, also close to $2.5\omega_n = 2.4$ MHz. We had predicted that the peaking partially arises because $\omega_u \approx 2\pi(2\,\text{MHz})$ is not much lower than the third pole frequency of $2\pi(2.8\,\text{MHz})$. The plot for $C_2 = 0.93$ pF suggests a peaking of 1.7 dB and a -3-dB bandwidth of 2.8 MHz. A lower C_2 reduces the area under the phase noise profile and hence the random jitter due to the reference, but yields a greater ripple on the VCO control line and hence a higher deterministic jitter. Some optimization is therefore necessary if the reference phase noise is significant.

With the simplified PLL topology in Fig. 9.14(c), pss and pnoise simiulations take less than 1 minute, allowing many design iterations for optimum loop behavior.

9.6 Effect of VCO Phase Noise

9.6.1 VCO Phase Noise Model

We wish to include the VCO phase noise in the simplified model of Fig. 9.14(c), which employs a noiseless behavioral VCO model. From the VCO design in Chapter 3, we can express the VCO phase noise as

$$S_{\phi n} = \frac{\alpha}{f^3} + \frac{\beta}{f^2} \tag{9.15}$$

and select α and β for a best fit [Fig. 9.24(a)]. Now, we surmise that $S_{\phi n}$ can be "referred" to V_{cont} if it is

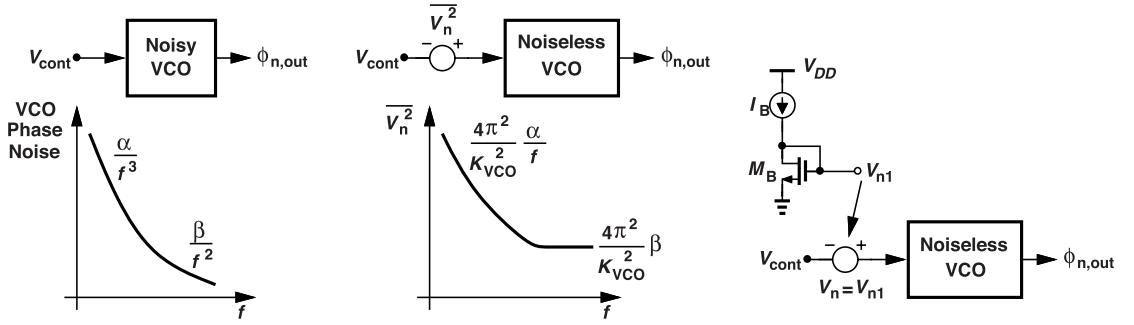


Figure 9.24 (a) Phase noise model of a VCO, (b) representation using a voltage source, and (c) generation of noise voltage.

divided by $K_{VCO}^2/(4\pi^2 f^2)$ [Fig. 9.24(b)]. The spectrum of the input-referred voltage is given by

$$V_n^2 = \frac{S_{\phi n}}{K_{VCO}^2} 4\pi^2 f^2 \tag{9.16}$$

$$= \frac{4\pi^2 \alpha}{K_{VCO}^2 f} + \frac{4\pi^2}{K_{VCO}^2}\beta. \tag{9.17}$$

To generate the two noise terms on the right-hand side, we note that they correspond to $1/f$ and white noise, respectively, and construct the circuit shown in Fig. 9.24(c). Here, the diod-connected transistor acts as a $1/f$ and thermal noise generator. The bias current and the transistor's dimensions are chosen so that the noise voltage produced by M_B matches the above expression for $\overline{V_n^2}$. We then insert a voltage-controlled voltage source, $V_n = V_{n1}$, in series with V_{cont}. In Fig. 9.14(c), V_n precedes the divider.

9.6.2 VCO Phase Noise Suppression

The PLL developed thus far suppresses the VCO phase noise up to an offset frequency of about $\omega_n = 2\pi(1\ \text{MHz})$ (because $\zeta \approx 1$). Let us return to the approximation described in Chapter 8 for calculating the jitter: as illustrated in Fig. 9.25, we assume a flat profile for $f < f_{BW}$ and compute the area under the shaped spectrum. However, since the phase noise of the oscillator under consideration here is dominated by flicker noise upconversion for offsets as high as tens of megahertz, we must revise our derivation. The total area is given by

$$\phi_{out,rms}^2 = 2S_1 \cdot f_{BW} + 2\int_{f_{BW}}^{\infty} \frac{\alpha}{f^3} df, \tag{9.18}$$

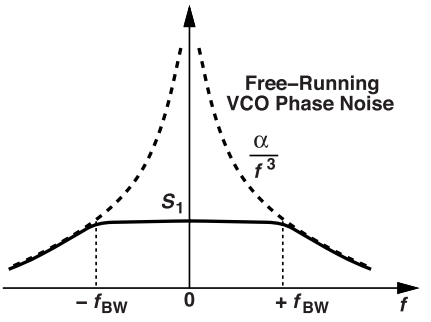


Figure 9.25 Approximation of closed-loop phase noise.

where S_1 denotes the free-running VCO phase noise at f_{BW} and is equal to α/f_{BW}^3. It follows that

$$\overline{\phi_{out,rms}^2} = 2S_1 \cdot f_{BW} + \frac{\alpha}{2f_{BW}^3} \tag{9.19}$$

$$= 3S_1 \cdot f_{BW}. \tag{9.20}$$

In comparison to the expression $4S_1 f_{BW}$ for thermal phase noise, flicker noise yields a factor of 3 because of the sharper drop of α/f^3. Of course, S_1 is higher in the presence of flicker noise. For the 2.4-GHz ring oscillator, $S_1 \approx -80$ dBc/Hz at 1-MHz offset, yielding $\phi_{out,rms} = 0.18$ rad. For an output period of 417 ps, this translates to a jitter of

$$\Delta t_{rms} = 11.6 \text{ ps, rms}. \tag{9.21}$$

Unfortunately, the present design does not meet our 2-ps random jitter specification.

Phase Noise Simulation Using the simplified loop of Fig. 9.14(c) and the free-running VCO noise modeling depicted in Fig. 9.24, we can simulate the closed-loop phase noise, arriving at the plots shown in Fig. 9.26. Since we examine the 20-MHz divider output here, the maximum frequency offset is 10 MHz.

These plots offer interesting insights. The highest profile corresponds to the model in Fig. 9.24 and is matched to that of the actual VCO designed in Chapter 3. The divider output phase noise is lower by $20 \log 120 = 41.6$ dB. The closed-loop profile is shown for $C_2 = 0.2C_1$ (the original design) and $C_2 = 0.1C_1$. Recall from Chapter 8 that the closed-loop phase noise reaches a peak at $\omega_n/\sqrt{3}$ that is 12 dB below that of the free-running phase noise if flicker noise is dominant. With $\omega_n/\sqrt{3} \approx 2\pi(577 \text{ kHz})$, we see a reasonable agreement between theory and simulations. We also observe that $C_2 = 0.1C_1$ yields a relatively flat peak from $\omega_n/\sqrt{3}$ to about $1.2\omega_n$, whereas $C_2 = 0.2C_1$ raises the peak by about 2.5 dB and pushes it out by about 700 kHz. That is, $C_2 = 0.2C_1$ leads to greater random jitter but a smaller ripple on the control line.

Jitter Reduction How can we reduce this random jitter? We can increase the loop bandwidth, but we must bear in mind that (1) ζ must not decrease, and (2) the ripple on the control voltage and hence the deterministic jitter must not become excessively large. Fortunately, we know from Example 9.2 that the latter is about 4 ps,pp, about half of the specification. Let us ponder this task as follows: with $\zeta = (R_1/2)\sqrt{I_p K_{VCO} C_1/(2\pi M)}$, $\omega_n = \sqrt{I_p K_{VCO}/(2\pi C_1 M)}$, and the loop bandwidth $\approx 2.1\omega_n$ (for $\zeta = 1$), we can increase the bandwidth by reducing C_1, but we must also increase I_p proportionally so that ζ remains unchanged. For example, if C_1 is halved, I_p must be doubled. Such a scenario, however, would *quadruple* the ripple because the charge-pump transistors must double in width so as to maintain the same output voltage compliance. Instead, we reduce C_1 by a factor of $\sqrt{2}$ and raise I_p by the same factor. With $C_1 = 6.6$ pF and $I_p = 140\ \mu$A, ω_n increases by a factor of $\sqrt{2}$, reaching $2\pi(1.34 \text{ MHz})$. We now have a loop bandwidth of $2.1\omega_n \approx 2\pi(2.8 \text{ MHz})$, slightly violating our condition of $\omega_u \approx \omega_{in}/10$. But let us proceed with this modification.

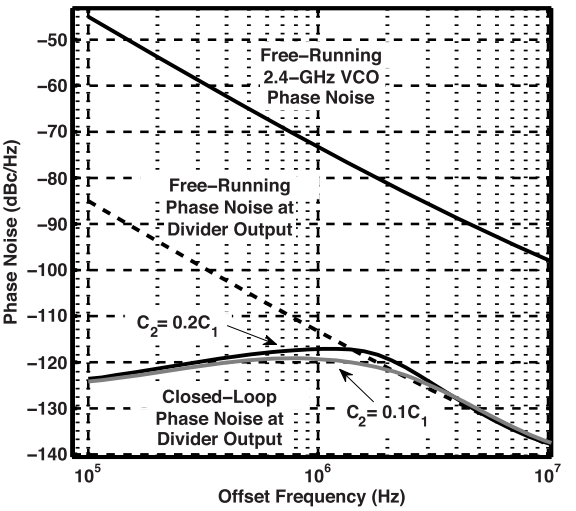


Figure 9.26 Free-running and closed-loop phase noise profiles.

We return to Eq. (9.20) and note that S_1 now denotes the phase noise at 1.34-MHz offset, which is less than that at 1-MHz offset by a factor of $(1.34)^3$ (why?). Since f_{BW} has increased by a factor of $\sqrt{2}$, $\overline{\phi_{out,rms}^2}$ falls by a factor of 1.7, yielding

$$\Delta t_{rms} = 6.8 \text{ ps, rms.} \tag{9.22}$$

To bring this value down to our specification of 2 ps,rms, we must draw upon the trade-off between the phase noise and power consumption in the VCO. That is, we employ linear scaling (Chapter 2) by a factor of $(6.8/2)^2 = 12$, raising the oscillator power to $(42\ \mu\text{A}) \times 12 \times 0.95\ \text{V} = 0.48$ mW, still a reasonable value and well below the budget. Of course, the other penalty is the large area occupied by the VCO.

Interestingly, the simple CP of Fig. 9.11 provides adequate performance without the need for an integrator. This is because, in this particular design, the loop filter capacitors are large enough to suppress the ripple.

In summary, our PLL design is now modified to $C_1 = 6.6$ pF, $C_2 = 1.3$ pF, $I_p = 140\ \mu$A, and $\omega_n = 2\pi(1.34\ \text{MHz})$, with the free-running VCO phase noise scaled down to $-80\ \text{dBc/Hz} - 10\log 19 = -93$ dBc/Hz at 1-MHz offset while drawing 0.48 mW.

9.7 Loop Filter Noise

Recall from Chapter 8 that the loop filter resistor's noise modulates the VCO frequency (Fig. 9.27). From our derivations in that chapter, we obtain a peak output phase noise of

$$S_{\phi n,p} = \frac{16kT\pi^2 M^2}{R_1 I_p^2} \tag{9.23}$$

at $\omega = \omega_n$ for $\zeta = 1$.

For our PLL, $M = 120$, $R_1 = 35.9\ \text{k}\Omega$, and $I_p = 140\ \mu$A, yielding

$$S_{\phi n,p} = -109\ \text{dBc/Hz} \tag{9.24}$$

at 1-MHz offset. The VCO contribution, on the other hand, is around -93 dBc/Hz at this offset (Section 9.6.2). The noise peaking due to R_1 is therefore negligible in this case. Of course, this issue must be revisited in more demanding applications.

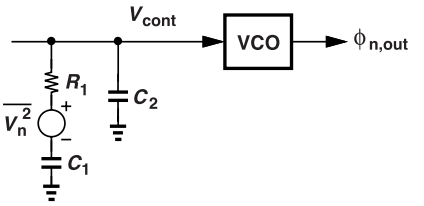


Figure 9.27 Phase noise due to loop filter noise.

9.8 Doubling the Reference Frequency

We can generally improve the performance of PLLs by increasing their reference frequency, f_{REF}. The reference is typically provided by a crystal oscillator, and the frequency of the crystal is specified by the application environment. For example, in a mobile phone, f_{REF} is around 20 MHz. Thus, we can increase f_{REF} only by interposing a frequency multiplier between the crystal oscillator and the PLL. In this section, we employ a frequency doubler and redesign the PLL accordingly.

9.8.1 Doubler Design

In order to double the frequency of a periodic waveform, we can XOR it with its delayed copy [Fig. 9.28(a)]. The delay, ΔT, determines the output pulsewidth and must be long enough for proper PFD operation.

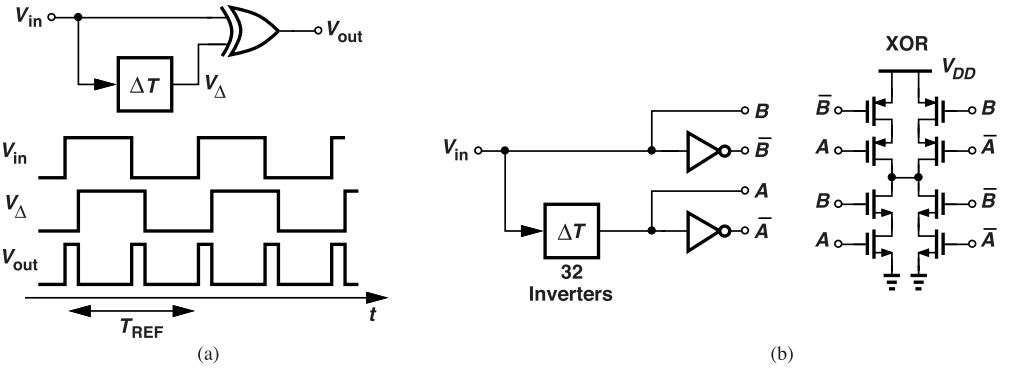


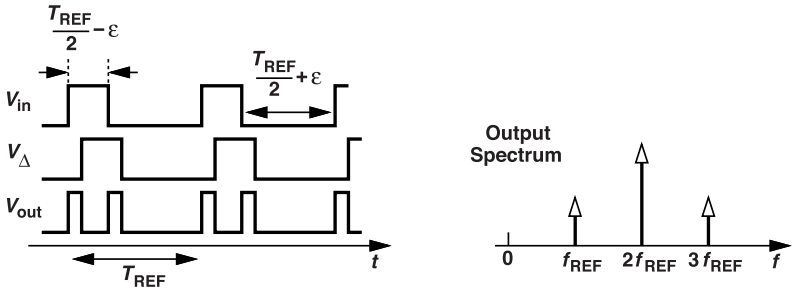
Figure 9.28 (a) Doubler circuit, and (b) its implementation.

Figure 9.28(b) shows the transistor-level implementation of the doubler. The XOR gate senses V_{in} $(= B)$, its delayed copy (A), and their complements. The delay stage includes 32 inverters to yield a delay of about 350 ps. For all inverters, $(W/L)_N = 500\ \text{nm}/40\ \text{nm}$ and $(W/L)_P = 1000\ \text{nm}/40\ \text{nm}$. In the XOR circuit, $W_N = 300\ \text{nm}$, $W_P = 1200\ \text{nm}$, and $L = 40\ \text{nm}$.

9.8.2 Frequency Doubling Issues

The doubling action entails two issues. First, it raises the phase noise, a natural effect here: a 1-ps edge displacement in V_{in} translates to a phase disturbance of $2\pi(1\ \text{ps}/T_{REF})$ radians at the input and a phase disturbance of $2\pi(1\ \text{ps}/T_{REF}/2) = 4\pi(1\ \text{ps}/T_{REF})$ radians at the output. However, (1) since the divide ratio is halved while $f_{out} = 2.4$ GHz, the output phase noise due to the reference does not change; and (2) since we plan to double the loop bandwidth, the *integrated* jitter due to the reference will rise. Nonetheless, in our design, this is negligible.

Second, the output exhibits unwanted frequency components if the input duty cycle departs from 50%. This can be seen from Fig. 9.29(a), where the high level is less than $T_{REF}/2$ by ϵ seconds. We recognize that V_{out} now repeats itself every T_{REF} seconds rather than every $T_{REF}/2$ seconds. In other words, the spectrum of V_{out} contains small, undesired components at f_{REF}, $3f_{REF}$, etc. [Fig. 9.29(b)].



(a) (b)

Figure 9.29 (a) Effect of duty cycle error on doubler output, and (b) the resulting spectrum.

We return to this point in Chapter 13 and prove that, with a duty cycle error of ϵ, the doubler's output first harmonic amplitude normalized to the second harmonic amplitude is equal to $\pi\epsilon/T_{REF}$ if $\Delta T \ll T_{REF}$. In this context, the first and third harmonics are often called "spurs" to indicate that they are undesirable.

Another method of quantifying the doubler imperfections is to examine its output jitter. Plotted in Fig. 9.30 is the doubler output eye diagram, exhibiting a peak-to-peak jitter of about 10 ps. The jitter arises because the odd-numbered low levels are slightly narrower than the even-numbered ones [Fig. 9.29(a)].

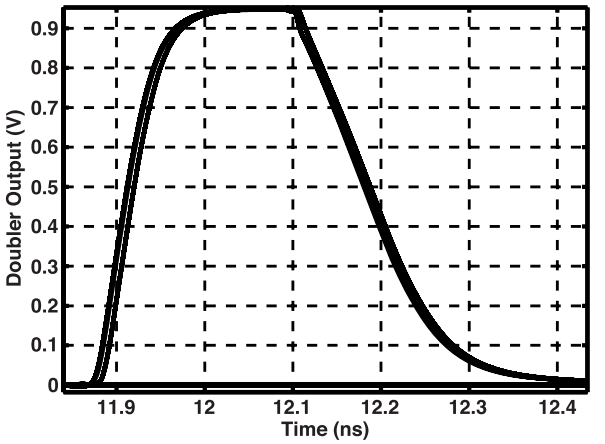


Figure 9.30 Doubler output eye diagram.

How much duty cycle error is tolerable? We make the following observations. First, the new PLL design preferably selects a loop bandwidth as large as $2\omega_{REF}/10$, providing only some attenuation for the first harmonic at ω_{REF}. The reader can show that this attenuation is about 14 dB for our design. Second, the normalized level of this spur rises by the feedback divide ratio, M, as it travels to the PLL output. For example, with $M = 120/2 = 60$, the input spur experiences a "gain" of $20\log 60 - 14$ dB ≈ 22 dB. Third, the tolerable spur level at the output depends on the application: if only the time-domain jitter matters, a magnitude of -35 to -40 dBc suffices, but in some RF applications, levels below -60 dBc are sought.

Let us target a PLL output spur level of -40 dBc and hence an input level of -62 dBc. This means that $\pi\epsilon/T_{REF}$ in Fig. 9.29(a) must be less than $1/1260$ and hence $\epsilon < 12.6$ ps. Such a tight requirement necessitates duty cycle correction prior to the doubler, which we describe in Chapter 11.

9.8.3 PLL Redesign with Doubled Reference

We repeat the calculations in Section 9.1 for $f_{REF} = 40$ MHz. We note from

$$2.1\sqrt{\frac{I_p K_{VCO}}{2\pi M C_1}} = \frac{2\pi \times (40\text{ MHz})}{10} \tag{9.25}$$

that M is now equal to 60 while the right hand side is doubled. Thus, C_1 must be halved:

$$C_1 = 4.7\text{ pF}. \tag{9.26}$$

We choose $C_2 = 0.94$ pF. Since both M and C_1 are halved, (9.1) implies that R_1 does not change:

$$R_1 = 35.9\text{ k}\Omega. \tag{9.27}$$

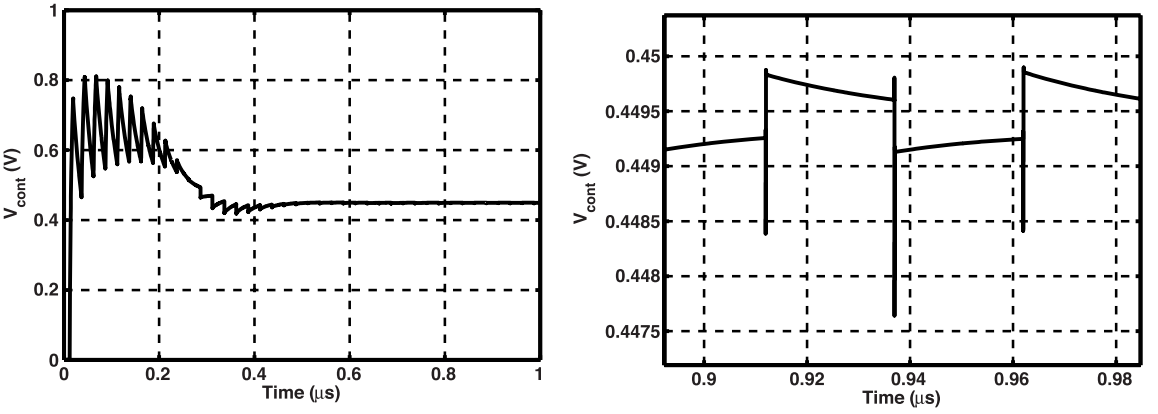
The charge pump current need not change and remains equal to 100 μA. The key benefit here is that the VCO phase noise suppression bandwidth, ω_n, is doubled and equal to $2\pi(1.9)$ MHz. Dominated by flicker noise, the VCO phase noise at this offset is a factor of 8 ($\equiv$ 9 dB) less than at 0.95 MHz. Since in Eq. (9.20), S_1 falls by this factor and f_{BW} is doubled, we obtain a jitter of

$$\Delta t_{rms} = 5.8\text{ ps, rms}. \tag{9.28}$$

If we apply to the VCO the linear scaling described in Section 9.6.2, this jitter decreases by another factor of $\sqrt{19} = 4.4$, falling to about 1.3 ps,rms. But we must also examine the ripple and ensure that the deterministic jitter is sufficiently small.

9.8.4 PLL Simulations

Using the simplified loop of Fig. 9.14(c), we simulate the doubler/PLL chain. Figure 9.31(a) plots the VCO control voltage as a function of time, revealing a lock time of about 0.5 μs.[4] The close-up in Fig. 9.31(b) shows a ripple amplitude of 0.6 mV_{pp}, about three times that of the previous design.



(a) (b)

Figure 9.31 (a) Simulated transient behavior of V_{cont} in the doubler/PLL chain, and (b) its close-up view.

[4]The large-signal lock behavior depends on the loop's initial conditions. Thus, the lock time with different initial conditions is different and does not linearly scale with the bandwidth.

It is instructive to compare the two different approaches that we have adopted in Section 9.6.2 and this section to reduce the VCO phase noise. In the former, the loop bandwidth was raised by a factor of $\sqrt{2}$ and the ripple (i.e., the deterministic jitter) increased by a factor of 2. In the latter, these factors are equal to 2 and 3, respectively, (the latter obtained from simulations).

9.9 Feedback Divider Design

We deal with the design of frequency dividers in Chapter 15, but to complete the PLL under consideration, we develop a simple $\div 120$ circuit here.

9.9.1 Topology Selection

Writing 120 as $8 \times 3 \times 5$, we recognize that the divider can be configured as a cascade of three $\div 2$ stages, one $\div 3$ stage, and one $\div 5$ stage (Fig. 9.32).

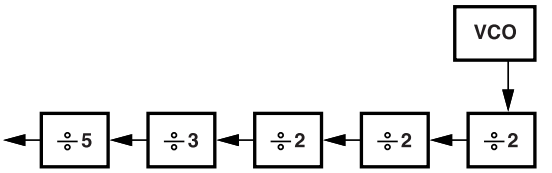


Figure 9.32 Feedback divider architecture.

A $\div 2$ circuit is readily realized as a master-slave D flipflop placed in a negative-feedback loop [Fig. 9.33(a)]. Here, the master output, A, begins to track its input, B, when CK goes high and is frozen when

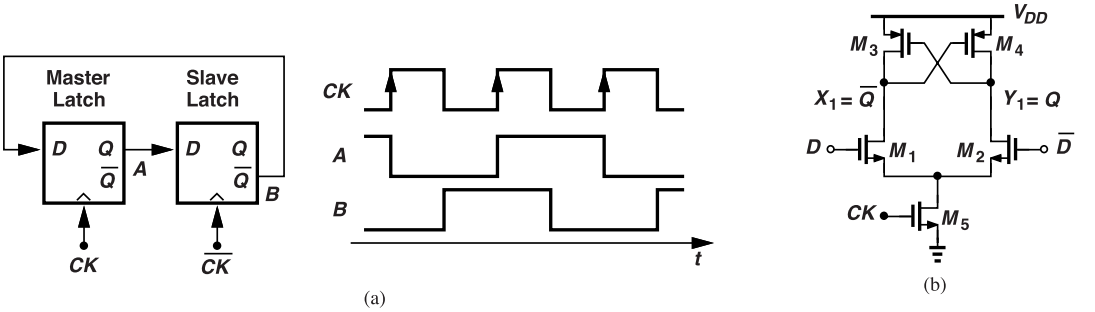


Figure 9.33 (a) Simple divide-by-2 circuit, and (b) possible realization of each latch.

CK goes low. Similarly, the slave output, B, follows $\overline{A}$ when CK falls and is stored when CK rises. Consequently, A and B toggle at half of the clock frequency while bearing a phase difference of 90°.

The latches in Fig. 9.33(a) must operate at 2.4 GHz and can have various implementations (Chapter 15). Shown in Fig. 9.33(b) is an example using complementary, rail-to-rail inputs and outputs. Suppose X_1 is high, M_3 is on, D is high, and CK rises. The series combination of M_1 and M_5 draws current from M_3, pulling X_1 down. As X_1 drops to $V_{DD} - |V_{THP}|$, M_4 turns on and raises Y_1. The regenerative action of M_3 and M_4 thus continues until X_1 falls to 0 and Y_1 rises to V_{DD}.

In the foregoing latch, the series combination of M_1 (or M_2) and M_5 must overcome M_3 (or M_4). The circuit is thus designed as "ratioed logic." We typically choose $W_1 = \cdots = W_4$ and $W_5 \approx 2W_{1,2}$.

The $\div 3$ circuit in Fig. 9.32 requires two D flipflops so as to support three states. As depicted in Fig. 9.34(a), we detect the state $AB = 11$ by a NAND gate and accordingly apply a zero to the input of FF_1. Thus, the circuit begins with $AB = 11$ and goes to 01, 10, and back to 11. The feedback NAND gate can be merged with the master latch in FF_1 to improve the speed and make the circuit more compact [Fig. 9.34(b)]. In this case, M_1, M_1', and M_5 must be wide enough to overcome M_3. The reader is encouraged to plot the waveform at node C.

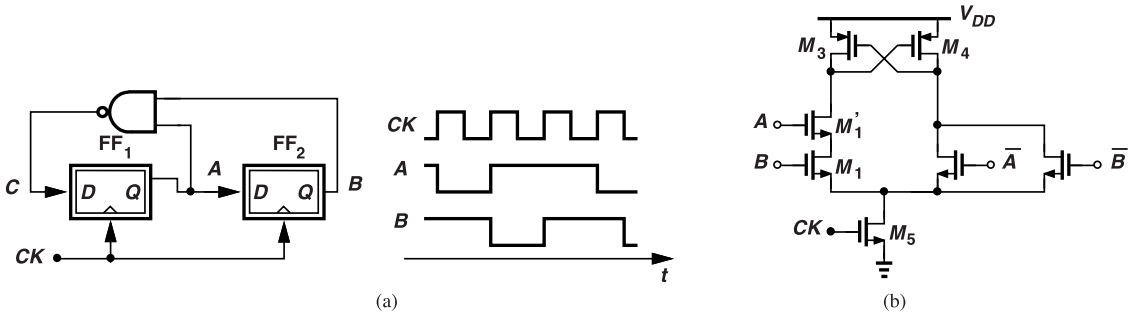


Figure 9.34 (a) Divide-by-3 circuit, and (b) merging the NAND gate with the first latch.

For a $\div 5$ circuit, we incorporate three flipflops with one feedback NAND gate (Fig. 9.35). The reader can readily show that this topology cycles through five states: $ABC = 100, 101, 001, 011$, and 110. As with the scheme in Fig. 9.34(b), the NAND gate can be merged with FF_1.

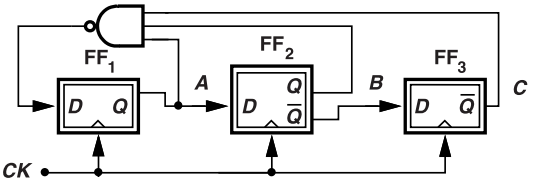


Figure 9.35 Divide-by-5 circuit.

9.9.2 Divider Circuit Design

The first $\div 2$ stage in Fig. 9.32 senses the VCO output and must operate at high speeds, demanding extensive simulations across corners. How fast must this circuit toggle? Recall from Chapter 3 that the ring oscillator of interest here runs as fast as 3.3 GHz in the FF, 0 °C corner—and so must the divider. This is because, while the PLL searches for the proper continuous tuning range (Section 9.10), the VCO/divider cascade must not fail. Another issue is that the three-stage ring oscillator does not provide complementary outputs whereas the flipflops introduced in the previous section require such clocks.

For initial simulations, we create the arrangement shown in Fig. 9.36, where the inverters and the pass

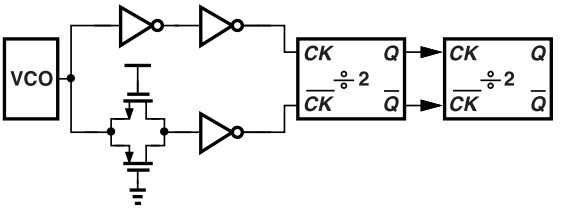


Figure 9.36 Interface between ring VCO and divider.

gate, all using $W/L = 1\ \mu\mathrm{m}/40$ nm, convert the single-ended output of the VCO to complementary clocks. The first divider is loaded by an identical stage. In Fig. 9.33(b), we choose $(W/L)_{1-4} = 0.5\ \mu\mathrm{m}/40$ nm and $(W/L)_5 = 1\ \mu\mathrm{m}/40$ nm. According to simulations, the divider cascade operates properly up to about 8 GHz in the SS, 75 °C corner, allaying our concern in regards to the initial search for the tuning range.

For the NAND latch shown in Fig. 9.34(b), we choose $W = 1\ \mu$m for M_1, M_1', and M_5, and 0.5 μm for the remaining transistors. The overall $\div 120$ circuit draws about 200 μW at a VCO frequency of 3 GHz.

9.10 Use of Lock Detectors for Calibration

As described in Chapter 3, the use of discrete tuning helps reduce K_{VCO} and hence the effect of noise coupled to the oscillator control. However, it also complicates the PLL design. How does the PLL select the proper continuous tuning range? A calibration procedure must run at the startup or if the temperature or supply voltage change significantly. We have described an analog method in Chapter 8 where two comparators decide whether a given tuning curve is acceptable. The following example revisits that method.

Example 9.3

Suppose the PLL begins with the lowest VCO tuning curve. If the frequency is not high enough, the loop fails to lock and V_{cont} rises to V_{DD}. How can we exploit this property for calibration?

Solution

We can use a simple comparator that detects whether V_{cont} eventually exceeds a certain, unacceptable level, e.g., $V_{REF} = V_{DD} - 100$ mV (Fig. 9.37). If such an event occurs, the logic reconfigures the VCO for the

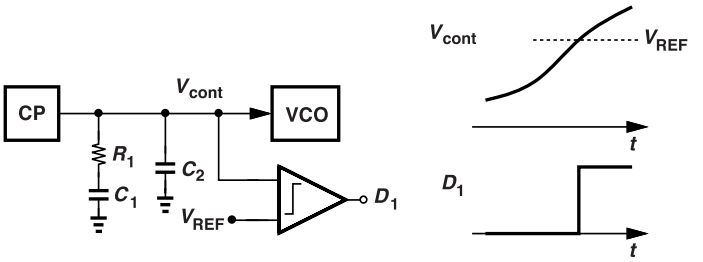


Figure 9.37 Using a comparator to determine whether a VCO tuning curve is acceptable.

next higher tuning curve. The loop is then allowed to settle again and the comparator output is monitored.

The difficulty with this approach is that the comparator must be clocked properly (or be realized as a high-gain amplifier) and its offset must be small enough. We thus seek a more digital solution. We envision that the calibration should proceed as follows: (1) we measure the VCO frequency, (2) we compare it with the desired value, and (3) if needed, we switch the tuning to a higher or lower tuning characteristic. For the first and second steps, it is simpler to compare the divider output frequency, f_{div}, with f_{REF}. That is, we need a circuit that computes $f_{div} - f_{REF}$ and accordingly decides whether it is necessary to switch to another tuning curve. Such a circuit is called a "lock detector." Figure 9.38 depicts the idea.

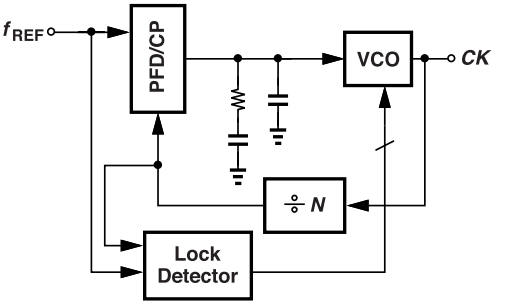


Figure 9.38 Use of a lock detector for VCO calibration.

To construct a lock detector, let us begin with the arrangement shown in Fig. 9.39(a), where two identical counters are clocked by f_{REF} and f_{div}. Suppose, for example, that $f_{REF} > f_{div}$ and the circuit starts at a reset state. As a result, Counter 1 fills up first, i.e., D_{out1} reaches $11...1$ before D_{out2} does. Now, as shown in

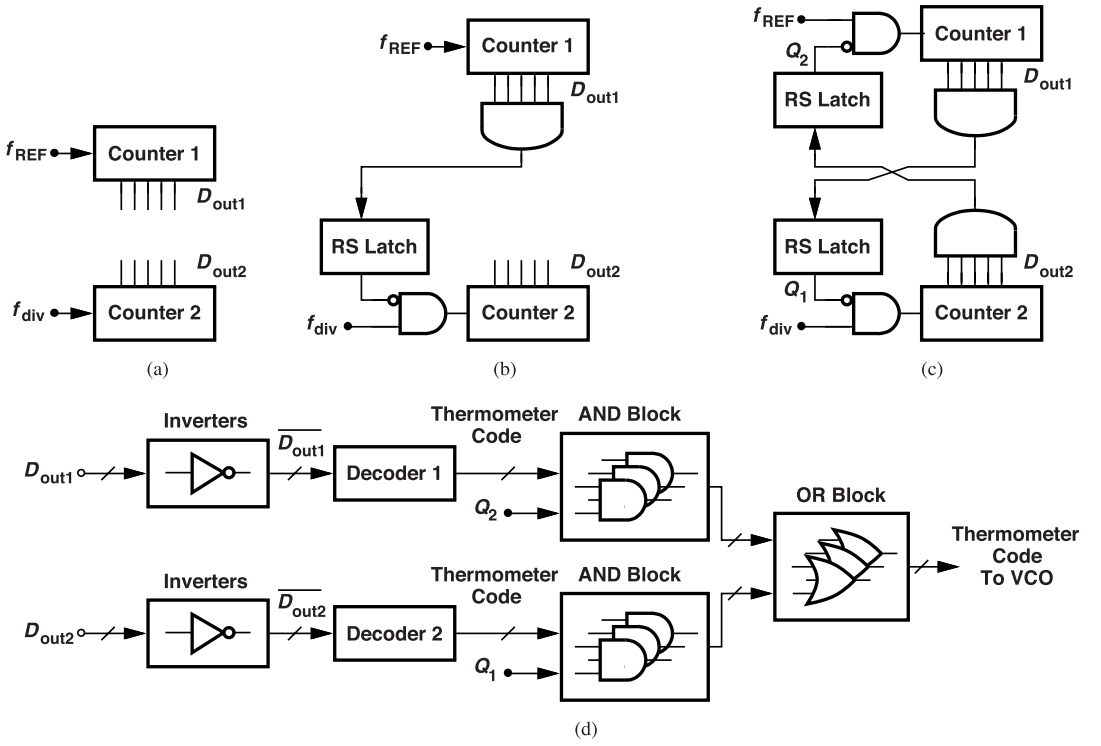


Figure 9.39 (a) Two counters driven by f_{REF} and f_{div}, (b) logic to detect whether $f_{REF} > f_{div}$, (c) conceptual lock detector, and (d) complete lock detector.

Fig. 9.39(b), we add an AND gate to monitor D_{out1} and an RS latch to control Counter 2. Once D_{out1} reaches $11...1$, the AND gate flips the RS latch, disabling the clock path of Counter 2.[5] The difference between D_{out1} and D_{out2} is now proportional to $f_{REF} - f_{div}$.

Since either Counter 1 or Counter 2 can fill up first, i.e., since $f_{REF} - f_{div}$ can be positive or negative, we implement the logic for Counter 2 as well [Fig. 9.39(c)]. Here, whichever counter fills up first also flips the corresponding RS latch and freezes the other counter. But, how do we utilize the result? The situation can be summarized as follows. (1) After one RS latch changes state, the complement of D_{out1} or D_{out2} represents the difference between f_{REF} and f_{div}. (2) The VCO discrete tuning circuit operates with a thermometer code. We therefore surmise that a decoder must convert D_{out1} or D_{out2} to a thermometer code such that the proper VCO tuning curve is selected. The decoding must, of course, be performed on the *frozen* counter output, which is identified by Q_1 or Q_2 in Fig. 9.39(c). These thoughts lead to the logic depicted in Fig. 9.39(d). Decoders 1 and 2 convert $\overline{D_{out1}}$ and $\overline{D_{out2}}$, respectively, to thermometer codes. The resulting values are then gated by Q_1 and Q_2 before they reach the VCO. For example, if Counter 1 fills up first, Q_1 goes high, allowing the complement of D_{out2} to drive the VCO.

The lock detector of Fig. 9.39(c) merits some remarks. First, the counters should have a higher resolution than the VCO thermometer code. For example, the ring oscillator employed in this chapter has five tuning curves but we should choose 4- or 5-bit counters to obtain a sufficiently accurate value for $|D_{out1} - D_{out2}|$. The decoders then map this value to one of the tuning curves.

[5] Since Counter 1 subsequently produces $00...0$, the RS latch is necessary to keep the f_{div} path disabled.

Example 9.4

In Fig. 9.39(b), suppose D_{out1} is equal to 111100 when D_{out2} reaches 111111. Explain how the thermometer code is generated. Assume the VCO has five discrete tuning curves.

Solution

Since $\overline{D_{out1}} = 000011$, we simply convert the LSBs to a thermometer code value of 11100 ($\equiv 3$), enabling the first three capacitor banks.

Second, the lock detector must count f_{REF} and f_{div} for a long time so as to suppress the ambiguity during the initial transient (while $f_{REF} - f_{div}$ fluctuates considerably). Third, the circuit can continue to monitor $f_{REF} - f_{div}$ and repeat the search if the selected tuning curve goes "out of range" with temperature or supply variations. Fourth, the counters and the latches must begin from a reset state.

An alternative approach is to perform "open-loop" calibration: we tie V_{cont} to approximately $V_{DD}/2$, measure f_{div}, compare it with f_{REF} (using a lock detector), and accordingly switch the tuning capacitors into or out of the VCO. For this purpose, we can use a binary search or a linear search.

9.11 Design Summary

In this chapter, we have developed a basic PLL at the transistor level. Based on the three-stage ring oscillator designed in Chapter 3, the circuit operates at 2.4 GHz with a 20-MHz reference and exhibits random and deterministic jitters equal to 2 ps,rms and 5.5 ps,pp, respectively. The supply current consists of about 300 μA for the charge pump and its reference branches,[6] 660 μA for the VCO, and 200 μA for the feedback divider. With $V_{DD} = 0.95$ V, this amounts to 1.1 mW, about half of the budget. We can, for example, scale the VCO up by another factor of 2 and reduce the random jitter by a factor of $\sqrt{2}$, but at the cost of greater area occupation by the VCO.

We have also observed the extreme difficulties in low-voltage charge pump design. In our particular case, the low output impedance of the CP transistors proves the most troublesome. The reader is encouraged to review the improved CP topologies in Chapter 8 and determine whether any alleviate this issue.

Our loop design effort extensively draws upon behavioral simulations. With the VCO and the divider modeled by a simple equation, we have reduced both the number of transistors and the maximum signal frequencies, allowing much faster simulations.

References

[1]. A. Homayoun and B. Razavi, "Analysis of phase noise in phase/frequency detectors," *IEEE Trans. Circuits and Systems - Part I*, vol. 60, pp. 529-539, March 2013.

[6] The CP current actually flows for a short period of time (about five gate delays) in the locked condition.

Problems

9.1. We wish to halve the bandwidth of the PLL designed in Section 9.1. If K_{VCO} is given, how do we adjust I_p, R_1, and C_1 so that ζ remains constant?

9.2. Suppose an application requires the divide ratio, M, in Section 9.1 to range from 120 to 60 so as to provide an output frequency from 2400 MHz to 1200 MHz. Assuming that K_{VCO} is proportional to the output frequency, explain how I_p, R_1, and C_1 must be adjusted to ensure that both ζ and ω_u remain constant.

9.3. Repeat the previous problem but assume that the output frequency and K_{VCO} are constant while f_{REF} can be either 20 MHz or 40 MHz and $\omega_u = 0.1\omega_{REF}$.

9.4. What does the slope of the plot in Fig. 9.2(b) signify?

9.5. From the CP noise current plot in Fig. 9.4, determine the in-band phase noise of the PLL designed in this chapter. We have $f_{REF} = 20$ MHz, $M = 120$, and $I_p = 100$ μA. Assume the PFD reset time is 50 ps.

9.6. Assuming all transistors operate in saturation, compute the output noise current in Fig. 9.3(a) due to M_5.

9.7. In Fig. 9.5, does M_5 pull the gate of M_3 up after S_1 turns off? Explain why or why not.

9.8. The simulation results depicted in Fig. 9.11 correspond to the SS, 75 °C corner. Explain how the waveforms change in the FF, 0 °C corner.

9.9. Suppose we insert an inverter between the divider and the PFD in Fig. 9.14(c). Explain what changes in the loop.

9.10. Repeat Example 9.2 but assume that the control voltage falls exponentially. From Fig. 9.15(b), the time constant is around 25 ns.

9.11. How does the peak-to-peak jitter (in seconds) change in Eq. (9.13) as ω_m varies?

9.12. For a given Δ_1 in Fig. 9.21, explain how Δ_2 varies as f_m goes from zero to large values.

9.13. A mysterious VCO exhibits a phase noise profile given by $S_{\phi n}(f) = \alpha/f^4$. Determine the output phase noise of a type-II PLL exmploying such an oscillator. How does the result in Eq. (9.20) change in this case?

9.14. In Section 9.6.2, we performed linear scaling by a factor of 15.6 to reduce the VCO phase noise. Suppose we have a power budget of 3 mW for the VCO and can scale it by another factor of $3\ \text{mW}/0.62\ \text{mW} = 4.8$. How much is the VCO jitter contribution?

9.15. In the previous problem, by how much can we reduce the loop bandwidth if the VCO jitter contribution must not exceed 4 ps, rms?

9.16. Equation (9.23) is independent of K_{VCO}. Suppose we double K_{VCO}. How do we adjust the loop parameters so that both ζ and $S_{\phi,p}$ remain unchanged?

9.17. Consider Eq. (9.23). We halve M and the PLL output frequency while K_{VCO} is constant. Explain how the loop parameters must be adjusted so that ζ remains constant. Can we also keep $S_{\phi n,p}$ constant?

9.18. Suppose V_{in} in Fig. 9.28(a) has a 50% duty cycle. Explain what happens to V_{out} if the PMOS devices in the XOR gate are weaker than the NMOS devices. Does this effect produce spurs in V_{out}?

9.19. Repeat the previous problem for if one of the inverters in Fig. 9.28 exhibits unequal rise and fall times.

9.20. Consider the latch shown in Fig. 9.33(b). We wish to determine how the NMOS devices must be sized to ensure that they can change the state. For simplicity, assume $W_1 = W_2 = W_5$. When both D and CK are at V_{DD}, node X_1 must be pulled down to $V_{DD} - |V_{THP}|$ so that M_4 turns on and begins the regeneration. Using square-law equations and modeling M_1 and M_5 by one transistor in saturation, find the minimum acceptable value for W_1.

10

Digital Phase-Locked Loops

The PLL architectures studied in the previous chapters can be viewed as "analog" implementations: the PFD output information, the charge pump output current, the control voltage, and the VCO output phase are all analog quantities. In other words, no signal within the loop is *digitized*. By contrast, "digital PLLs" (DPLLs), also known as "all-digital PLLs," incorporate a combination of analog and digital functions. Such realizations present certain advantages over their analog counterparts and have become popular.

In this chapter, we study the principles of digital PLL design. The reader is encouraged to review Chapters 7 and 8 before beginning this chapter.

10.1 Basic Idea

Some applications require PLLs that operate with a low reference frequency and/or a narrow loop bandwidth. This bandwidth, around $2.2\omega_n = 2.2\sqrt{I_p K_{VCO}/(2\pi M C_1)}$ (for $\zeta = 1$), can be reduced by increasing C_1 (and decreasing I_p if ζ must remain constant), demanding large on-chip capacitors.

To appreciate the difficulty, suppose a PLL must generate an output frequency of 2.4 GHz for a Bluetooth receiver from a reference frequency of 1 MHz. Let us assume $K_{VCO} \approx 400$ MHz/V and note that $M = 2400$. For $2.2\omega_n \approx 2\pi(100\,\text{kHz})$, we can choose $I_p = 50\ \mu$A, arriving at a value of 102 pF for C_1.

As an alternative approach, we can implement the loop filter in the digital domain, expecting potentially smaller area consumption. We therefore ask how the PLL should be modified so as to accommodate a digital filter. We predict that (1) the filter's input must be digital, dictating that the phase error generated by the PFD be digitized, and (2) the filter's output is digital, requiring that the oscillator's frequency be controlled digitally.

The foregoing thoughts lead to the transformation depicted in Fig. 10.1(a). In the digital PLL, the PFD output information is digitized (by a "quantizer") and then applied to the filter. Also, the filter's output is converted to analog form by a digital-to-analog converter (DAC) before reaching the VCO. The feedback divider need not be modified.

The DPLL of Fig. 10.1(a) is often realized in the more compact form shown in Fig. 10.1(b), where the PFD and the quantizer are merged and called a "time-to-digital converter" (TDC), and the DAC and the VCO are merged and called a "digitally-controlled oscillator" (DCO). The TDC digitizes the phase difference, $\Delta\phi$, and delivers the digital representation to the filter.

Digital PLLs offer some advantages. First, their loop filter can occupy less area. The absence of capacitors also means that their leakage current is no longer a concern. Second, the charge pump is eliminated, a significant benefit in view of the design difficulties that we faced in Chapter 9. Third, as explained in Chapter 12, the $\Delta\Sigma$ modulator noise in a fractional-N loop can be cancelled more accurately. Fourth, the loop stability can be maintained for a wide range of M values by simply adjusting the loop filter parameters. This proves useful, for example, in microprocessors, which must run at vastly different frequencies depending on the computation at hand, requiring the clock generation PLL to accommodate a wide range.

Digital PLLs suffer from some disadvantages as well. As explained later in this chapter, they contain

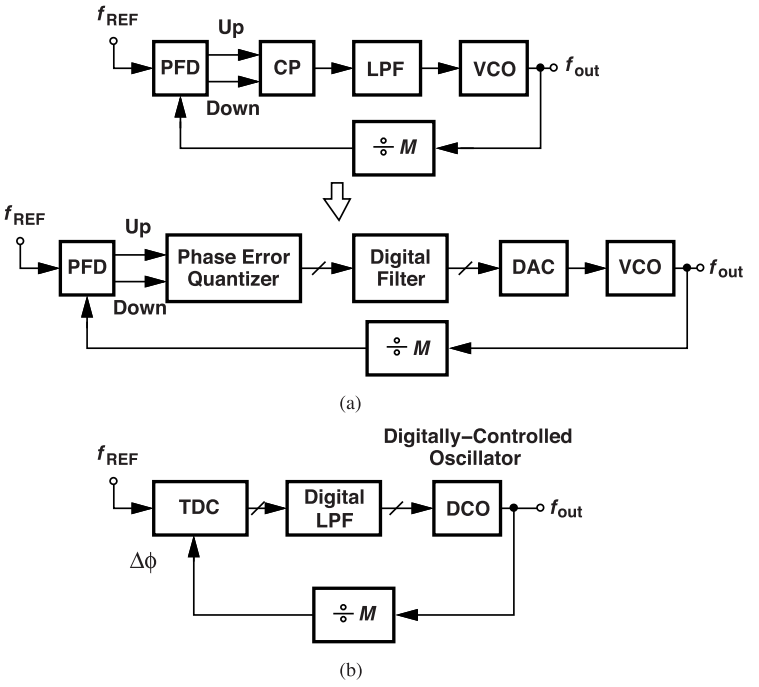


Figure 10.1 (a) Conceptual transformation of analog PLL to digital PLL, and (b) actual implementation.

additional sources of phase noise and jitter. Moreover, the circuit is not as "digital" as it may appear; in fact, the TDC and the DCO in Fig. 10.1(b) present serious analog design challenges.

As evident from Fig. 10.1(b), DPLLs involve both analog-to-digital conversion (within the TDC) and digital-to-analog conversion (within the DCO). The loop design thus heavily draws upon ADC and DAC concepts. We briefly review the former in the next section and the latter in Section 10.7.2.

10.2 ADC Basics

10.2.1 Quantization

An ADC, or more specifically, a quantizer, produces a digital output in proportion to its analog input. We say the quantizer converts a continuous-amplitude signal to a discrete-amplitude signal. Plotted in Fig. 10.2(a) is the input/output characteristic of an ideal quantizer. As the analog input increments by $\Delta = 1$ LSB, so does the digital output. We can also associate a *gain* with the quantizer, defined as the slope of the characteristic.

Also depicted in Fig. 10.2 is the transformation of a sinusoidal input with a frequency of f_{in} as it travels through the quantizer. How do we view the output imperfections? Can we say that the quantizer has added *noise* to the input signal? Strictly speaking, no. Since the input is periodic and the system is time-invariant, the output is also periodic, lending itself to a Fourier series expansion. That is, the quantized sinusoid consists of only harmonics—to very high orders, in fact [Fig. 10.2(b)].

Let us now assume that the quantized signal in Fig. 10.2(a) is sampled at a rate of f_S. If f_S is slightly greater than $2f_{in}$ (so as to satisfy Nyquist's criterion for f_{in}), the harmonics experience significant aliasing [Fig. 10.2(c)]. The resulting signal contains a large number of closely-spaced tones and can be approximated by a continuous, flat spectrum from $-f_S/2$ to $+f_S/2$ [Fig. 10.2(d)]. The spectral density is equal to $\Delta^2/(12f_S)$ and the total quantization noise power is given by $\Delta^2/12$.[1] It is interesting to note that increasing f_S lowers the spectral density but not the total noise power.

[1]This approximation assmues that the input signal fluctuations are occasionally large enough to reach 3Δ or 4Δ.

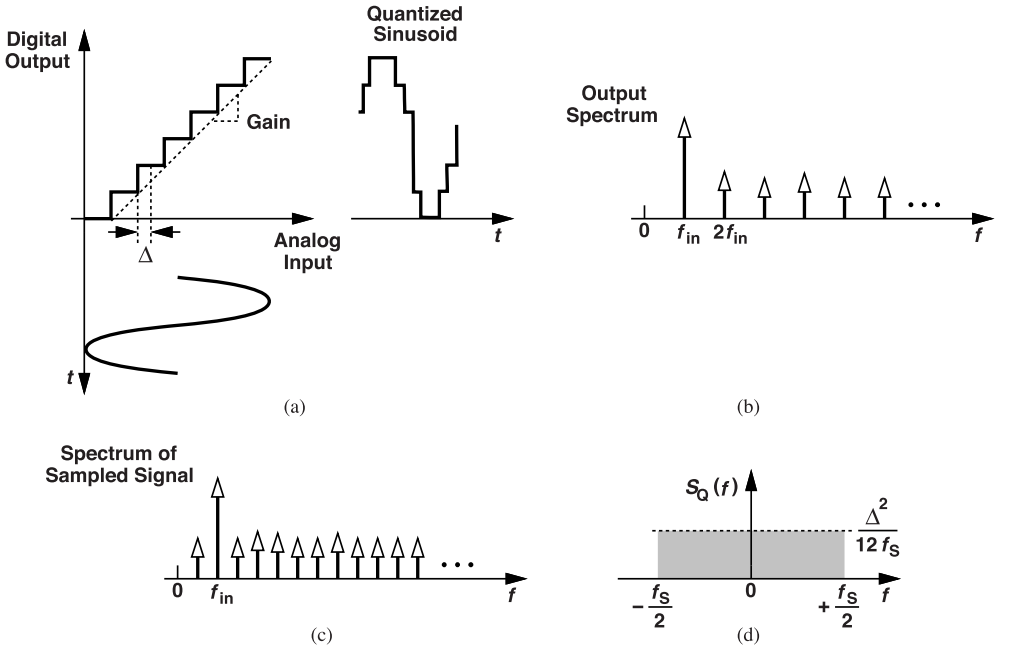


Figure 10.2 (a) A quantizer sensing a sinusoid, (b) quantizer output spectrum, (c) spectrum of sampled signal (for simplicity, f_S is not shown), and (d) approximation by a flat, continuous spectrum.

10.2.2 Flash ADC

Recall from Fig. 10.1(b) that the time-to-digital converter digitizes the *phase* difference. Nonetheless, it is simpler to begin our study of quantizers with *voltage* quantities. How do we generate a digital value corresponding to an analog voltage? The characteristic shown in Fig. 10.2(a) suggests that a set of equally-spaced reference voltages, $V_1, ..., V_n$, is necessary so that, as illustrated in Fig. 10.3(a), when V_{in} exceeds V_1, D_{out} jumps to 0001, when V_{in} exceeds V_2, D_{out} jumps to 0010, etc. We also need means of comparing V_{in} to the reference voltages.

From the foregoing observations emerges the "flash" ADC shown in Fig. 10.3(b). Here, a resistor ladder tied between V_{FS} and 0 generates $2^m - 1$ equally-spaced references, and 2^m comparators compare V_{in} to $V_1, ..., V_n$. Called the "full-scale" range, V_{FS} represents the maximum input value that the ADC can sense without saturation. The comparators are simultaneously clocked so as to perform sampling as well. We recognize that, as V_{in} varies from 0 to V_{FS}, the outputs of the comparators go from $B_1B_2...B_n = 00...0$ to 11...1. For example, if $V_j < V_{in} < V_{j+1}$, then $B_1...B_j$ are equal to 1 and the remaining outputs equal to 0. We say B_1-B_n form a "thermometer" code because the number of ONEs increases like the height of the alcohol in a thermometer. The decoder converts this output to binary (or Gray) code. The input voltage levels at which the comparators "flip" are called "decision thresholds." The comparators are designed to have an input-referred offset voltage less than Δ.

Example 10.1

Suppose comparator A_{j+1} in Fig. 10.3(b) has an input-referred offset equal to $\Delta/2$. Explain how the input-output characteristic changes. What happens to the thermometer code if the offset is more negative than $-\Delta$?

Solution

With an offset of $\Delta/2$, this comparator's decision threshold changes from V_{j+1} to $V_{j+1} + \Delta/2$ [Fig. 10.4(a)]. If the offset is more negative than $-\Delta$, the threshold shifts to *below* V_j. That is, as V_{in} increases from zero, A_{j+1} flips *before* A_j does. The thermometer code thus appears as ...11110100... for $V_j < V_{in} < V_{j+1}$ (Fig.

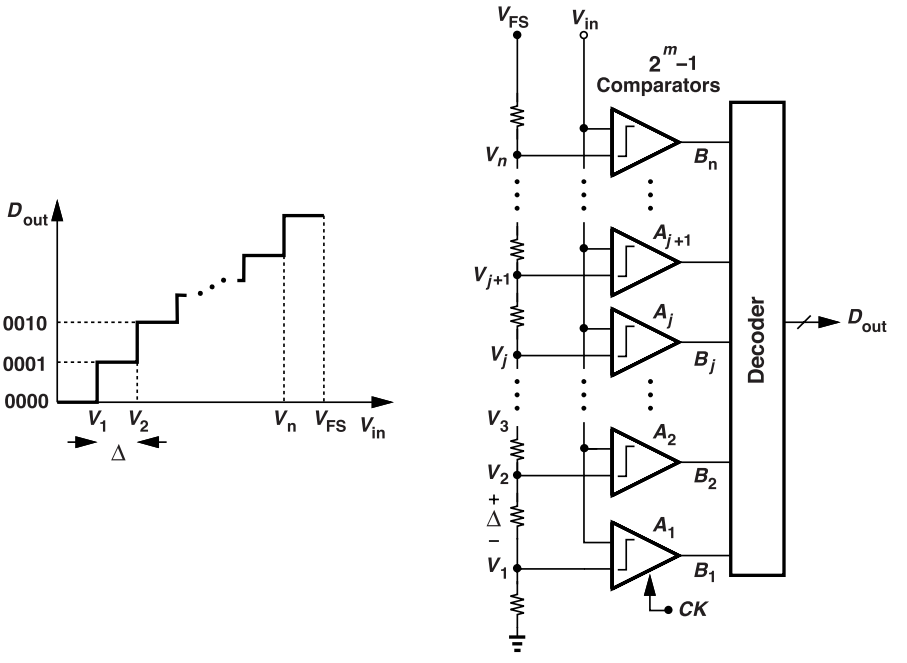


Figure 10.3 (a) Input-output characteristic of a voltage quantizer, and (b) flash ADC architecture.

10.4). Called a "bubble," the zero located between the 1's can cause large errors in the decoder's interpretation of the thermometer code. For this reason, some bubble removal logic should precede the decoder. As explained later, we can simply count the number of ONEs in the thermometer code to obtain a reasonable value in the presence of the bubble.

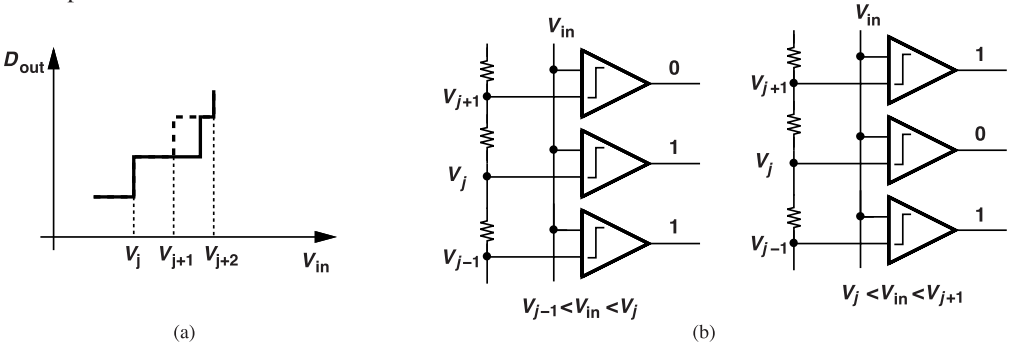


Figure 10.4 (a) Effect of $0.5-$LSB offset, and (b) bubble in flash ADC output as a result of a large comparator offset.

10.2.3 Interpolation

The comparators in the flash stage of Fig. 10.3(b) typically employ a differential pair along with a latch [Fig. 10.5(a)]. Latch 1 generates a logical output, B_1, according to the polarity of $V_{X1} - V_{Y1}$, and Latch 2 operates in a similar manner. For a resolution of m bits, the ADC thus loads the analog input with 2^m differential

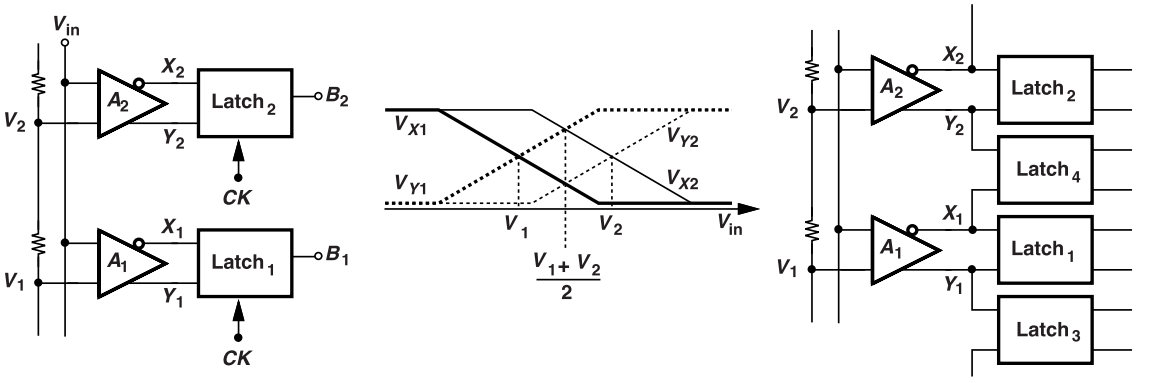


Figure 10.5 (a) One section of a flash ADC showing the signal path in each comparator, (b) internal voltages plotted as a function of V_{in}, and (c) interpolation network.

pairs.[2] It is possible to reduce the number of differential pairs through the use of "interpolation."

To understand this concept, we first examine the differential outputs X_1, Y_1, X_2, and Y_2 as V_{in} varies from below V_1 to above V_2. We know that V_{X1} and V_{Y1} cross at $V_{in} = V_1$, and V_{X2} and V_{Y2} at $V_{in} = V_2$ [Fig. 10.5(b)]. But additional information can be obtained by noting that V_{X1} and V_{Y2} cross at $(V_1 + V_2)/2$ (and so do V_{Y1} and V_{X2}). In other words, if we sense V_{X1} and V_{Y2} by means of a latch, we can determine whether V_{in} is less or greater than $(V_1 + V_2)/2$. Illustrated in Fig. 10.5(c) and called an "interpolating" flash ADC, the resulting architecture provides twice the resolution without doubling the number of differential pairs. For example, a 5-bit interpolating quantizer can employ 16 differential pairs and 31 latches. In Problem 10.2, we study the effect of comparator offset in this architecture.

10.3 Time-to-Digital Conversion

A digital PLL must convert the input analog phase error to a digital value. Figure 10.6 depicts the operation conceptually. The TDC senses two periodic inputs having a phase difference of T_D and generates a digital

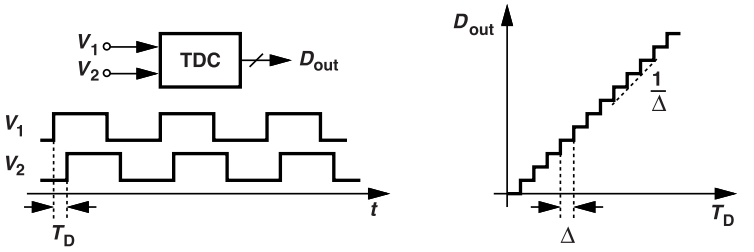


Figure 10.6 Time-to-digital converter and its input-output characteristic.

output, D_{out}, in proportion to it. We denote the analog LSB size by Δ seconds. The slope of the characteristic is equal to $1/\Delta$.

The similarity between the characteristics in Figs. 10.3(a) and 10.6 suggests that voltage-domain quantization techniques can be extended to the time domain as well. The next section builds upon this idea.

[2] Or with $2^m - 1$ differential pairs if comparison with 0 or V_{FS} is not required.

10.3.1 Basic TDC Topology

How do we quantize the phase difference between V_1 and V_2 in Fig. 10.6? Recall from Section 10.2.2 that a flash stage compares the input voltage to a set of equally-spaced reference voltages. In the TDC case, we must generate equally-spaced reference *times*, i.e., edges. To this end, we apply V_1 to a delay line having a unit delay equal to Δ [Fig. 10.7(a)]. The edges of the delayed replicas, V_{1a}, etc., can serve as references for "flash" conversion. Next, we must compare the phase of V_2 with these references; in other words, we must determine whether the rising edge of V_2 occurs between those of V_1 and V_{1a}, or between those of V_{1a} and V_{1b}, etc. This can be accomplished by a set of D flipflops, as shown in Fig. 10.7(b). In this TDC, when V_2 rises, the FFs sample the logical values of V_1 and its delayed replicas, generating a thermometer code $Q_1Q_2...Q_n$.

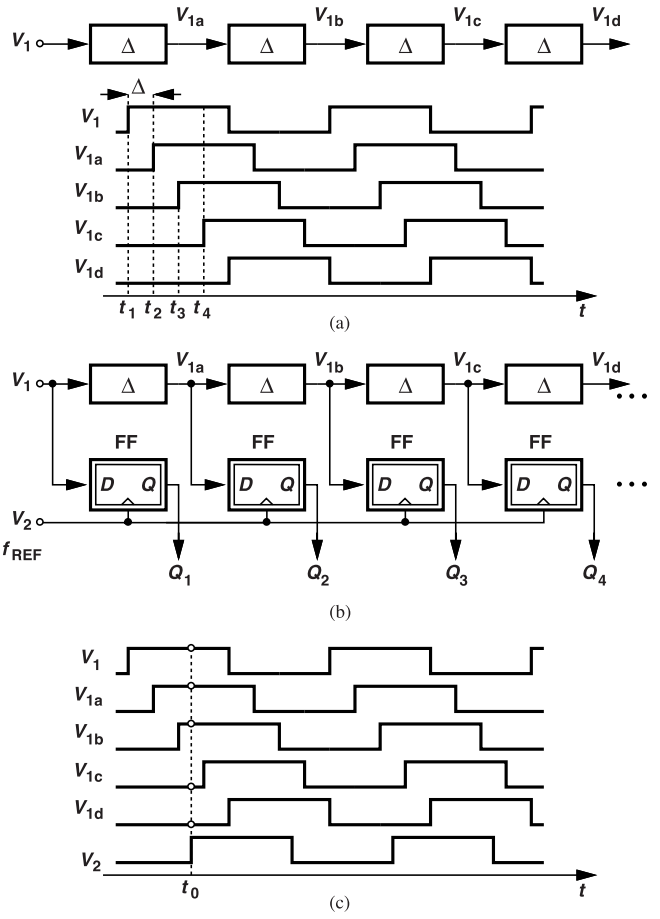


Figure 10.7 (a) A simple delay line, (b) a TDC wherein flipflops driven by V_2 sample delayed copies of V_1, and (c) circuit's waveforms.

Figure 10.7(c) shows, as an example, an output code of 11100 if V_2 rises at $t = t_0$.

We observe that the phase difference between the input signals decreases as they propagate to the right in Fig. 10.7(b). Specifically, suppose V_1 and V_2 have a skew of T_1. After m delay stages, V_1 is shifted by $m\Delta$ seconds but V_2 is not shifted at all, and hence the delay difference seen by the $(m + 1)$th flipflop is equal to $T_1 - m\Delta$.

It is interesting to recognize the similarity between the voltage and time quantizers shown in Figs. 10.3(b) and 10.7(b), respectively. Both generate a set of references, compare the input to these references, and produce a thermometer code. Note that the resolution, Δ, of this TDC architecture cannot be less than one gate delay, e.g., $\Delta > 10$ ps in 40-nm CMOS technology.

Example 10.2

An engineer designs the TDC of Fig. 10.7(b) with two identical inverters serving as buffers for V_1 and V_2 (Fig. 10.8). However, the two inverters see different loads. Explain the effect of this asymmetry on the performance.

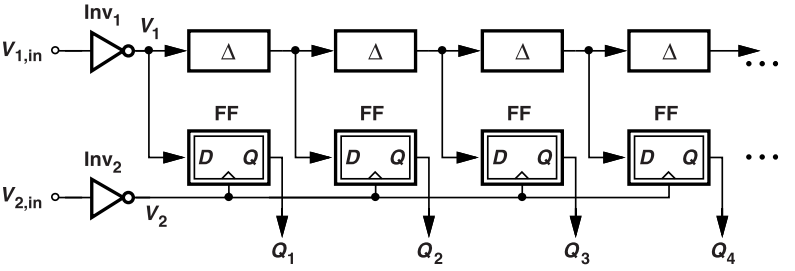


Figure 10.8 Effect of fanout imbalance on TDC.

Solution

We observe that Inv_1 is loaded by one delay element and one flipflop, and Inv_2 by many flipflops. As a result, V_2 incurs a greater delay. This skew manifests itself as a static phase offset between $V_{1,in}$ and $V_{2,in}$.

Example 10.3

Assume the phase error between V_1 and V_2 in Fig. 10.7(c) has the opposite sign but is relatively small, i.e., the rising edge of V_2 occurs earlier than that of V_1. Examine the TDC's output and sketch the overall TDC characteristic.

Solution

As shown in Fig. 10.9(a), V_2 now samples only zero values. Plotted in Fig. 10.9(b) is the overall characteristic, displaying a zero gain for negative phase errors.

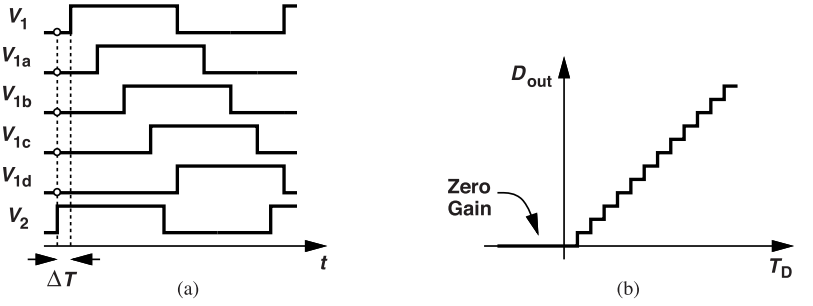


Figure 10.9 (a) TDC waveforms for a negative phase error, and (b) resulting characteristic.

The zero TDC gain illustrated in Fig. 10.9(b) is undesirable as it leads to a zero loop gain. In other words, if, for example, the VCO jitter accumulates and moves T_D in the negative direction, the PLL offers no feedback to correct it. To eliminate this effect, we can add another TDC that swaps the roles of V_1 and V_2; i.e., V_2 is delayed and V_1 performs sampling [Fig. 10.10(a)]. The second digital output is then subtracted from the first, yielding a code that can assume both negative and positive values.

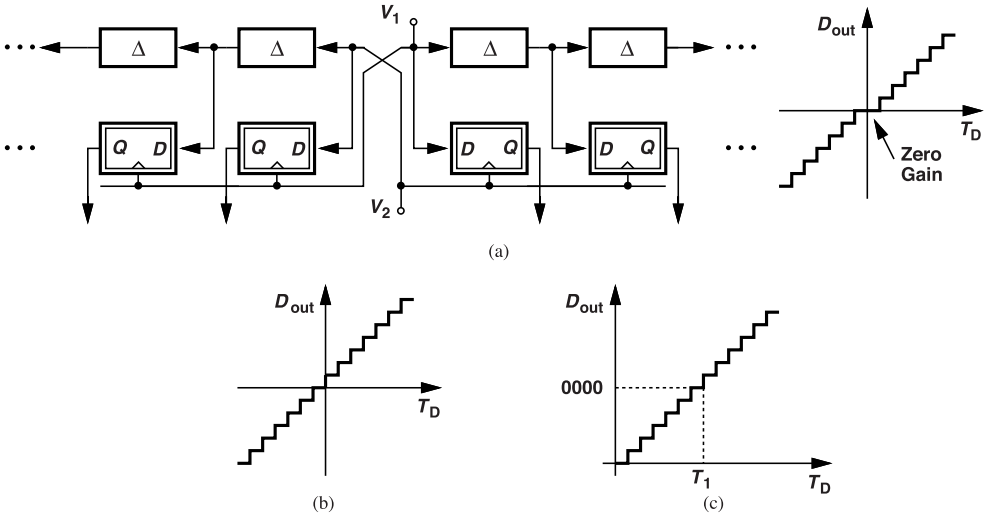


Figure 10.10 (a) Use of two TDCs to avoid the zero-gain region shown in Fig. 10.9(b), (b) improved characteristic avoiding zero gain around $T_D = 0$, and (c) alternative approach.

The TDC characteristic obtained in Fig. 10.10(a) still exhibits a zero gain around $T_D = 0$, allowing the PLL output to accumulate a jitter as large as $\pm\Delta$ without any correction. To avoid this effect, we wish to create gain at exactly $T_D = 0$. Suppose we simply add 1 LSB to the digital code produced by the TDC on the right. The first quadrant of the characteristic is then shifted up by 1 LSB, creating an upward transition at $T_D = 0$. Similarly, we can subtract 1 LSB from the output of the TDC on the left. Depicted in Fig. 10.10(b), the resulting behavior exhibits a high gain around the origin.

Another approach to avoiding the zero-gain region is shown in Fig. 10.10(c), where the digital output is assigned a value of zero for a certain phase error T_1. The infinite loop gain thus forces the system to lock with $T_D \approx T_1$ so as to minimize the apparent error, D_{out}. Now, if the phase fluctuates around T_1 by less than T_1, the TDC characteristic remains linear. This solution does create a phase offset of T_1 between the input and the feedback signal, an important issue in timing applications but not in RF synthesizers.

The TDC of Fig. 10.7(b) acts as a phase detector but provides little information regarding the frequency difference between V_1 and V_2. The issue can be resolved by incorporating a frequency acquisition circuit. For example, the lock detector described in Chapter 9, which provides a digital output in proportion to the frequency error, can be placed in parallel with the TDC to deliver this information to the loop filter.

Example 10.4

Given that the TDC output assumes only discrete values, explain how a DPLL locks.

Solution

The loop prefers to lock with a phase difference of zero. In reality, the oscillator jitter causes fluctuations in T_D, allowing D_{out} to reach a nonzero value occasionally. As illustrated in Fig. 10.11, if the oscillator phase fluctuation exceeds $+\Delta$ or $+2\Delta$, D_{out} reaches 001 or 010, respectively. Similarly, T_D randomly fluctuates to negative values so that the average T_D is zero.

10.3.2 Effect of Quantization Noise

Since the TDC approximates the analog phase error by a digital quantity, it introduces quantization noise in the phase domain. Returning to the noise calculations in Section 10.2.1, we observe that an LSB size of Δ seconds translates to a total noise power of $\Delta^2/12$. We also recognize that the TDC of Fig. 10.7(b) samples

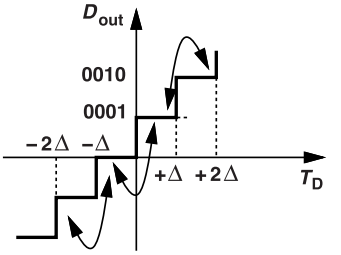


Figure 10.11 Jumps in TDC digital output when a DPLL is locked.

the phase at a rate of f_{REF}, thus exhibiting a noise spectral density of $\Delta^2/(12f_{REF})$ from $-f_{REF}/2$ to $+f_{REF}/2$.[3] To compute the corresponding spectrum for the phase, we must divide Δ by T_{REF} and multiply the result by 2π, obtaining $(2\pi\Delta/T_{REF})^2/(12f_{REF})$ (Fig. 10.12). This noise is generated at the PLL input and experiences the loop's low-pass transfer function as it appears in the output phase. That is, the in-band TDC phase noise is multiplied by the square of the divide ratio.

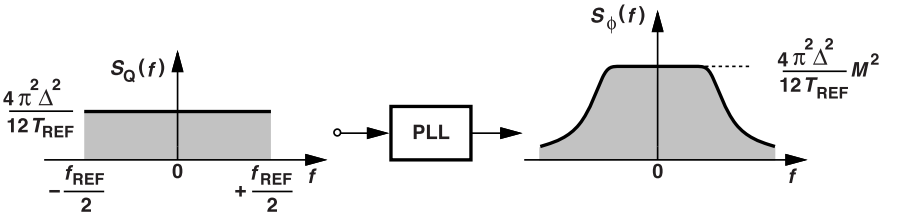


Figure 10.12 Effect of TDC quantization noise in a PLL.

Example 10.5

A 5-GHz WiFi frequency synthesizer operates with $f_{REF} = 20$ MHz. Determine the output in-band phase noise due to the TDC if $\Delta = 10$ ps.

Solution

The TDC phase noise spectrum has a height of $4\pi^2\Delta^2/(12T_{REF}) = 6.58 \times 10^{-15}$ Hz^{-1} $= -142$ dBc/Hz. With $M = 250$, we have

$$S_{\phi,out}(f) = -94 \text{ dBc/Hz}. \tag{10.1}$$

This calculation assumes that the TDC input error occasionally reaches several Δ seconds, which would occur if the VCO jitter has a peak around this value. In a simple PLL having a constant M, our result may overestimate the phase noise. The reader can see from Fig. 10.10(b) that for small jitter in V_1 and V_2, T_D does not exceed $\pm\Delta$.

Example 10.6

If the synthesizer in the above example has a closed-loop bandwidth of 2 MHz, determine the total output jitter arising from the TDC quantization noise. Assume $\zeta \geq 1.5$.

[3] These calculations assume that the TDC input phase error varies randomly, occasionally reaching a value of several Δ seconds.

Solution

Recall from Equation (8.44) that with $\zeta \geq 1.5$, the loop can be approximated by a one-pole system, generating from the input reference noise an output jitter given by

$$\Delta t_{rms} = \frac{1}{2\pi}\sqrt{\pi M^2 S_0 f_{-3dB}}T_{out}. \tag{10.2}$$

In this case, $M^2 S_0 = 4\pi^2\Delta^2 M^2/(12T_{REF})$, yielding

$$\Delta t_{rms} = \sqrt{\frac{\pi}{12}\frac{f_{-3dB}}{f_{REF}}}\Delta. \tag{10.3}$$

With $f_{-3dB} = 0.1 f_{REF}$, we have

$$\Delta t_{rms} \approx 0.162\Delta \tag{10.4}$$

$$\approx 1.62 \text{ ps}. \tag{10.5}$$

The foregoing examples reveal that $\Delta = 10$ ps may be inadequate for WiFi applications because the VCO and reference contributions further raise Δt_{rms}. In systems such as GSM or wideband CDMA transceivers, the issue becomes even more serious. The principal difficulty here is that Δ has a lower bound equal to the minimum gate delay afforded by the technology. We address this issue in Section 10.5.

10.3.3 TDC Dynamic Range

In addition to the LSB size, we are also interested in the total input range that the TDC can handle without saturation (the full scale). As shown in Fig. 10.13, the ratio of this range, T_{FS}, and the LSB, Δ, determines

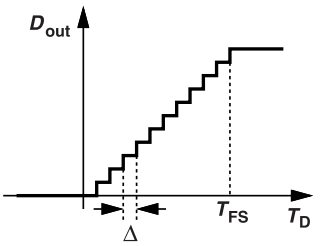


Figure 10.13 TDC characteristic exhibiting saturation beyong T_{FS}.

the number of delay elements necessary in the circuit of Fig. 10.7(b). The minimum acceptable value of T_{FS} depends on the PLL architecture: (1) in simple (integer-N) loops, T_{FS} must accommodate the expected phase offset as well as the peak-to-peak oscillator jitter, and (2) in fractional-N loops, T_{FS} must be larger than this amount by a value equal to several oscillator cycles (Chapter 12).

Example 10.7

Suppose the 5-GHz synthesizer of Example 10.5 requires a TDC input range of three oscillator cycles plus 10 ps,pp for the oscillator jitter. How many delay elements are required in the TDC?

Solution

The value of T_{FS} in Fig. 10.13 is equal to $3 \times 200 \text{ ps} + 10 \text{ ps} = 610$ ps. With $\Delta = 10$ ps, the TDC requires at least 122 delay elements and flipflops. (Recall that the TDC must detect both positive and negative phase errors.) Since a 10-ps resolution is inadequate in most cases, this number must be even greater.

10.3.4 TDC Imperfections

In addition to a resolution limited by the technology, the simple TDC of Fig. 10.7(b) suffers from several nonidealities that must be taken into account in the design process.

First, the delay, Δ, varies considerably with PVT, proportionally affecting the TDC's characteristics. For example, the delay of an inverter varies by about a factor of 2 across PVT corners. The difficulty here is twofold: (a) if Δ falls in the FF corner, the total range, T_{FS} in Fig. 10.13, may not suffice, and (b) if Δ rises in the SS corner, so do the phase noise and the jitter. Thus, the TDC must be "overdesigned" with (a) a *minimum* T_{FS}, in the FF corner, equal to the necessary input range, and (b) a *maximum* Δ, in the SS corner, not exceeding the desired value. The former means that the number of delay elements and flipflops in Example 10.7 must be roughly doubled to accommodate the FF corner. Alternatively, one can employ calibration to ensure that the unit delay remains around the desired value. This is possible by placing a replica of the delay chain in a DLL (Chapter 13).

Second, delay mismatches within the TDC distort the conversion. For example, the Δ values in Fig. 10.7(b) experience mismatches (as do resistors in a flash ADC's ladder), introducing nonlinearity in the input-output characteristic (Fig. 10.14). This effect is particularly problematic in fractional-N synthesizers as it folds down

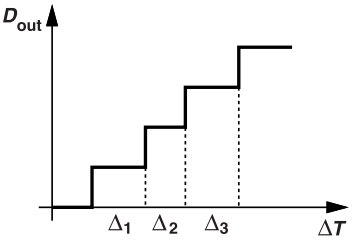


Figure 10.14 Effect of delay mismatch on the TDC characteristic.

high-frequency noise (Chapter 12) and even generates spurs.

Third, the signal propagating through the delay line accumulates phase noise. The intrinsic phase noise of inverters is negligible if their input and output transitions are fast enough, but the effect of supply noise can be significant, as illustrated by the following example.

Example 10.8

How does the supply noise, $V_{DD}(t)$, affect the delay chain in a TDC?

Solution

Supply noise modulates the unit delay, Δ. We can define a supply sensitivity, $K_{DD} = \partial\Delta/\partial V_{DD}$, for one delay element, where K_{DD} is expressed in rad/V. For a chain of n elements, the sensitivity is given by nK_{DD}, converting the supply noise to phase modulation as follows

$$V_{out} = V_0 \cos[\omega_{REF} t + nK_{DD}V_{DD}(t)]. \tag{10.6}$$

That is, the effect of supply noise linearly accumulates along the delay chain and directly adds phase noise to the input. Another consideration is that $V_{DD}(t)$ modulates the TDC full scale, T_{FS} in Fig. 10.13, equivalently changing the *gain* of the TDC. As illustrated in Fig. 10.15, as $V_{DD}(t)$ decreases, Δ and T_{FS} increase, and the TDC gain falls and vice versa.

Let us also consider a quasi-differential delay line [Fig. 10.15(b)]. We note that supply variations change the delays of both chains by the same amount. Thus, quasi-differential operation does not alleviate the problem of supply noise.

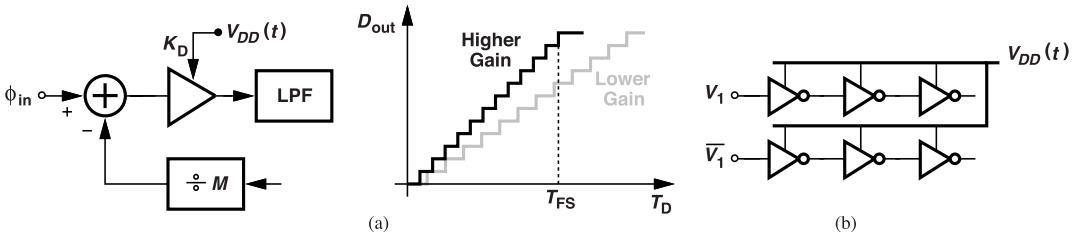


Figure 10.15 (a) Variation of TDC gain with supply voltage, and (b) quasi-differential delay line experiencing supply noise.

The fourth TDC nonideality relates to the offset of the delay elements and the flipflops in Fig. 10.7(b); these two effects are similar to the offset of the comparators in the flash ADC of Fig. 10.3(b). To understand this point, consider the situation depicted in Fig. 10.16, where V_{OS1} and V_{OS2} denote the input-referred offset

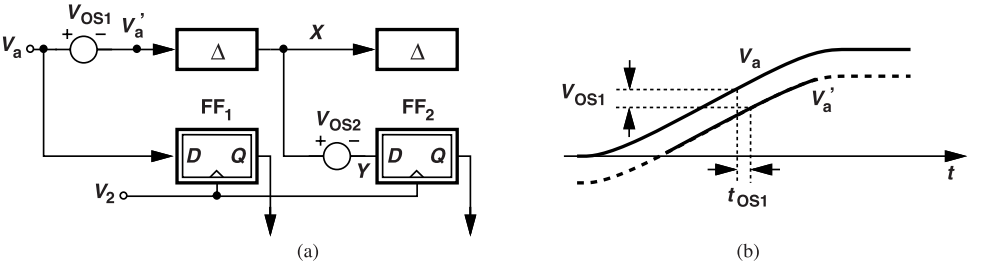


Figure 10.16 (a) TDC including offsets, and (b) translation of offset voltage to delay mismatch.

voltages of one delay element and one flipflop, respectively. Note that offsets can be defined even for single-ended circuits: e.g., V_{OS1} simply signifies that this delay element's input threshold (trip point) is shifted with respect to that of another one. With finite rise and fall times, the voltage offsets translate to time offsets. For example, as shown in Fig. 10.16(b), since V_a is shifted down by V_{OS1}, V_a' makes a transition t_{OS1} seconds later, where $t_{OS1} = V_{OS1}/r$ and r is the input slew rate. Thus, this delay element exhibits a unit delay equal to $\Delta + V_{OS1}/r$. For a swing of 1 V, a rise time of 10 ps, and $V_{OS1} = 20$ mV, we have $t_{OS1} = 0.2$ ps.

The key point here is that the TDC must not be viewed as a simple digital circuit: the slew rates not only affect the intrinsic phase noise but also convert voltage offsets to time offsets. The TDC's device dimensions and layout must therefore be designed carefully so as to manage these imperfections.

Example 10.9

In analogy with the flash ADC of Fig. 10.3(b), can a TDC suffer from bubbles?

Solution

Yes, it can. Consider the TDC shown in Fig. 10.17(a), where V_{OS2} is negative, causing V_Y to rise later than desired. If $t_{OS2} > \Delta$, then V_b rises *before* V_Y [Fig. 10.17(b)], yielding an output thermometer code equal to 10100. For this reason, we typically follow the TDC with bubble correction logic (Section 10.4).

10.4 Transistor-Level TDC Design

In this section, we implement the simple TDC of Fig. 10.7(b) at the transistor level. It is generally desirable to use complementary signals and quasi-differential circuits, although Example 10.8 suggests that the delay modulation due to supply noise persists.

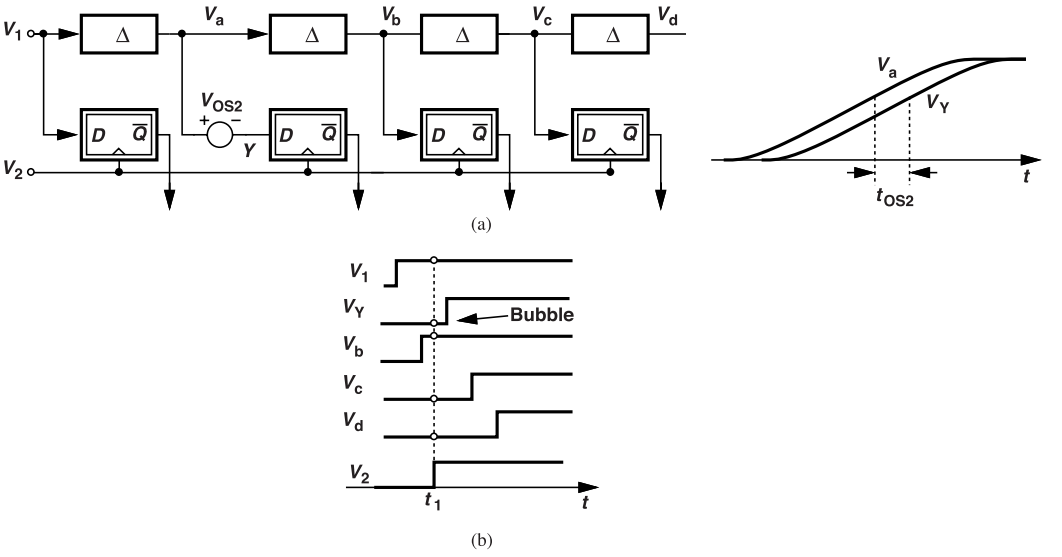


Figure 10.17 (a) Effect of FF offset in a TDC, and (b) resulting bubble in output code.

The delay elements can be realized by simple, fast inverters. The design of the delay chain must deal with two issues: (1) the random and deterministic mismatches must remain below 1 LSB, requiring proper sizing and careful layout, and (2) the effect of supply noise must be reduced through the use of an LDO or the choice of a more noise-tolerant topology.

For the sampling flipflops, we exploit the StrongArm latch as it provides a high sensitivity and a small offset [1, 2]. Shown in Fig. 10.18(a), this latch consists of an input differential pair (M_1-M_2), two cross-coupled pairs (M_3-M_4 and M_5-M_6), and four precharge switches (S_1-S_4). We study the circuit in four phases, assuming that $V_{in1} > V_{in2}$. When CK is low (Phase 1), all three pairs are off and nodes X, Y, P, and Q are precharged to V_{DD}. Thus, M_3-M_6 are off. When CK goes high (Phase 2), M_{CK} turns on and M_1 and M_2 begin to draw current from the parasitic capacitances C_X and C_Y while M_3-M_6 remain off [Fig. 10.18(b)]. Now, V_X and V_Y fall at different rates because $I_{D1} > I_{D2}$. With proper choice of g_{m1}, g_{m2}, C_X, and C_Y, the circuit can provide a voltage gain, $(V_X - V_Y)/(V_{in1} - V_{in2})$, in this phase (Problem 10.9). This is possible because V_X and V_Y begin from V_{DD} and can fall by an amount equal to $V_{DD} - V_{in1} + V_{THN}$ before M_1 (or M_2) enters the triode region. In other words, the input transistors reside in saturation for some time.

As V_X and V_Y decline in Fig. 10.18(b) and fall below $V_{DD} - V_{THN}$, M_3 and M_4 turn on (Phase 3), thereby drawing current from C_P and C_Q [Fig. 10.18(c)]. In this phase, M_5 and M_6 are still off but V_P and V_Q drop at different rates. This continues until these voltages fall to about $V_{DD} - |V_{THP}|$ and M_5 and M_6 turn on (Phase 4) [Fig. 10.18(d)]. The regenerative action of this pair then restores V_Q to V_{DD} while allowing V_P to drop to zero. As a result, M_4 also turns off. The reader can show that, without M_3 and M_4, the circuit would consume static power in this phase (Problem 10.10).

The StrongArm latch acts as a robust sampler with a fairly small input offset. We note that the voltage gain provided in Fig. 10.18(b) at $t = t_1$ equivalently attenuates the contributions of M_3-M_6 to the input offset. Similarly, since M_5 and M_6 turn on only after V_X and V_Y have diverged considerably, their contribution is negligible.

Example 10.10

Some designers omit the precharge switches tied to nodes X and Y in Fig. 10.18(a). Explain the drawbacks of such an arrangement.

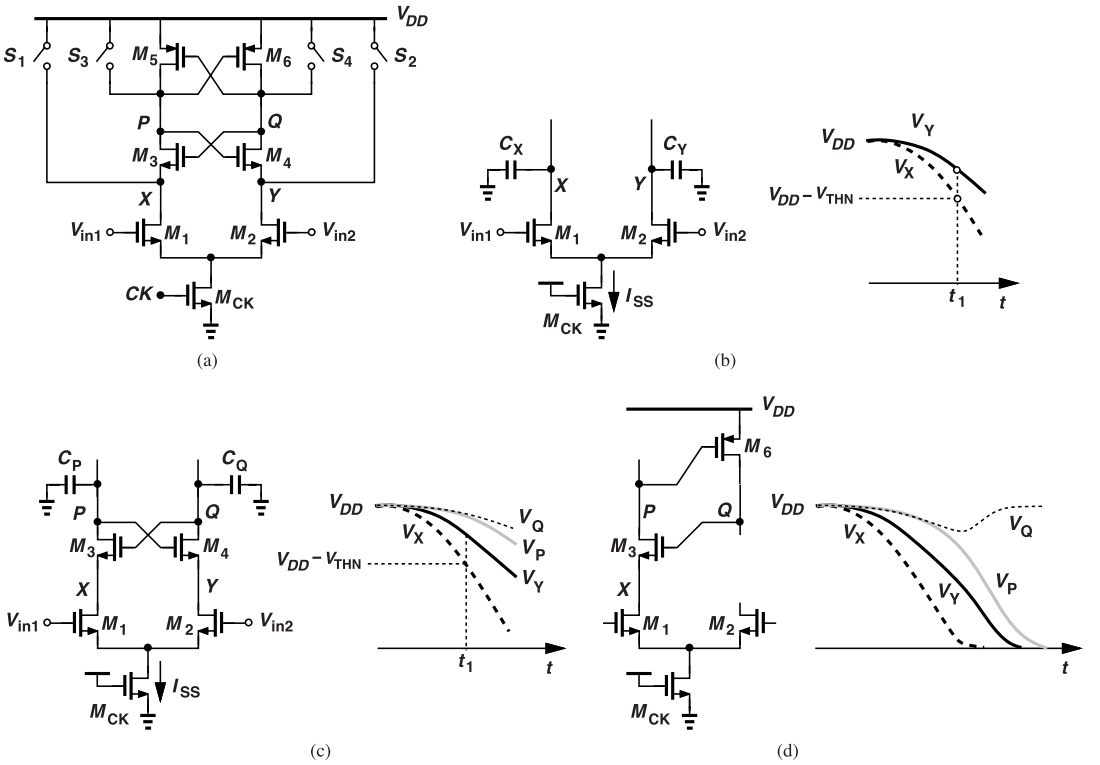


Figure 10.18 (a) StrongArm latch, (b) behavior of circuit during first amplification phase, (c) behavior of circuit during second amplification phase, and (d) final regeneration by PMOS devices.

Solution

Without these switches, X and Y reach $V_{DD} - V_{THN}$ rather than V_{DD} in the precharge mode (why?). Consequently, V_X and V_Y have little margin in Fig. 10.18(b) before they fall enough to drive M_1 or M_2 into the triode region. That is, the circuit cannot provide as much gain in this phase, making the offsets contributed by M_3-M_6 more pronounced. Additionally, without a reset action at X and Y, these nodes maintain some of the previous state and create a "dynamic offset."

In the precharge mode, the StrongArm latch produces a high level at both of its outputs, which cannot be considered valid logical levels representing the input difference. To ensure that the subsequent logic receives valid signals at all times, we insert an RS latch as shown in Fig. 10.19. Here, the RS latch stores a valid state and changes it only when P or Q falls to zero. The inverters serve as buffers and also generate low levels during precharge, allowing the use of NMOS devices at the RS latch input. That is, as the StrongArm latch enters the precharge mode, M_7 and M_8 turn off, and the RS latch retains its state. This state can change only after the StrongArm latch evaluates and P or Q goes low.

Figure 10.20 shows the TDC designed thus far. The quasi-differential delay line incorporates inverters and senses complementary inputs, but V_2 is single-ended. As mentioned in Example 10.8, Δ is still sensitive to V_{DD} in this case because the inverters' delay is inversely proportional to V_{DD}.

The thermometer code provided by the RS latches must now be subjected to bubble correction and then converted to binary code. An efficient approach is to merge these two functions and design the logic such that it generates a binary output equal to the number of ONEs in the thermometer code (Problem 10.12). For example, if the TDC of Fig. 10.7(b) suffers from a bubble in the form of 1110100 (Fig. 10.21), then the logic simply yields the binary equivalent of 4 by counting the ONEs and ignoring the ZEROs.

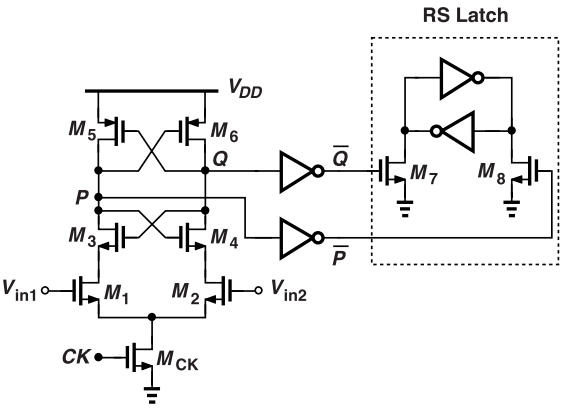


Figure 10.19 Use of an RS latch after StrongArm latch.

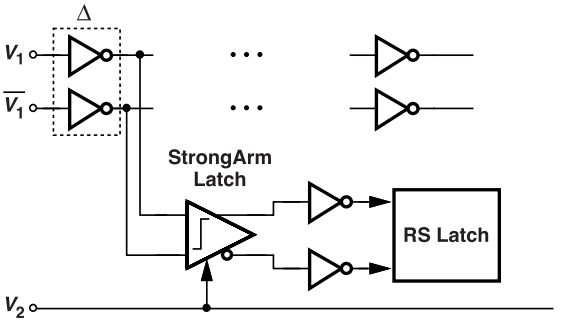


Figure 10.20 TDC design example.

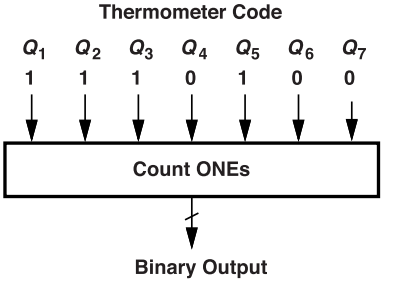


Figure 10.21 Bubble correction and thermometer-to-binary conversion.

10.5 Improved TDCs

As observed in previous sections, the resolution of the simple TDCs shown in Figs. 10.7(b) and 10.20 is limited by the technology's gate delay. For applications requiring smaller Δ values, we must consider alternative architectures.

10.5.1 Vernier TDC

A popular approach to obtaining a resolution finer than one gate delay employs the "vernier" TDC topology, depicted in Fig. 10.22 [3]. Here, both inputs travel through delay lines, but with different unit delays, Δ_1 and Δ_2. We observe that after m stages, V_1 and V_2 are delayed by $m\Delta_1$ and $m\Delta_2$, respectively. That is, the delay difference seen by the corresponding flipflop is equal to the original skew between V_1 and V_2 minus

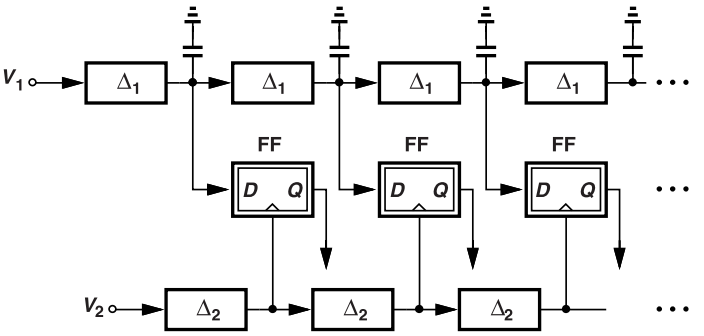


Figure 10.22 Vernier TDC architecture.

$m(\Delta_1 - \Delta_2)$. We conclude that the TDC provides a resolution of $\Delta_1 - \Delta_2$ seconds, a value not limited by the technology's gate delay. For example, if $\Delta_1 = 15$ ps and $\Delta_2 = 10$ ps, the timing resolution is 5 ps. Note that the actual value of $\Delta_1 - \Delta_2$ should account for the delay difference between the flipflops' data path and clock path. In the StrongArm latch of Fig. 10.18(a), for example, the delay from $V_{in1} - V_{in2}$ to the output is slightly less than that from CK. If this difference is denoted by Δ_e, then the actual vernier delay is equal to $\Delta_1 - (\Delta_2 + \Delta_e)$.

How small can $\Delta_1 - \Delta_2$ be? In the presence of random mismatches between Δ_1 and Δ_2, their difference can change *sign*, an undesirable situation. In Problem 10.14, we study how such an error distorts the output thermometer code. We conclude that $\Delta_1 - \Delta_2$ must be chosen greater than the expected mismatch.

It is difficult to establish Δ_1 and Δ_2 such that their difference is the desired value. Consider, for example, the arrangement shown in Fig. 10.23, where the two inverters are identical and C_1 is added to create the

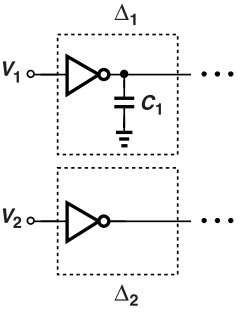


Figure 10.23 Vernier TDC exmple.

difference. With PVT variations, the inverters' drive strength changes and so does $\Delta_1 - \Delta_2$.

Providing a finer resolution, the vernier TDC necessitates a proportionally larger number of delay stages and flipflops to accommodate a certain full scale. For example, if $\Delta_1 = 15$ ps and $\Delta_2 = 10$ ps, then each delay line in Fig. 10.22 contains twice as many as these in Fig. 10.7(b).

It is instructive to extend our study in Section 10.3.4 of imperfections to the vernier TDC as well. First, the resolution, $\Delta_1 - \Delta_2$, varies with PVT, demanding overdesign. The number of stages must provide the necessary dynamic range even in the FF corner. Moreover, $\Delta_1 - \Delta_2$ must be small enough in the SS corner.

Second, as explained above, delay mismatches must not change the sign of $\Delta_1 - \Delta_2$. Third, with two delay chains accumulating phase noise, the intrinsic and supply noise become more serious. The following example elaborates on this point.

Example 10.11

Repeat Example 10.8 for the vernier TDC of Fig. 10.22.

Solution

Let us denote the supply sensitivities of the Δ_1 and Δ_2 delay cells by K_{D1} and K_{D2}, respectively, recognizing that, for inverters, $K_{D1} > K_{D2}$ if $\Delta_1 > \Delta_2$. To understand this point, note that a change in V_{DD} primarily affects the driving strength of an inverter. Thus, between two identical inverters loaded by different capacitors, the slower one produces a greater delay change in response to a V_{DD} change.

Suppose Figs. 10.7(b) and 10.22 incorporate n and m delay stages, respectively. In the latter, the total phase modulation is given by $m(K_{D1} - K_{D2})V_{DD}(t)$, which should be compared to $nK_DV_{DD}(t)$ in the former. We assume $\Delta_2 \approx \Delta$ and $K_{D2} \approx K_D$. We must have $m(\Delta_1 - \Delta_2) = n\Delta$ so as to obtain the same full-scale range for the two designs. Since $m = n\Delta/(\Delta_1 - \Delta_2)$, the two phase modulation values can be expressed as $[n\Delta_2/(\Delta_1 - \Delta_2)](K_{D1} - K_{D2})V_{DD}(t)$ and $nK_{D2}V_{DD}(t)$. Equivalently, we can compare $(K_{D1}/K_{D2}) - 1$ and $(\Delta_1/\Delta_2) - 1$ (why?). Interestingly, these two values are approximately equal because $K_{D1}/K_{D2} \approx \Delta_1/\Delta_2$. In other words, the two TDCs exhibit the same supply sensitivity.

The effect of TDC and flipflop offsets in the vernier TDC is similar to that in the simple TDC of Fig. 10.7(b).

10.5.2 Multi-Path TDCs

The challenge that we have faced in TDC design is to generate edges that are spaced by less than one gate delay. Let us return to the simple TDC of Fig. 10.7(b), duplicate its delay line, and increase the delay of the first stage in the second delay line by 50% [Fig. 10.24(a)], for example, by adding some capacitance [6]. We observe from the waveforms in Fig. 10.24(b) that the top and bottom paths provide signals that respectively have a delay of $m\Delta$ and $1.5\Delta + (m-1)\Delta$ with respect to V_1. Since the edges created by the bottom chain appear halfway between those of the top one, the TDC resolution can be doubled. Figure 10.24(c) shows the overall implementation [6]. To combine the two thermometer codes, we simply add up the number of ONEs in both.

The reader may recognize that, in Fig. 10.24(b), the minimum delay between V_1 and the next available edge, V_a, is still Δ rather than $\Delta/2$, as illustrated in the input-output characteristic. This issue can be avoided if we simply delay V_2 by Δ (why?).

Example 10.12

What happens if the first delay in the bottom chain of Fig. 10.24(c) deviates from 1.5Δ?

Solution

The edges are then not uniformly spaced, and the resolution degrades. This is illustrated in the input-output characteristic of Fig. 10.25 for the case of the delay being equal to 1.7Δ.

The TDC illustrated in Fig. 10.24(c) can achieve finer resolutions if we extend the duplicate-and-delay concept [6]. Shown in Fig. 10.26 is a four-path example utilizing delay values of 1.25Δ, 1.5Δ, and 1.75Δ in the first stage. The last three paths create three edges between every two edges generated by the first chain.

Example 10.13

We wish to examine the effect of process variations on the vernier and four-path TDCs. Neglect random mismatches and assume that both are designed for the same nominal resolution of $\Delta/4$. What happens to their input-output characteristics if the load capacitors are 10% less than expected?

Solution

In the vernier topology of Fig. 10.22, we nominally have $\Delta_1 - \Delta_2 = \Delta_2/4$ and hence $\Delta_1 = 5\Delta_2/4$. With a 10% reduction in the load capacitance, $\Delta_1 = 45\Delta_2/40 = 9\Delta_2/8$, and hence $\Delta_1 - \Delta_2 = \Delta_2/8$. (We assume Δ_2 is constant.) The LSB size is thus halved.

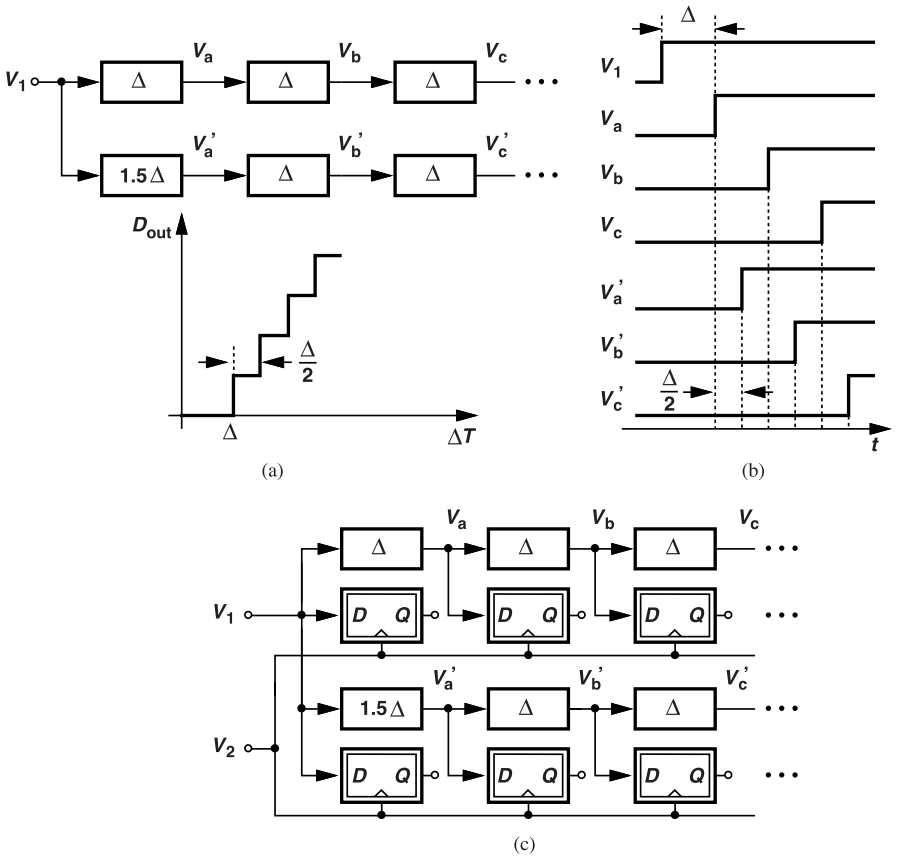


Figure 10.24 (a) Two delay lines with a delay offset equal to 0.5Δ, (b) internal waveforms, and (c) use in a TDC.

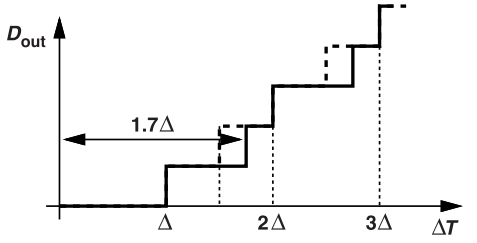


Figure 10.25 Effect of departure of delay offset in Fig. 10.24(a) from 0.5Δ.

We now plot the edges in the time domain for the nominal load capacitances and the smaller values [Fig. 10.27(a)], concluding that the full scale is halved and the TDC gain is doubled.

In the four-path architecture of Fig. 10.26(a), on the other hand, the three front-end delays fall from $5\Delta/4$, $6\Delta/4$, and $7\Delta/4$ to $45\Delta/40$, $54\Delta/40$, and $63\Delta/40$, respectively, while Δ itself remains constant. As a result, the edges provided by the additional delay chains occur earlier than expected, leading to the waveforms and the characteristic shown in Fig. 10.27(b). In this case, the full scale does not change (why?). It is possible to use calibration and adjust the capacitors here so as to obtain uniform edge spacing [6].

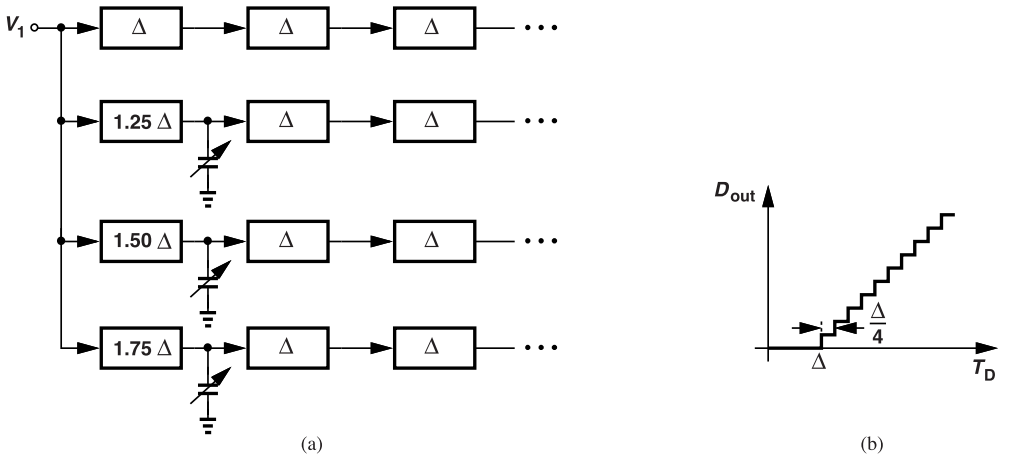


Figure 10.26 (a) TDC consisting of four delay lines with offset increments equal to 0.25Δ, and (b) resulting characteristic.

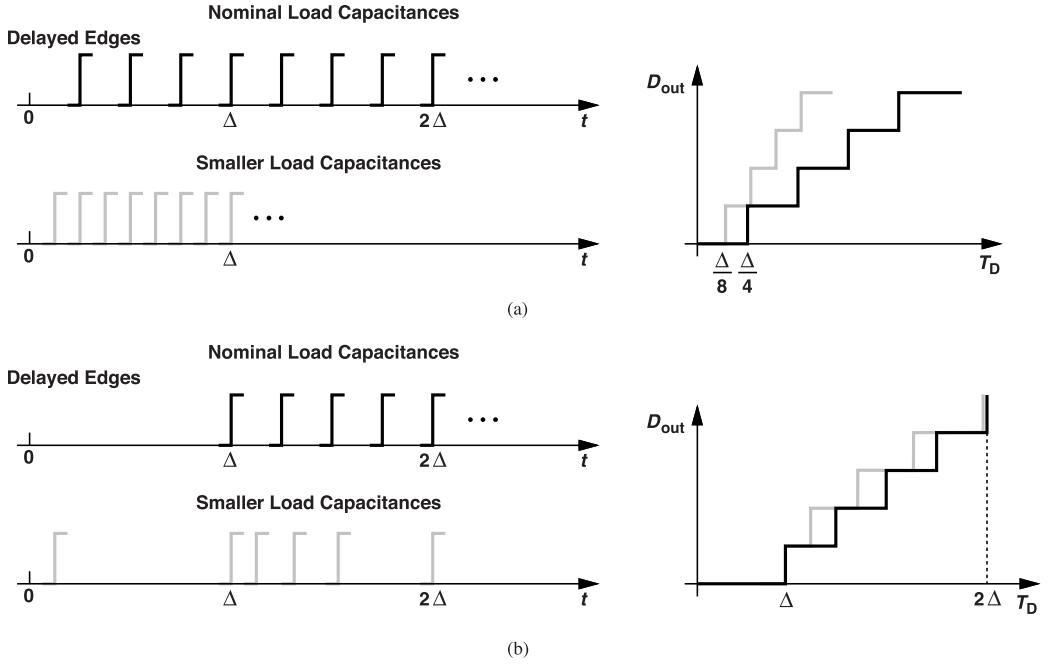


Figure 10.27 Effect of delay variation on (a) vernier TDC, and (b) four-path TDC.

10.6 TDC/Oscillator Combinations

If the oscillator in a digital PLL provides multiple phases, it can be merged with the TDC [5]. To develop this concept, let us return to the simple TDC of Fig. 10.7(b), assuming that V_1 is the PLL feedback signal and V_2 the reference. This example creates delayed edges of V_1 by means of a delay chain. But what if these edges are already available?! Suppose, for example, that the oscillator generates quadrature outputs. As illustrated in Fig. 10.28, we use the reference to sample these outputs, recognizing that the digital code, D_1D_2, varies with T_D. The combination therefore quantizes the phase difference between one oscillator output and the

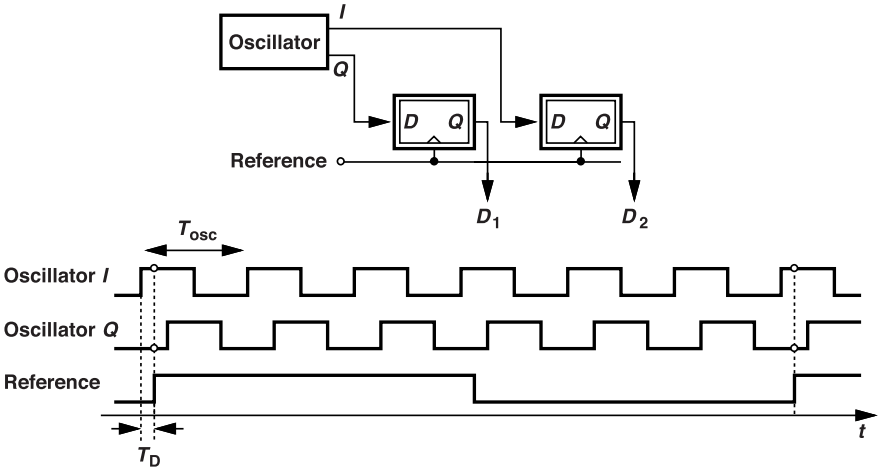


Figure 10.28 Use of oscillator quadrature phases to clock a TDC.

reference with a resolution equal to $T_{osc}/4$ (why?).

An important advantage of the above topology is that the TDC resolution is independent of PVT. Also, since the LSB scales with the output period, the PLL can accommodate a wide frequency range while maintaining a constant normalized TDC resolution. The drawback is that the oscillator must generate at least quadrature phases.

Example 10.14

Exploiting the concept of interpolation in flash ADCs (Section 10.2.3), explain how the resolution of the circuit in Fig. 10.28 can be doubled. Assume the quadrature phases are available in differential form.

Solution

In addition to sampling (I, I) and (Q, Q), the reference can also sample (I, Q) (Fig. 10.29). Here, FF_3 senses the difference between $\overline{I}$ and Q, providing opposite outputs if the reference edge occurs before or after

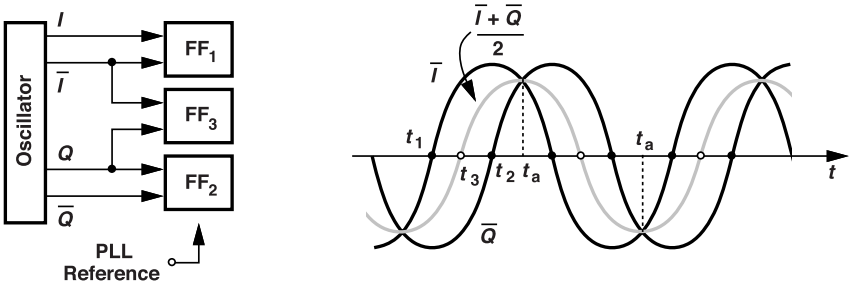
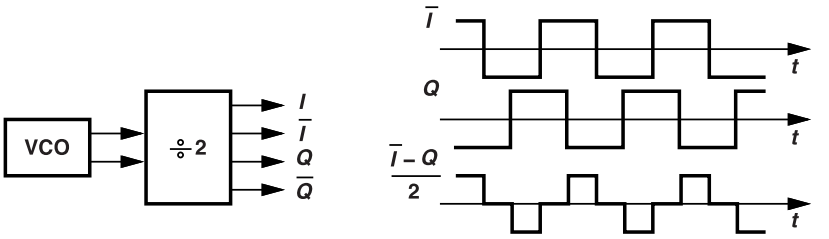


Figure 10.29 Use of oscillator quadrature phases in an interpolating TDC.

$t_3 = (t_1 + t_2)/2$. We can say that FF_3 senses the difference between $\overline{I}$ and Q and detects the polarity of this difference. For simplicity, we assume sinusoidal waveforms, observing that $\overline{I} - Q = \overline{I} + \overline{Q}$ has a phase difference of 45° with respect to I and Q. However, the arrangement does not interpolate at t_a and t_b. The reader is encouraged to attach another flipflop to I and Q and see how zero crossings at these two times can be created. Interpolation is then performed in all four quadrants of I and Q, and the TDC resolution is equal to $(T_{osc}/4)/2 = T_{osc}/8$.

Example 10.15

An engineer designs an LC oscillator for twice the required frequency and follows it with a $\div 2$ stage to generate quadrature phases [Fig. 10.30(a)]. (Such an arrangement generally provides a higher performance than a quadrature LC oscillator does.) Can one use interpolation with these phases?



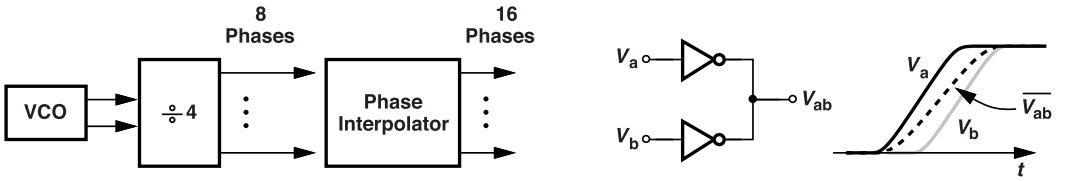
(a) (b)

Figure 10.30 (a) Use of a divider to generate quadrature waveforms, and (b) interpolation between I and Q with distortion at t_3.

Solution

It depends. If the divider output transitions are fast, interpolation suffers from a "kink" and becomes sensitive to noise and offset. Consider the extreme situation depicted in Fig. 10.30(b). The difference sensed by the interpolating flipflop (FF_3 in Fig. 10.29) exhibits a zero slope in the vicinity of t_3. This means that the offset voltage of FF_3 translates to a large deviation from t_3. We conclude that the edges subjected to interpolation cannot be arbitrarily fast.

It is possible to increase the TDC resolution if more phases are available. Shown in Fig. 10.31(a) is an example [11]. An oscillator running at four times the desired frequency drives a $\div 4$ circuit, generating eight



(a) (b)

Figure 10.31 (a) Interpolation among eight phases generated by a divide-by-4 circuit, and (b) interpolation network.

phases, which are then interpolated by a factor of 2. The interpolation network consists of nominally identical inverters whose outputs are combined, in effect performing averaging between V_a and V_b. That is, V_{ab} is the inverted version of $(V_a + V_b)/2$. The reader can repeat Example 10.15 for this scheme as well and show that V_{ab} exhibits a kink if V_a and V_b have sharp transitions. For phase interpolation techniques, the reader is referred to Chapter 11.

The design in [11] follows the arrangement shown in Fig. 10.31(a) with additional interpolators and obtains 64 phases. The difficulty is that the interpolators operate at high speeds, consuming significant power.

It is instructive to determine the interpolation factor necessary in a typical application. Suppose $f_{osc} = 2$ GHz and a TDC resolution of 5 ps is required. The VCO's I and Q phases provide an edge spacing of 125 ps. We thus need to interpolate by a factor of 25. Note that the interpolation must be realized in all four quadrants of I and Q, i.e., between I and Q, between I and $\overline{Q}$, etc.

Another approach to increasing the number of oscillator phases is to injection-lock a ring oscillator to an LC oscillator [12]. Illustrated in Fig. 10.32, the idea is to create many phases by the former while benefiting

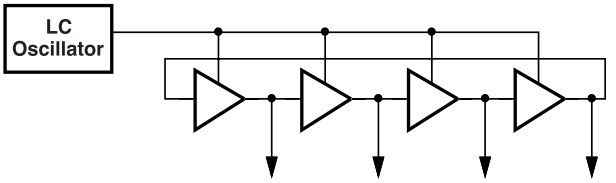


Figure 10.32 Injection-locking a ring oscillator to an LC oscillator so as to produce multiple phases.

from the low phase noise of the latter. This occurs because injection locking reduces the ring's phase noise to approximately that of the LC oscillator. Of course, the number of stages in the ring is limited by the oscillation frequency that it must achieve to match that of the LC circuit. Thus, this method may still need to employ interpolation after the ring oscillator to create a finer phase spacing. The reader is encouraged to compare the hardware and power requirements of the interpolating and injection locking topologies for a given TDC resolution.

10.7 Digitally-Controlled Oscillators

A digital PLL requires that its oscillator be controlled by the digital output of the loop filter. Shown conceptually in Fig. 10.33 along with its characteristic, such an oscillator generates only discrete frequency values,

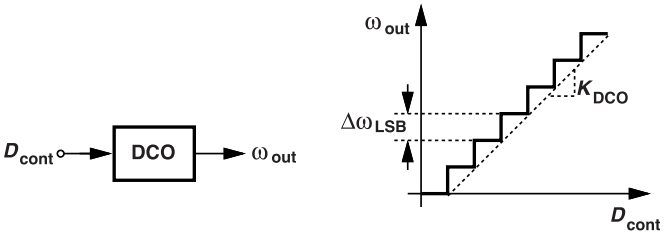


Figure 10.33 Conceptual view of a DCO along with its characteristic.

with an LSB size of $\Delta\omega_{LSB}$ and a gain of K_{DCO} rad/s/LSB. Of course, the discrete tuning techniques described in Chapters 3-6 can be applied here as well. The DCO behaves as a digital-to-analog converter, encouraging us to briefly review DAC basics. But we must first examine the effect of discrete frequency values.

10.7.1 Problem of Discrete Frequencies

Let us ask the following question: if the DCO cannot generate any arbitrary frequency, how does a DPLL lock? Our definition of lock in Chapter 7 inevitably led to $f_{out} = Mf_{in}$, a condition impossible to meet here if none of the discrete ω_{out} values in Fig. 10.33 is equal to $2\pi(Mf_{in})$. If noiseless, the DPLL must toggle between two states: as shown in Fig. 10.34(a), the DCO alternately runs at ω_1 and ω_2 so that the "average" frequency is the desired value. This, of course, means that the output spectrum consists of two strong tones [Fig. 10.34(b)], and the time-domain waveform suffers from substantial jitter.

Example 10.16

If the DCO dwells at ω_1 for T_{REF} seconds and at ω_2 for another T_{REF} seconds, determine the peak-to-peak phase variation.

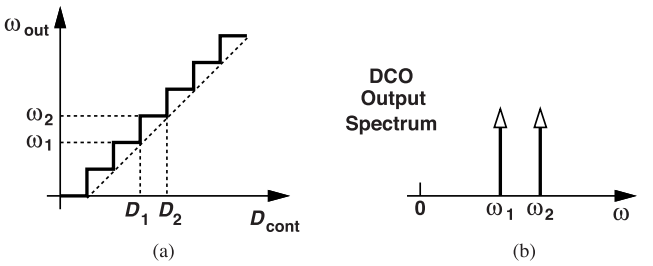


Figure 10.34 (a) A DCO toggling between two frequencies, and (b) resulting spectrum.

Solution

We surmise that the oscillator output can be expressed as

$$V_{out}(t) = V_0 \cos\left[\omega_c t + \int KS(t)dt\right], \tag{10.7}$$

where ω_c and K must be calculated and $S(t)$ toggles between -1 and $+1$ every T_{REF} seconds. When $S(t) = +1$, the output frequency, $\omega_c + K$, is equal to ω_2. Similarly, $\omega_c - K = \omega_1$. It follows that $\omega_c = (\omega_2 + \omega_1)/2$ and $K = (\omega_2 - \omega_1)/2$. As shown in Fig. 10.35, the peak-to-peak excess phase is given by

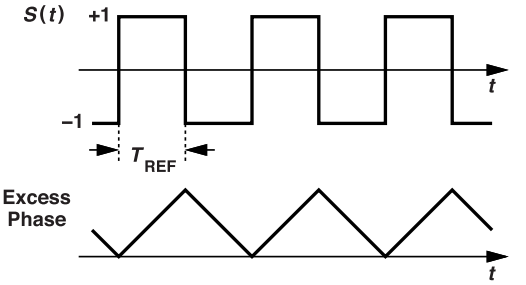


Figure 10.35 Phase excursions of a DCO during frequency toggling.

$$\phi_{ex} = \int_0^{T_{REF}} KS(t)dt \tag{10.8}$$

$$= \frac{T_{REF}(\omega_2 - \omega_1)}{2}. \tag{10.9}$$

In practice, the oscillator inherent phase noise and other sources of noise in a PLL tend to randomize the frequency toggling depicted in Fig. 10.34, thereby converting the spurs to phase noise. We approximate the result by associating quantization noise with the DCO. In analogy with the quantizer of Fig. 10.2(a), we view the DCO of Fig. 10.33 as a quantizer having an LSB size of $\Delta\omega_{LSB} = 2\pi\Delta f_{LSB}$ and hence a *frequency* noise spectrum that is flat and given by

$$S_f(f) = \frac{\Delta f_{LSB}^2}{12}\frac{1}{f_{REF}}. \tag{10.10}$$

Here, $\Delta^2/12$ is divided by f_{REF} because the DCO is updated ("sampled") every T_{REF} seconds. Recall that the spectrum of phase noise is equal to that of frequency noise divided by f^2:

$$S_\phi(f) = \frac{\Delta f_{LSB}^2}{12} \frac{1}{f_{REF}} \cdot \frac{1}{f^2}. \tag{10.11}$$

Equation (10.11) prescribes a frequency resolution, Δf_{LSB}, that can yield an acceptably small phase noise—only from this effect. The DCO, of course, also suffers from the various phase noise mechanisms described in Chapters 3-5.

Example 10.17

A 2-GHz DPLL operates with $f_{REF} = 20$ MHz and a DCO resolution of 1 MHz. Compute the quantization-induced phase noise.

Solution

We have $\Delta f_{LSB} = 1$ MHz, obtaining from (10.11):

$$S_\phi(f) = -84 \text{ dBc/Hz} \tag{10.12}$$

at 1-MHz offset. This value is unacceptably high for many applications.

Exhibiting a high phase noise, the DCO of Example 10.17 must deal with another critical issue as well. Suppose a tuning range of $\pm 5\%$ is necessary. Then, with a 1-MHz LSB, the DCO must accommodate 100 steps. For example, a differential LC implementation would need 200 unit capacitors and switches, facing considerable complexity.As explained later, we can reduce the complexity by grouping the capacitors into a fine array and a coarse array.

Example 10.18

Sketch the floor plan for the LC DCO described above.

Solution

We must place 200 unit switching branches by the inductor. As shown in Fig. 10.36, the two inductor terminals are stretched by a distance, d, to reach all of the capacitors. Consequently, the resistance of the two legs reduces the Q and must be carefully taken into account. (The inductances of the legs mostly cancel due to their strong mutual coupling.)

10.7.2 DAC Principles

In order to convert a digital input to an analog voltage or current, we can employ one of three topologies, namely, resistor ladders, current-source arrays, or capacitor arrays. All three have also been used in DCOs.

Figure 10.37 shows a resistor ladder serving as a DAC. The ladder generates, from V_{REF}, equally-spaced tap voltages, with one of the switches turned on to select the analog value corresponding to D_{in}. The switches are therefore driven by a "1-of-n" code, e.g., 000100, so that only one switch conducts. The decoder converts D_{in} to such a code. Resistor ladders display a monotonic characteristic because, as D_{in} increases, a higher tap on the ladder is selected and V_{out} rises. While obvious here, this property is not guaranteed in some other DAC architectures.

Example 10.19

Design a DCO using a resistor-ladder DAC.

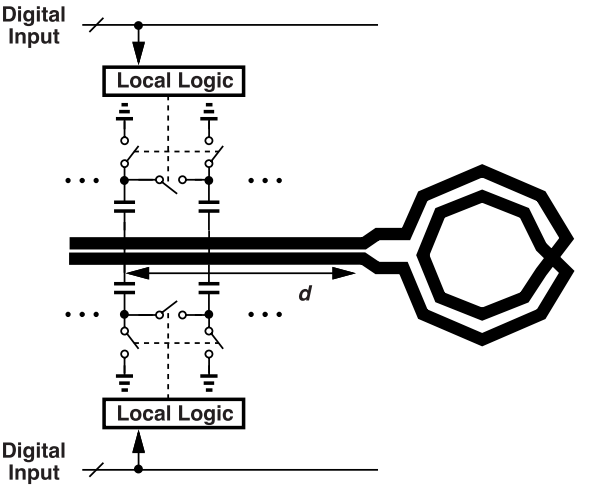


Figure 10.36 Floor plan of an LC DCO.

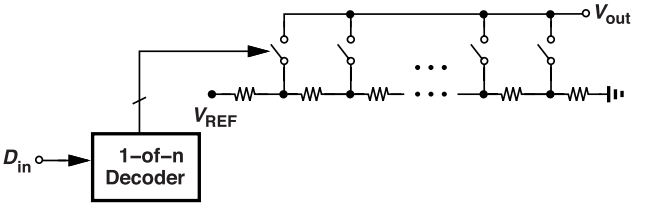


Figure 10.37 Resistor-ladder DAC.

Solution

We can simply drive a VCO by a ladder, as shown in Fig. 10.38. The DAC full scale, $V_2 - V_1$, is chosen to

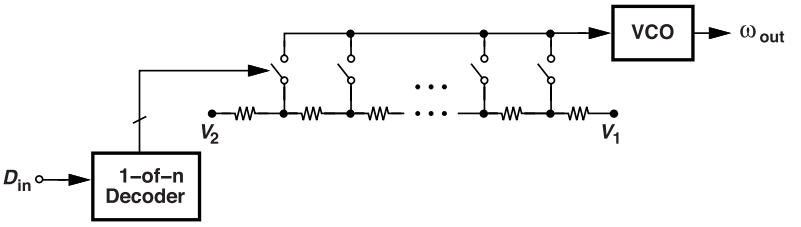


Figure 10.38 DCO using a resistor-ladder DAC.

match the input voltage range of the VCO, e.g., $V_1 = 0$ and $V_2 = V_{DD}$. In this case, the thermal noise of the ladder and the conducting switch modulates the VCO and must be sufficiently small.

DACs can incorporate current sources rather than resistors. Depicted in Fig. 10.39(a) is a 3-bit example consisting of "binary-weighted" current sources. As the input binary code, D_{in}, varies from 000 to 111, the output current goes from 0 to $7I_u$ in steps of I_u. For proper matching, of course, each current source employs an integer number of unit "slices" in parallel. Figure 10.39(b) illustrates this method for $2I_u$.

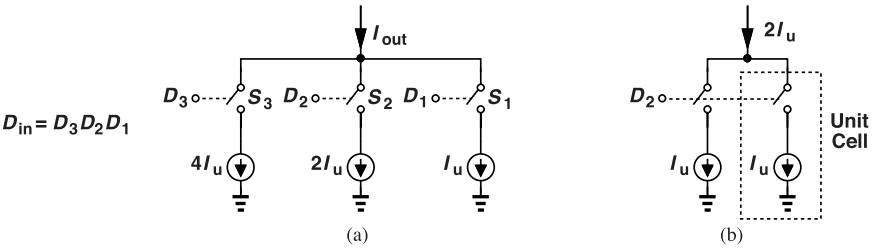


Figure 10.39 (a) Binary-weighted current-steering DAC, and (b) actual implementation of the $2I_u$ unit.

Example 10.20

Design a DCO using a current-switching DAC.

Solution

Among the oscillators studied in Chapters 3-5, the current-controlled ring topology in Fig. 3.47 is a good candidate for this purpose. As shown in Fig. 10.40, a current-switching DAC can tune the oscillation frequency in discrete steps. In this case, the DAC output noise current and the noise of M_1 are amplified if

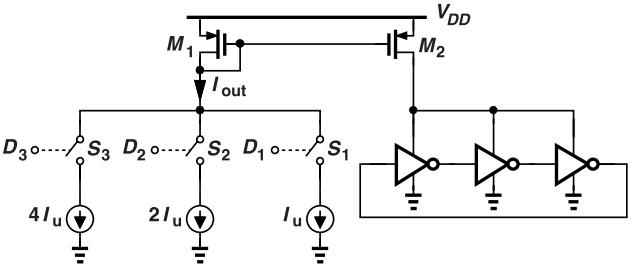


Figure 10.40 Current-controlled DCO.

$(W/L)_2 > (W/L)_1$. In a typical design, these noise contributions and the noise of M_2 exceed that of the ring itself.

Binary-weighted current-switching DACs can suffer from a nonmonotonic behavior. Consider the 4-bit system in Fig. 10.41(a), where $D_{in} = 0111$ and $I_{out} = I_3 + I_2 + I_1$. If D_{in} now rises to 1000, we have $I_{out} = I_4$ [Fig. 10.41(b)], and hence $\Delta I_{out} = I_4 - (I_3 + I_2 + I_1)$. In the ideal case, all units match and $\Delta I_{out} = I_u$, but if I_4 and $I_3 + I_2 + I_1$ suffer from mismatches, ΔI_{out} can be *negative*, leading to the nonmonotonic characteristic shown in Fig. 10.41(c). While illustrated here for a 4-bit DAC, this behavior actually arises for higher resolutions, around 8 bits or above, as the matching requirements become very stringent.

Another undesirable phenomenon in binary-weighted DACs relates to their transient response. Suppose the digital input in Fig. 10.40 jumps from 011 to 100. If the lines carrying the bits to the switches have different delays, I_{out} does not monotonically change from $3I_u$ to $4I_u$. For example, if S_1 and S_2 turn off slightly before S_3 turns on, then I_{out} drops to zero before returning to $4I_u$ (Fig. 10.42). Such a behavior translates to a glitch in the oscillator's output frequency, causing large spurs. In particular, if the DCO must toggle between ω_1 and ω_2, significant glitches appear in the output frequency. (Recall from Section 10.7.1 that in most cases, the DCO does need to toggle.)

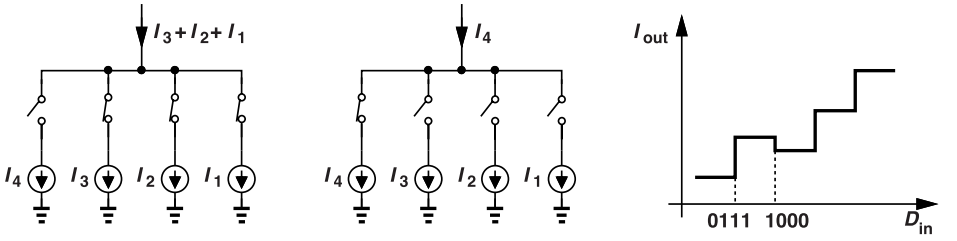


Figure 10.41 (a) A DAC with its first three LSB currents enabled, (b) the DAC with its MSB current enabled, and (c) input-output characteristic in the presence of mismatch.

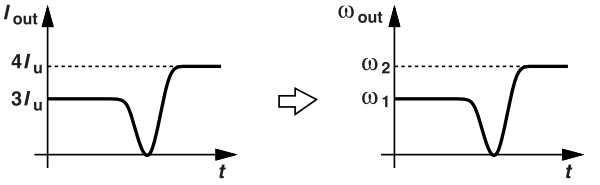


Figure 10.42 Translation of current glitch to frequency glitch in a DCO.

The foregoing DAC's nonmonotonic behavior and output glitches can be alleviated through the use of "segmentation." Illustrated in Fig. 10.43, a segmented DAC (also called a "thermometric" DAC or a "unary" DAC) consists of nominally identical (rather than binary-weighted) slices that are controlled by a thermometer

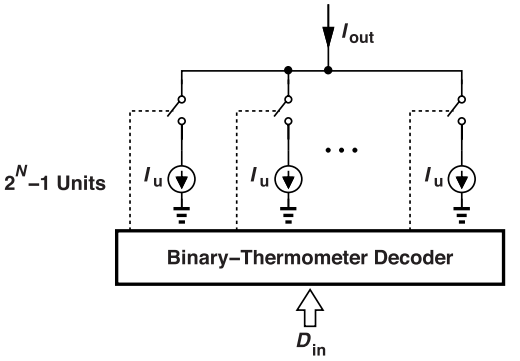


Figure 10.43 Segmented current-steering DAC.

code. To see why I_{out} varies monotonically, suppose D_{in} goes from 011 to 100. Then, the thermometer code changes from 1110000 to 1111000. In the former case, three unit current sources are turned on, and in the latter, four. That is, I_{out} increases regardless of the mismatches. The glitch can also be avoided, at least in principle, because only one unit current source (the fourth one in our example) switches during this transient. However, as explained below, the glitch may still occur in some cases.

10.7.3 Matrix Architecture

As explained in Section 10.7, DCOs must typically incorporate hundreds of unit slices so as to achieve both a fine resolution and an adequate tuning range. How do we arrange these slices and their associated logic in

a layout-friendly floor plan? Shown in Fig. 10.44(a), the "matrix" architecture is an efficient solution. Here, the input binary code is decomposed into an LSB word (D_3-D_1) and an MSB word (D_6-D_4), which are applied to a column decoder and a row decoder, respectively. The decoders generate separate thermometer

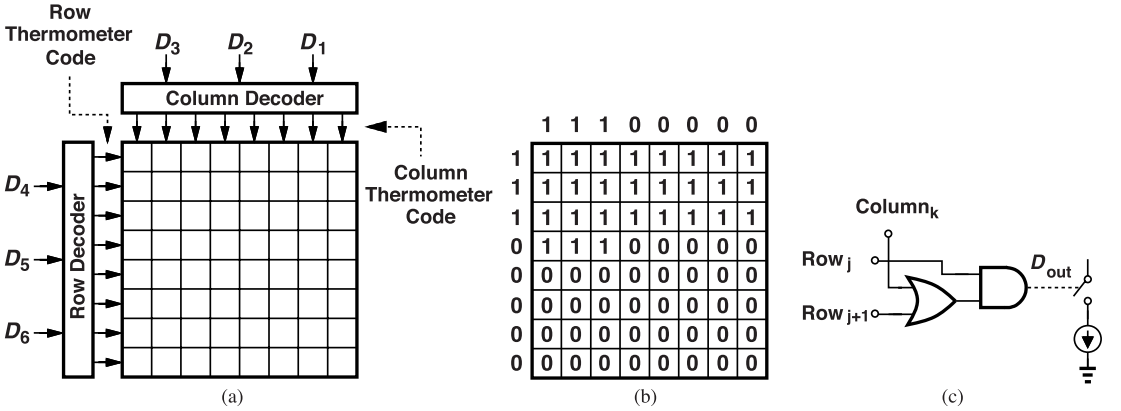


Figure 10.44 (a) Matrix DAC architecture, (b) example of logical values within the matrix, and (c) local decoder implementation within each cell.

codes, and these codes travel within the matrix, reaching each cell. The cells contain a local decoder and a unit current-switching slice.

The local decoder in Fig. 10.44(a) must turn the current sources on or off according to the column and row thermometer codes. We recognize three possibilities: (1) a row is completely on, (2) a row is completely off, and (3) a row is partially on. Figure 10.44(b) depicts an example. Based on these observations, we construct the logic shown in Fig. 10.44(c). The current source is switched on if $\text{Row}_j = 1$ and $\text{Row}_{j+1} = 1$ or $\text{Column}_k = 1$. That is, if both Row values are equal to 1, $D_{out} = 1$, and if $\text{Row}_{j+1} = 0$, then Column_k determines the output.

The matrix architecture of Fig. 10.44(a) affords a compact, modular design for segmented DACs. Nonetheless, the control lines traveling across the array can experience unequal delays, causing glitches at the output. In practice, each cell includes a latch [after D_{out} in Fig. 10.44(c)] so as to enable the current source control only when the clock arrives. Of course, the assumption is that the clock can be distributed with smaller skews than the control lines can.

The third type of DAC consists of capacitors. As with current-switching topologies, such DACs can be realized as binary or segmented arrays. DCOs employing binary-weighted capacitor DACs suffer from nonmonotonicity and output glitches. We therefore prefer segmentation. The following example illustrates the idea.

Example 10.21

Design a differential LC DCO with 6-bit control.

Solution

The circuit requires 64×2 unit switched capacitors, which we configure as an 8×8 matrix. Shown in Fig. 10.45(a) is the implementation. Note that, as illustrated in Fig. 10.45(b), the inductor legs must extend both vertically and horizontally across the array, thereby incurring substantial resistance.

An interesting issue in large arrays is the effect of "gradients" on the chip, i.e., when one device parameter varies monotonically from one side of the chip to the other. For example, the gate oxide thickness of the current sources in the matrix may decrease from left to right, yielding a monotonic increase in their drain currents. The error thus accumulates as we turn on more current sources in a row. We return to this point later.

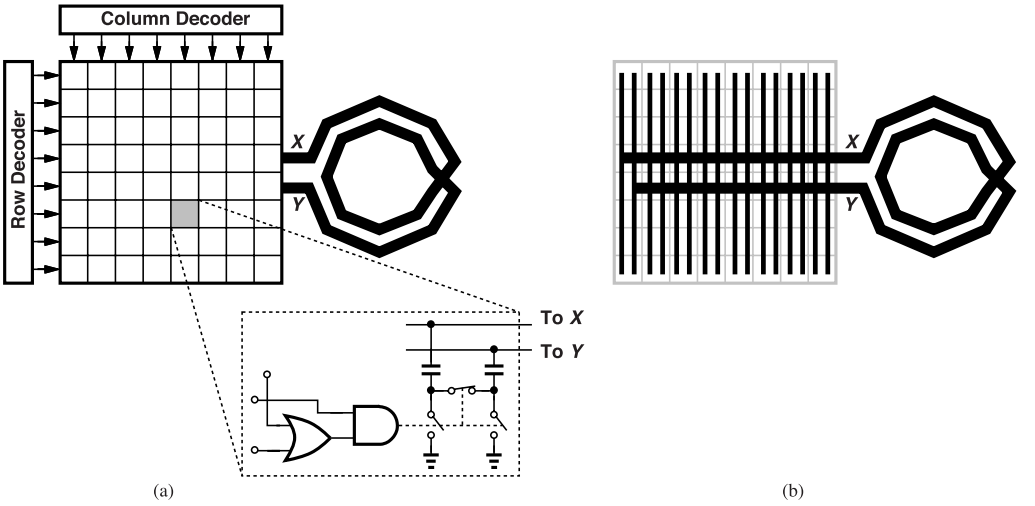


Figure 10.45 (a) Matrix architecture for a capacitor DAC, and (b) long interconnects necessary to reach all of the cells.

10.7.4 Coarse/Fine DACs

The complexity of segmented current-switching or capacitor arrays reaches prohibitive levels as higher resolutions are sought. We can add a coarse DAC to alleviate this issue. Figure 10.46 illustrates the concept: the input MSBs control the coarse elements, creating *overlapping* tuning characteristics, while the LSBs control

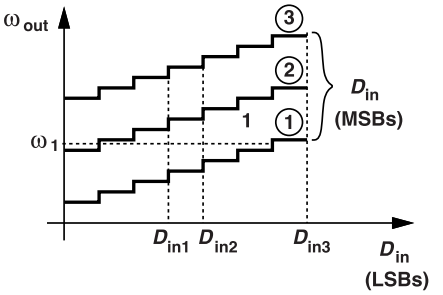


Figure 10.46 Coarse and fine frequency control in a DCO.

the fine units, providing small steps. The overlap between successive curves is necessary to avoid a blind zone in the output frequency values. The key point here is that, if D_{in} toggles by a few LSBs while the PLL is locked, ω_{out} changes *monotonically*. In this case, the MSB elements need not be segmented, but the LSB units are. For example, if D_{in} changes between D_{in1} and D_{in2}, a monotonic change is guaranteed on all of the characteristics. The reader is encouraged to see what happens if D_{in} is allowed to reach a value such as D_{in3}. In practice, we wish to avoid locking near D_{in3} and hence switch to the next curve. For example, to operate at ω_1, we prefer curve 2 to curve 1.

Figure 10.47 shows how coarse-tuning capacitors can be added to the LC DCO of Fig. 10.45(a).

10.7.5 DCO Topologies

The DAC fundamentals studied in the previous sections naturally point to certain DCO topologies. Table 10.1 summarizes our findings: (1) both LC and ring DCOs can be tuned by varactors that receive discrete voltage levels from a DAC, or by switched-capacitor arrays, and (2) ring DCOs can also be tuned by current-switching DACs. It is possible to combine these techniques as well. The use of capacitor arrays in ring oscillators poses an area penalty but generally introduces less phase noise than other tuning methods.

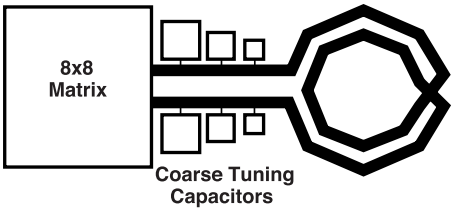


Figure 10.47 LC DCO using coarse tuning capacitors.

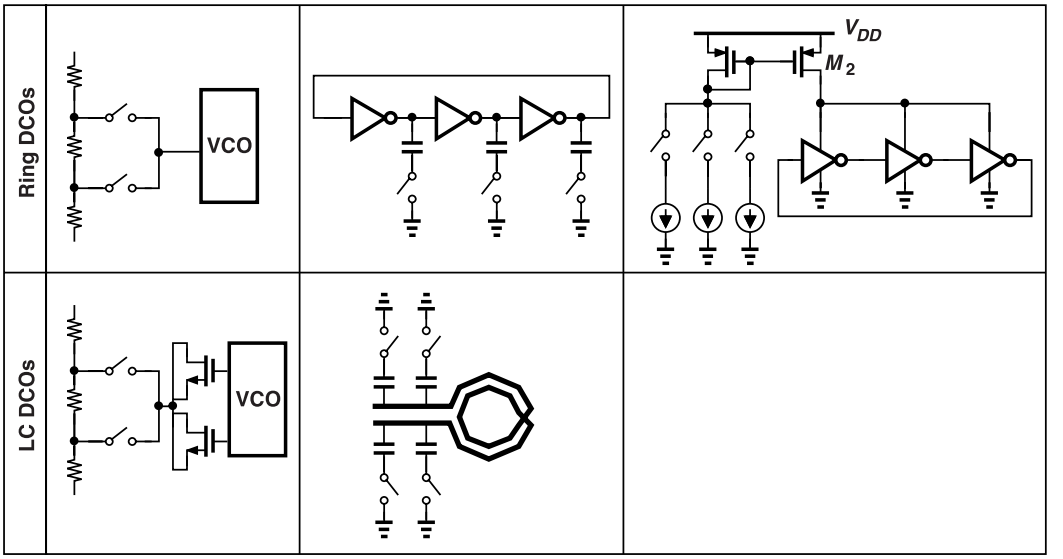


Table 10.1 Summary of different DCO implementations.

LC DCOs Figure 10.48 shows an LC DCO example incorporating a resistor ladder and switched capacitors [7]. The former provides 5 bits of resolution for achieving a fine step size and the latter allows coarse tuning. The filter consisting of R_1 and C_1 suppresses the ladder's transient glitches.

Example 10.22

Explain how the range overlaps shown in Fig. 10.46 can be guaranteed in Fig. 10.48.

Solution

We must run a number of simulations on the overall circuit. The simulations are only necessary for the extreme points of the ladder voltage. We choose the digital input to set V_{lad} to its minimum and determine the values of ω_{out} as the coarse input is incremented (Fig. 10.49). We repeat this process with the ladder voltage set to its maximum: $V_{lad} = V_{max}$. If some of the characteristics do not exhibit overlap, we must either reduce the coarse steps or increase the tuning range provided by the varactors. We typically opt for the former.

Figure 10.50 shows another LC DCO using both coarse and fine tuning [4]. Configured according to the matrix architecture of Fig. 10.44, the fine array consists of 1024 cells, each including two varactors. Depending on the local decoder's output, the varactors' gates assume a high or low value, thus tuning the oscillator

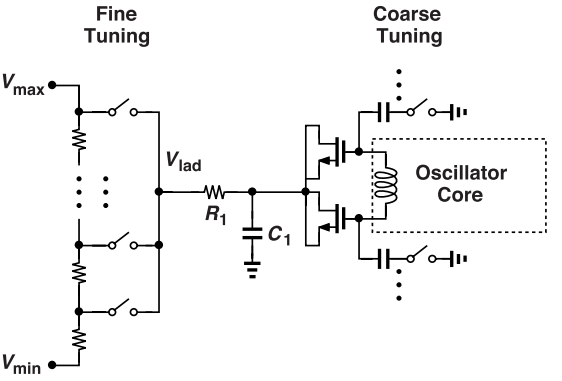


Figure 10.48 LC DCO using a resistor-ladder DAC and varactor tuning.

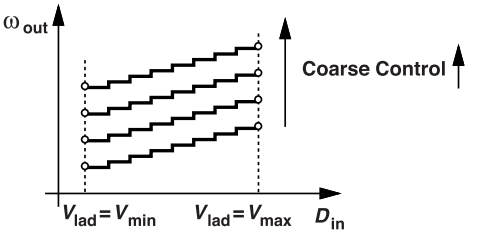


Figure 10.49 Illustration of how range overlap can be created.

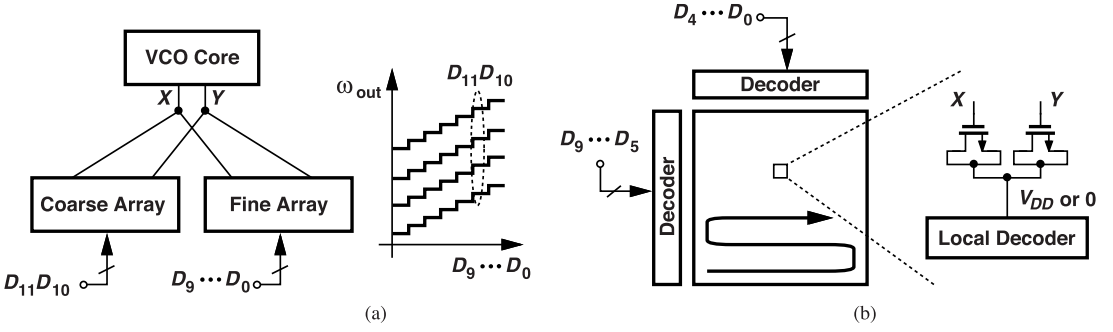


Figure 10.50 (a) LC DCO using coarse and fine capacitor arrays, and (b) matrix architecture showing zig-zag capacitor selection.

in small steps. Alternatively, we can replace these varactors with constant capacitors and switches. As mentioned previously, the long X and Y lines traveling across the array can degrade the tank Q significantly. Moreover, we must ensure some overlap between the successive fine-tuning curves to avoid blind zones. This means that the coarse jumps must be sufficiently small.

Figure 10.50(b) shows that the switching sequence within the array follows a zig-zag pattern.[4] That is, the cells in the bottom row turn on one by one from left to right, but those in the next row from right to left, etc. In Problem 10.19, we study whether this method alleviates the effect of gradients in the horizontal direction.

[4] For simplicity, we assume here that the rows turn on from the bottom of the matrix rather than from the top [as was the case in Fig. 10.44(b)].

Example 10.23

The supply noise of the local decoder in Fig. 10.50(b) modulates the varactors. Is this a serious issue?

Solution

It can be. As illustrated in Fig. 10.51, if the average varactor V_{GS} places the operating point around A, then the

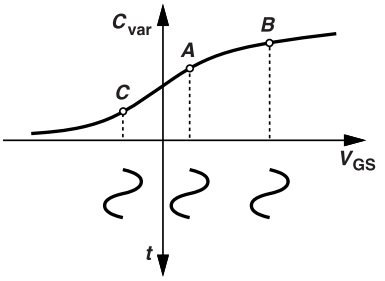


Figure 10.51 MOS varactor characteristic.

high varactor slope produces substantial modulation. That is, a given supply change leads to a large change in the average capacitance. On the other hand, points B or C are less troublesome.

Is it possible to combine the fine tuning techniques of Figs. 10.48 and 10.50? Yes, rather than apply just the high or low levels at the varactor gates in Fig. 10.50(b), we can provide finer voltage steps by means of a ladder [13]. Depicted in Fig. 10.52, such an arrangement achieves a resolution given by both DACs.

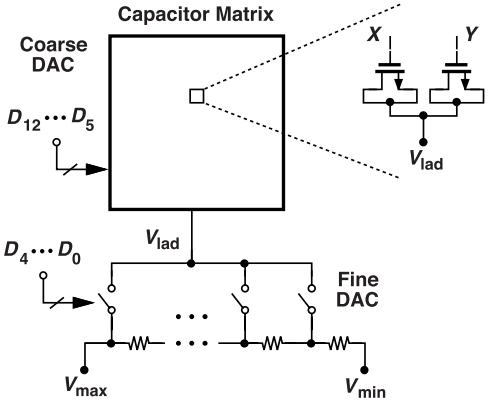


Figure 10.52 DCO combining tuning techniques of Figs. 10.48 and 10.50.

For example, [13] employs 8 bits for the segmented array and 5 bits for the ladder, thus reaching an overall resolution of 13 bits. Since V_{lad} varies monotonically from V_{min} to V_{max}, the frequency tuning characteristic is also monotonic for the fine DAC.

Example 10.24

Suppose the ladder in Fig. 10.52 generates eight voltage steps, each equal to ΔV. Study the nonlinearities that can appear in the overall DCO characteristic.

Solution

The first source of nonlinearity are the varactors themselves. As V_{lad} in Fig. 10.52 varies, the varactors tied to X and Y traverse their C-V characteristic, producing the behavior shown in Fig. 10.53(a).

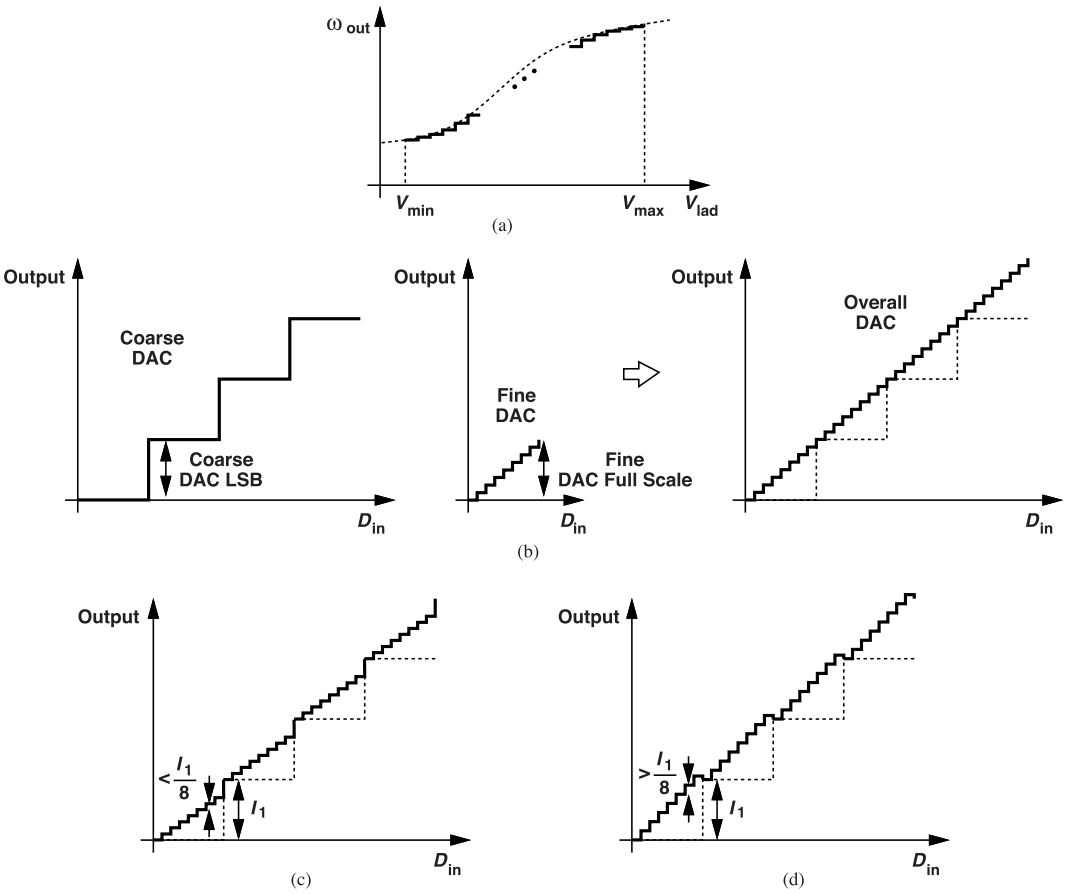


Figure 10.53 (a) DCO nonlinearity due to varactor nonlinearity, (b) ideal coarse and fine DACs yielding a linear overall characteristic, (c) effect of negative gain error in fine DAC, and (d) effect of positive gain error in fine DAC.

The second source relates to the "agreement" between the coarse and fine DACs, more specifically, between their gains. Consider first the ideal case depicted in Fig. 10.53(b), where two general coarse and fine DACs have equal gains, yielding a linear overall characteristic.[5] For example, the coarse DAC employs four unit current sources, each equal to I_1, and the fine DAC consists of eight unit current sources, each equal to $I_1/8$. In this case, the fine DAC full scale matches the coarse DAC LSB. On the other hand, if the fine DAC gain is lower (its unit source is less than $I_1/8$), then its full scale is less than the coarse DAC LSB and the characteristic exhibits positive jumps at the MSB transitions [Fig. 10.53(c)]. Similarly, if the fine DAC gain is higher, negative jumps appear [Fig. 10.53(d)]. In the latter case, the DCO gain becomes negative at these transitions.

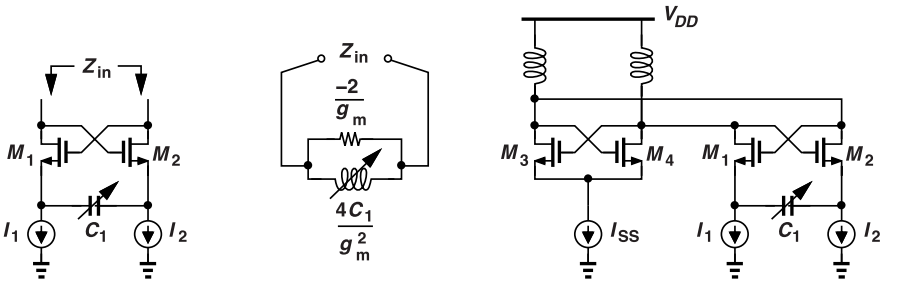
These jumps manifest themselves in the DCO of Fig. 10.52 if the full-scale provided by the ladder does not exactly match the LSB of the capacitor matrix. Returning to Fig. 10.53(a), we note that, as V_{lad} reaches V_{max}, ω_{out} assumes a certain value. The next frequency step is created by setting V_{lad} to V_{min} and turning

[5] This can be visualized more easily for current-steering DACs but is also applicable to capacitor DACs.

off one more varactor cell in the matrix of Fig. 10.52. It is difficult to guarantee that the resulting step is positive and equal to the desired value. In other words, we should avoid lock around the coarse/fine boundary.

The frequency jumps described above become problematic if the DCO locks in their vicinity. Then, as the DCO toggles, the output frequency makes large transitions, creating significant jitter. If the downward jumps in Fig. 10.53(d) are excessively large, the change in the DCO gain *polarity* may even prohibit the PLL from locking.

The need for extremely fine frequency steps in LC DCOs has led to interesting developments. In addition to combining DAC techniques, another idea is to "attenuate" the effect of switched unit capacitors on the frequency so that they need not be excessively small.[6] Shown in Fig. 10.54(a) is such an example [13]. The circuit negates the degeneration capacitance, yielding



(a) (b) (c)

Figure 10.54 (a) A topology generating a negative capacitance, (b) equivalent impedance, and (c) use of the circuit for fine tuning in a DCO.

$$Z_{in}(s) = -\frac{2}{g_m} - \frac{1}{C_1 s}. \tag{10.13}$$

If $s = j\omega$, the admittance is given by

$$Y_{in}(j\omega) = \frac{-g_m C_1 j\omega}{2C_1 j\omega + g_m}. \tag{10.14}$$

Multiplying the numerator and the denominator by the conjugate of the latter and assuming $4C_1^2\omega^2 \gg g_m^2$, we have

$$Y_{in}(j\omega) \approx -\frac{g_m}{2} - j\frac{g_m^2}{4C_1\omega}. \tag{10.15}$$

Since the admittance of an inductor L reads $1/(jL\omega) = -j/(L\omega)$, we recognize that the second term in (10.15) represents a parallel inductance [Fig. 10.54(b)]. If this inductance is much greater than that of the tank, then changes in C_1 translate to small frequency steps. An alternative perspective is to rewrite (10.15) as

$$Y_{in}(j\omega) \approx -\frac{g_m}{2} - j\omega C_1 \left(\frac{g_m}{2\omega C_1}\right)^2, \tag{10.16}$$

and consider $[g_m/(2\omega C_1)]^2$ as a factor by which the value of C_1 is shrunk [13]. For example, $[g_m/(2\omega C_1)]^2 = 1/10$ yields a tenfold attenuation for changes in C_1.

[6] Unit capacitors well below 1 fF are difficult to implement because they experience significant fringe fields that terminate on the neighboring units.

The transforming circuit of Fig. 10.54(a) can be readily combined with an LC oscillator, as shown in Fig. 10.54(c) [13]. Nonetheless, the differential noise injected by I_1 and I_2 must be managed.

Ring DCOs As illustrated in Table 10.1, a ring DCO can be simply realized by preceding a ring VCO with a resistor-ladder or current-switching DAC. The table illustrates how a current-switching DAC controls an inverter-based ring oscillator. To arrive at another example, let us ask, how do we tune a differential ring oscillator by means of a *current*? Consider the stages shown in Figs. 10.55(a) and 10.55(b): in the former, I_{SS} has little effect on the delay, and in the latter, changing I_1 also changes the output swing, an undesirable

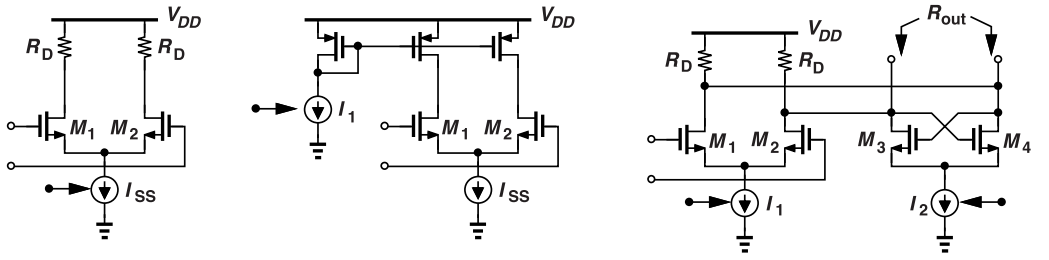


Figure 10.55 Frequency tuning by means of (a) the tail current source, (b) variable load resistors, and (c) a negative resistance.

effect. Now, we examine the circuit in Fig. 10.55(c), where I_1 and I_2 vary in opposite directions so that $I_1 + I_2$ remains relatively constant, thereby producing constant output swings. We note that as I_2 increases, the regenerative pair presents a "stronger" negative resistance, $-2/g_{m3,4}$, to the output port, raising the total resistance and hence the delay [8]:

$$R_{out} = (2R_D)||\frac{-2}{g_{m3,4}} \tag{10.17}$$

$$= \frac{2R_D}{1 - g_{m3,4}R_D}. \tag{10.18}$$

In the limit, $g_{m3,4}R_D$ approaches unity and M_3 and M_4 latch up. A ring oscillator can incorporate, for example, three such stages and can be tuned by adjusting I_1 and I_2 in each stage. Of course, the noise current in I_1 and I_2 appears in the control path and directly modulates the frequency. Based on this idea, the DCO in [9] employs a 10-bit segmented current-switching DAC for adjusting I_1 and I_2. Illustrated in Fig. 10.56 is one stage of the DCO. This work also proposes a decoding scheme for the DAC matrix that minimizes the glitch due to timing skews.

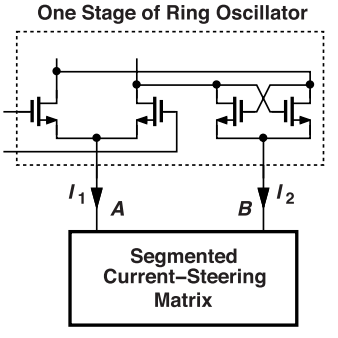
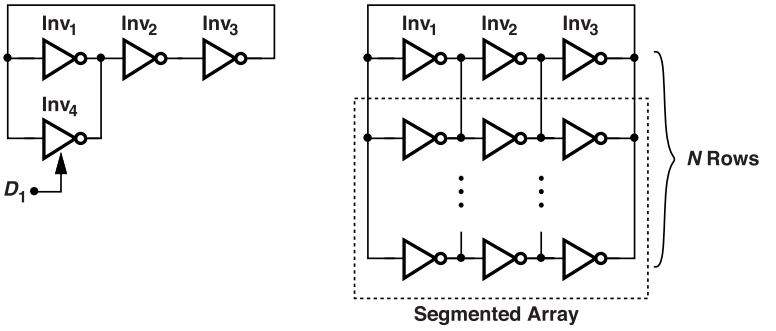


Figure 10.56 Digital control of the currents in a ring oscillator.

Another approach to discrete tuning of ring oscillators draws upon the principle of "array filling." Suppose, as shown in Fig. 10.57(a), we add an inverter in parallel with one stage in a ring oscillator. What happens if



(a) (b)

Figure 10.57 (a) Tuning a ring oscillator frequency by turning an inverter on and off, and (b) generalization to obtain a wide tuning range.

this inverter is disabled or enabled? We observe that the input capacitance of Inv_4 changes by a small amount in the off and on states, but its drive strength directly adds to that of Inv_1. Thus, when Inv_4 is turned on, the oscillation frequency increases [10].

Figure 10.57(b) extends the idea to a large number of inverters in the form of a segmented array. As more inverters within the array are turned on (as the array is "filled"), the frequency rises. In contrast to the topology in Fig. 10.55(c), this arrangement is free from noise in the control path, but its frequency resolution is limited. The following example elaborates on this point.

Example 10.25

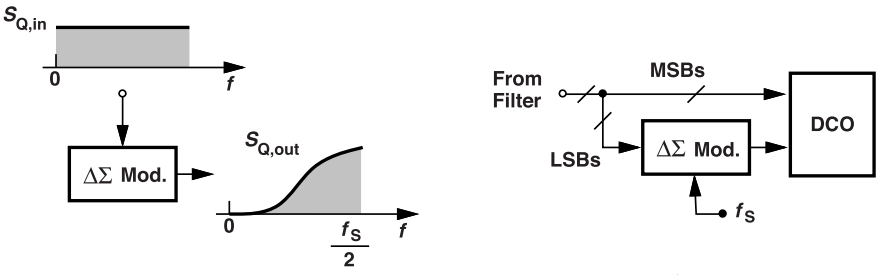
Estimate the frequency resolution that the array filling technique shown in Fig. 10.57(b) can achieve. Assume N is large.

Solution

Suppose the array consists of N rows, each having three inverters. When all inverters are enabled, the circuit reduces to a simple three-stage ring, operating at a frequency of $f_{max} = 1/(6T_D)$, where T_D is the gate delay. We can view the situation as if each inverter in a simple three-stage ring were scaled up by a factor of N; since both the strength and the capacitances of the inverters are scaled by this factor, the oscillation frequency is the same as that of an unscaled ring.

When all rows but the top one are disabled, the inverters still see a load capacitance roughly equal to that of N inverters in parallel, generating an oscillation frequency of about f_{max}/N. Thus, the total tuning range is given by $f_{max} - f_{max}/N \approx f_{max}$. With $3(N-1)$ switchable inverters, this translates to a frequency resolution of $(f_{max} - f_{max}/N)/[3(N-1)] = f_{max}/(3N)$.

The DCO resolution can also be improved at the DPLL architecture level. Suppose the DCO resolution is insufficient, i.e., its quantization noise is unacceptably high. We can draw upon the concept of "noise shaping" to suppress this noise. We describe this concept in detail in Chapter 12, but point out here that a "$\Delta\Sigma$ modulator" can receive a digital input containing quantization noise and *shape* its spectrum by attenuating its low-frequency components [Fig. 10.58(a)]. The peak of the shaped spectrum occurs at $f_S/2$, where f_S denotes the $\Delta\Sigma$ modulator's clock frequency. We can then apply the LSBs controlling the DCO to a $\Delta\Sigma$ modulator [Fig. 10.58(b)] so as to push the quantization noise to high frequencies. The DCO transfer function, K_{DCO}/s, suppresses these noise components—provided that $S_{Q,out}K_{DCO}^2/(4\pi f)^2$ is actually small enough as f approaches f_S. This requires that f_S be as high as possible, e.g., $f_S = f_{VCO}/4$.



(a) (b)

Figure 10.58 (a) Noise shaping by a Δ-Σ modulator, and (b) use in a digital PLL to improve DCO resolution.

10.8 Loop Dynamics

We have developed a design procedure for analog PLLs in Chapter 9 and wish to extend it to digital PLLs as well. In this section, we explore this possibility.

10.8.1 Digital Filter Implementation

Let us first implement the loop filter in the digital domain. We seek a transfer function similar to

$$H(s) = R_1 + \frac{1}{C_1 s}, \tag{10.19}$$

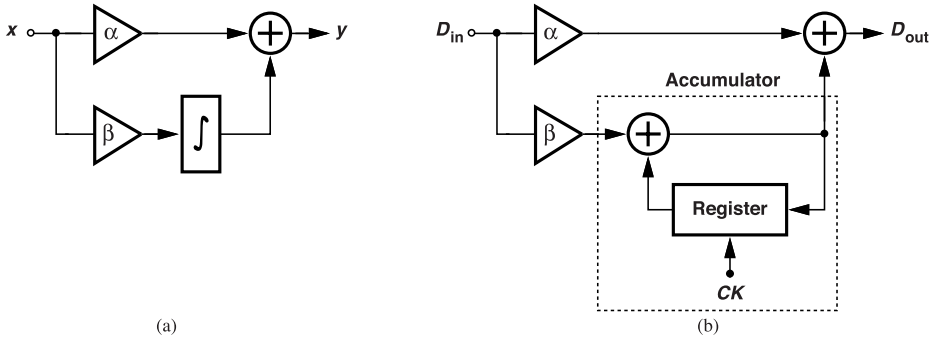
i.e., one with a proportional path and an integral path (called a PI realization in control theory). This brings us to the conceptual topology shown in Fig. 10.59(a), where

$$y = \alpha + \beta \int x dt \tag{10.20}$$

for continuous-time quantities. Here, $\alpha = R_1$ and $\beta = 1/C_1$. To arrive at a digital counterpart, we change x to D_{in} and y to D_{out} and write this equation as a discrete-time expression:

$$D_{out}(k) = \alpha + \beta \sum_{k=0}^{\infty} D_{in}(k). \tag{10.21}$$

Taking the z transform of both sides and noting that a discrete-time integrator is represented by $1/(1 - z^{-1})$,



(a) (b)

Figure 10.59 (a) Filter using proportional and integral paths, and (b) its digital implementation.

we obtain the following transfer function:

$$\frac{D_{out}}{D_{in}}(z) = \alpha + \frac{\beta}{1 - z^{-1}}. \tag{10.22}$$

Depicted in Fig. 10.59(b), the actual realization incorporates an accumulator consisting of an adder and a register that returns the output to the adder after one clock delay. The clock is typically the same as the PLL reference. It is tempting to create a correspondence between the analog values, R_1 and C_1, and the digital counterparts, α and β. The following example makes such an attempt.

Example 10.26

Using a time-domain analysis, estimate α and β in terms of R_1, C_1, and T_{CK}.

Solution

Let us examine the step responses of the analog and digital loop filters. As shown in Fig. 10.60(a) for $I_{in} = u(t)$, the former exhibits a jump equal to $I_{in}R_1$ followed by a ramp having a slope of I_{in}/C_1. From

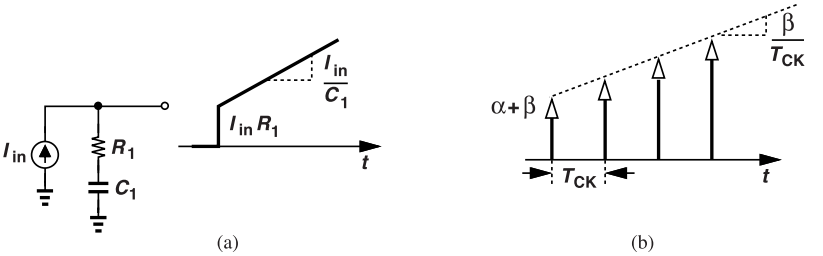


Figure 10.60 (a) Response of an RC branch to a current step, and (b) response of digital filter counterpart.

Fig. 10.59(b), the latter instantaneously generates an output equal to $(\alpha + \beta)u(k)$, where $u(k)$ is the unit step, followed by a digital ramp having a slope of β/T_{CK}. This is because the accumulator output in Fig. 10.59(b) changes by an amount equal to βD_{in} in every clock cycle. We equate the jumps and the slopes in Figs. 10.60(a) and (b), obtaining $\beta = T_{CK}/C_1$ and $\alpha = R_1 - T_{CK}/C_1$. This intuitive approach finds a correspondence between the two systems, but the results are not correct! We elaborate on this point below.

10.8.2 Correspondence between Analog and Digital PLLs

Before establishing proper correspondence between analog and digital PLLs, let us make two observations. First, recall from Fig. 10.6 that an LSB size of Δ for a TDC leads to a slope of $1/\Delta$, which has a unit of $(\text{second})^{-1}$. To change the unit to $(\text{radian})^{-1}$, we divide Δ by the input period, T_{REF}, and multiply the result by 2π. That is, the TDC gain is equal to $T_{REF}/(2\pi\Delta)$. Second, the digitally-controlled oscillator in a DPLL can be approximately modeled by K_{DCO}/s, where K_{DCO} denotes the gain and is equal to the change in the output frequency (in rad/s) for a 1-LSB change in the DCO's digital input.

We now construct the analog and digital PLL models shown in Fig. 10.61. The DPLL model includes both discrete-time and continous-time transfer functions; so we seek a CT approximation for $1-z^{-1}$. Representing a one-clock delay, z^{-1} can be modeled by the Laplace transform of a delay line, $\exp(-sT_{REF})$. Since the loop bandwidth is typically much less than $1/T_{REF}$, for frequencies of interest we can write $|sT_{REF}| \ll 1$ and hence $\exp(-sT_{REF}) \approx 1 - sT_{REF}$. That is, $\beta/(1 - z^{-1}) \approx \beta/(sT_{REF})$. We can now equate the two loops' forward transfer functions:

$$\frac{I_p}{2\pi}\left(R_1 + \frac{1}{C_1 s}\right)\frac{K_{VCO}}{s} = \frac{T_{REF}}{2\pi\Delta}\left(\alpha + \frac{\beta}{sT_{REF}}\right)\frac{K_{DCO}}{s}. \tag{10.23}$$

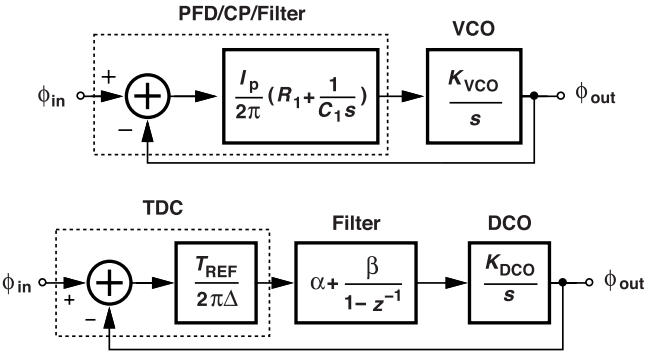


Figure 10.61 Models of analog and digital PLLs.

Equating the coefficients of $1/s$ yields

$$\frac{I_p R_1 K_{VCO}}{2\pi} = \frac{T_{REF} K_{DCO}}{2\pi\Delta}\alpha. \tag{10.24}$$

Similarly, if the coefficients of $1/s^2$ are set equal:

$$\frac{I_p K_{VCO}}{2\pi C_1} = \frac{T_{REF}\beta K_{DCO}}{2\pi\Delta T_{REF}}. \tag{10.25}$$

It follows that

$$\alpha = I_p R_1 \frac{\Delta}{T_{REF}} \frac{K_{VCO}}{K_{DCO}} \tag{10.26}$$

$$\beta = \frac{I_p \Delta}{C_1} \frac{K_{VCO}}{K_{DCO}}. \tag{10.27}$$

These equations allow us to design the digital filter based on the analog loop parameters (I_p, R_1, K_{VCO}, and C_1) and those of the digital loop (Δ and K_{DCO}). The results remain unchanged in the presence of a feedback divider.

It is also possible to design the digital PLL directly. From the forward transfer function in Eq. (10.23), the reader can find the closed-loop parameters:

$$\zeta = \frac{\alpha T_{REF}}{2}\sqrt{\frac{K_{DCO}}{2\pi\beta\Delta}} \tag{10.28}$$

$$\omega_n = \sqrt{\frac{\beta K_{DCO}}{2\pi\Delta}}. \tag{10.29}$$

References

[1]. T. Kobayashi et al., "A current-mode latch sense amplifier and a static power saving input buffer for low-power architecture," *VLSI Circuits Symp. Dig. Tech. Papers,* pp. 28-29, June 1992.

[2]. J. Montanaro et al., "A 160-MHz 32-b 0.5-W CMOS RISC microprocessor," *IEEE J. Solid-State Circuits,* vol. 31, pp. 1703-1714, Nov. 1996.

[3]. T. Rahkonen and J. Kostamovaara, "The use of stabilized CMOS delay lines for the digitization of short time intervals," *IEEE J. Solid-State Circuits*, vol. 28, pp. 887-894, Aug., 1993.

[4]. T. Pittorino et al., "A UMTS-compliant fully digital controlled oscillator withh 100-MHz fine-tuning range in 0.13-μm CMOS," *ISSCC Dig. Tech. Papers,* pp. 210-211, Feb. 2006.

[5]. M. Chen, D. Su, and S. Mehta, "A calibration-free 800-MHz fractional-N digital PLL with embedded TDC," *ISSCC Dig. Tech. Papers*, pp. 472-473, Feb. 2010.

[6]. J. Borremans et al., "An 86-MHz-to-12-GHz digital-intensive phase-modulated fractional-N PLL using a 15-pJ/shot 5-ps TDC in 40-nm digital CMOS," *ISSCC Dig. Tech. Papers*, pp. 480-481, Feb. 2010.

[7]. D. Tasca et al., "A 2.9-to-4.0-GHz fractional-N digital PLL with bang-bang phase detector and 560 fs,rms integrated jitter at 4.5 mW power," *ISSCC Dig. Tech. Papers*, pp. 88-89, Feb. 2011.

[8]. B. Razavi, Ed., *Monolithic Phase-Locked Loops and Clock Recovery Circuits,* New York, IEEE Press, 1996.

[9]. J. Lin et al., "A PVT-tolerant 0.18 MHz to 600 MHz self-calibrated digital PLL in 90-nm CMOS process," *ISSCC Dig. Tech. Papers*, pp. 488-489, Feb. 2004.

[10]. A. V. Rylyakov et al., "A wide power supply range (0.5 V to 1.3 V) wide tuning range (500 MHz to 8 GHz) all-static CMOS AD PLL in 65 nm SOI," *ISSCC Dig. Tech. Papers*, pp. 172-173, Feb. 2007.

[11]. D. Miyashita et al., "A -104dBc/Hz in-band phase noise 3-GHz all-digital PLL with phase interpolation based hierarchical time-to-digital convertor," *Dig. Symp. VLSI Circuits*, pp. 112-113, June 2011.

[12]. F. Opteynde, "A 40-nm CMOS all-digital fractional-N synthesizer without requiring calibration," *ISSCC Dig. Tech. Papers*, pp. 346-347, Feb. 2012.

[13]. L. Fanori, A. Liscidini, and R. Castello, "3.3GHz DCO with a frequency resolution of 150Hz for all-digital PLL," *ISSCC Dig. Tech. Papers*, pp. 48-49, Feb. 2010.

Problems

10.1. Plot the quantizer characteristic shown in Fig. 10.3(a) if comparator A_j in Fig. 10.3(b) has an offset less than 1 LSB. Consider both positive and negative offsets. What happens as the offset approaches $+1$ LSB or -1 LSB.

10.2. We wish to study the effect of offset in the interpolating flash ADC of Fig. 10.5(c). First, assume differential pair A_1 in Fig. 10.5(a) has an offset $V_{os} < 1$ LSB. What happens to the characteristics in Fig. 10.5(b)? Now, suppose the differential pair at the input of comparator A_j in Fig. 10.3(b), where j is an odd number, has an offset $V_{os} < 1$ LSB, and plot the characteristic. Finally, assume we remove the even-numbered differential pairs in Fig. 10.3(b) and use interpolation to create the missing decision thresholds. Explain how the offset A_1 affects the characteristic.

10.3. Suppose Latch 4 in Fig. 10.5(c) has a small offset. What is the effect on the overall characteristic?

10.4. Looking at the TDC concept in Fig. 10.6, a student reasons that an XOR acting as a PD generates an analog output, which can be digitized by a flash ADC so as to provide a digital representation. Is this a good idea? (Recall that only the *average* value of the XOR output is proportional to the phase error.)

10.5. Suppose t_0 in Fig. 10.7(c), i.e., the phase difference between V_2 and V_1, varies from near zero to above 180°. Plot the TDC input-output characteristic.

10.6. An engineer mistakenly includes a small delay in V_2 in Fig. 10.7(b). Explain what happens to the characteristic obtained in the previous problem.

10.7. Consider the TDC noise shaping depicted in Fig. 10.12. What happens to $S_\phi(f)$ if the gain of the oscillator (similar to K_{VCO}) is doubled? Approximate the PLL by an analog, continuous-time loop.

10.8. In Fig. 10.18(b), how long do V_X and V_Y take to reach $V_{DD} - V_{THN}$? Assume $V_X \approx V_Y$ for simplicity.

10.9. Using the result from the previous problem and assuming that $I_{D1} - I_{D2} = g_{m1,2}(V_{in1} - V_{in2})$, determine $V_X - V_Y$ at $t = t_1$ and hence obtain the voltage gain. How do we choose the circuit's parameters for a gain of 3?

10.10. Draw the StrongArm latch of Fig. 10.18(d) without M_3 and M_4 and assume V_{in1} is high and V_{in2} is low. Explain why the circuit incurs static power consumption in this phase.

10.11. In Fig. 10.20, does the delay of the inverters following the StrongArm latch introduce a phase offset in the TDC?

10.12. Design the logic necessary for the bubble correction scheme in Fig. 10.21. A truth table proves helpful here.

10.13. In the vernier TDC of Fig. 10.22, we have $\Delta_1 = 15$ ps and $\Delta_2 = 10$ ps. Determine the code generated by the flipflops as we sweep the phase difference between V_1 and V_2 from 0 to 20 ps.

10.14. In the vernier TDC of Fig. 10.22, Δ_1 and Δ_2 are nominally equal to 15 ps and 10 ps. Suppose the mismatch between the third delay stages causes $\Delta_2 = 17$ ps. Explain what happens to the output code as we sweep the phase difference between V_1 and V_2 from 0 to 20 ps.

10.15. Plot the input-output characteristic of the TDC in Fig. 10.24(a) if the second delay in V_1 path is mistakenly changed from Δ to 1.2Δ.

10.16. Explain what happens to the output code generated by the flipflops in Fig. 10.24(a) if the delay of 1.5Δ is mistakenly changed to 0.8Δ.

10.17. Suppose each unit I_u in Fig. 10.40 is accompanied with a noise current $\overline{I_n^2}$. That is, as more current sources are switched into M_1, their noise also increases. How does this affect the oscillator's phase noise as we tune it from low frequencies to high frequencies?

10.18. Consider the thermal noise generated by the ladder in Fig. 10.48. Explain for what digital input this noise causes maximum phase noise. Assume only V_{lad} is changing with the code.

10.19. Suppose the matrix in Fig. 10.50(b) suffers from a linear horizontal gradient in the value of the varactors of the form $C_k = C_0 + (k-1)\Delta C$, where C_0 denotes the value of the leftmost cell in a row and ΔC is the increment per cell. Determine whether the zig-zag switching suppresses the accumulation of this error.

10.20. Repeat the previous problem while allowing the digital input to change the coarse-tuning capacitors as well.

10.21. Derive Eq. (10.14), multiply the numerator and the denominator by the conjugate of the latter, and prove Eq. (10.15).

11

Delay-Locked Loops

Delay-locked loops (DLLs) are newer development than PLLs, dating back to the 1980s. In some applications, DLLs are necessary or preferable over PLLs, offering such advantages as lower sensitivity to supply noise, lower phase noise, and the ability to generate multiple phases. The reader is encouraged to study the PLL fundamentals in Chapters 7 and 8 before beginning this chapter.

11.1 Basic Idea

Suppose, as shown in Fig. 11.1(a), an input clock travels through a buffer, B_1, a long interconnect, and another buffer, B_2, experiencing a significant skew, ΔT. How do we align CK_{out} with CK_{in}? Since the clock is periodic, we surmise that an *additional* delay can be introduced in B_2 so as to make the total delay equal to one clock cycle [Fig. 11.1(b)]. To set the delay properly, we view ΔT as an error that can be suppressed by means of negative feedback. That is, if the phase of CK_{out} is compared to that of CK_{in}, the resulting error can be used to adjust the delay and remove the skew. This conjecture leads us to the arrangement depicted in Fig. 11.1(c). Here, the phase detector measures the skew and adjusts the delay of B_2 so as to reduce ΔT. As with PLLs, the low-pass filter attenuates the high-frequency components generated by the PD. This circuit exemplifies a simple delay-locked loop.

The residual phase error in Fig. 11.1(c) depends on the loop gain, i.e., the gain of the PD, K_{PD}, and the gain of the variable-delay stage. The latter is defined as $K_{DL} = \partial\phi/\partial V_{cont}$, where ϕ is the stage's delay in radians. Rather than attempt to maximize $K_{PD}K_{DL}$, we can add an integrator to the loop. Drawing upon our knowledge of PLLs, we thus construct the architecture shown in Fig. 11.1(d), where the PFD/CP/capacitor combination provides an infinite gain, driving the skew toward zero. The variable-delay stage is realized as a voltage-controlled delay line (VCDL). While the loop does not require frequency detection (why?), the PFD offers a convenient interface with the CP. As explained below, no resistor is necessary in series with C_1. This DLL architecture is commonly used in high-speed systems.

Since the VCDL responds (almost) instantaneously to a change in V_{cont}, the DLL of Fig. 11.1(d) is of first order, facing no stability issues. Moreover, it benefits from the lower phase noise and supply sensitivity of delay lines compared to oscillators. We elaborate on these points in the following sections.

In contrast to PLLs, delay-locked loops do not *generate* a frequency; rather, they simply delay the input. As such, DLLs are less versatile than PLLs. For example, a DLL would not be able to generate a 5-GHz clock from a 20-MHz reference.

Another drawback of DLLs is that they allow the input duty cycle error to propagate to the output. In fact, the delay line may further increase this error. Thus, the VCDL is typically preceded and/or followed by a duty cycle correction stage (Section 11.10). A third drawback of DLLs is that they operate the PFD and the CP at high speeds. We return to this point in Section 11.9. This issue can be alleviated by inserting two frequency dividers at the PFD inputs.

The dynamic behavior of DLLs determines how they respond to such effects as input phase noise, supply noise, etc. We therefore study this behavior in the next section.

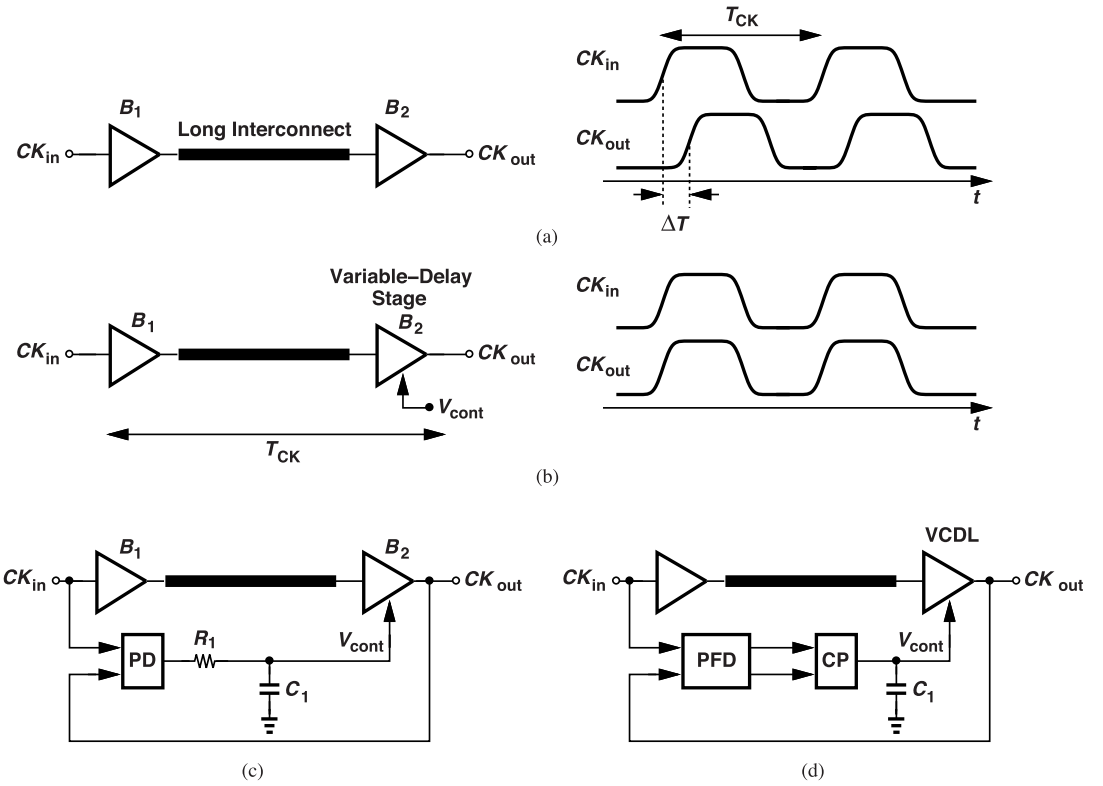


Figure 11.1 (a) Clock distribution suffering from skew, (b) correction of skew by a variable-delay stage, (c) simple feedback system for controlling the delay, and (d) basic DLL.

11.2 Loop Dynamics

We wish to analyze the dynamic behavior of the DLL shown in Fig. 11.2(a), where, for simplicity, the entire clock path is modeled by a VCDL. In the locked state, the phase difference between CK_{in} and CK_{out} is constant and, in principle, equal to zero. Thus, the VCDL provides a delay of one clock period, T_{CK}. The VCDL can be realized using the ring oscillator tuning techniques studied in Chapter 3. For example, a chain of varactor-controlled inverters can serve this purpose [Fig. 11.2(b)].

How do we express the static behavior of a VCDL? In analogy with VCOs, we may write $\phi_{out} = \phi_0 + K_{DL}V_{cont}$, where ϕ_0 is the intercept and K_{DL} the gain. But this is incorrect because the VCDL shifts the *input* phase. We must therefore write

$$\phi_{out} = \phi_0 + K_{DL}V_{cont} + \phi_{in}. \tag{11.1}$$

For small-signal analysis, we can view the VCDL as a circuit that multiplies V_{cont} by K_{DL} to create a phase shift of $\phi_{out} - \phi_{in}$.

Before delving into the overall loop dynamics, let us understand those of the VCDL itself. The circuit has a clock input and a control input. What happens if CK_{in} in Fig. 11.2(b) incurs a phase step? This step propagates through the chain and emerges at the output T_{CK} seconds later [Fig. 11.2(c)]. That is, the transfer function associated with this path can be expressed as $\exp(-T_{CK}s)$. In practice, T_{CK} is much less than the overall DLL time constant, allowing the approximation $\exp(-T_{CK}s) \approx 1$.

The second path is from V_{cont} to CK_{out}. If we apply a step at V_{cont} in Fig. 11.2(b), how long does it take to affect the output phase? From the waveforms shown in Fig. 11.2(d), we recognize that this path, too, has a

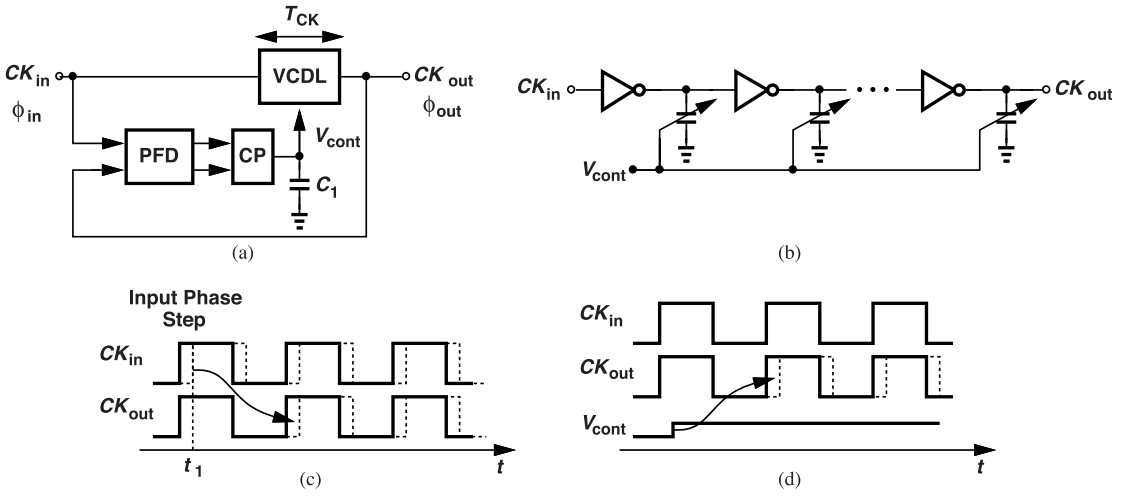


Figure 11.2 (a) DLL with voltage-controlled delay line, (b) simple VCDL implementation, (c) propagation of input phase step to output, and (d) propagation of step on V_{cont} to output.

delay of at most one T_{CK}. Thus, the static model, $\phi_{out} = \phi_0 + K_{DL}V_{cont} + \phi_{in}$, still proves fairly accurate.

It is instructive to first examine the overall DLL's response qualitatively. If the phase of CK_{in} in Fig. 11.2(a) fluctuates slowly, the DLL maintains a high loop gain, keeping CK_{out} aligned with CK_{in}. That is, the closed-loop transfer function has a unity magnitude for slow phase variations. Now suppose CK_{in} experiences very fast phase changes. Then, the DLL has little loop gain, V_{cont} does not change, and CK_{in} simply propagates to CK_{out}. In this case, too, the closed-loop response is around unity because the input phase changes appear at the output with only a delay of T_{CK} seconds. We thus conclude that DLLs exhibit an all-pass response, a point of contrast to the low-pass behavior of PLLs.

The all-pass nature of DLLs can also be confirmed mathematically. For the DLL of Fig. 11.2(a), we draw the phase model as shown in Fig. 11.3(a), where the PFD is represented by a phase subtractor, the CP and the capacitor by an integrator, and the VCDL gain by K_{DL}. Since V_{cont} is given by $(\phi_{in} - \phi_{out})[I_p/(C_1 s)]$, and

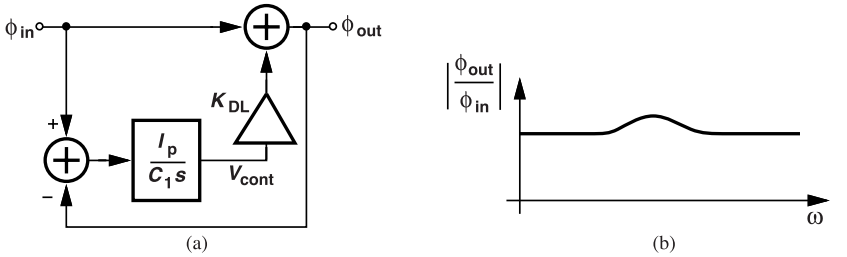


Figure 11.3 (a) Linear model of DLL, and (b) DLL response.

since the VCDL phase shift, $\phi_{out} - \phi_{in}$, is equal to $K_{DL}V_{cont}$, we have

$$\phi_{out} - \phi_{in} = (\phi_{in} - \phi_{out})\frac{I_p}{C_1 s}K_{DL}. \tag{11.2}$$

which implies $\phi_{in} = \phi_{out}$. In practice, the response exhibits a small amount of peaking due to the finite delay through the VCDL [Fig. 11.3(b)] [1].

The above study reveals two points. (1) DLLs do not generally face stability issues and can operate with a wide range of values for I_p and C_1, and (2) the lack of filtering ability precludes the use of DLLs in applications where the input jitter must be removed.

11.3 Choice of Number of Delay Stages

A delay line is typically formed as a chain of inverters, exhibiting trade-offs among the delay, the phase noise, and the power consumption. For a given total delay, T_D, we can choose (1) a sufficient number of inverters with a fanout of 1, or (2) a smaller number of inverters, each loaded by additional capacitance. The question is which method provides a more favorable trade-off between the phase noise and the power consumption. This point becomes particularly relevant if the ratio of T_D and the gate delay is large, for example, if $T_D = 2$ ns and the gate delay in the FF, low-temperature corner is 5 ps.

To answer this question, we make two observations. First, according to our studies in Chapter 3, the phase noise of a ring oscillator is relatively independent of the number of stages and the load capacitance seen by each stage. Second, if we open a ring oscillator loop to create a delay line, the phase noise profiles of the two are related by [2]:

$$S_{\phi,ring}(f) = \left(\frac{f_0}{\pi f}\right)^2 S_{\phi,DL}(f), \tag{11.3}$$

where f denotes the offset from f_0. We then conclude that $S_{\phi,DL}$, too, is independent of the number of stages and the load capacitance. In other words, for a given T_D, the noise-power trade-off of a delay line is relatively rigid.

Example 11.1

A student reasons that adding capacitances to the inverters' outputs lowers the supply sensitivity by overwhelming the (voltage-dependent) drain-bulk capacitances. Is this correct?

Solution

No, it is not. As we discovered in Chapter 3, additional capacitances do not reduce the supply sensitivity because this sensitivity is dominated by the inverters' driving strength, which is still modulated by the supply.

11.4 Effect of Nonidealities

The imperfections encountered in PLLs are also present in DLLs, but most have a less serious effect on the performance. In this section, we deal with these issues.

11.4.1 PFD/CP Nonidealities

As described in Chapters 7 and 8, the most critical consequence of the PFD/CP imperfections is the ripple on the control voltage as it causes jitter and sidebands in PLLs. In DLLs, on the other hand, it proves benign, as shown in the following example.

Example 11.2

Determine the effect of the control voltage ripple in Fig. 11.2(a).

Solution

The ripple arising from PFD/CP nonidealities repeats at the input frequency, f_{in}. That is, the phase difference between CK_{in} and CK_{out} is modulated at a rate equal to f_{in}. As shown in Fig. 11.4, this simply means that CK_{out} experiences a *constant* phase offset, exhibiting no sidebands or jitter; regardless of the shape of the ripple, the rising edge of CK_{out} is advanced (or retarded) in every period. We can verify our intuition analytically by approximating the phase-modulated output as

$$V_{out}(t) = V_0 \cos[\omega_{in} t + K_{DL} V_{cont}(t)]. \tag{11.4}$$

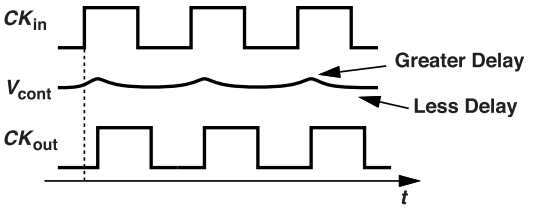


Figure 11.4 Ripple on control voltage of a DLL.

If we substitute $t + 2\pi/\omega_{in}$ for t, we observe that $V_{out} = V_0 \cos[\omega_{in}t + 2\pi + K_{DL}V_{cont}(t + 2\pi/\omega_{in})]$ does not change. That is, V_{out} is still periodic. This property may appear to be another advantage of DLLs over PLLs, but if a PLL operates with a feedback divide ratio equal to 1, it, too, produces a sideband-free output. The phase offset observed above proves problematic as it introduces a skew between CK_{in} and CK_{out}.

The foregoing example suggests that in Fig. 11.2(a), C_1 need not be so large or I_p so small as to yield a low ripple. As explained in the next section, this flexibility helps suppress the effect of supply noise. However, the ripple becomes problematic in frequency-multiplying DLLs (Section 11.6).

11.4.2 Supply Noise

The principal effect of supply noise, $V_{DD}(t)$, in DLLs is to modulate the delay of the VCDL. How does the DLL of Fig. 11.2(a) respond to $V_{DD}(t)$? If the noise varies slowly, the loop has enough "strength" to keep ϕ_{out} close to ϕ_{in}, i.e., V_{cont} opposes $V_{DD}(t)$ and ϕ_{out} is not much affected. For high-frequency noise, on the other hand, the loop gain drops and ϕ_{out} is directly modulated by V_{DD}.

Let us define for the VCDL a gain from V_{DD} to ϕ_{out} as $K_{VDD} = \partial\phi_{out}/\partial V_{DD}$. Shown in Fig. 11.5(a) is the DLL model with supply noise and $\phi_{in} = 0$. Beginning from the output, we can write V_{cont} as $-\phi_{out}[I_p/(C_1 s)]$ and hence

$$-\phi_{out}\frac{I_p}{C_1 s}K_{DL} + V_{DD}K_{VDD} = \phi_{out}. \tag{11.5}$$

It follows that

$$\frac{\phi_{out}}{V_{DD}}(s) = \frac{K_{VDD}C_1 s}{C_1 s + I_p K_{DL}}. \tag{11.6}$$

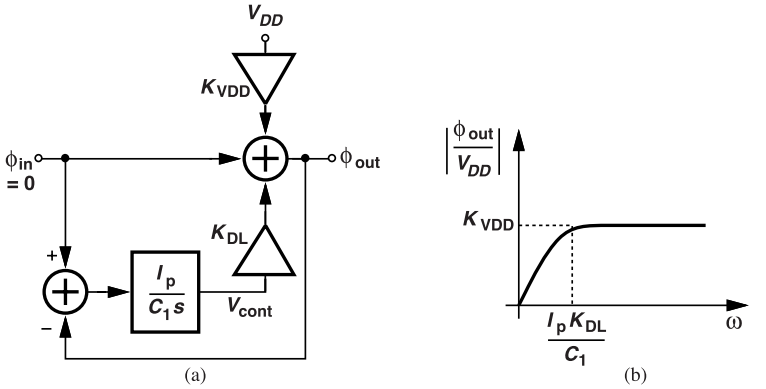


Figure 11.5 (a) Supply noise in a DLL, and (b) DLL phase response.

Plotted in Fig. 11.5(b), the response begins to flatten out beyond the pole frequency, $\omega_p = I_p K_{DL}/C_1$. We must therefore choose a high value for I_p/C_1 so as to maximize the supply rejection.

In contrast to the PLL supply noise response studied in Chapter 8, the above results reveal that (1) the DLL has no ability to suppress high-frequency supply noise, but (2) the corner frequency, $I_p K_{DL}/C_1$, can be chosen much greater than the loop bandwidth of a PLL because of the much higher DLL input frequency and its more stable behavior.

11.4.3 Phase Noise

DLLs are generally considered to generate much less phase noise than do PLLs, but the comparison must be done carefully. Let us begin with the input (reference) phase noise. As exemplified by the transfer function plotted in Fig. 11.3(b), this noise experiences no attenuation and simply propagates to the output.

The case of the VCDL phase noise is more interesting. We make two observations. First, as formulated in Chapter 3 and mentioned in Section 11.3, the phase noise of a delay line, $S_{\phi,DL}$, and that of a ring oscillator using such a line, $S_{\phi,ring}$, are related as follows:

$$S_{\phi,ring}(f) = S_{\phi,DL}\left(\frac{f_0}{\pi f}\right)^2. \tag{11.7}$$

Figure 11.6 depicts this relationship. One interpretation of this result is that an edge in a ring continues to accumulate phase noise as it circulates around the loop, whereas an edge in a VCDL experiences only once the phase noise of the delay stages before it reaches the output. We thus conclude that the VCDL produces much lower phase noise than the ring does.

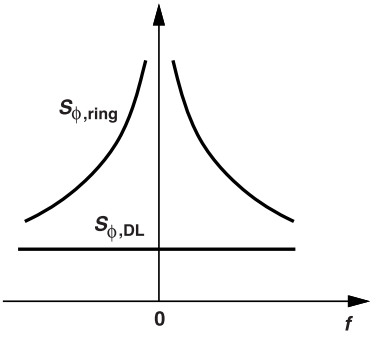


Figure 11.6 Comparison of ring oscillator phase noise and delay line phase noise.

Second, we model the VCDL phase noise as shown in Fig. 11.7(a) and write

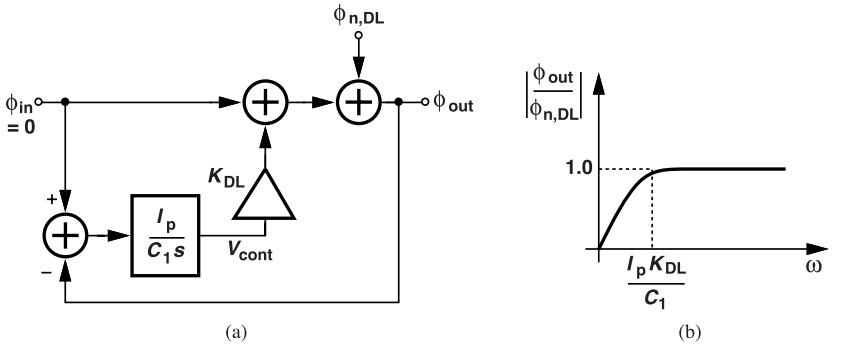


Figure 11.7 (a) DLL model including delay line phase noise, and (b) DLL response.

$$-\phi_{out}\frac{I_p}{C_1 s}K_{DL} + \phi_{n,DL} = \phi_{out}, \tag{11.8}$$

obtaining

$$\frac{\phi_{out}}{\phi_{n,DL}}(s) = \frac{C_1 s}{C_1 s + I_p K_{DL}}. \tag{11.9}$$

Similar to the effect of supply noise, this result indicates a first-order high-pass behavior [Fig. 11.7(b)]. As expected, the loop rejects slow phase fluctuations caused by the VCDL. In general, the dominant source of phase noise in VCDLs is the supply noise.

11.5 Generation of Multiple Phases

In addition to the deskewing function illustrated in Fig. 11.1(d), DLLs also find application in systems requiring multiple clock phases. For example, some clock and data recovery circuits demand 32 or 64 equally-spaced clock phases, a difficult challenge for ring oscillators as their operation frequency is inversely proportional to the number of stages that they employ.

Figure 11.8 depicts a DLL that delivers multiple clock phases. Incorporating N nominally identical delay

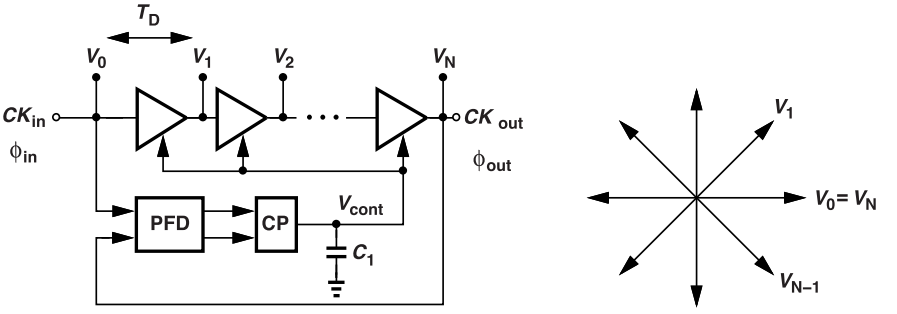


Figure 11.8 A DLL generating multiple phases.

stages, the VCDL provides N phases with a minimum spacing equal to the delay of one stage, T_D. The key point here is that $T_D = T_{CK}/N$ because the loop locks such that CK_{out} and CK_{in} have a phase difference of T_{CK}. In other words, by virtue of the feedback around the loop, T_D remains well-defined and relatively precise even with PVT variations. By comparison, a "free-running" delay line can experience a nearly twofold change in its delay as a function of PVT.

Example 11.3

Is it possible to design a 5-GHz DLL that directly generates 32 phases?

Solution

In this case, we have $T_D = 200\ \text{ps}/32 = 6.25$ ps. If the CMOS technology node can guarantee such a small gate delay in the worst PVT corner, then this DLL design is feasible. Otherwise, we must resort to phase interpolation (Section 11.8).

The multiphase DLL shown in Fig. 11.8 faces several issues. First, due to PFD/CP imperfections in Fig. 11.8, a phase offset is bound to exist and V_0 and V_N are not exactly aligned. This offset, $\phi_{out} - \phi_{in} = \Delta\phi$, leaves only $2\pi - \Delta\phi$ radians for the delay of the VCDL, thereby shifting V_1 by $\Delta\phi/N$, V_2 by $2\Delta\phi/N$, etc. The phase nonuniformity proves challenging and must be minimized through the use of high-performance CP design techniques (Chapters 7 and 8).

Second, due to unequal loading, the phase spacings at the boundaries of the VCDL can be different from those in the middle. To understand this point, consider the situation illustrated in Fig. 11.9(a), where inverters

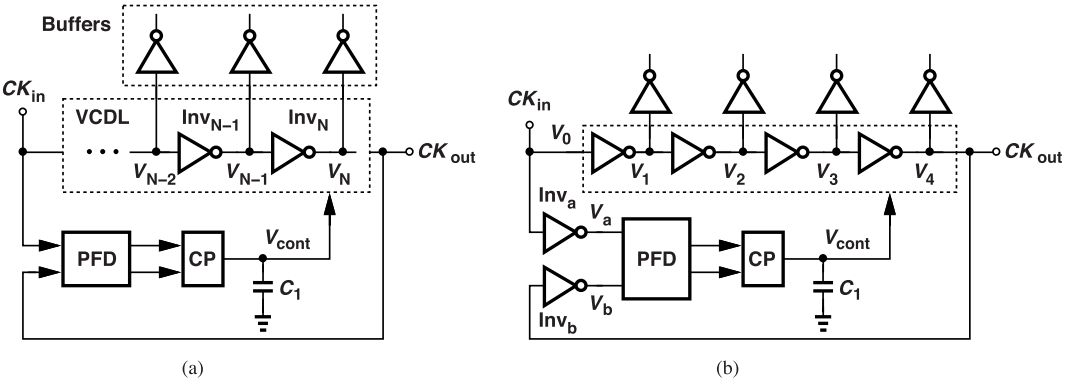


Figure 11.9 (a) DLL of Fig. 11.8 with output buffers added, and (b) correction for boundary errors.

Inv_{N-1} and Inv_N see different fanouts: the former drives two inverters but the latter, an inverter and a PFD (which can present a heavier load). As a result, the phase difference between V_{N-2} and V_{N-1} is less than that between V_{N-1} and V_N. As shown in Fig. 11.9(b) for $N = 4$, this issue is overcome by inserting two inverters at the PFD inputs. If all of the inverters are identical, we observe that (1) the fanout at V_4 is equal to that at V_1, V_2, and V_3, and (2) the loop drives the phase difference between V_a and V_b to zero, thus aligning V_0 and V_4 as well. One assumption here is that the waveform arriving at V_0 has approximately the same rise and fall times as that at V_4; otherwise, the delays through Inv_a and Inv_b are slightly different. Of course, random mismatches between these inverters contribute additional errors.

Another issue in the DLL of Fig. 11.8 is the problem of "false lock." Let us assume that the VCDL is designed to provide a total delay of T_{CK} at the TT, 27 °C corner. Now, suppose the DLL operates in the SS, high-temperature corner, and, upon startup, the total VCDL delay is slightly *greater* than $2T_{CK}$ (Fig. 11.10). Then, the DLL simply attempts to align V_0 and V_N, and it can do so if the phase difference between these

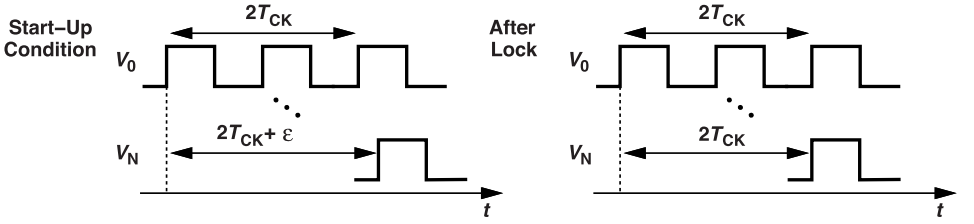


Figure 11.10 Problem of false lock.

two signals reaches $2T_{CK}$ rather than T_{CK}. As a result, the phase spacings will be equal to $2T_{CK}/N$, twice as much as the desired value.

We should remark that false lock proves problematic in nonmultiplying DLLs as well. With a total delay of $2T_{CK}$, the circuit accumulates greater phase noise and is more sensitive to the supply.

Avoiding false lock generally requires substantial added complexity, especially if the DLL must operate across a wide frequency range. Depicted in Fig. 11.11 is a solution employing a PLL. We employ an exernal control, V_{coarse}, to tune the VCDL delay so as to bring it close to T_{CK}. The DLL only controls V_{fine} and makes a small adjustment to the delay. The circuit operates as follows. A replica of the VCDL is configured as a ring oscillator and phase-locked to the main input, thus guaranteeing that the delay through the replica VCDL from A to B is equal to T_{CK} because A and B carry the same waveform. Consequently, V_{cont1} reaches the desired value. Now, this voltage serves as the coarse control for the main VCDL, allowing the DLL to provide only a fine adjustment through V_{fine}. For example, if the two VCDLs have a delay mismatch of 10%, then V_{fine} must vary the delay by only this amount, thereby avoiding false lock. The filter preceding

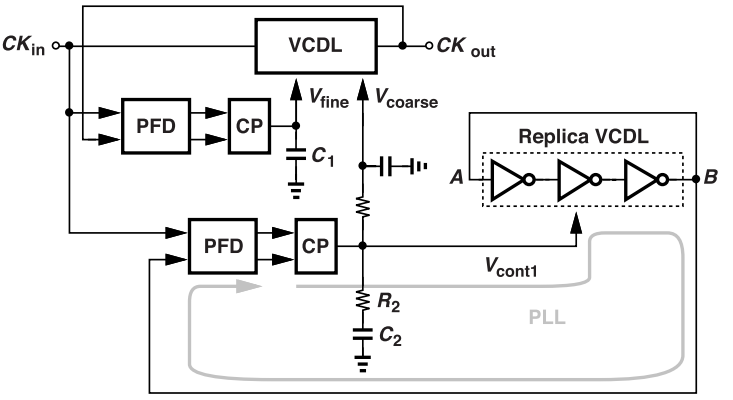


Figure 11.11 Addition of a PLL to a DLL to avoid false lock.

V_{coarse} suppresses the ripple and the noise present in V_{cont1}. This architecture approximately doubles the area and the power consumption. Another method is described in Section 11.6.

The fifth issue in the multiphase DLL of Fig. 11.8 relates to the mismatches between the delay units, which translate to departures in the phase spacings from T_{CK}/N. Present in both the driving strength of the stages and their load capacitances, the mismatches must be managed by proper sizing and careful layout so as to keep the phase spacing errors acceptably small. This issue proves particularly problematic in frequency-multiplying DLLs (Section 11.6).

Another source of error is the ripple on the control voltage: since a jump on V_{cont} in Fig. 11.8 alters the total input-output delay, a phase offset appears between the input and the output (Example 11.2), and the phase spacings are also affected. That is, all of the phase spacings are slightly less or more than the nominal value. Capacitor C_1 must be chosen large enough to minimize this effect.

11.6 Frequency-Multiplying DLLs

11.6.1 Basic Topologies

An important shortcoming of DLLs is their inability to perform frequency synthesis, i.e., generate arbitrary output frequencies. In some applications, a DLL can multiply the input frequency by an integer, thereby acting as a "poor man's" frequency synthesizer.

Recall that the DLL of Fig. 11.8 produces N equally-spaced clock edges with a resolution of T_{CK}/N seconds. If we "combine" these edges, we can generate an output having a higher frequency. Shown in Fig. 11.12(a) is an eight-phase delay line, with its outputs applied to XOR gates. Noting that the DLL aligns V_8 and V_0, we observe a phase difference of $45°$ between adjacent taps. That is, the PFD senses V_8 and V_0.

As illustrated in Fig. 11.12(b), the XOR result of V_1 and V_2 (V_a) exhibits pulses every $T_{CK}/2$ seconds, and so does the XOR result of V_3 and V_4 (V_b). Since V_a and V_b have a phase difference of $90°$, their XOR result, V_X, has a period of $T_{CK}/4$ seconds. The other output, V_Y, is the complement of V_X. From another perspective, the first rank of XOR gates doubles the frequency and the second rank doubles it again. Of course, the operation assumes a precise 50% duty cycle for V_0-V_8. The eight-stage DLL thus multiplies the input frequency by a factor of 4. Note that the XORs introduce uniform loading along the delay line.[1] The XOR stages comprise an "edge combiner" here. Figure 11.12(c) depicts a possible implementation of the XOR gate. The reader can readily show that swapping A and $\overline{A}$ creates an XNOR gate, which is necessary

[1] Only if the XOR gates themselves are symmetric with respect to their two inputs.

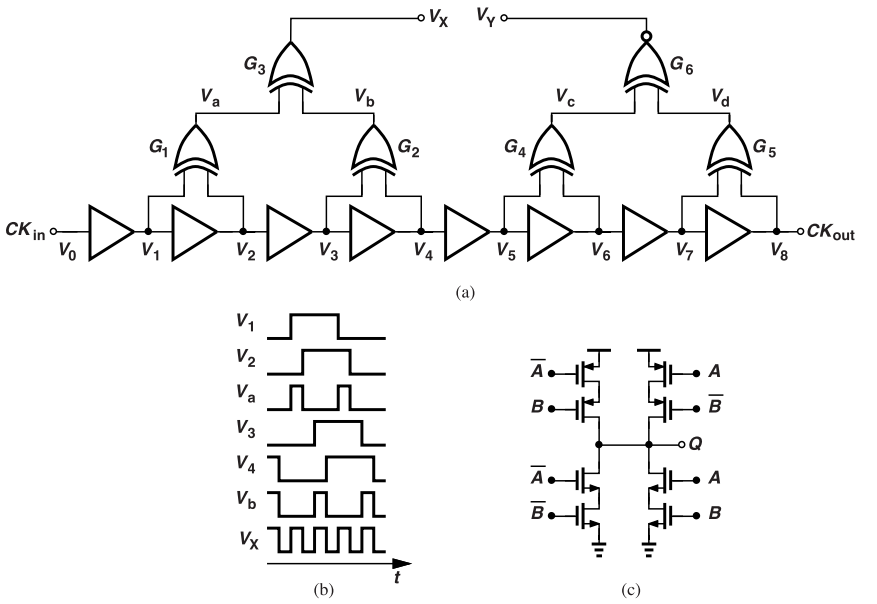


Figure 11.12 (a) Edge-combining circuit for frequency multiplication, (b) its waveforms, and (c) XOR implementation.

for V_Y.

The topology of Fig. 11.12(a) merits two remarks. First, we recognize that, in fact, $V_c = V_a$ and $V_d = V_b$ (why?). It then appears that G_4 and G_5 are redundant, and G_6 can simply be tied to V_a and V_b. However, the first rank of XORs also ensures that the delay stages have equal loading and create uniform 45° phase separations. If G_4 and G_5 are omitted, the phases associated with V_5-V_8 are distorted.

Second, suppose we keep only the first five delay stages and G_1-G_3. How should the DLL operate to ensure that V_0 and V_4 are still 180° out of phase? This is possible if the phase detector sensing V_0 and V_4 generates a zero output for a phase error of 180°. Such a PD is described in Section 11.9.

Example 11.4

Design an edge combiner using only AND and OR gates.

Solution

The delay line shown in Fig. 11.12(a) provides delayed signals and their complements because each stage delays the signal by 45°. For example, $V_5 = \overline{V_1}$, $V_6 = \overline{V_2}$, etc. If we perform an AND function on V_1 and V_2, we obtain one pulse of width $T_{CK}/8$ every T_{CK} seconds [Fig. 11.13(a)]. Similarly, $V_3 \cdot V_4$ exhibits the same shape but shifted by $T_{CK}/4$. It follows that $V_1 \cdot \overline{V_2} + V_3 \cdot \overline{V_4}$ yields two such pulses every T_{CK} seconds. Noting that $V_5 \cdot \overline{V_6} + V_7 \cdot \overline{V_8}$ has the same behavior but shifted by $T_{CK}/2$, we conclude that $V_1 \cdot \overline{V_2} + V_3 \cdot \overline{V_4} + V_5 \cdot \overline{V_6} + V_7 \cdot \overline{V_8}$ is a signal with four times the input frequency. The implementation is shown in Fig. 11.13(b). In reality, the AND and OR gates are replaced by NANDs. The XOR implementation is still preferable as it requires only short interconnects between V_1-V_8 and the first rank of XORs.

Another type of frequency multiplication employs *addition* rather than logic gates for edge combining [3]. As a simple example, suppose a differential delay line produces three edges (and their complements) with a phase spacing of 120° [Fig. 11.14(a)]. What happens if we add these waveforms? We recognize that the sum makes a transition for every edge in V_1, V_2, and V_3, providing a frequency multiplication factor of 3. The implementation is shown in Fig. 11.14(b), where three differential pairs convert the tap voltages to current, and the addition is performed in the current domain.

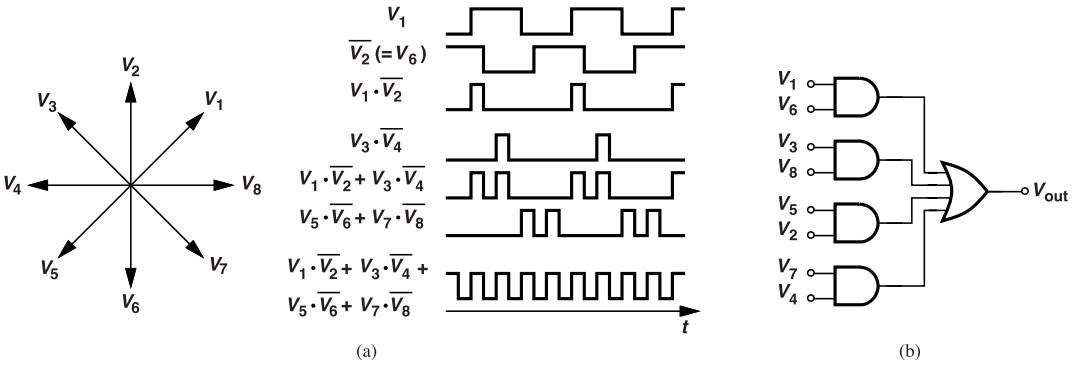


Figure 11.13 (a) Use of AND and OR operations for frequency multiplication, and (b) logical implementation.

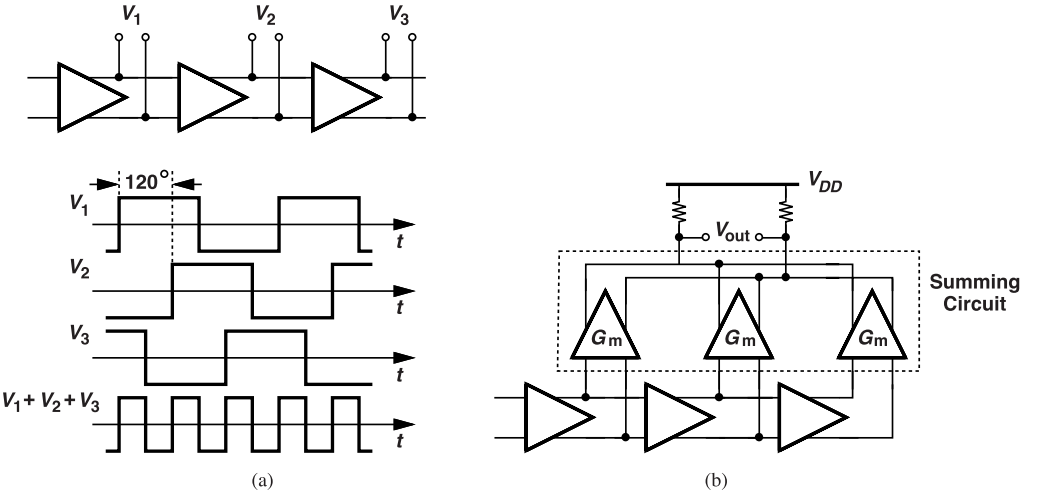


Figure 11.14 (a) Circuit generating 120° phases, and (b) summing circuit performing frequency multiplication.

Example 11.5

Using Fourier series, explain how the combining operation in Fig. 11.14(a) multiplies the frequency.

Solution

Writing

$$V_1 = V_0 \cos(\omega_0 t) + (V_0/3)\cos(3\omega_0 t) + \cdots \tag{11.10}$$

$$V_2 = V_0 \cos(\omega_0 t + 120^\circ) + (V_0/3)\cos(3\omega_0 t + 360^\circ) + \cdots \tag{11.11}$$

$$V_3 = V_0 \cos(\omega_0 t + 240^\circ) + (V_0/3)\cos(3\omega_0 t + 720^\circ) + \cdots, \tag{11.12}$$

where $\omega_0 = 2\pi/T_{CK}$, we note that the sum is free from the first harmonic (why?) and exhibits a fundamental at $3\omega_0$ [3]. This calculation shows that the currents summed in Fig. 11.14(b) must contain a strong third harmonic. For example, if the differential stages in the main path do not experience hard switching, the third harmonic amplitude is less than $V_0/3$.

The topology depicted in Fig. 11.14(b) can be extended to N stages so as to multiply the input frequency by N. This approach limits N to odd integers because differential signals, such as V_1, V_2, and V_3 in the above example, nominally contain only odd harmonics.

11.6.2 Design Issues

The multiplication factor in Figs. 11.12 and 11.14 is difficult to change, a point of contrast to PLLs. Moreover, delay mismatches among the stages give rise to jitter in the time domain and spurs in the frequency domain. For example, if the phase shifts in Eqs. (11.10)-(11.12) depart from integer multiples of 120°, then the first harmonic terms do not completely cancel, acting as a spurious component for the third harmonic [Fig. 11.15(a)]. Suppose that the stage generating V_2 incurs a delay error of ΔT. Illustrated in Fig. 11.15(b),

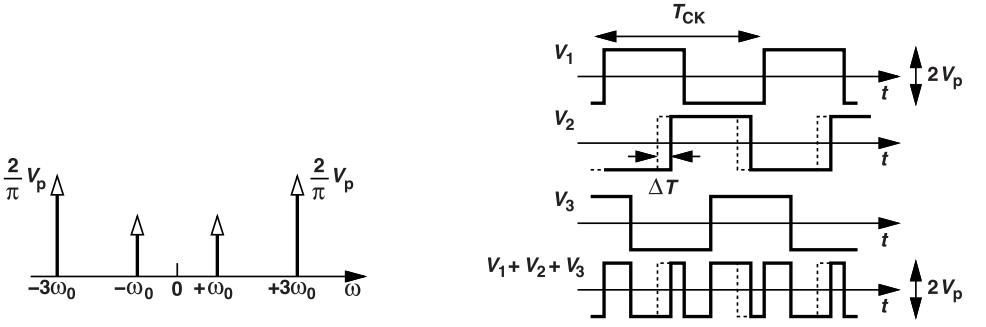


Figure 11.15 (a) Unwanted component at ω_0 due to mismatches, and (b) time-domain waveforms illustrating the effect.

the time-domain waveforms also reveal that this mismatch manifests itself on the rising and falling edges of $V_1 + V_2 + V_3$ every T_{CK} seconds, producing a jitter equal to ΔT.

Example 11.6

Determine the relative level of the spurs in Fig. 11.15(a) in terms of ΔT in Fig. 11.15(b). Assume a peak swing of V_p for $V_1 + V_2 + V_3$.

Solution

The sum waveform in Fig. 11.15(b) can be viewed as an ideal square wave at $3f_{CK}$ plus two other waveforms consisting of positive and negative pulses of width ΔT that occur every T_{CK} seconds (Fig. 11.16). Approximating these pulses with impulses having an area of $2V_p \cdot \Delta T$, we can express them as

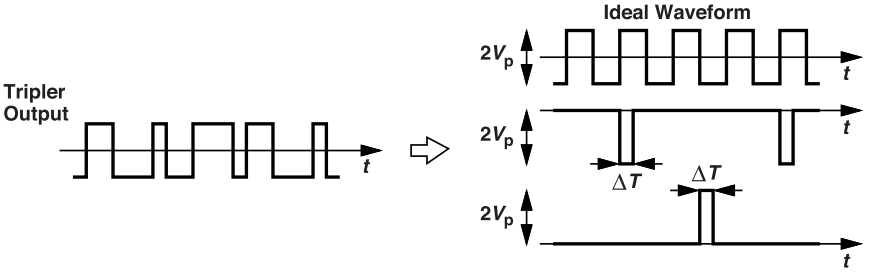


Figure 11.16 Decomposition of frequency-multiplied output to an ideal waveform and mismatch-induced components.

$$V_{mis}(t) = 2\Delta T \cdot V_p \sum_{k=-\infty}^{+\infty} \left[\delta\left(t - kT_{CK} - \frac{T_{CK}}{2}\right) - \delta(t - kT_{CK})\right]. \tag{11.13}$$

Taking the Fourier transform gives

$$V_{mis}(f) = \frac{2\Delta T \cdot V_p}{T_{CK}} \sum_{k=-\infty}^{+\infty} \left[(e^{-j\pi f T_{CK}} - 1)\delta(f - kf_{CK})\right]. \tag{11.14}$$

For the component at $\omega_0 = 2\pi f_{CK}$, we set k to 1 and note that $\exp(-j\pi f T_{CK})\delta(f - f_{CK}) = \exp(-j\pi f_{CK} T_{CK})\delta(f - f_{CK}) = -\delta(f - f_{CK})$. That is, the magnitude of this component is equal to $(2\Delta T \cdot V_p/T_{CK}) \times 2$. Normalized to the first harmonic of $V_1 + V_2 + V_3$, this value translates to $2\pi\Delta T/T_{CK}$. For example, a ΔT of 1 ps at 5 GHz produces a spur level of -30 dBc.

The foregoing example underscores the stringent matching required of the delay stages if a low spur level is necessary. For example, to obtain a level of -60 dBc with $f_{CK} = 5$ GHz, we must have $\Delta T = 32$ fs, an extremely low value. We then ask whether it is possible to suppress the mismatches by means of calibration. In principle, we can sense and compare the delays and adjust their values so as to minimize their differences. We note, however, that any circuit sensing the delays (e.g., a PD) suffers from its own mismatches, which are possibly greater than 32 fs, thus limiting the accuracy of the calibration. Of course, adjusting the delays in 32-fs steps presents another challenge as well.

As mentioned in Section 11.1, DLLs do not correct the input duty cycle error. In Problem 11.10, we investigate whether this error introduces jitter and spurs in frequency-multiplying DLLs.

11.6.3 Use of Frequency Multiplication in False Lock Detection

The frequency multiplication ability of DLLs can be exploited to detect false locking. Consider the architecture shown in Fig. 11.17, where an edge combiner multiplies the frequency by a factor of N. This result,

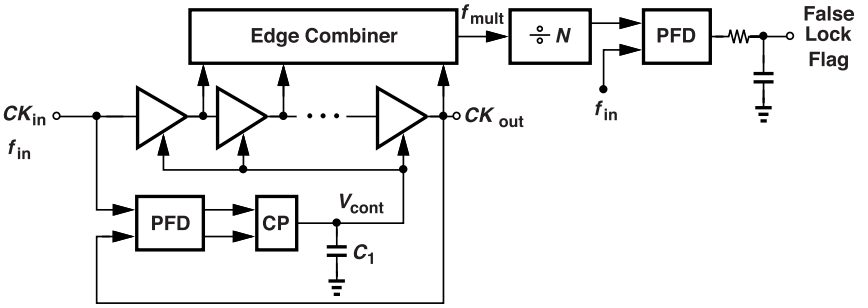


Figure 11.17 False lock detector using a frequency multiplier.

f_{mult}, is then divided by N and compared to f_{in}. With correct locking, $f_{mult} = Nf_{in}$, leading to a low average value for the PFD output. In the presence of false lock, on the other hand, the total delay from CK_{in} to CK_{out} is equal to $2T_{CK}$, and $f_{mult} = Nf_{in}/2$. Consequently, the CP output ramps up to V_{DD}. The false lock flag can then be used to *reduce* the tuning range of the delay line so that the total delay remains less than $2T_{CK}$. In Problem 11.12, we explore whether we can save power by turning off the edge combiner and the stages following it after the DLL is properly locked.

11.7 DLL/PLL Hybrids

We have thus far observed that, compared to ring oscillators, delay lines inherently produce less jitter but cannot easily generate new frequencies. We then wonder whether the two, and hence their advantages, can be combined.

Suppose a ring oscillator operates at a desired frequency f_0, accumulating jitter as the edges travel around the loop [Fig. 11.18(a)]. Of course, a PLL can counteract this effect, but only over a large number of cycles.

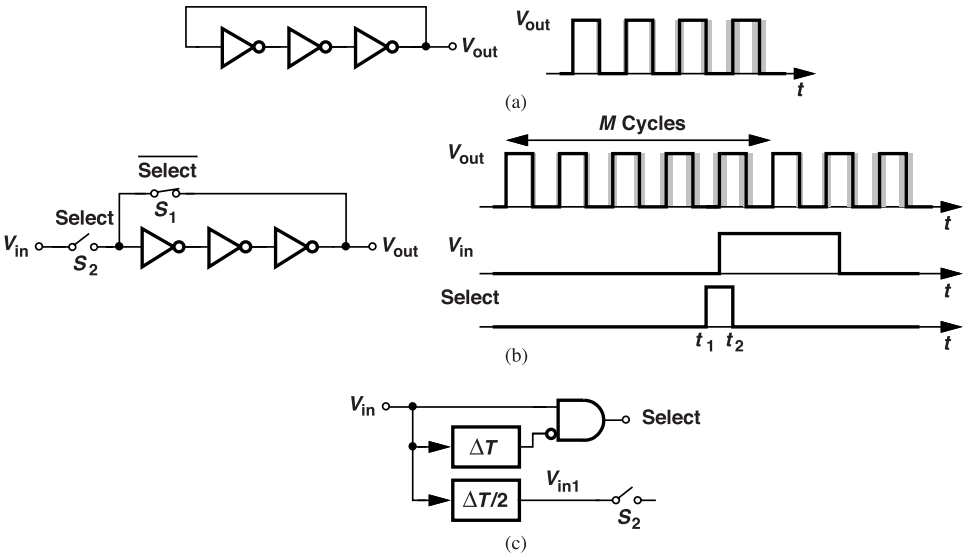


Figure 11.18 (a) Jitter accumulation in a ring oscillator, (b) resetting the phase by an input edge, and (c) logic for generating the select command.

(Recall that the maximum PLL bandwidth is around one-tenth of the reference frequency.) We surmise that the ring's jitter accumulation can be *reset* if one of the edges in Fig. 11.18(a) is periodically replaced by a "clean" edge. That is, if we wish to multiply the input frequency by a factor of M, we design the ring oscillator for operation at Mf_{in}, but replace one of its edges with an input edge every M output cycles [5].[2] Illustrated in Fig. 11.18(b) is an example where the inverters are configured as a ring oscillator until t_1 (when S_1 is on) and as a delay line from t_1 to t_2 (when S_2 is on). The output jitter can thus accumulate for only M cycles [5]. By comparison, a simple PLL having a bandwidth equal to one-tenth of f_{in} would allow greater accumulation.

How is the Select command in Fig. 11.18(b) generated? This signal must rise slightly before the rising edge of V_{in} and fall slightly after. Consider the circuit shown in Fig. 11.18(c), where V_{in} is delayed by ΔT, inverted, and ANDed with itself. The output exhibits pulses ΔT seconds wide on the rising edges of V_{in}. However, Select does not go high *before* V_{in} does. To resolve this issue, we delay V_{in} by $\Delta T/2$ to obtain V_{in1} and utilize this signal as the input in Fig. 11.18(b). We choose ΔT less than one half of the oscillator period.

Let us now construct a complete loop using this arrangement. As depicted in Fig. 11.19, we follow the oscillator with a $\div M$ circuit and apply the result to the PFD. The circuit locks so as to minimize the phase difference between V_{in1} and V_{div}. Here, the chain of inverters acts as a VCO most of the time—and the overall system as a PLL. The operation is briefly interrupted once every input cycle when Select goes high. Note that the PLL requires R_1, C_1, and C_2 for proper operation while a simple DLL does not.

The architecture in Fig. 11.19 entails a critical issue, namely, the misalignment between the oscillator edges and the input edges. Suppose the $\div M$ circuit incurs a delay of T_{div}. If the phase difference between V_{in1} and V_{div} is forced to zero by the PLL, the transitions in V_{out} are *not* aligned with those in V_{in1} [Fig. 11.20(a)]. Consequently, as shown in Fig. 11.20(b), the circuit removes the rising edge of V_{out} at t_a and replaces it with that of V_{in1} at t_b. The output periodically exhibits an edge displacement equal to T_{div}, which translates to jitter.

[2] The clean input edge appears in V_{out} after one cycle.

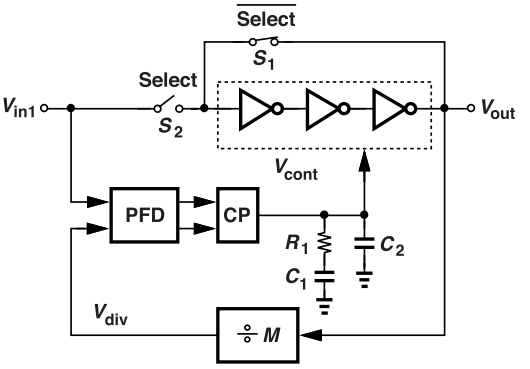


Figure 11.19 Hybrid DLL/PLL circuit.

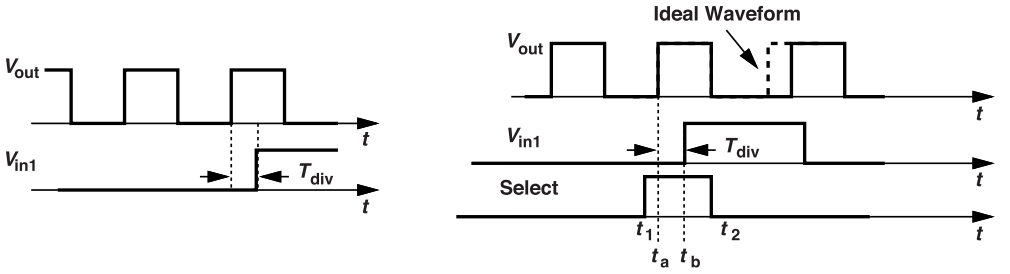


Figure 11.20 (a) Edge misalignment due to divider delay, and (b) effect on hybrid DLL/PLL.

Example 11.7

Explain how the hybrid loop of Fig. 11.19 responds after $t = t_2$ in Fig. 11.20.

Solution

The edge displacement experienced by V_{out} between t_a and t_b propagates through the divider, manifesting itself as a phase error at the PFD input. Consequently, the PFD, the CP, and the loop filter create a change in V_{cont} so as to correct this error. As expected, the phase-locked loop attempts to counteract this disturbance.

It is possible to eliminate the effect of the divider delay: as illustrated in Fig. 11.21, we utilize V_{out} to "retime" V_{div} by means of a flipflop. Now, V_{FF} makes transitions only on the rising edges of V_{out}. In this case, the phase error between V_{FF} and V_{out} is equal to the FF delay, ΔT_{FF}, which is typically much less than the divider delay. Nonetheless, the PFD and CP imperfections still produce a phase offset at the PFD input, inevitably creating the phase discontinuity shown in Fig. 11.20.

11.8 Phase Interpolation

The DLLs studied in the previous sections do not provide a phase spacing less than one gate delay, a limitation reminiscent of the simple TDC presented in Chapter 10. For higher resolutions, we can incorporate phase interpolation, as we did in the same chapter.

Figure 11.22(a) illustrates, as an example, interpolation by a factor of 2, where the threshold crossing at t_{12} occurs halfway between t_1 and t_2. Ideally, we have $V_{12} = (V_1 + V_2)/2$. How do we realize this function with rail-to-rail swings? As shown in Fig. 11.22(b), two identical inverters can perform interpolation by virtue of their finite output impedances. We note that at $t = t_{12}$, the NFET in Inv_1 and the PFET in Inv_2 are heavily

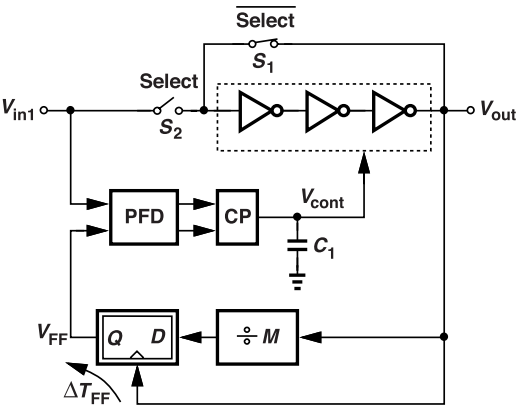
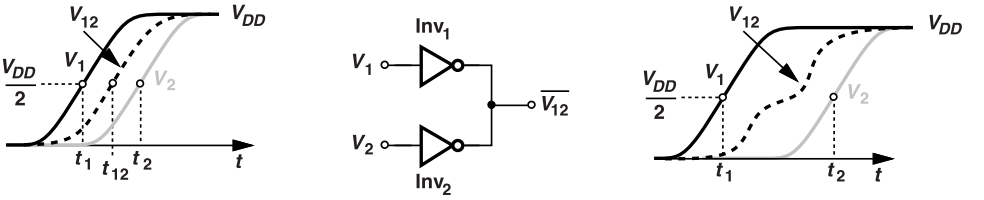


Figure 11.21 Correction for divider delay by means of a retimer.



(a) (b) (c)

Figure 11.22 (a) Phase interpolation between two waveforms, (b) interpolator implementation, and (c) appearance of a kink if the transitions are excessively fast.

on, fighting each other. If these two devices have equal strengths, then $V_{12}(t_{12}) \approx (V_1 + V_2)/2$. The phase resolution is therefore doubled. The output is denoted by $\overline{V_{12}}$ to include the inversion.

It is important to note that phase interpolation fundamentally requires that the transitions be *slow*! As shown in Fig. 11.22(c), if the original edge spacing, $t_2 - t_1$, is greater than the transition times of V_1 and V_2, then the interpolated waveform exhibits a "kink," suffering from a low slope and becoming prone to higher jitter.

Example 11.8

Delay lines that are followed by interpolation experience a higher fanout and hence a greater unit delay than those without interpolation. Explain how the circuit can achieve a greater phase resolution despite the heavier load.

Solution

As shown in Fig. 11.23, each stage drives three inverters. To minimize the unit delay along the main path,

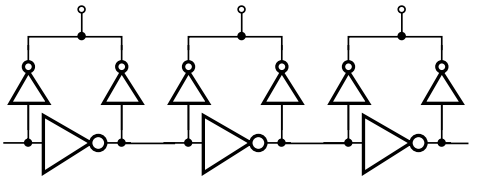


Figure 11.23 Use of large inverters in a delay line to lessen the effect of fanout.

the stages in this path should be chosen much stronger than the interpolating inverters. The unit delay thus approaches the basic gate delay with a fanout of unity. Of course, this remedy raises the power consumption.

The interpolation network in Fig. 11.22(b) provides an interpolated edge, but it also incurs a delay with respect to the original edges, V_1 and V_2. For this skew to be removed, V_1 and V_2 must experience the same delay. Shown in Fig. 11.24 is an arrangement utilizing inverters in parallel to emulate the drive strength of

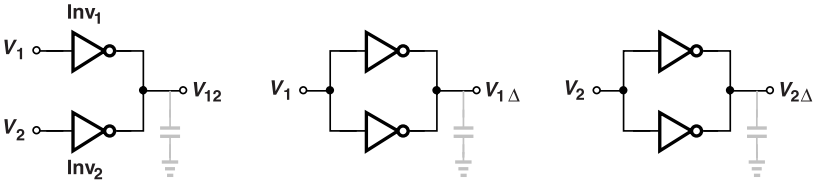


Figure 11.24 Phase interpolation along with delayed signals for proper alignment.

the interpolating circuit consisting of Inv_1 and Inv_2. Here, the edges of V_{12} nominally lie halfway between those of $V_{1\Delta}$ and $V_{2\Delta}$. This approach merits two remarks. First, the delay in the path of V_1 and V_2 is, in fact, less than that in the interpolating stage: during transitions, the NFET and PFET in the latter fight for a longer time, providing less net current for charging the load capacitance. Consequently, the V_{12} edges do not exactly occur halfway between the $V_{1\Delta}$ and $V_{2\Delta}$ edges. Second, the fanout seen by V_1 or V_2 is now higher, leading to a greater unit delay along the main path (Example 11.8).

The interpolation concept can be extended to higher factors. Figure 11.25 depicts how another factor of

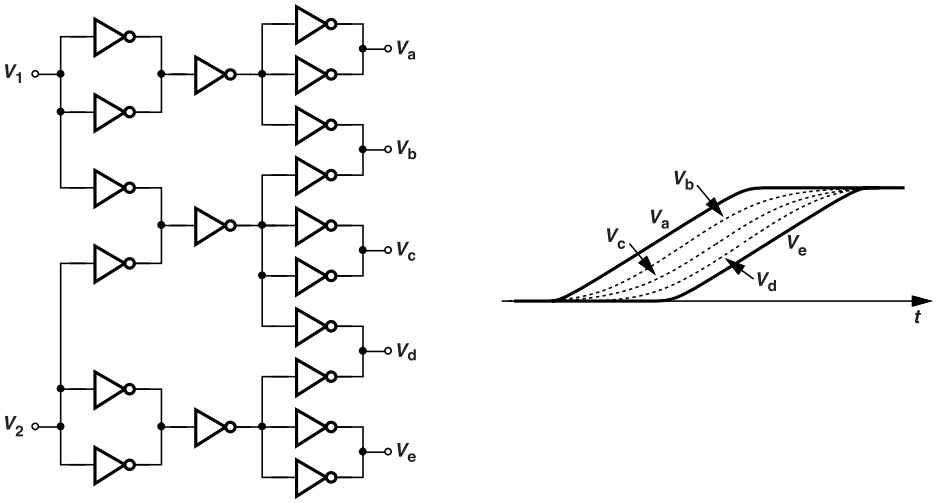


Figure 11.25 Phase interpolation by a factor of 4.

2 is realized [6]. Repeating the structure shown in Fig. 11.24, this circuit provides between V_a and V_e the interpolated waveforms V_b, V_c, and V_d. As can be seen, such a network potentially consumes a high power. For example, if the DLL provides eight original phases and we wish to interpolate by a factor of 4 between them, the circuit contains a very large number of inverters running at high speeds.

As an alternative approach, let us consider differential pairs for the interpolation network. Depicted in Fig. 11.26(a) is an arrangement for a factor of 2, with identical pairs $M_{1,2}$ and $M_{3,4}$ converting V_1 and V_2, respectively, to current, performing the summation at their drain nodes, and injecting the results into the load

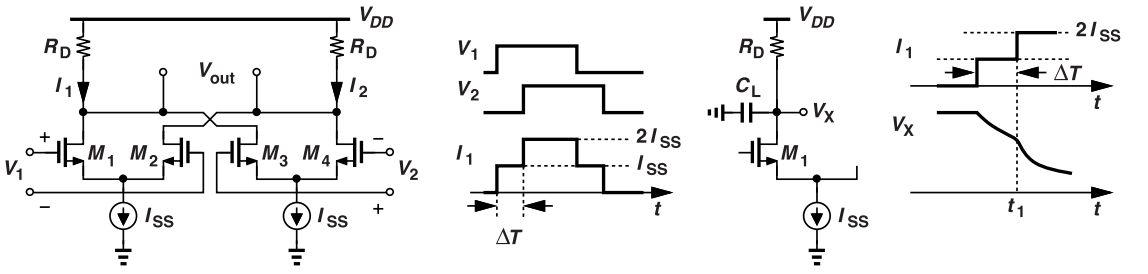


Figure 11.26 (a) Phase interpolation by means of differential pairs, (b) circuit's waveforms, (c) addition of capacitance at the output node, and (d) resulting waveforms.

resistors. We assume V_1 and V_2 have a phase spacing of ΔT. From a small-signal perspective, we can write $V_{out} \propto (G_m V_1)R_D + (G_m V_2)R_D = G_m R_D(V_1 + V_2)$, where G_m denotes the transconductance of each pair. Since V_{out} is proportional to $(V_1 + V_2)/2$, this result implies proper interpolation.

But what happens if V_1 and V_2 have sharp edges? As shown in Fig. 11.26(b), then $I_1 = I_{D1} + I_{D3}$ exhibits a kink as it goes from 0 to I_{SS} and then to $2I_{SS}$ (and so does I_2). In order to avoid the kink, we have two options. (1) Slow down the edges of V_1 and V_2; but even then the nonlinearity of the differential pairs tends to sharpen the edges. In other words, this approach requires both slow input edges (which are more sensitive to noise) and moderately linear differential pairs. (2) Apply low-pass filtering at the output nodes so that the kink in the currents does not appear in the output voltages [Fig. 11.26(c)]. In reality, the time constant, $\tau = R_D C_L$, must be several times ΔT for the kink to be negligible, leading to slow transitions at the output. To appreciate this point, suppose $\tau \approx 2\Delta T$. Then, as shown in Fig. 11.26(d) and formulated in Problem 11.19, the slope of V_X changes, at t_1, by about 40%, still causing a kink.

11.9 High-Speed PD Design

Another point of contrast between PLLs and DLLs is that the latter require that the PD (or the PFD) and the CP operate at high speeds. In this section, we deal with this issue.

The simple PFD implementation studied in Chapter 7 begins to fail at high frequencies. To improve the speed, we can seek a faster D flipflop implementation. A good candidate is the true single-phase clock (TSPC) logic family [8], which is described in Chapter 15. Shown in Fig. 11.27 is a TSPC PFD design [4]. Considering the top path, suppose both A and Reset are low. Thus, X is high, M_3 is off, and Up maintains its present state on the output node capacitance. Now, if A goes high, M_1 turns on, bringing Up to zero. The bottom path behaves similarly, causing Down to fall and the reset path to be activated. The minimum width of the up and down pulses, T_{RST}, is thus equal to three gate delays (why?). In this case, Up and Down are at zero when they are "active." Thus, Up must drive the PMOS branch in the charge pump and Down must be inverted to drive the NMOS branch.

In some DLL designs, we wish to have a phase difference of 180° between the input and the output. An RS latch can serve this purpose, as shown in Fig. 11.28. Suppose the inputs, A and B, have a phase difference close to 180°. Before $t = t_1$, $A = 0$, $B = 1$, and hence $Q_A = 1$ and $Q_B = 0$, keeping the CP off. When A rises, the outputs remain unchanged and when B falls, both Q_A and Q_B are activated, still yielding $I_{CP} = 0$ if the up and down currents match. At $t = t_3$, A falls and Q_A rises, turning off the up current, and at $t = t_4$, B rises and Q_B falls. The charge drawn by the CP from t_3 to t_4 represents the phase error. The DLL thus locks so as to align the rising edges of A with the falling edges of B and vice versa.

The foregoing topology faces difficulties if the duty cycles of A and B are not 50%. To ensure phase comparison between only the rising edges of A and the falling edges of B (or the other way around), one can precede the circuit with edge detectors [7].

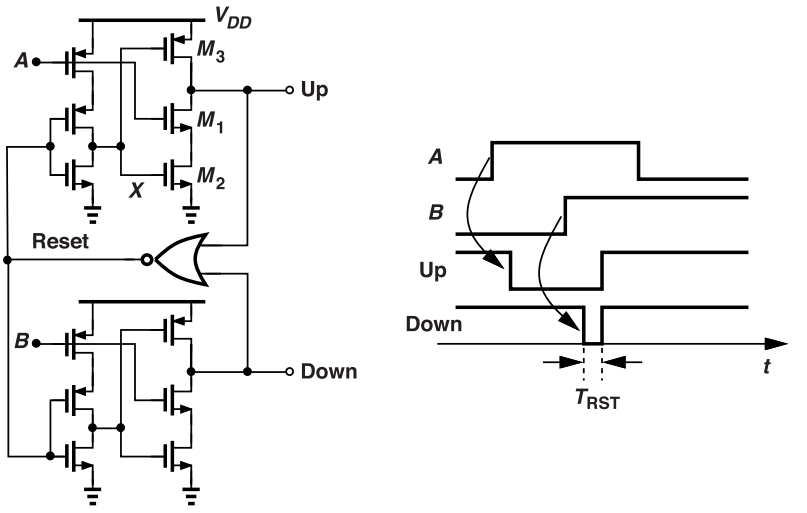


Figure 11.27 PFD realization using TSPC logic.

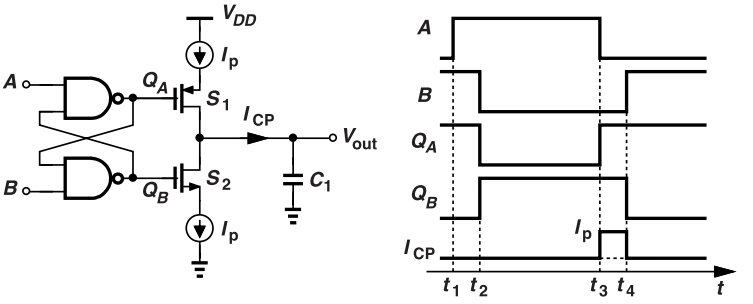


Figure 11.28 RS latch serving as a phase detector.

Example 11.9

Compare the effect of CP mismatches in Fig. 11.28 to that in a PFD/CP combination.

Solution

In Fig. 11.28, the CP remains on for *half* of the input period, allowing the mismatch between the up and down currents to flow through the capacitor and cause substantial ripple. In a PFD/CP cascade, on the other hand, the CP is on for only a short amount of time (e.g., for three gate delays in the circuit of Fig. 11.27). Another disadvantage is the much higher power consumption of the CP here (why?).

11.10 Duty Cycle Correction

In contrast to PLLs, DLLs generate an output whose duty cycle depends on that of the input. Since the input duty cycle can substantially deviate from 50% in some systems, the timing margins for the circuits driven by the DLL degrade, necessitating duty cycle correction (DCC).

Consider the input waveform shown in Fig. 11.29(a), where the duty cycle is given by $(T_{CK}/2 + \Delta T)/T_{CK}$. DCC requires (1) measuring ΔT, or the difference between the high and low times, and (2) driving this error toward zero. The former task can be performed by comparing the areas under the high and

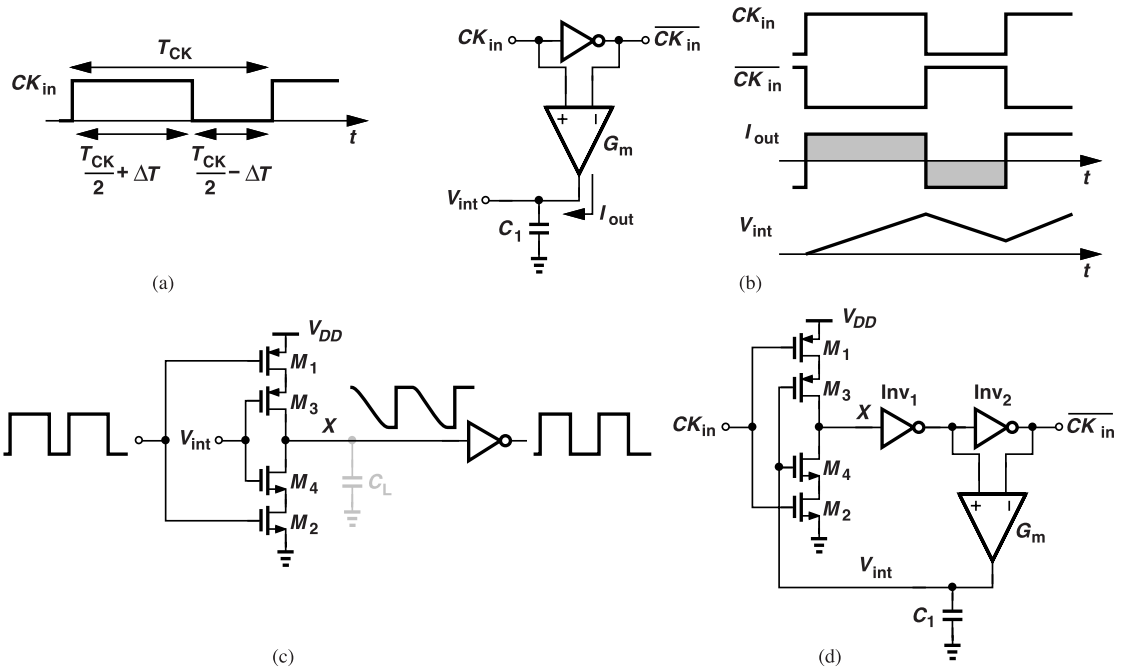


Figure 11.29 (a) Duty cycle error, (b) circuit for measuring this error, (c) adjustment of duty cycle, and (d) overall duty cycle correction circuit.

low levels, as exemplified by the circuit shown in Fig. 11.29(b) [9]. Here, when CK_{in} is high, the G_m stage charges C_1 and vice versa. Thus, if $\Delta T \neq 0$, the average value of V_{int} continues to rise (or fall), serving as a measure of the duty cycle error.

The result obtained above must be applied to a circuit that adjusts the duty cycle. Shown in Fig. 11.29(c) is an example based on an inverter [9]. Transistors M_3 and M_4 can create an *arbitrary* difference between the rise and fall times at X. If V_{cont} is relatively high, M_3 slows down the rising edge. Similarly, M_4 controls the falling edge. The second inverter sharpens the transitions and produces a duty cycle that depends on the rise and fall times at X. We now attach the circuit in Fig. 11.29(b) to that in Fig. 11.29(c), arriving at the DCC topology shown in Fig. 11.29(d) [9].

The principal issue in the foregoing DCC circuit is that it slows down the edge at X, making the waveform more susceptible to supply noise. Viewing M_1 and M_2 as an inverter, we observe that, if V_{int} in Fig. 11.29(d) is relatively constant, then noise on V_{DD} modulates the delay of this stage and generates jitter.

The G_m stage in Fig. 11.29(d) must provide an output voltage range from V_{THN} to $V_{DD} - |V_{THP}|$ (why?). Also, the output noise of this stage modulates the delay of the main path, especially at low frequencies. Capacitor C_L must be chosen large enough to suppress most of this noise. The reader is encouraged to examine the circuit's operation if the delay of Inv_2 is not negligible.

References

[1]. M.-J. E. Lee et al., "Jitter transfer characteristics of delay-locked loops - theories and design techniques," *IEEE J. Solid-State Circuits*, vol. 38, pp. 614-621, April 2003.

[2]. A. Homayoun and B. Razavi, "Relation between delay line phase noise and ring oscillator phase noise," *IEEE J. Solid-State Circuits*, vol. 49, pp. 384-391, Feb. 2014.

[3]. B. Razavi, and J. Sung, "A 6-GHz 60-mW BiCMOS phase-locked loop," *IEEE Journal of Solid-State Circuits*, vol. 29, pp. 1560-1565, Dec. 1994.

[4]. W.-H. Lee, J.-D. Cho, and S.-D. Lee, "A high speed and low power phase-frequency detector and charge-pump," *Proc. Asia and South Pacific Design Automation Conf.*, pp. 269-272, Jan. 1999.

[5]. R. Farjad-Rad et al., "A low-power multiplying DLL for low-jitter multigigahertz clock generation in highly integrated digital chips," *IEEE J. Solid-State Circuits*, vol. 37, pp. 1804-1813, Dec. 2002.

[6]. B. W. Garlepp et al., "A portable digital DLL for high-speed CMOS interface circuits," *IEEE J. Solid-State Circuits*, vol. 34, pp. 632-644, May 1999.

[7]. S. Sidiropoulos and M. A. Horowitz, "A semidigital dual delay-locked loop," *IEEE J. Solid-State Circuits*, vol. 32, pp. 1683-1692, Nov. 1997.

[8]. J. Yuan and C. Svensson, "High-speed CMOS circuit technique," *IEEE J. of Solid-State Circuits,* vol. 24, pp. 62-70, Feb. 1989.

[9]. E. Song et al., "A reset-free anti-harmonic delay-locked loop using a cycle period detector," *IEEE J. of Solid-State Circuits,* vol. 39, pp. 2055-2061, Nov. 2004.

Problems

11.1. Determine the static phase offset in the DLL of Fig. 11.1(c). Assume a gain of K_{PD} for the PD and a gain of $K_{DL} = \partial\phi_{out}/\partial V_{cont}$ for B_2.

11.2. Consider the thermal noise of R_1 in Fig. 11.1(c). What type of response (low-pass, band-pass, or high-pass) does this noise experience as it translates to phase noise at the output?

11.3. Repeat the phase step analysis of Fig. 11.2(c) for the DLL of Fig. 11.1(c). Does the static phase offset change after this transient?

11.4. If B_1 and B_2 in Fig. 11.1(c) are realized as inverters, explain which one of the following changes the static phase offset: the supply voltage, the process, the temperature, or the load capacitance driven by B_2.

11.5. Repeat the previous problem for the DLL in Fig. 11.2(a).

11.6. Sketch the output phase of the DLL in Fig. 11.2(a) as a function of time if a small step is applied to the supply voltage of the VCDL. Verify that your result agrees with the analysis depicted in Fig. 11.5.

11.7. What happens in the DLL of Fig. 11.8 if the up and down inputs of the CP are swapped?

11.8. What happens in the DLL of Fig. 11.9(b) if Inv_b is omitted?

11.9. The replica VCDL in Fig. 11.11 is a copy of that used in the DLL and has two controls, V_{fine} and V_{coarse}. How are these controls connected in the PLL? For example, do we short V_{fine} to V_{coarse} or do we simply remove V_{fine} from the replica VCDL? How do these choices affect the systematic mismatch between the two VCDLs?

11.10. Draw the waveforms in Fig. 11.12(b) carefully with a small, but equal, duty cycle error in V_1-V_4. Explain how V_a, V_b, and V_X are affected.

11.11. Suppose V_2 and V_1 in Fig. 11.12(b) suffer from a small phase mismatch of ΔT. Determine the normalized level of the spurs in V_X. (The analysis in Example 11.6 can be repeated here.)

11.12. In Fig. 11.17, we turn off the edge combiner and the stages following it after the DLL is properly locked. (To do so, the flag generated by the PFD must be stored as a logical level to ensure that the VCDL setting does not vanish when the PFD is disabled.) Now, suppose the VCDL delay varies with the temperature. Is it possible for the DLL to enter false lock?

11.13. Suppose the stages in the VCDL in Fig. 11.17 suffer from mismatches. Does the output of the $\div N$ stage contain spurs? Explain why or why not.

11.14. If the CP up and down currents in Fig. 11.19 have a mismatch, does V_{out} suffer form phase discontinuity?

11.15. Suppose the commands $Select$ and $Select$ in Fig. 11.19 are slightly misaligned. Explain what happens if $Select$ rises (a) before or (b) after $\overline{Select}$ does.

11.16. Suppose the top two AND gates in Fig. 11.13(b) see slightly different load capacitances. Explain what happens to V_{out}.

11.17. A student mistakenly replaces the bottom AND gate in Fig. 11.13(b) with a NAND gate. Draw the output waveform for this case.

11.18. If we add two more stages to the circuit of Fig. 11.14(b) while keeping the load resistors unchanged, how does the output amplitude (at the fifth harmonic) compare to that of the third harmonic in Example 11.5?

11.19. Assuming $\tau = R_D C_4 = 2\Delta T$ in Fig. 11.26(c), determine the change in the slope of V_X at $t = t_1$. Model the differential pair by an ideal current source that steps from zero to $I_{SS}/2$ and, ΔT seconds later, from $I_{SS}/2$ to I_{SS}.

11.20. Suppose V_1 and V_2 in Fig. 11.26(a) are the quadrature phases of a 2-GHz clock. How should the time constant $R_D C_L$ in Fig. 11.26(c) be chosen to ensure the slope of V_X changes by less than 20% at $t = t_1$? Estimate the rise and fall times of V_X.

12

RF Synthesis

An important application of phase-locking is in RF synthesis, namely, the generation of a periodic waveform that drives RF transmitters and receivers. RF synthesizers are broadly classified as "integer-N" and "fractional-N" loops. In this chapter, we first study the requirements that RF synthesizers must meet. Next, we deal with integer-N loops, recognizing that they build upon the PLL concepts described in the previous chapters. We then introduce fractional-N synthesis and the concept of noise shaping. Finally, we address various imperfections in fractional-N designs. The reader is assumed to have some basic knowledge of RF design and is encouraged to review Chapters 7-9.

12.1 RF Synthesis Requirements

Figure 12.1(a) shows a generic RF transceiver. In the transmit (TX) path, the baseband signal to be trans-

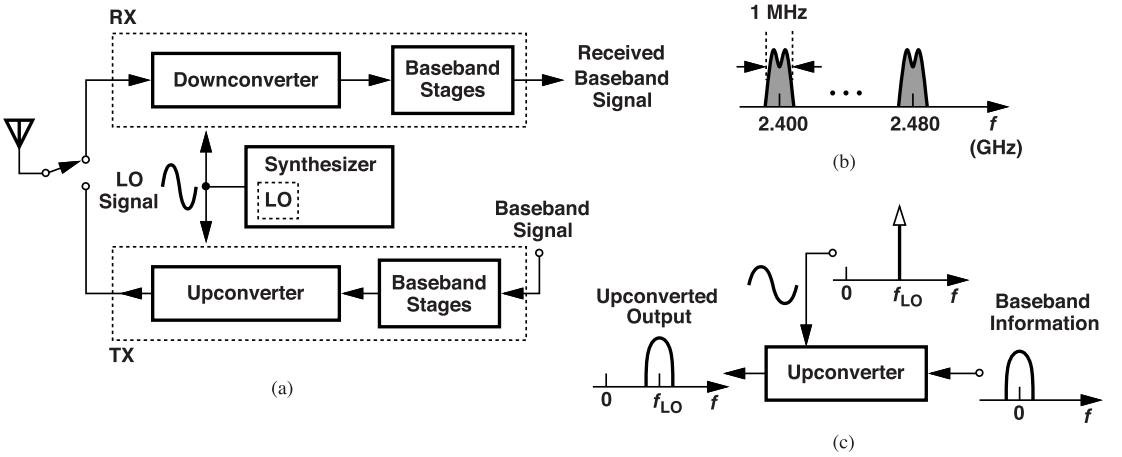


Figure 12.1 (a) Generic RF transceiver showing the synthesizer's role, (b) RF channels in Bluetooth, and (c) use of LO in upconversion.

mitted, e.g., voice or data, is applied to an "upconverter," which is also driven by the periodic output of the synthesizer. The upconverter impresses the baseband information on the synthesizer output (in an operation called "modulation"), delivering to the antenna a signal whose spectrum is centered around a carrier frequency, f_C. The receiver senses the signal picked up by the antenna and, with the aid of the synthesizer output, downconverts it to the baseband. The oscillator within the synthesizer is called the "local oscillator" (LO) and the waveform(s) driving the upconverter and the downconverter, the "LO signal(s)."

A simple example illustrates the functionality required of the synthesizer. A Bluetooth transceiver operates from 2.400 GHz to 2.480 GHz, providing approximately 80 1-MHz channels for communication [Fig. 12.1(b)]. In the transmit mode, the system shown in Fig. 12.1(a) sets the LO frequency to one of these values so that the transmitted signal occupies the corresponding channel [Fig. 12.1(c)]. Similarly, in the receive mode, the LO frequency is chosen equal to the center of the desired received channel. In this example, the synthesizer must provide an output frequency that can vary from 2.400 GHz to 2.480 GHz in 1-MHz steps. We say the "channel spacing" is 1 MHz.

How do we construct a PLL that can operate in this system if the reference frequency is *constant*? With a $\div M$ circuit in the feedback path, we have $f_{out} = Mf_{REF}$. Thus, for f_{out} to have 1-MHz steps, one solution is to select $f_{REF} = 1$ MHz and design the divider so that M can change in steps of 1. For example, $M = 2400$ for the first Bluetooth channel and $M = 2480$ for the last. We construct the abstract model shown in Fig. 12.2, where the digital control sets the value of M. The reference frequency is derived from a low-noise crystal oscillator.

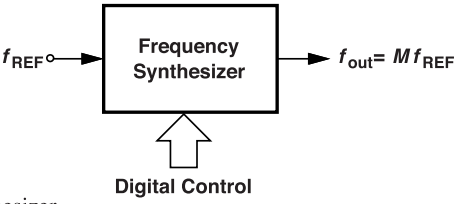


Figure 12.2 Abstract view of a synthesizer.

The foregoing example reveals two important aspects of RF synthesizers: they typically employ (1) a large divide ratio, making it more difficult to achieve proper stability, i.e., to obtain $\zeta \approx (R_1/2)\sqrt{I_p C_1 K_{VCO}/(2\pi M)} \approx 1$, and (2) a low reference frequency, thereby suffering from a narrow loop bandwidth. (Recall that the PLL bandwidth is about one-tenth of the reference frequency.)

RF synthesizers must generally satisfy stringent performance requirements, specifically, in terms of phase noise, spur levels, lock time, supply rejection, and power consumption. As studied in Chapter 2, phase noise corrupts both the transmitted and received signals, and spurs can translate unwanted frequency components onto the desired channel during downconversion. Let us briefly consider the other issues.

Example 12.1

Upon leaving home, a cell phone user experiences a temperature change of $40\ ^\circ$C. Explain what happens if the VCO within the synthesizer no longer operates at the desired frequency.

Solution

The VCO must now switch to another tuning characteristic so as to reach the desired frequency. This means that the synthesizer must search for the proper characteristic (Chapter 9) and eventually relock. The search and relock must be short enough not to bother the user.

Example 12.2

A cellular phone operates at an RX LO frequency of f_1 in a given cell (Fig. 12.3), and as it moves away from the base station, it receives a weaker signal. In such a situation, the base station in another cell, cell B, may be providing a *stronger* signal and can better serve this user. The user is thus switched to the base station in cell B (in a process called "hand-off"), but now the phone's synthesizer must operate at a different LO frequency, f_2. Explain how this is accomplished.

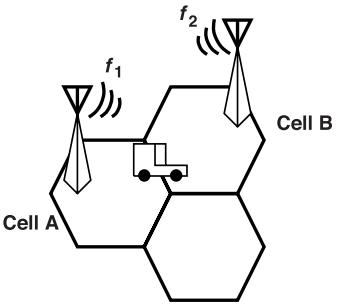


Figure 12.3 Handover of a call from one base station to another.

Solution

The new base station specifies which frequency is assigned to the user. The user's synthesizer must then search for a proper tuning range and relock. Again, this process must occur fast enough to avoid dropping the call during the hand-off process.

The accuracy of f_{out} in Fig. 12.2 also proves critical. Consider the Bluetooth example illustrated in Fig. 12.1(b) and suppose one user's TX LO frequency incurs an error of 100 ppm, approximately 240 kHz. This user's signal therefore partially overlaps another user's channel, causing corruption. The accuracy of f_{out} directly depends on that of f_{REF}, which is generated by a crystal oscillator. The difficulty is that crystals offering a higher accuracy are also more expensive.

12.2 Integer-N Synthesizers

The frequency-multiplying PLLs studied in Chapter 7 are also called integer-N synthesizers as they provide an *integer* ratio between the input and output frequencies. The reader is referred to Chapters 7 and 8 for the phase noise, spur, and lock behavior of such PLLs.

Shown in Fig. 12.4, an integer-N synthesizer generates $f_{out} = Mf_{REF}$, requiring that f_{REF} be equal

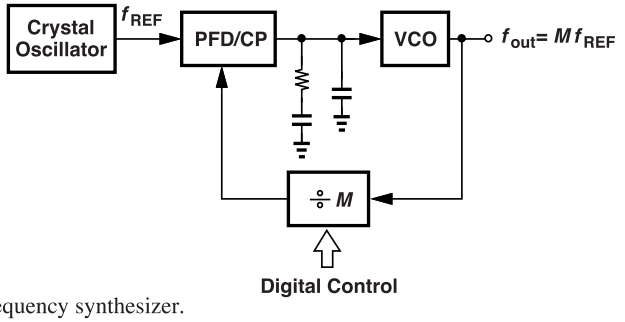


Figure 12.4 Integer-N frequency synthesizer.

to the channel spacing and M vary in steps of unity (Section 12.1). This means that the loop suffers from a narrow bandwidth in wireless standards that use a small channel spacing. For example, the 200-kHz channels in GSM impose a loop bandwidth less than 20 kHz, leading to little oscillator phase noise suppression and a long lock time. Additionally, the phase noise due to the reference and the PFD/CP cascade is multiplied by a large value as it propagates to the output, e.g., by a factor of 4500 in 900-MHz GSM.

Our study of integer-N synthesizers in this chapter is brief because their design follows the principles described in Chapters 2-9. The next example illustrates the starting point.

Example 12.3

A 2.4-GHz Bluetooth radio employs a VCO with $K_{VCO} = 150$ MHz/V. Study the range of values for the synthesizer's loop parameters.

Solution

As explained above, $f_{REF} = 1$ MHz and $M = 2400$. With $\zeta = 1$, we have

$$1 = \frac{R_1}{2}\sqrt{\frac{I_p(2\pi \times 150 \times 10^6)C_1}{2\pi \times 2400}} \tag{12.1}$$

and hence

$$R_1\sqrt{I_pC_1} = \frac{1}{125} \quad \text{V/rad/s}. \tag{12.2}$$

Also, if the loop bandwidth $2\zeta\omega_n \approx (2\pi f_{REF})/10$, then

$$2\sqrt{\frac{I_p(2\pi \times 150 \times 10^6)}{2\pi C_1 \times 2400}} = 2\pi(100 \text{ kHz}), \tag{12.3}$$

i.e.,

$$\sqrt{\frac{I_p}{C_1}} = 2\pi(200 \text{ Hz}). \tag{12.4}$$

Equations (12.2) and (12.4) contain three variables (and also yield $R_1I_p = 10$ V). Let us select $I_p = 0.5$ mA as an example, obtaining $C_1 = 317$ pF from the latter and $R_1 = 20$ kΩ from the former. The capacitor value is rather large; instead, we can choose $I_p = 0.25$ mA, $C_1 = 159$ pF, and $R_1 = 80$ kΩ.

The performance of integer-N synthesizers is directly affected by the PFD/CP imperfections and the VCO phase noise. In general, the CP and the VCO constitute the two challenging designs. These issues have been addressed in the previous chapters.

Example 12.4

A 20-MHz crystal oscillator is followed by a $\div 20$ circuit and the resulting 1-MHz reference drives a Bluetooth synthesizer. If the oscillator has phase noise of -150 dBc/Hz at 50-kHz offset, determine the synthesizer output phase noise at the same offset. Assume a loop bandwidth of 100 kHz.

Solution

The divider output phase noise is equal to $-150 - 20\log 20 = -176$ dBc/Hz. This value rises by $20\log(2400) \approx 68$ dB as it reaches the synthesizer output. Alternatively, we can say the oscillator phase noise goes up by $20\log(2400/20) = 42$ dB. These calculations have assumed that the divider phase noise falls well below -176 dBc/Hz, a challenge in the design of the divider circuit.

As mentioned previously, the feedback divider in Fig. 12.4 must have a modulus that can change in steps of unity. Such a divider topology is described in Chapter 15.

12.3 Fractional-N Synthesizers

Our study of integer-N synthesizers in the previous section points to a severe limitation, namely, the narrow loop bandwidth due to the relationship between the channel spacing and the reference frequency. Another

difficulty is that some systems must support several standards while using a single crystal oscillator. For example, a cell phone must synthesize frequencies for GSM, GPS, Bluetooth, WiFi, etc., from a 19.8-MHz crystal. Most of these channel frequencies do not have an integer ratio with respect to this reference or its integer submultiples.

These issues have motivated work on fractional-N synthesizers, in which the output and reference frequencies do not necessarily bear an integer ratio. The basic fractional-N loop resembles the integer-N topology shown in Fig. 12.4, except that the divide ratio (the modulus) is not a *constant* integer. We surmise that the modulus can be changed with time so that its *average* value is a noninteger number. For example, if the modulus is equal to M for half of the time and $M+1$ for the other half, then the average divide ratio is $M+0.5$, yielding $f_{out} = (M+0.5)f_{REF}$. These thoughts lead to the architecture depicted in Fig. 12.5, where $b(t)$ toggles the modulus.

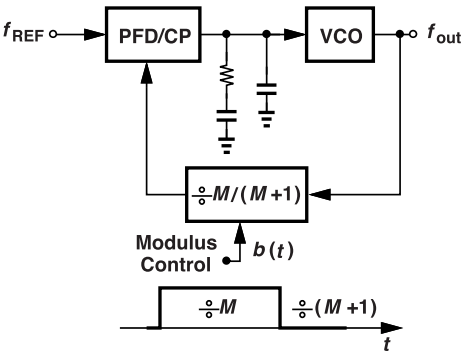


Figure 12.5 Simple implementaion of a fractional divide ratio.

This approach appears to provide several advantages. (1) If the average modulus can be anywhere between M and $M+1$, an arbitrarily narrow channel spacing is obtained for a given f_{REF}. For example, choosing the modulus equal to M for 99% of the time and $M+1$ for 1%, we can realize an average value of $M+0.01$, achieving an output frequency resolution of $0.01f_{REF}$. (2) With the channel spacing and f_{REF} "decoupled," we can opt for a high f_{REF} and a wide loop bandwidth so as to suppress the VCO phase noise. (3) A single crystal oscillator can serve different standards if the frequency resolution is chosen high enough. For example, we can synthesize $f_{out} = 5.2$ GHz from $f_{REF} = 19.5$ MHz if the average divide ratio is 266.667. (4) With the value of f_{REF} higher than the channel spacing, the loop's frequency multiplication factor is smaller, allowing less "amplification" of the reference phase noise.

Example 12.5

A student reasons that the fourth advantage mentioned above cannot be fulfilled because a wider loop bandwidth also leads to greater integrated reference noise. Is this conclusion correct?

Solution

No, it is not. Recall from Chapter 7 that, for $\zeta \geq 1.5$, a flat reference phase noise of S_0 yields at the output

$$\phi_{out,rms}^2 \approx \frac{M^2 S_0 \omega_{-3dB}}{2}. \tag{12.5}$$

Thus, even though ω_{-3dB} is increased by the same factor that M is reduced, $\phi_{out,rms}^2$ still decreases. Our assumption here is that the crystal oscillator phase noise, S_0, is relatively independent of its frequency, f_{REF}.

Fractional-N synthesizers must deal with a myriad of issues that are absent in integer-N topologies. For this reason, tremendous effort has been expended on the former. We build the foundation in this section and refer the reader to the vast literature on the subject for more advanced techniques.

12.3.1 The Need for Modulus Randomization

In the fractional-N loop of Fig. 12.5, we are tempted to toggle the modulus periodically between M and $M+1$. Let us derive this periodic control from the reference as shown in Fig. 12.6(a), e.g., using a $\div 10$ circuit. If the modulus is equal to M for nine reference cycles, $9T_{REF}$, and $M+1$ for one, we expect an average divide ratio

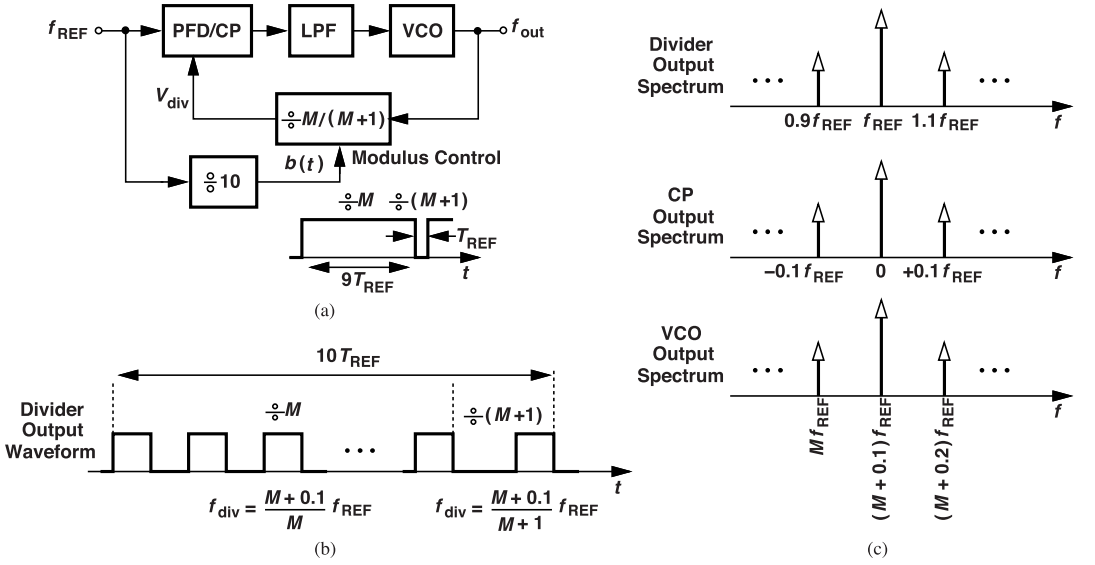


Figure 12.6 (a) Generation of modulus control for a fractional divide ratio, (b) divider output waveform, and (c) spurs at the divider output, PFD/CP output, and VCO output.

of $M + 0.1$. But how does the loop react to these modulus changes? Since the time scales are much shorter than the PLL settling time (our rule of thumb for the lock time is about 100 reference cycles), the loop is slow and the VCO frequency does not find enough time to lock to Mf_{REF} or $(M + 1)f_{REF}$. That is, f_{out} settles around the average of these two, $(M + 0.1)f_{REF}$. However, the divider output, V_{div}, is not periodic [Fig. 12.6(b)]: it has a frequency of $(M+0.1)f_{REF}/M$ for nine reference periods and $(M+0.1)f_{REF}/(M+1)$ for one. We observe that the divider output experiences *periodic* phase modulation, thereby exhibiting sidebands. Since the period of modulation is equal to $10T_{REF}$, we expect V_{div} to contain sidebands at frequency offsets equal to $\pm 0.1f_{REF}$, $\pm 0.2f_{REF}$, etc. [Fig. 12.6(c)].

We must now ask how the PFD/CP cascade responds to an unmodulated reference and a modulated feedback signal. Viewing the PFD as a mixer (an analog multiplier), we recognize that the sidebands in V_{div} are shifted up and down by f_{REF} as they mix with the reference, landing at frequencies equal to $\pm 0.1f_{REF}$, $\pm 0.2f_{REF}$, etc., at the CP output [Fig. 12.6(c)]. Since the PFD measures the phase difference between its two inputs, we conclude that the spurs present in V_{div} can be "referred" to the reference input, i.e., as if V_{div} had none but the reference were phase-modulated. This indicates that the spurs experience a low-pass response as they travel to the synthesizer output. The VCO is modulated by the CP output components and produces sidebands at the same offset frequencies around $f_{out} = (M + 0.1)f_{REF}$. Distinct from reference-induced sidebands, these components are called "fractional spurs."

Fractional spurs prove troublesome because they are large and they can fall well within the synthesizer's loop bandwidth and hence are not attenuated. For example, if we choose the modulus duty cycle such that $f_{out} = (M + 0.01)f_{REF}$, these spurs appear at $\pm 0.01f_{REF}$, $\pm 0.02f_{REF}$, etc., at the CP output.

We recognize that fractional spurs arise from the periodic nature of $b(t)$. In order to reduce the magnitude of these spurs, we can switch the divide ratio between the two values *randomly* while maintaining a desirable average. As illustrated in Fig. 12.7, the modulus is controlled by a random binary sequence, $b(t)$, and the divider output frequency is randomly modulated so that it is free from spurs. The modulation now produces

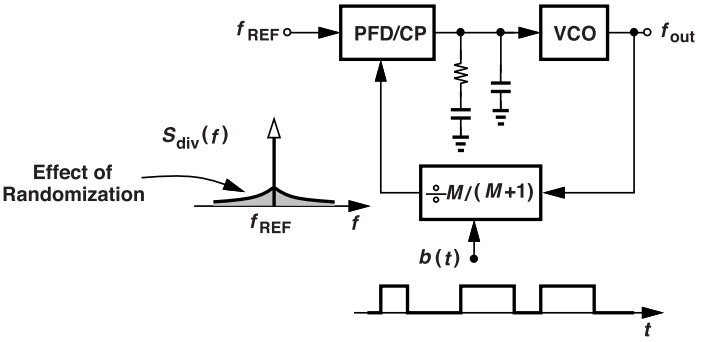


Figure 12.7 Randomization of divide ratio to convert spurs to noise.

phase noise, which travels through the PFD and the CP and emerges as a noise voltage across the loop filter, raising the phase noise at the synthesizer output.

Let us summarize our thoughts. Our objective is to create a fractional divide ratio so that the synthesizer output frequency can vary in small steps even with a relatively high f_{REF}. This is accomplished by switching the divide ratio between M and $M + 1$ so that the average is the desired value. However, this toggling must occur randomly to avoid fractional periodicity in the divider output.

The additional phase noise arising from modulus randomization poses its own issues. We must therefore analyze this effect and determine under what conditions it can be tolerated.

Example 12.6

Examine the output waveform of the frequency divider in the above architecture.

Solution

Figure 12.8 plots the VCO and divider outputs as the divide ratio jumps from 3 to 4. We denote the VCO

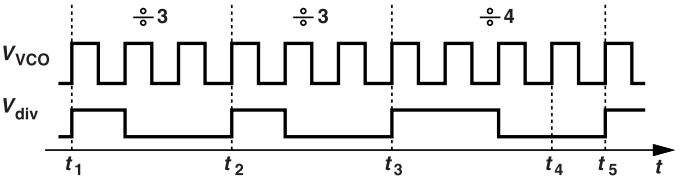
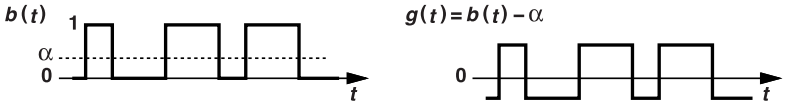


Figure 12.8 Phase jump at divider output when modulus changes.

period by T_{VCO}. We can say that the divider output frequency is modulated. From another viewpoint, suppose only the rising edges of V_{div} are of interest (e.g., because the PFD responds to such edges). We note the uniformly-spaced transitions at t_1, t_2, and t_3. But the next rising edge does not occur at t_4 and is instead shifted by T_{VCO} to t_5. In other words, the divider output *phase* jumps by T_{VCO} at t_4. We therefore say that the modulus randomization causes phase jumps equal to $\pm T_{VCO}$, $\pm 2T_{VCO}$, etc., in the divider output.

Phase Noise Calculation We wish to formulate the phase noise resulting from the random modulation of the divide ratio. The divide ratio can be expressed as $M + b(t)$, where $b(t)$ randomly toggles between 0 and 1 such that its average is the desired value, α [Fig. 12.9(a)]. As a binary (one-bit) waveform approximating α, $b(t)$ contains significant quantization noise, $g(t)$. That is, $b(t) = \alpha + g(t)$ with $g(t)$ shown in Fig. 12.9(b). Of course, if we could have a multilevel waveform for $b(t)$, then the quantization noise would be smaller, but the difficulty is that $b(t)$ must be a one-bit sequence because the feedback divider can divide by only M or $M + 1$ and by no other value in between.



(a) (b)

Figure 12.9 (a) Random binary data having an average of α, and (b) waveform showing the quantization noise.

Our approach to phase noise calculation proceeds as follows. We first compute the divider output frequency as a function of time. Next, we integrate this quantity to obtain the divider output phase. The random component then represents the phase noise.

We can write the divider output frequency as

$$f_{div} = \frac{f_{out}}{M + b(t)} \tag{12.6}$$

$$= \frac{f_{out}}{M + \alpha + g(t)}. \tag{12.7}$$

In general, $g(t) \ll M + \alpha$, leading to

$$f_{div} \approx \frac{f_{out}}{M + \alpha}\left[1 - \frac{g(t)}{M + \alpha}\right] \tag{12.8}$$

$$\approx \frac{f_{out}}{M + \alpha} - \frac{f_{out}}{(M + \alpha)^2} g(t). \tag{12.9}$$

We can consider the second term as frequency noise since it directly corrupts f_{div}. The output waveform of the divider is obtained by converting f_{div} to phase and using it as the argument of a sinusoid:

$$V_{div}(t) \approx V_0 \cos\left[\frac{2\pi f_{out}}{M + \alpha} t - \frac{2\pi f_{out}}{(M + \alpha)^2} \int g(t)dt\right]. \tag{12.10}$$

Since $f_{out}/(M + \alpha) = f_{REF}$, the first term in the argument represents the desired component, and the second, the phase noise. It is important not to confuse the modulus control quantization noise, $g(t)$, with the phase noise in the divider output; the latter is the integral of the former multiplied by $2\pi f_{out}/(M + \alpha)^2$.

We are primarily interested in the spectrum of the phase noise, $S_{\phi n}(f)$, and can write from (12.10):

$$S_{\phi n}(f) = \left[\frac{2\pi f_{out}}{(M + \alpha)^2}\right]^2 S_g(f) \cdot \frac{1}{4\pi^2 f^2}, \tag{12.11}$$

where $S_g(f)$ is the modulus quantization noise spectrum [similar to the spectrum of $b(t)$] and the factor $(4\pi^2 f^2)^{-1}$ accounts for the time-domain integral necessary to go from frequency to phase. It follows that

$$S_{\phi n}(f) = \frac{S_g(f)}{(M + \alpha)^4}\left(\frac{f_{out}}{f}\right)^2 \tag{12.12}$$

$$= \frac{S_g(f)}{(M + \alpha)^2}\left(\frac{f_{REF}}{f}\right)^2. \tag{12.13}$$

Representing the divider output phase noise, this general result holds for any random sequence $b(t)$ that is applied to the divider's modulus control. Note from Fig. 12.9 that the spectra of $b(t)$ and $g(t)$ differ by only an impulse as the former contains a dc value of α.

The phase noise in the divider output can be simply referred to the reference input, traveling to the synthesizer output according to the PLL's low-pass transfer function. Within the closed-loop bandwidth, $S_{\phi n}(f)$ is

simply multiplied by $(M+\alpha)^2$ and yields

$$S_{\phi,out}(f) = \frac{S_g(f)}{(M+\alpha)^2}\left(\frac{f_{out}}{f}\right)^2. \tag{12.14}$$

For example, if $S_g(f)$ is white, then the output phase noise spectrum appears as shown in Fig. 12.10.

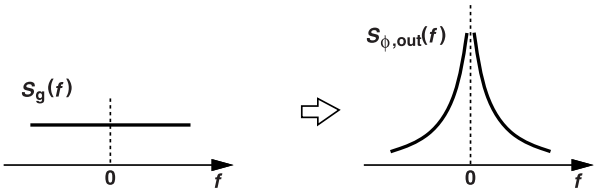


Figure 12.10 Conversion of white quantization noise to phase noise.

Example 12.7

The output phase noise spectrum depicted in Fig. 12.10 exhibits very *high* values within the PLL bandwidth, an unacceptable situation. Explain generally how this effect can be avoided.

Solution

The high values arise from the inevitable multiplication of $S_g(f)$ by $(f_{out}/f)^2$ as we go from frequency noise to phase noise. To avoid this effect, we must ensure that the signal randomizing the divide ratio does *not* have a white spectrum. In fact, as shown in Fig. 12.11, this spectrum must fall steeply enough to zero as $f \to 0$ so that its multiplication by $(f_{out}/f)^2$ still does not cause high values.

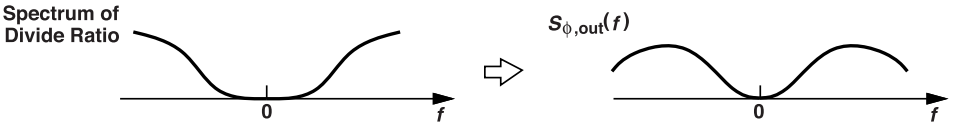


Figure 12.11 General shape of divide ratio spectrum for acceptable phase noise.

The principal challenge in fractional-N synthesizer design is to minimize the phase noise due to modulus randomization. Numerous techniques have been developed to address this issue. An important method is "noise shaping."

12.3.2 Noise Shaping

Basic Concepts Figure 12.10 suggests that the modulus quantization noise can considerably contribute to the output phase noise. Thus, we must avoid a *white* spectrum for $g(t)$. As in Example 12.7, we can ask, what is a desirable shape for $S_g(f)$? As illustrated in Fig. 12.12(a), we prefer the noise to appear only at high frequencies so that it is filtered upon propagation through the PLL. In other words, we wish to randomize the divide ratio such that the resulting phase noise within the loop bandwidth is small. This means that the spectrum of $b(t)$ is "shaped" as shown in Fig. 12.12(a). Of course, if we could, we would also avoid the higher noise away from $f = 0$, but this rise is an inevitable result of quantization and shaping. We loosely say that the noise energy is pushed to high frequencies, i.e., the spectrum has a high-pass shape.

These thoughts lead to the scenario depicted in Fig. 12.12(b), where the modulus control is randomized according to a specific sequence, $b(t)$, whose spectrum is not white but shaped. Consequently, the divider output spectrum begins from zero at f_{REF} and rises as $|f - f_{REF}|$ increases.

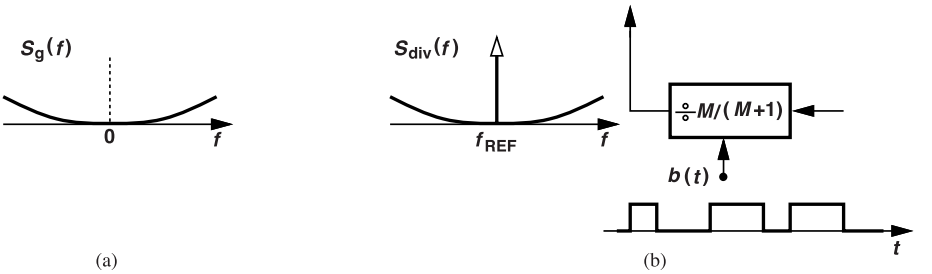


Figure 12.12 (a) Desirable quantization noise spectrum, and (b) random binary data controlling the divide ratio for desired phase noise spectrum.

The reader may have several questions at this point. (1) How exactly do we choose the random sequence of ONEs and ZEROs in $b(t)$ to obtain a high-pass shaped spectrum and an average value equal to α? (2) Given that, according to Eq. (12.9), the spectrum of $g(t) = b(t) - \alpha$ appears in the *frequency* of the divider output, how do we ensure that the *phase* is also shaped as desired? (3) How fast should (or can) $b(t)$ toggle? (4) What happens to the noise peaks in the divider output as the signal travels through the PFD/CP/LPF/VCO chain? We answer these questions as we follow the developments below.

Noise Shaping by Feedback Before creating a random sequence of ONEs and ZEROs with a high-pass spectrum, we consider the negative-feedback system shown in Fig. 12.13(a). Here, in addition to $X(s)$, the loop senses another signal, $Q(s)$, "near" the output. The transfer function experienced by $Q(s)$ (with $X = 0$)

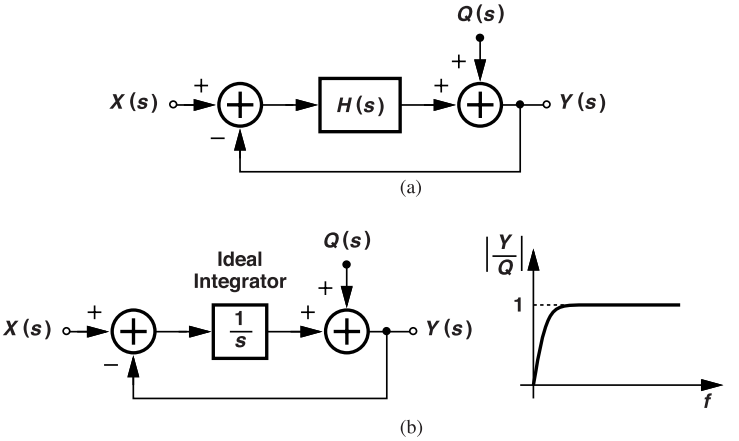


Figure 12.13 (a) Noise injection into a feedback system, and (b) behavior if $H(s)$ is an ideal integrator.

is given by

$$\frac{Y(s)}{Q(s)} = \frac{1}{1 + H(s)}. \tag{12.15}$$

Proper choice of $H(s)$ can yield a *high-pass* response; for example, if $H(s)$ is an ideal integrator [Fig. 12.13(b)],

$$\frac{Y(s)}{Q(s)} = \frac{s}{1 + s}. \tag{12.16}$$

Note that $Y/X = 1/(1+s)$ in this case. We recognize that low-frequency components in $Q(s)$ are suppressed. This is because the high loop gain opposes the time variation of $Q(s)$ by providing, at the integrator output,

a negative replica of Q. At very high frequencies, on the other hand, the integrator gain approaches zero, allowing $Q(s)$ to reach the output with a gain close to unity. This behavior is similar to that of VCO phase noise in a PLL (Chapter 8). If we assume that $X(s)$ in Fig. 12.13(b) is constant and equal to α, then the low-pass transfer, $1/(1+s)$, implies that the output also has an average value equal to α. In other words, the output contains some components of Q riding atop a dc value of α. If Q represents unwanted noise, we observe that the system high-pass filters it.

Example 12.8

A student decides to replace the ideal integrator in Fig. 12.13(b) with a first-order low-pass filter having a low-frequency gain of A_0. Explain how Q is shaped in this case.

Solution

Shown in Fig. 12.14(a), such a system provides a noise-shaping function given by

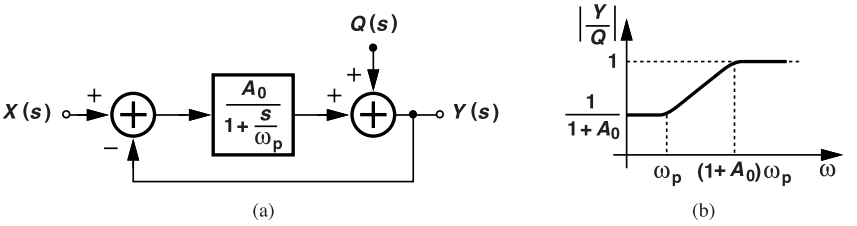


Figure 12.14 Noise injection into a feedback system with $H(s)$ realized as a first-order LPF.

$$\frac{Y}{Q} = \frac{1}{1+\dfrac{A_0}{1+\dfrac{s}{\omega_p}}} \tag{12.17}$$

$$= \frac{1+\dfrac{s}{\omega_p}}{1+A_0+\dfrac{s}{\omega_p}}. \tag{12.18}$$

We observe a zero at $-\omega_p$ and a pole at $-(1+A_0)\omega_p$. Thus, as plotted in Fig. 12.14(b), $|Y/Q|$ does *not* begin from zero at $\omega = 0$, and the flat noise from $\omega = 0$ to ω_p corrupts Y.

Let us now construct a digital counterpart of the system shown in Fig. 12.13(b). In the first step, we apply a digital representation of α to a loop consisting of an adder and a digital integrator as shown in Fig. 12.15(a). Note that, owing to its finite resolution, the digital integrator adds quantization noise as it processes C. This noise is similar to the additive term, Q, in Fig. 12.13(b). What can we say about Y? This output is a digital quantity having an average equal to α. Viewing the system as that in Fig. 12.13(b), approximating its closed-loop response by $Y/X = 1/(1+s)$, and assuming $y(t=0) = 0$, we observe that the output grows as in a first-order system [Fig. 12.15(b)] and, in the steady stage, toggles around α by 1 least significant bit (LSB). This output incurs little quantization noise with respect to α if the word length is large enough, say, 16 bits. We thus envision that Y can drive the modulus control of the feedback divider [Fig. 12.15(c)]. But the difficulty is that a 16-bit Y cannot be directly applied to the modulus control.

What if we use only the most significant bit (MSB) of Y to toggle the divide ratio between M and $M+1$ [Fig. 12.15(d)]? Then, acting as a one-bit approximation, the MSB exhibits high quantization noise. Unfortunately, this noise is *unshaped* because the quantization occurs outside the loop, i.e., Q is injected after the point sensed by the feedback network. Instead, we can perform the quantization within the loop, as shown

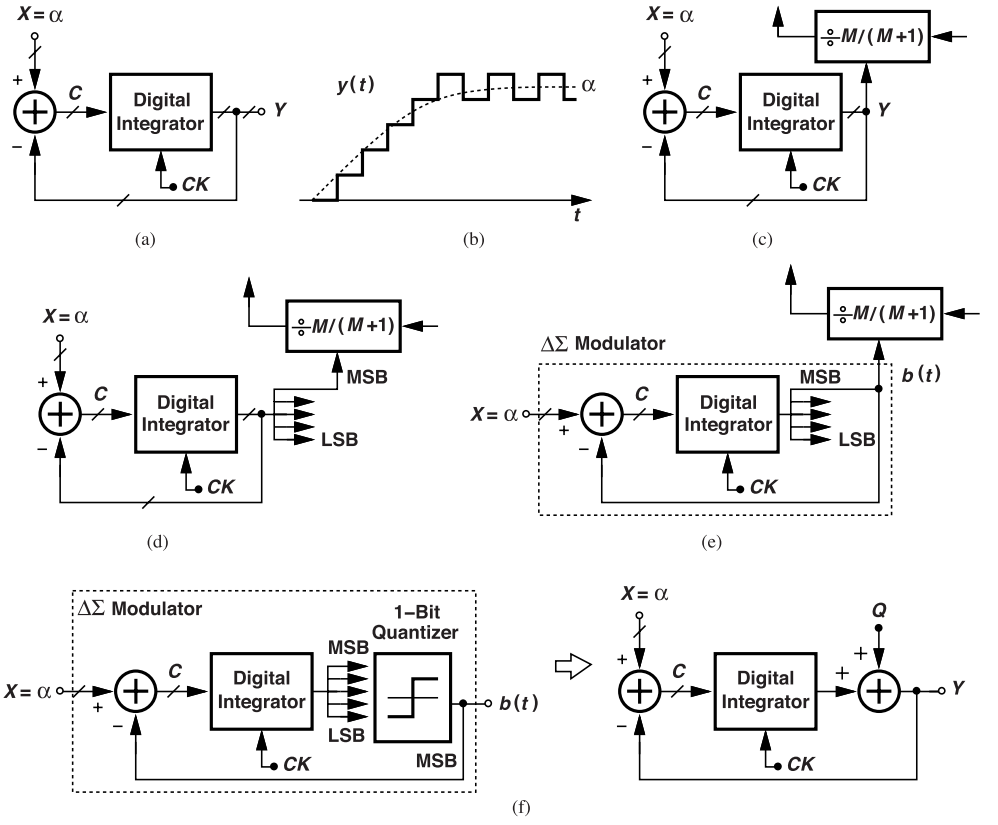


Figure 12.15 (a) Feedback system using a digital integrator, (b) system's output as a function of time, (c) system driving a feedback divider, (d) using only MSB output to control the divider, (e) quantization within the feedback loop, and (f) equivalent quantization operation.

in Fig. 12.15(e). Discarding all of the output bits but the MSB is equivalent to passing the signal through a one-bit quantizer [Fig. 12.15(f)] and hence adding quantization noise to it, but the feedback now shapes this noise's spectrum and suppresses its low-frequency components.

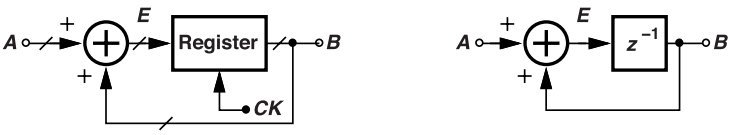
It is instructive to summarize what we have accomplished thus far. The digital feedback system in Fig. 12.15(e) senses an accurate (e.g., 16-bit) representation of α and generates a one-bit sequence, $b(t)$, whose quantization noise is shaped into a high-pass spectrum by virtue of the loop containing an integrator. If applied to the divider modulus control, $b(t)$ imparts its shaped noise to the output frequency of the divider. The loop generating $b(t)$ in Fig. 12.15(e) is called a "$\Delta\Sigma$ modulator."[1]

12.3.3 Discrete-Time Model

The digital feedback architecture shown in Fig. 12.15(e) operates as a discrete-time system, i.e., it updates the results around the loop once per clock cycle. As such, the $\Delta\Sigma$ modulator must be analyzed using a discrete-time model—even though we have intuitively approximated the transfer function from Q to Y by $s/(1+s)$. Of course, we still expect a high-pass response for Y/Q.

Let us first examine the digital integrator in Fig. 12.15(e). Depicted in Fig. 12.16(a), this function is realized as an "accumulator" in the digital domain, employing an adder and a register in a loop. In each clock

[1] Also known as a $\Sigma\Delta$ modulator.



(a) (b)

Figure 12.16 (a) Simple accumulator, and (b) its equivalent discrete-time model.

cycle, the register reads and stores $A + B$. Thus, the content of the register is equal to its previous value plus the present value of A, which signifies discrete-time integration. Since the register simply provides a one-cycle delay, we can model it by a z^{-1} block, arriving at the discrete-time loop shown in Fig. 12.16(b) and writing

$$\frac{B}{A}(z) = \frac{z^{-1}}{1 - z^{-1}}. \tag{12.19}$$

The topology therefore integrates the input and delays the result by one clock cycle.

Example 12.9

Is the signal at E in Fig. 12.16(b) also the integral of A?

Solution

Yes, it is. The signals at E and B only differ by one clock delay. Redrawing the structure as shown in Fig. 12.17, we have

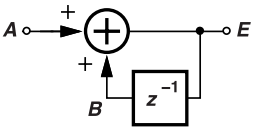


Figure 12.17 Accumulator with delay in feedback path.

$$\frac{E}{A}(z) = \frac{1}{1 - z^{-1}}. \tag{12.20}$$

The topologies in Figs. 12.16(b) and 12.17 are called "delaying" and "nondelaying" integrators, respectively. We also conclude that a discrete-time differentiator can be represented by $1 - z^{-1}$ because the cascade of an integrator and a differentiator must have a transfer function equal to 1.

In order to formulate the shaping of the quantization noise in Fig. 12.15(e), we construct its discrete-time model as shown in Fig. 12.18(a). Setting X to 0, we have

$$\frac{Y}{Q}(z) = \frac{1}{1 + \dfrac{z^{-1}}{1 - z^{-1}}} \tag{12.21}$$

$$= 1 - z^{-1}. \tag{12.22}$$

From our continuous-time analysis leading to Fig. 12.13(b), we expect this response to represent a high-pass function. To see this, we visualize Eq. (12.22) as in Fig. 12.18(b), where a delayed replica of Q is subtracted from it. We recognize that, if Q varies slowly as a function of time, it resembles the delayed copy, Q_Δ, and

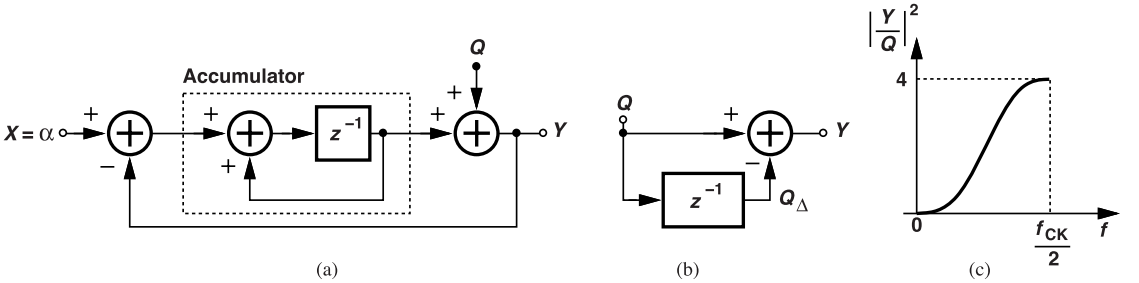


Figure 12.18 (a) Discrete-time model of $\Delta\Sigma$ modulator, (b) equivalent behavior for the quantization noise, and (c) spectral shaping function from Q to Y.

their difference is small. Conversely, if varying quickly, Q bears little similarity to Q_Δ, thus producing large excursions in Y. From another perspective, $1 - z^{-1}$ is simply a differentiator. The system therefore exhibits a high-pass response from Q to Y. Also, Y/X is a low-pass function, allowing Y to have an average value of α.

We wish to compute the output noise spectrum in Fig. 12.18(a). Since $z = \exp(j2\pi fT_{CK})$, where T_{CK} denotes the clock period, we first write the magnitude squared of Eq. (12.22) as

$$\left|\frac{Y}{Q}(z = e^{j2\pi fT_{CK}})\right|^2 = |1 - e^{-j2\pi fT_{CK}}|^2 \tag{12.23}$$

$$= 4\sin^2(\pi fT_{CK}) \tag{12.24}$$

$$= 2[1 - \cos(2\pi fT_{CK})]. \tag{12.25}$$

Plotted in Fig. 12.18(c), this spectral shaping function reaches a maximum of 4 at $f = 1/(2T_{CK}) = f_{CK}/2$.

Example 12.10

Does the result given by (12.24) agree with our intuitive approximation expressed as $s/(1 + s)$?

Solution

From the continuous-time approximation, we have $|j\omega/(1 + j\omega)|^2 = \omega^2/(1 + \omega^2)$, which, for low noise frequencies, reduces to ω^2. Equation (12.24) too yields $|Y/Q|^2 \approx 4\pi^2 f^2 T_{CK}^2 = \omega^2 T_{CK}^2$ for $f \ll (\pi T_{CK})^{-1}$. The results are therefore similar. The reader is encouraged to explain the factor of T_{CK}^2 discrepancy.

Denoting the unshaped spectrum of Q in Fig. 12.18(a) by $S_q(f)$, we write from (12.24)

$$S_Y(f) = 4\sin^2(\pi fT_{CK})S_q(f). \tag{12.26}$$

The architectures shown in Figs. 12.15(e) and 12.18(a) are called a "first-order $\Delta\Sigma$ modulator," with the order defined by the number of integrators [1].

Example 12.11

Examine the noise spectrum at different points within the $\Delta\Sigma$ modulator of Fig. 12.19(a). Assume X is constant.

Solution

The spectrum at A is similar to that of Y because $A = Y - \alpha$. To determine the spectrum at B, we first set X to zero and write $B(z) = -Y(z)[z^{-1}/(1 - z^{-1})] = -(1 - z^{-1})Q(z)[z^{-1}/(1 - z^{-1})] = -z^{-1}Q(z)$.

That is, B is simply a delayed negative replica of Q, containing *unshaped* quantization noise. Assuming a flat spectrum for Q, we obtain the results shown in Fig. 12.19(b).

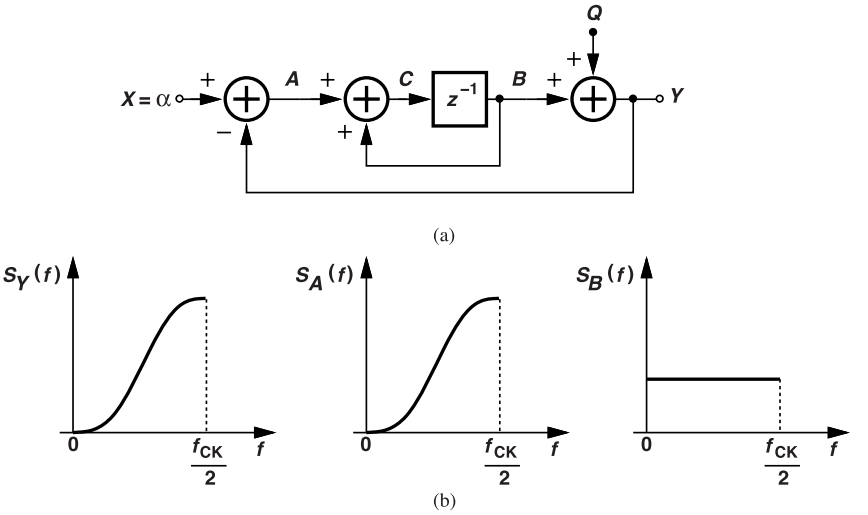


Figure 12.19 (a) $\Delta\Sigma$ modulator, and (b) noise spectra at different points within the loop.

In summary, the first-order $\Delta\Sigma$ modulator of Fig. 12.15(e) produces a one-bit random binary sequence having an average value of α and a noise spectrum whose energy is pushed to high frequencies. We typically approximate the unshaped quantization noise by $S_q(f) = 1/(12f_{REF})$.

Example 12.12

A student notes that both Q and B have a white spectrum in the previous example, but their difference, Y, does not. The student then doubts this result, reasoning that the difference between two white noise sources must also be white. Explain the flaw in this reasoning.

Solution

The difference is white if the two sources are uncorrelated. Here, on the other hand, B is derived from Q and is correlated with it. As another example, Q and Q_Δ in Fig. 12.18(b) can be white but their difference is not.

12.3.4 $\Delta\Sigma$ Fractional-N Synthesizers

The foregoing developments lead to the overall fractional-N synthesizer depicted in Fig. 12.20 [2]. Here, the $\Delta\Sigma$ modulator generates the binary waveform $b(t)$ with an average value equal to α and a high-pass quantization noise spectrum. This waveform toggles the modulus between M and $M+1$, creating an average divide ratio of $M+\alpha$ and hence an output frequency of $(M+\alpha)f_{REF}$. Note that the modulator is clocked by the divider output. The digital representation of α can be changed in very small steps (for example, α can have a word length of 18 bits), and so can f_{out}. The step size is dictated by the maximum allowable frequency error in a given application.

From Eq. (12.7),

$$f_{div} = \frac{f_{out}}{M + \alpha + g(t)} \tag{12.27}$$

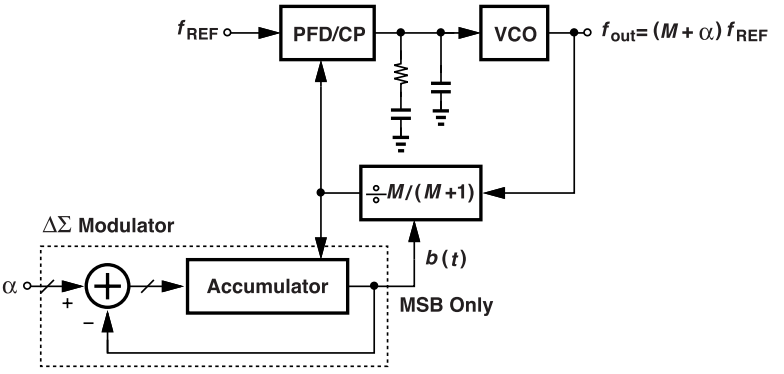


Figure 12.20 Basic $\Delta\Sigma$ fractional-N synthesizer.

$$\approx \frac{f_{out}}{M+\alpha} - \frac{f_{out}}{(M+\alpha)^2} g(t), \tag{12.28}$$

where $g(t)$ represents the noise in $b(t)$ (the $\Delta\Sigma$ quantization noise) and produces frequency noise in the divider output. How do we compute the corresponding phase noise? In Section 12.3.1, we related frequency and phase quantities by a factor of $1/s$ and derived Eq. (12.13). We use this equation, noting that $S_g(f)$ is now shaped by the $\Delta\Sigma$ modulator and equal to $4\sin^2(\pi f T_{REF}) S_q(f)$:

$$S_{\phi n}(f) = \frac{1}{(M+\alpha)^2} \left(\frac{f_{REF}}{f}\right)^2 [4\sin^2(\pi f T_{REF})] S_q(f) \tag{12.29}$$

$$= \frac{f_{REF}^2}{(M+\alpha)^2} \frac{4\sin^2(\pi f T_{REF})}{f^2} S_q(f), \tag{12.30}$$

where $S_q(f)$ is the unshaped spectrum of $b(t)$. This shaping function is relatively flat for $f \ll (\pi T_{REF})^{-1} = f_{REF}/\pi$ and approximately equal to $[4\pi^2/(M+\alpha)^2] S_q(f)$. That is, the phase noise at low offset frequencies is *not* zero. We resolve this issue in Section 12.3.5 by raising the order of the $\Delta\Sigma$ modulator, i.e., by implementing a sharper noise shaping function. Note that the divider output noise is indistinguishable from the reference noise and experiences the PLL's low-pass response as it reaches the VCO output. Figure 12.21 summarizes our findings thus far. In the time domain, the divider output phase jumps by $\pm T_{VCO}$, $\pm 2T_{VCO}$, etc., where $T_{VCO} = 1/f_{out}$.

It is important not to confuse unshaped and shaped quantization noise quantities. When only the MSB of the integrator output in Fig. 12.15(e) is retained and the LSBs are dropped, we introduce unshaped quantization, $S_q(f)$ (similar to the free-running phase noise of a VCO). But the feedback loop shapes this noise, multiplying it by $4\sin^2(\pi f T_{REF})$. It is also important not to confuse frequency noise and phase noise.

Problem of Tones In addition to a high-pass spectrum, the output of $\Delta\Sigma$ modulators exhibits pure sinusoidal components, i.e., periodic waveforms. These spurs or "tones" arise because the output is, in fact, periodic. To understand this phenomenon, we return to the topology of Fig. 12.15(e) and redraw it as shown in Fig. 12.22(a), where the accumulator is modeled by an integrator having a multibit output, $w(t)$. The act of discarding all of the bits of $w(t)$ except for the MSB is represented by a 1-bit quantizer, i.e., $y(t)$ is equal to the MSB of $w(t)$. Suppose, for example, that $\alpha = 0.1$ and $w(t)$ begins from zero. The subtractor output, $\alpha - y(t)$, is then positive, in which case the accumulator output rises with time [Fig. 12.22(b)]. This continues until the MSB of $w(t)$ jumps to 1 and so does $y(t)$, producing a value of $\alpha - 1 = 0.1 - 1$ at the subtractor output. In the next clock cycle, the integrator adds this negative value to its output, causing $w(t)$ to fall to α. Such a behavior repeats every $10T_{CK}$ seconds so that the average value of $y(t)$ is equal to 0.1. In other words, the modulator's output is not random at all! We say the loop experiences a "limit cycle."

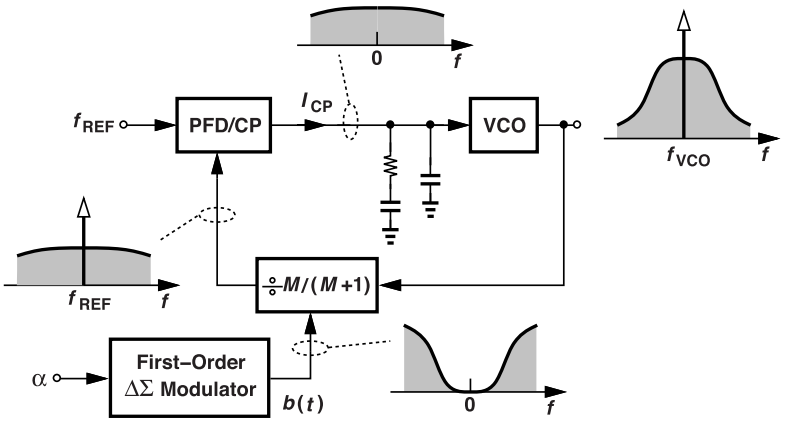


Figure 12.21 Fractional-N synthesizer showing quantization noise spectra at different points.

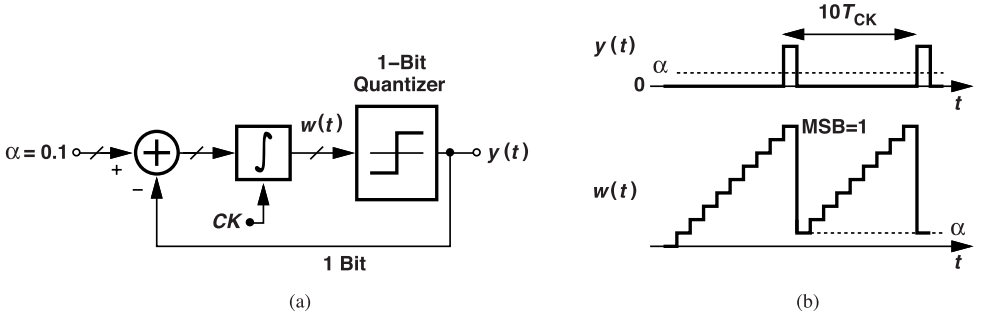


Figure 12.22 (a) $\Delta\Sigma$ modulator with an input equal to 0.1, and (b) waveforms within the loop revealing periodicity.

The $y(t)$ waveform in Fig. 12.22(b) has a fundamental equal to $0.1f_{CK} = 0.1f_{REF}$, giving rise to tones at this frequency and its harmonics. For smaller values of α, the tone frequencies are even lower, passing through the PLL unattenuated.

In order to suppress the tones, the periodicity observed in Fig. 12.22(b) must be broken. For example, if the LSB of α randomly jumps between 0 and 1 (in an operation called "dithering"), then the pulses in $y(t)$ occur randomly. As a result, the tones are spread into noise. The act of dithering requires its own random sequence generator. Another approach to reducing the spurs is to raise the order of the $\Delta\Sigma$ modulator.

12.3.5 Higher-Order $\Delta\Sigma$ Modulators

As discovered in the previous section, a first-order $\Delta\Sigma$ modulator leads to a finite phase noise in the divider output even at low offset frequencies. This is because the noise shaping function $4\sin^2(\pi f T_{REF})$ divided by $4\pi^2 f^2$ yields a constant, nonzero value for small f. For this reason and for their tone problems, we rarely use first-order modulators. We seek a sharper noise-shaping function, perhaps of the form $\sin^{2n}(\pi f T_{REF})$, where $n > 1$. The discrete-time equivalent would be $(1 - z^{-1})^n$.

Let us consider the first-order loop shown in Fig. 12.22(a) and ask how we can reduce its quantization noise—at least at low frequencies. We can utilize, for example, the two MSBs at the accumulator output (Fig. 12.23), and design the divider to have two moduli. However, this method does not reduce the phase noise because the minimum divider output phase jump is still T_{VCO}. In other words, it is the divider output phase resolution that must be improved.

How else can we reduce the quantization noise? Can we change the 1-bit quantizer in Fig. 12.22(a) to one having less noise and still providing a single-bit output? Yes, we can replace this quantizer by another $\Delta\Sigma$

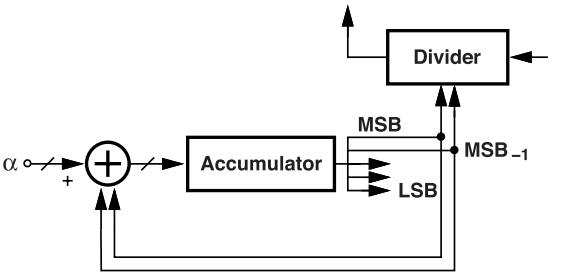


Figure 12.23 $\Delta\Sigma$ modulator with 2-bit output.

modulator, as depicted in Fig. 12.24(a). The circuit now contains two integrators and acts as a second-order

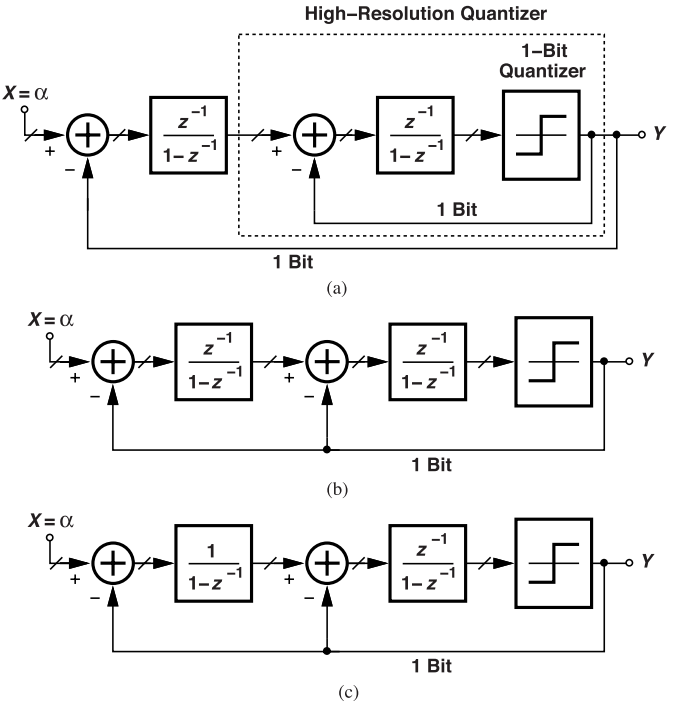


Figure 12.24 (a) Reduction of quantization noise by replacing the quantizer with another $\Delta\Sigma$ modulator, (b) simplfied architecture, and (c) change of first integrator to a nondelaying topology.

modulator. Note that the two lines carrying $y(t)$ to the subtractors can be merged into one. In Problem 12.14, we prove from the equivalent topology shown in Fig. 12.24(b) that this system provides a noise-shaping function given by

$$\frac{Y}{Q}(z) = \frac{(1-z^{-1})^2}{z^{-2}-z^{-1}+1}. \tag{12.31}$$

The numerator behaves as two cascaded differentiators, exhibiting sharper noise shaping and less noise at low frequencies. However, the poles introduced by the denominator prove undesirable. We thus replace the first delaying integrator with a nondelaying topology (Example 12.9), arriving at the architecture illustrated in Fig. 12.24(c). Setting X to 0, we can write

$$\left(-Y\frac{1}{1-z^{-1}}-Y\right)\frac{z^{-1}}{1-z^{-1}}+Q=Y \tag{12.32}$$

and hence

$$\frac{Y}{Q}(z) = (1 - z^{-1})^2. \tag{12.33}$$

In a manner similar to the calculations in Section 12.3.3, we obtain

$$S_Y(f) = |2\sin(\pi f T_{CK})|^4 S_q(f), \tag{12.34}$$

where $S_q(f)$ is the unshaped quantization noise spectrum. Figure 12.25 compares the noise-shaping functions

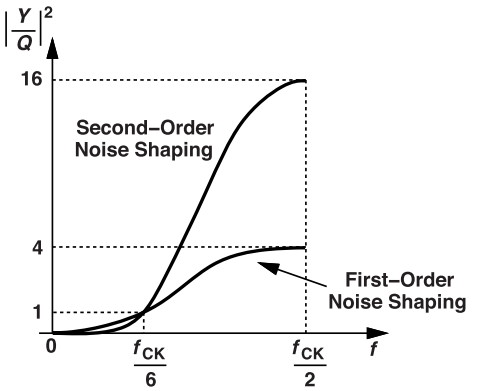


Figure 12.25 Quantization noise shaping functions in first- and second-order modulators.

of first- and second-order modulators, revealing two points: (a) the latter remains less than the former for frequencies up to $f_{CK}/6$, and (b) the latter exhibits a substantially higher peak at $f_{CK}/2$. The first property is desirable but the second is not.

Divider Output Phase Noise From (12.34) and following our derivations in Section 12.3.3, we can express the divider output phase noise as

$$S_{\phi n}(f) = \frac{1}{(M+\alpha)^2}\left(\frac{f_{REF}}{f}\right)^2 |2\sin(\pi f T_{REF})|^4 S_q(f), \tag{12.35}$$

which, for $f \ll f_{REF}/\pi$, reduces to

$$S_{\phi n}(f) \approx \frac{16\pi^4}{(M+\alpha)^2}\frac{f^2}{f_{REF}^2} S_q(f). \tag{12.36}$$

Thus, the phase noise begins at zero and rises quadratically with f. As mentioned in Section 12.3.3, this noise experiences the synthesizer's low-pass transfer function as it travels to the VCO output.

Cascaded Loops The idea of replacing the quantizer within a $\Delta\Sigma$ loop with another modulator can be extended to higher orders. For example, if we replace the 1-bit quantizer in Fig. 12.24(c) with a first-order $\Delta\Sigma$ loop, the overall system exhibits third-order noise shaping. However, feedback loops having an order of 3 or higher are potentially unstable.

An alternative approach is to cascade loops whose orders do not exceed 2. Let us first note that the noise introduced by a quantizer can be calculated by subtracting its input from its output [Fig. 12.26(a)]. This is because the quantizer output is equal to $A + Q_1$, where Q_1 denotes the noise, and hence $B = A + Q_1 - A = Q_1$. As shown in Fig. 12.26(b), therefore, we can reconstruct the unshaped quantization noise using a subtractor. Note that Y contains the shaped noise, but $Y - P = P + Q_1 - P = Q_1$ contains the unshaped

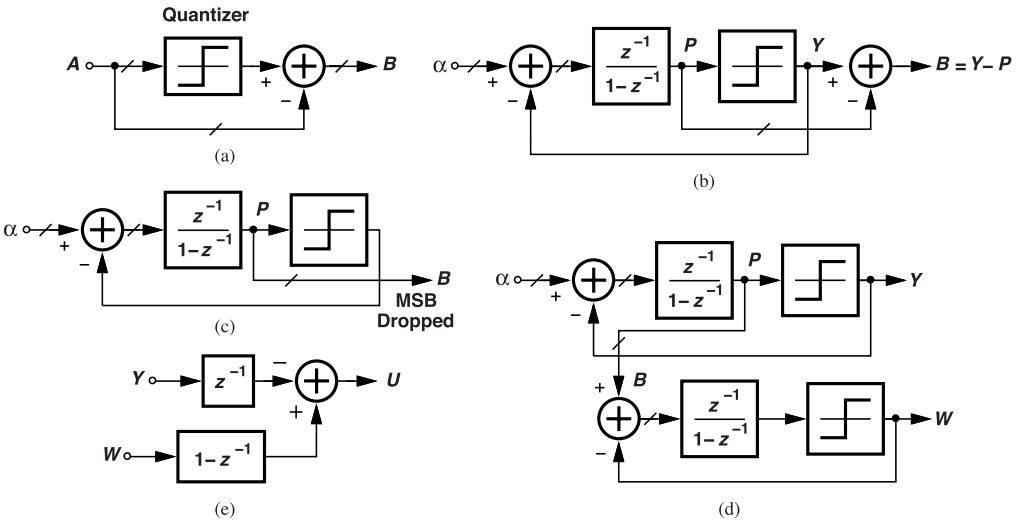


Figure 12.26 (a) Reconstruction of quantization noise, (b) use in a first-order loop, (c) taking the integrator output with MSB dropped, (d) using a second modulator to quantize B, and (e) combining Y and W to produce the final output.

noise. In reality, this subtractor is unnecessary because the difference between Y (the MSB) and P (the entire digital word) is simply equal to P with its MSB dropped [Fig. 12.26(c)].

The information provided by the reconstructed quantization noise can be utilized to improve the performance, but the difficulty is that this noise is expressed as a multibit digital signal. We proceed in two steps: (1) we convert the multibit B output to a single-bit stream by means of another $\Delta\Sigma$ modulator [Fig. 12.26(d)], and (2) we "combine" the result, W, with the output of the first loop, Y [Fig. 12.26(e)], so as to lower the quantization noise.

To determine the combiner's required function, we write, from Fig. 12.26(d), $W = z^{-1}B + (1 - z^{-1})Q_2$, where Q_2 denotes the unshaped noise contributed by the second quantizer. It follows that $W = z^{-1}Q_1 + (1 - z^{-1})Q_2$, and, from Fig. 12.26(b), $Y = (1 - z^{-1})Q_1$. (Since α is constant, we exclude it from Y.) How do we combine W and Y so as to eliminate Q_1? We multiply W by $1 - z^{-1}$ and Y by z^{-1} and subtract the results, arriving at

$$U = (1 - z^{-1})W - z^{-1}Y \tag{12.37}$$

$$= (1 - z^{-1})^2 Q_2. \tag{12.38}$$

The final output thus exhibits second-order noise shaping.

The cascaded topology in Fig. 12.26(d) is also called the "MASH" architecture and can be extended to higher orders with no concern for stability. This approach, however, generates a multibit output, requiring a multimodulus divider. We observe that U is a 2-bit signal if W and Y are 1-bit streams.

Example 12.13

As seen in this section, a multimodulus divider is necessary in two cases: if we utilize the two MSBs at the accumulator output (Fig. 12.23) or if we have a cascaded $\Delta\Sigma$ topology. Which case is preferable?

Solution

The former simply divides the first-order-shaped noise spectrum, $S_{\phi n}(f)$, by a factor of 4 whereas the latter raises the *order* of the shaping function, i.e., it creates a sharper response. The latter is therefore preferable.

Out-of-Band Noise As exemplified by Fig. 12.25, high-order $\Delta\Sigma$ modulators exhibit a high noise peak that may not be sufficiently suppressed as it travels around the synthesizer loop. Given by Eq. (12.35), the divider phase noise due to a second-order modulator begins from zero at $f = 0$ and reaches $[64/(M+\alpha)^2]S_q(f)$ at $f = f_{REF}/2$.

Example 12.14

Compare the phase noise at $f = f_{REF}/2$ in three cases: randomization but no shaping, a first-order modulator, and a second-order modulator. Assume the unshaped noise spectrum, $S_q(f)$, is relatively flat.

Solution

From Eq. (12.13), without shaping, we have $S_{\phi n}(f_{REF}/2) = [1/(M+\alpha)^2][4S_g(f)]$. For a first-order modulator, (12.29) yields a value of $[16/(M+\alpha)^2]S_q(f)$ at $f_{REF}/2$, the same as the first case. A second-order modulator, on the other hand, produces $S_{\phi n}(f_{REF}/2) = [64/(M+\alpha)^2]S_q(f)$. Figure 12.27 summarizes

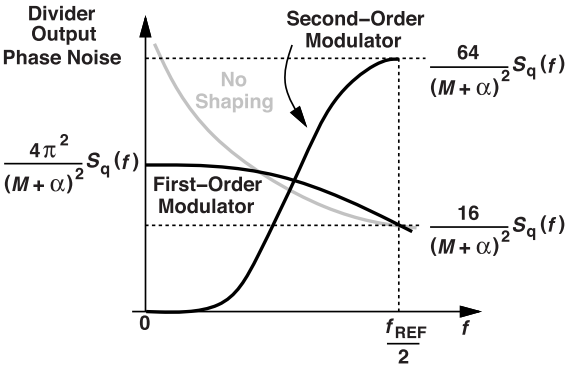


Figure 12.27 Summary of different noise profiles.

these results.

We can visualize the effect of the quantization noise peaking with the aid of Fig. 12.28, where the PLL

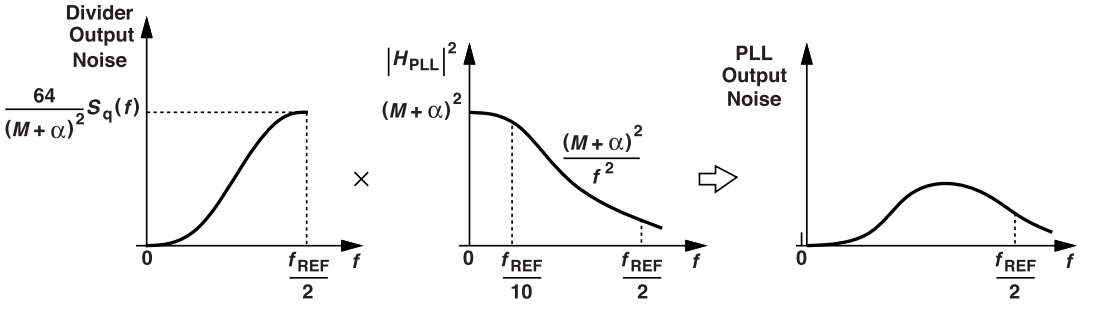


Figure 12.28 Phase noise peaking at PLL output due to $\Delta\Sigma$ noise.

bandwidth is assumed equal to $f_{REF}/10$. Here, Eq. (12.35) is multiplied by the squared magnitude of the PLL's low-pass response. At frequencies well above $f_{REF}/10$, $|H_{PLL}|^2$ can be approximated by a single pole and hence falls in proportion to $(M+\alpha)^2/f^2$.[2] The product of the $\Delta\Sigma$ noise and the PLL response is

[2]The PLL transfer function reduces to $2\zeta\omega_n(M+\alpha)/s$ at high frequencies.

zero at $f = 0$, rises to a maximum, and falls at high frequencies. If comparable to the VCO phase noise, this peak becomes problematic.

Example 12.15

Assuming a second-order type-II PLL with $\zeta = 1$ and a loop bandwidth of $2\zeta\omega_n \approx 2\pi f_{REF}/10$, estimate the output spectrum in Fig. 12.28 at $f_{REF}/2$.

Solution

The magnitude squared of the PLL transfer function derived in Chapter 7 reduces to approximately $0.04(M+\alpha)^2$ at $f = f_{REF}/2$. Also, as mentioned above, the $\Delta\Sigma$ phase noise spectrum is equal to $[64/(M+\alpha)^2]S_q(f)$ at this frequency. The PLL output spectrum is thus approximately equal to $2.6S_q(f)$ at $f = f_{REF}/2$. (But the peak occurs at lower frequencies.)

The foregoing studies suggest that the overall PLL output spectrum can exhibit peaking if a $\Delta\Sigma$ modulator of order 2 or higher is used. In particular, as shown in Fig. 12.29(a), the shape can change from a well-behaved

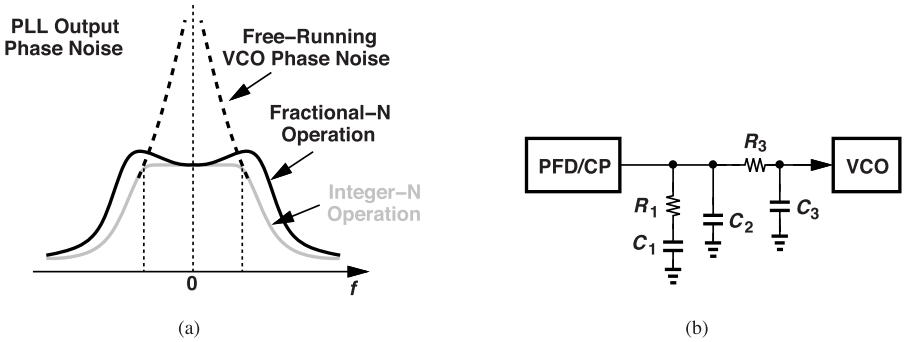


Figure 12.29 (a) Noise peaking at PLL output, and (b) use of additional filtering stage to reduce $\Delta\Sigma$ noise.

plateau in integer-N operation to a peaked response in fractional-N operation. To avoid this effect, the loop bandwidth is typically reduced to about $f_{REF}/20$ or lower. This is accomplished through the use of large loop filter capacitors. One can also add another filter section before the VCO so as to provide attenuation at $f_{REF}/2$ [Fig. 12.29(b)].

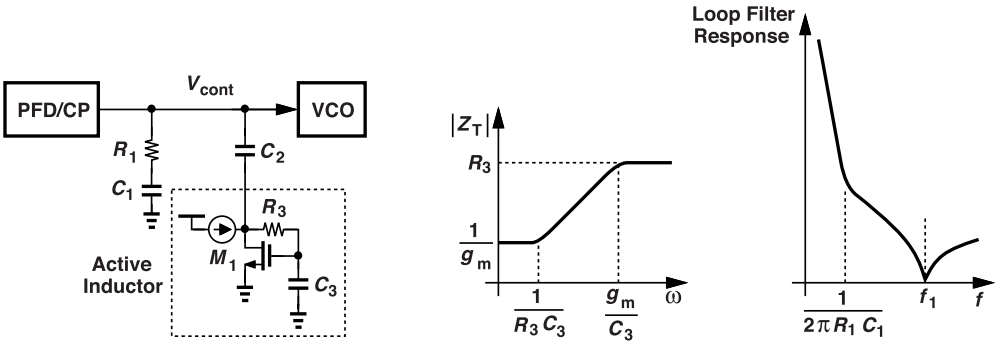
As a third possibility, we can insert before the VCO a notch filter. The filter should introduce a notch at the peak frequency in Fig. 12.28. We study two different implementations here. Illustrated in Fig. 12.30(a), the first approach is to construct a series-resonance circuit (a "trap") using a capacitor and an active inductor. The reader can show that the structure consisting of M_1, R_3, and C_3 exhibits an impedance given by

$$Z_T = \frac{R_3C_3s + 1}{g_m + C_3s} \tag{12.39}$$

if channel-length modulation is neglected. Sketched in Fig. 12.30(b), this impedance behaves inductively between $\omega_z = (R_3C_3)^{-1}$ and $\omega_p = g_m/C_3$. In fact, for $\omega_z \ll \omega \ll \omega_p$, we have $|Z_T| \approx R_3C_3\omega/g_m$ and hence an equivalent inductance of R_3C_3/g_m. This value and C_2 can be chosen for resonance at a desired frequency, f_1, leading to the overall filter response shown in Fig. 12.30(c).

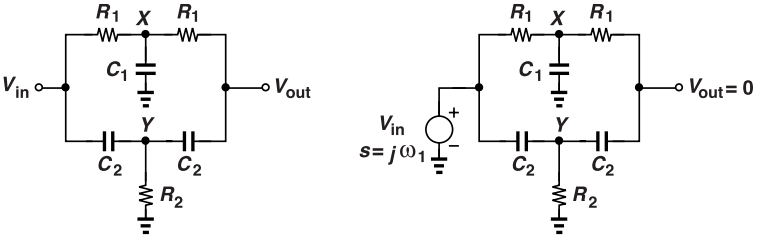
The notch filter of Fig. 12.30(a) entails two issues. First, the flicker noise of M_1 can contribute significant phase noise (Problem 12.17). Second, the nonlinearity of M_1 translates high-frequency quantization noise to low frequencies, as described in the next section.

Another notch filter implementation is depicted in Fig. 12.31(a). Called the "twin-T" notch filter, the circuit consists of two symmetric T sections in parallel. We wish to determine the resistor and capacitor values so as



(a) (b) (c)

Figure 12.30 (a) Use of active notch filter to reduce $\Delta\Sigma$ noise, (b) impedance of active inductor, (c) loop filter response.



(a) (b)

Figure 12.31 (a) Twin-T notch filter, and (b) setup for studying its frequency response.

to obtain a notch at $\omega = \omega_1$. That is, V_{out}/V_{in} must vanish at $s = j\omega_1$. At this frequency, the output voltage is zero and hence indistinguishable from the ground [Fig. 12.31(b)]. We can thus write

$$V_X = \frac{\dfrac{R_1}{R_1 C_1 s + 1}}{\dfrac{R_1}{R_1 C_1 s + 1} + R_1} V_{in} \tag{12.40}$$

$$= \frac{1}{R_1 C_1 s + 2} V_{in}, \tag{12.41}$$

and

$$V_Y = \frac{\dfrac{R_2}{R_2 C_2 s + 1}}{\dfrac{R_2}{R_2 C_2 s + 1} + \dfrac{1}{C_2 s}} V_{in} \tag{12.42}$$

$$= \frac{R_2 C_2 s}{2R_2 C_2 s + 1} V_{in}. \tag{12.43}$$

Since the currents flowing into the output node must be equal and opposite, we have $V_X/R_1 + V_Y C_2 s = 0$ and

$$-R_1 R_2 C_2^2 s^2 = \frac{2R_2 C_2 s + 1}{2\left(\dfrac{R_1 C_1 s}{2} + 1\right)}. \tag{12.44}$$

If we select $R_1C_1/2 = 2R_2C_2$, the right-hand side reduces to $1/2$ and, with $s = j\omega_1$, we obtain

$$\omega_1^2 = \frac{1}{2R_1R_2C_2^2}. \tag{12.45}$$

For minimum notch bandwidth, we choose $R_1 = 2R_2 = R$ and $C_1 = 2C_2 = 2C$, and

$$\omega_1 = \frac{1}{RC}. \tag{12.46}$$

The transfer function then reduces to

$$\frac{V_{out}}{V_{in}}(s) = \frac{R^2C^2s^2 + 1}{R^2C^2s^2 + 4RCs + 1}. \tag{12.47}$$

The twin-T network introduces a thermal noise density of $4kT(2R_1)$ at low frequencies.

Notch filters can also exhibit substantial phase shift within the loop bandwidth (Problem 12.18), degrading the loop stability. This issue proves serious if the synthesizer is designed for a wide bandwidth so as to suppress the VCO phase noise. The twin-T circuit, for example, suffers from a phase shift of about 25° at $\omega \approx 0.1\omega_1$.

The PVT dependence of RC in Eq. (12.46) proves problematic. As illustrated in Fig. 12.32, if the notch frequency shifts to $1/(R'C')$, then the rejection at the original ω_1 is limited to a. To deal with this issue,

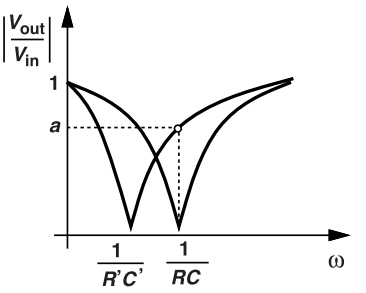


Figure 12.32 Shift of notch due to PVT.

we first observe that the notch frequency is independent of the source and load impedances connected to the twin-T network (why?). Now, we cascade three such sections, one providing the nominal notch frequency and two creating slightly higher and lower notch frequencies. Depicted in Fig. 12.33(a) is an example where

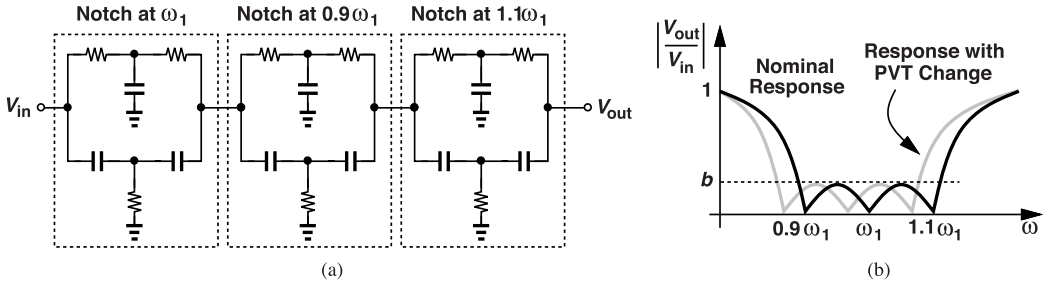


Figure 12.33 (a) Cascaded notches, and (b) their response in the presence of PVT variation.

the additional sections are skewed by 10%. This structure exhibits the nominal and shifted responses plotted in Fig. 12.33(b), guaranteeing a rejection factor of b at ω_1.

The cascaded networks in Fig. 12.33(a) suffer from two drawbacks: their thermal noise at low frequencies modulates the VCO, and their phase contributions affect the PLL stability.

12.4 Nonlinearities in Fractional-N Loops

The quantization noise peaking around $f_{REF}/2$ in higher-order $\Delta\Sigma$ modulators "folds down" to low frequencies if it travels through nonlinear stages (Fig. 12.34). As an example, suppose a circuit exhibits an input-output characteristic given by $y(t) = \beta_1 x(t) + \beta_2 x^2(t)$. We wish to determine the spectrum of $x^2(t)$ if

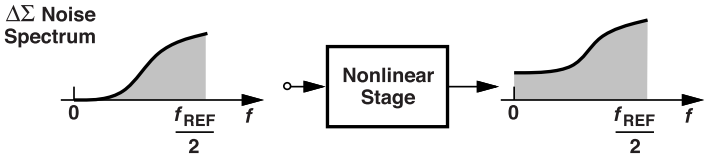


Figure 12.34 Noise folding due to nonlinearity.

$x(t)$ has a high-pass spectrum. The product $x(t) \times x(t)$ in the time domain (a mixing action) corresponds to convolution in the frequency domain. We begin with a simple case where $x(t)$ contains only two sinusoids at f_1 and f_2 [Fig. 12.35(a)], noting that $x^2(t)$ exhibits sinusoids at $f_2 \pm f_1$. Now, we extend this effect to noise:

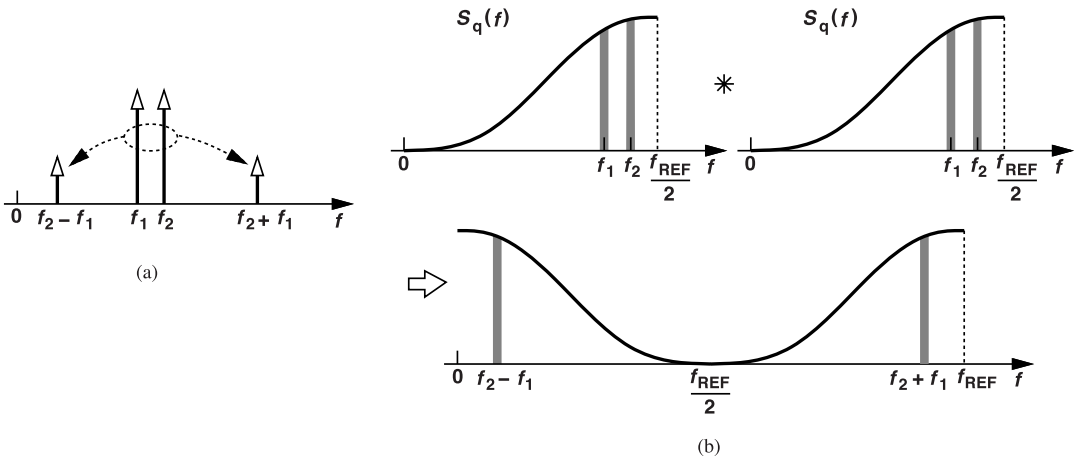


Figure 12.35 (a) Mixing of two components at f_1 and f_2, and (b) noise folding due to mixing.

as illustrated in Fig. 12.35(b), components around f_1 are mixed with those around f_2 and generate energy around $f_2 + f_1$ and $f_2 - f_1$. We say high-frequency noise is folded down. The spectrum of $x^2(t)$ therefore contains high levels of noise at low frequencies, defeating the purpose of noise shaping. The folding can occur for the tones described in Section 12.3.4 as well. We must therefore pay attention to all nonlinearities in the $\Delta\Sigma$ noise path.

Example 12.16

The control voltage in a fractional-N synthesizer contains large $\Delta\Sigma$ noise. Explain how the noise can fold down.

Solution

If the VCO characteristic is nonlinear (Fig. 12.36), then folding occurs. We can model the VCO as a cascade of a nonlinear function, $g(V)$, and an oscillator with linear control. That is, the high $\Delta\Sigma$ noise near $f_{REF}/2$ is translated to low frequencies by $g(V)$ and modulates the VCO, raising the close-in phase noise. Note that this issue is absent in integer-N loops.

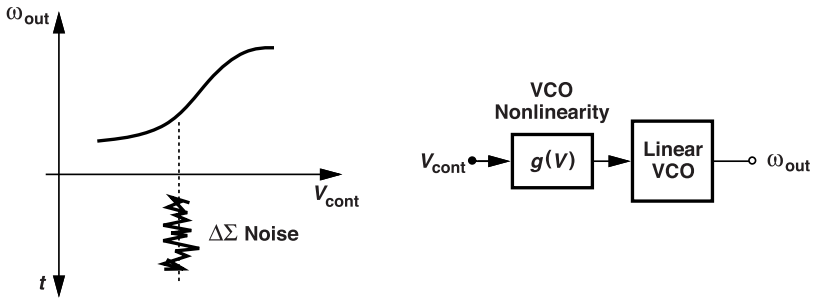


Figure 12.36 Noise folding due to VCO nonlinearity.

12.4.1 Charge Pump Nonlinearity

One particularly problematic source of nonlinearity is the up/down current mismatch within the charge pump. The following analysis shows how this mismatch gives rise to nonlinearity.

Consider the PFD/CP/filter cascade shown in Fig. 12.37(a), where $I_1 = I_2 - \Delta I$. The difference, ΔI, arises

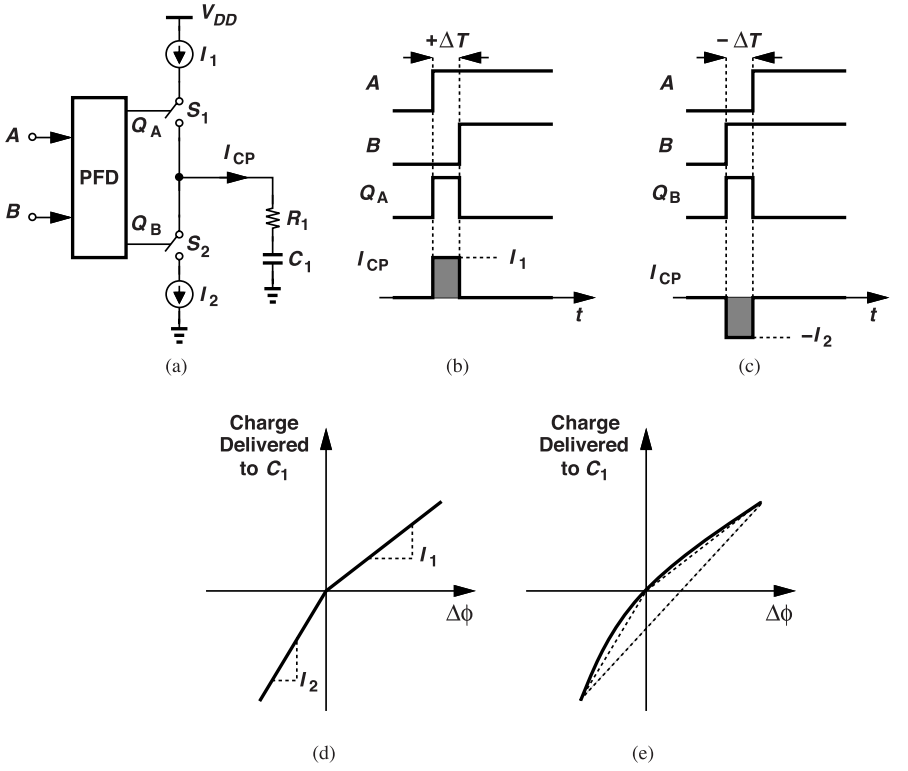


Figure 12.37 (a) PFD/CP sensing a finite phase error, (b) case of phase error $= +\Delta T$, (c) case of phase error $= -\Delta T$, (d) resulting characteristic with up and down current mismatch, and (e) approximation of the characteristic by a parabola.

from both random mismatches and channel-length modulation. Let us study two cases: the phase difference between A and B is $+\Delta T$ or $-\Delta T$. Depicted in Fig. 12.37(b), the former leads to charge delivered by I_1 to C_1 in the amount of $I_1 \Delta T$. Similarly, as illustrated in Fig. 12.37(c), the latter produces a charge packet equal to $I_2 \Delta T$ that is drawn from C_1. Plotting the charge deposited on C_1 as a function of the phase difference [Fig. 12.37(d)], we observe different slopes in the left and right half planes—and hence nonlinearity. This

point becomes clearer if we approximate this characteristic by a smooth curve [Fig. 12.37(e)] and recall from Fig. 12.34 that such nonlinearity folds the $\Delta\Sigma$ noise down.

In a fractional-N synthesizer, the divider produces large, random positive and negative phase fluctuations, e.g., by as much as three VCO cycles (Example 12.6). This causes positive and negative phase jumps at the PFD input, forcing the charge pump to operate, randomly, with two different gains. It can be shown [4] that the ratio of the folded noise floor to the peak is equal to $\Delta I/(4I_p)$, where $I_p = (I_1 + I_2)/2$.

The CP mismatches can be suppressed by means of the techniques described in Chapter 7. In addition, another approach proves useful here: we can ensure that the phase error seen by the PFD does not change *sign*, i.e., it remains in either of the half planes in Fig. 12.37(d). This is accomplished by deliberately introducing a skew in one of the PFD reset paths [5] [Fig. 12.38(a)]. We note that the bottom flipflop takes T_D seconds

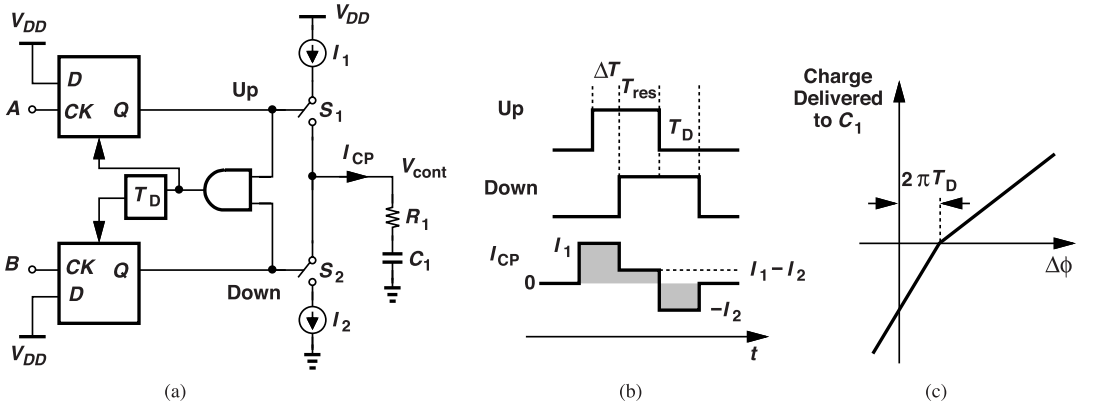


Figure 12.38 (a) PFD/CP with built-in skew, (b) circuit's waveforms, and (c) effect on overall characteristic.

longer to reset than does the top one. Owing to this difference, the loop locks with a finite phase error, ΔT, such that the net charge delivered to the loop filter is zero [Fig. 12.38(b)]. The reader can prove that $\Delta T \approx T_D$. Consequently, the characteristic is shifted to the right by approximately $2\pi T_D$ radians [Fig. 12.38(c)], i.e., the loop attempts to lock with this phase offset so that V_{cont} has no net change per input cycle. If the peak divider output phase fluctuation remains less than this value, the slope is constant and the CP linear. For example, if the divider output phase has a maximum deviation of $\pm 2T_{VCO}$, we choose $T_D > 2T_{VCO}$.

The principal difficulty with the foregoing solution is that the extra charge delivered to the loop filter introduces a large ripple, specifically a peak voltage disturbance of $I_1 \Delta T/C_1$. Note that the ripple in our previous studies was proportional to only the *mismatch* between the up and down currents, whereas it is proportional to I_1 here.

Example 12.17

The peak divider output phase fluctuation is equal to $3T_{VCO}$ for a certain $\Delta\Sigma$ modulator. If the gate delay of a 40-nm process is equal to 6 ps in the FF, 0 °C corner, determine how many gates are necessary for T_D in Fig. 12.38(a). Assume a 5-GHz synthesizer.

Solution

With $T_{VCO} = 200$ ps, we require $T_D = 3T_{VCO} = 600$ ps and hence 100 gates. The phase noise of these inverters modulates the down pulsewidth and can be problematic. Also, in the SS, 75 °C corner, T_D rises to about 1.2 ns, increasing both the phase noise and the ripple amplitude.

In order to remove the large ripple generated in the above circuit, we can employ the sampling loop filter described in Chapter 8, with the sampling switch turned on after the transient on V_{cont} has subsided.

Example 12.18

In a sampling circuit, the switch channel charge varies with the input level according to $Q_{ch} = WLC_{ox}(V_{GS} - V_{TH})$, potentially introducing nonlinearity. Does this phenomenon fold high-frequency $\Delta\Sigma$ noise in a sampling loop filter?

Solution

It does, but negligibly. This is because the ripple at the sampling switch's source and drain is small, and so is the change in the channel charge of the switch. From another perspective, this switch nonlinearity is similar to the VCO control nonlinearity, playing a negligible role if the $\Delta\Sigma$ quantization noise swing in the control path is small.

12.4.2 Charge Pump Settling Behavior

Charge pumps also suffer from *dynamic* nonlinearities, exacerbating the problem of noise folding. Let us first consider operation in the integer-N mode. The PFD produces output pulses that are approximately five gate delays wide [Fig. 12.39(a)]. The charge pump current sources are on for only part of this period because

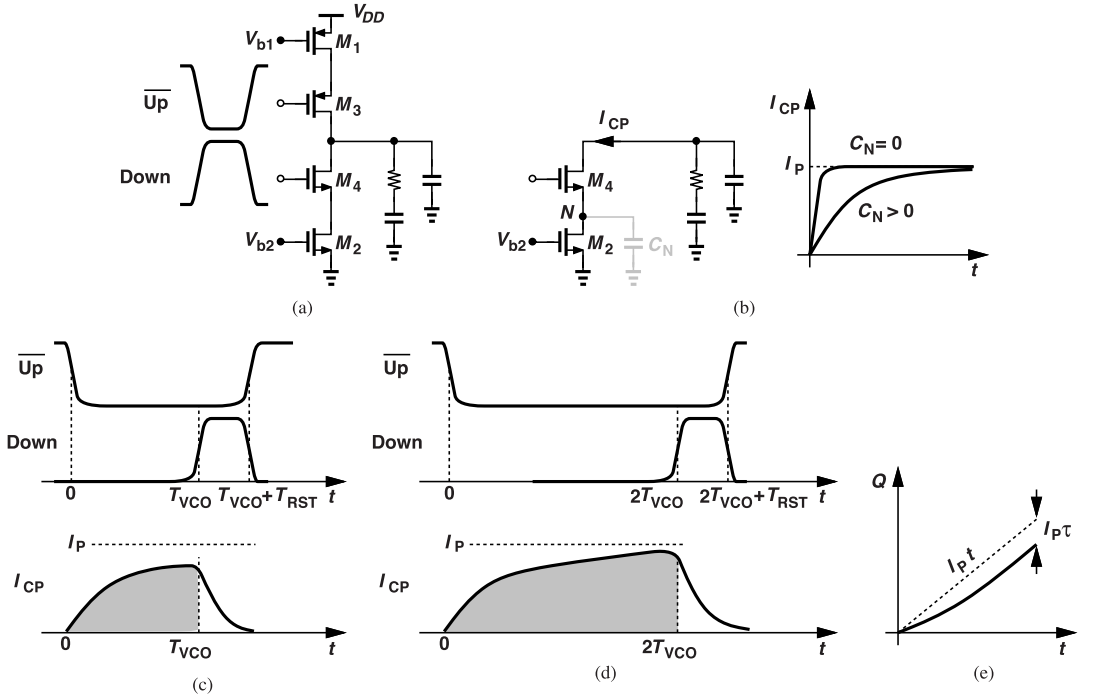


Figure 12.39 (a) CP sensing inputs, (b) settling behavior of output current, (c) settling for a phase error of T_{VCO}, (d) settling for a phase error of $2T_{VCO}$, and (e) output charge as a function of time.

of the finite rise and fall times and the intrinsic capacitances. For example, as shown in Fig. 12.39(b), the capacitance at node N slows down the settling of the charge pump current with a time constant, τ, given by $R_{on4}C_N$. In other words, it is possible that the current does not sufficiently settle before the CP turns off.

Now, suppose the circuit operates in the fractional-N mode, sensing a phase error that can assume a value of $\pm T_{VCO}$, $\pm 2T_{VCO}$, etc. The question is, does the charge delivered to the loop filter scale linearly as the phase error goes from zero to $+T_{VCO}$ to $+2T_{VCO}$, etc.? As an example, the PFD waveforms in Fig. 12.39(c) show that, due to an input phase error of T_{VCO}, the up current turns on T_{VCO} seconds before the down

pulse arrives. Consequently, the charge pump current, I_{CP}, rises from zero, reaches a certain value in T_{VCO} seconds, and falls back to zero when the down pulse is asserted. We observe two nonlinear mechanisms here. First, since I_{CP} does not settle to I_p instantaneously, the charge delivered to the filter is less than $I_p \cdot T_{VCO}$. If the phase error now changes to $2T_{VCO}$, giving the CP more time to settle, the total charge is closer to $I_p \cdot (2T_{VCO})$ [Fig. 12.39(d)]. In fact, approximating the settling behavior by a single time constant, τ, we can write $I_{CP} = I_p[1 - \exp(-t/\tau)]$ and the area under it as

$$Q = I_p t - \tau I_p \left(1 - \exp \frac{-t}{\tau}\right). \tag{12.48}$$

Plotted in Fig. 12.39(e), this result indicates that doubling the time does not double the charge, i.e., the exponential term introduces nonlinearity. Thus, τ must be small enough to minimize this error.

The second issue relates to the reset time, during which one current source that is already settled must cancel another that has just turned on and begun to settle. As depicted in Fig. 12.40, the amount of cancellation partially depends on how much the up current has settled, i.e., how many T_{VCO}'s have preceded the reset

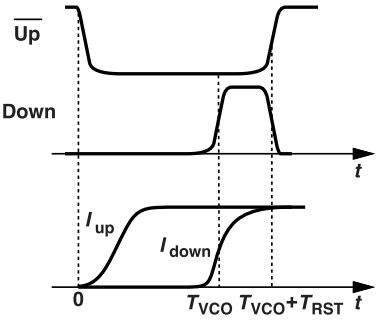


Figure 12.40 Effect of finite settling time during PFD reset.

period. We therefore conclude that this cancellation varies from one reference cycle to another because the phase error randomly jumps from zero to $\pm T_{VCO}$, $\pm 2T_{VCO}$, etc. To minimize this nonlinearity, we must reduce the reset time.

12.5 Reduction of Quantization Noise

Our study of $\Sigma\Delta$ modulators indicates that their low-frequency $\Delta\Sigma$ quantization noise can be reduced only at the cost of a higher peak at $f_{REF}/2$. We must therefore seek other methods of suppressing this noise in a synthesizer environment. In this section, we consider two.

12.5.1 DAC Feedforward

As explained in Section 12.3.5, the $\Delta\Sigma$ noise can be reconstructed. Given that this noise also appears at the divider output and propagates through the PFD and the CP, let us negate it and inject it into the loop filter (Fig. 12.41). Our hope is that the quantization noise components arriving at the filter through the two different paths cancel each other.

The conceptual approach shown in Fig. 12.41 must deal with a number of issues. (1) The reconstructed noise provided by the $\Delta\Sigma$ modulator is a digital quantity whereas the loop filter holds an analog value. We must therefore insert a digital-to-analog converter in this path. (2) The divider output contains phase noise whereas the reconstructed quantization noise represents frequency noise. Thus, the second path must incorporate an integrator, preferably in the digital domain. (3) If we reconstruct the noise by subtracting the input and output of the quantizer, we obtain the unshaped noise while the main $\Delta\Sigma$ modulator output contains shaped noise. For this reason, the noise reconstruction must be performed differently.

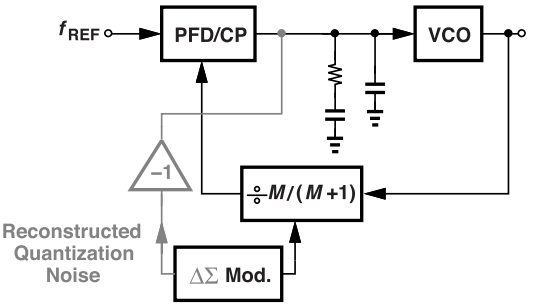
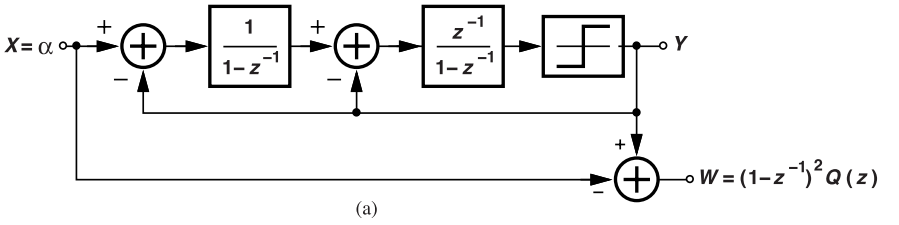
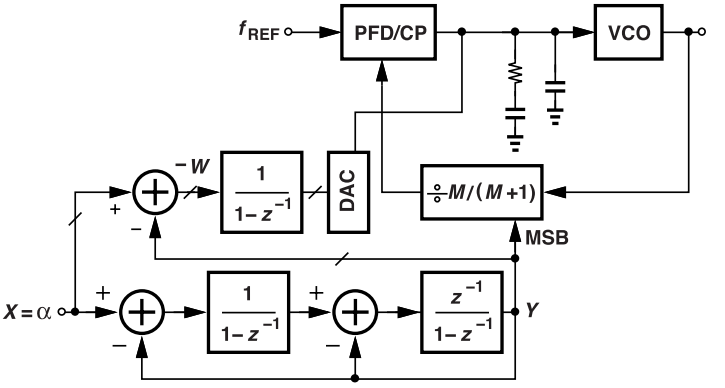


Figure 12.41 Conceptual illustration of quantization noise cancellation by feedforward.

Let us begin with the last issue. We return to Fig. 12.24(c) and ask how the shaped quantization noise, $(1-z^{-1})^2Q(z)$, can be reconstructed. Noting that $Y(z) = z^{-1}X(z) + (1-z^{-1})^2Q(z)$, we delay X by one clock cycle and subtract it from Y to obtain $(1-z^{-1})^2Q(z)$. In fact, since X is constant in the steady state, the delay is not necessary, and the implementation reduces to that shown in Fig. 12.42(a), where W denotes



(a)



(b)

Figure 12.42 (a) Construction of shaped quantization noise, and (b) fractional-N synthesizer with feedforward cancellation.

the shaped quantization noise. The reader is encouraged to repeat this exercise for a third-order modulator.

To duplicate the frequency-to-phase conversion of the divider, we now subject W to integration by multiplying it by $(1-z^{-1})^{-1}$. As illustrated in Fig. 12.42(b), this result is then converted to analog form and injected into the loop filter. The DAC output is naturally chosen to be a current quantity to match the CP output. Note that we subtract Y from X to generate $-W$ and hence create the negation necessary in the second path.

The "DAC feedforward" architecture of Fig. 12.42(b) presents other interesting challenges. First, if α is, for example, a 16-bit digital quantity, then $-W$ has at least 16 bits. Does this mean that the DAC must also

have such a a high resolution?! No, we need not cancel the $\Delta\Sigma$ noise with such high accuracy. In other words, we can simply convert the first 5 or 6 MSBs of $-W$ to an analog current, expecting proportional noise reduction. However, the truncation of $-W$ is equivalent to subjecting it to a sharp nonlinearity (a limiting function), causing the high-frequency noise to fold down.

We therefore seek a system that can approximate $-W$ with a lower resolution but with no folding. For example, we wish to convert 16 bits to 6 bits. We have, in fact, dealt with a similar situation previously: the $\Delta\Sigma$ modulators studied before represent α by a 1-bit output. In this spirit, let us apply $-W$ to a "multibit" modulator (Fig. 12.43) [6], where, for example, the 6 MSBs participate in the feedback loop. In this case,

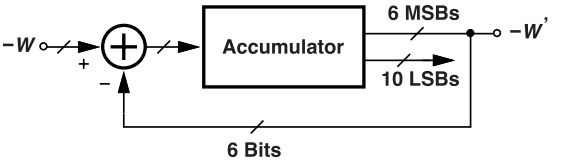


Figure 12.43 $\Delta\Sigma$ modulator using 6 MSBs.

discarding the 10 LSBs introduces quantization noise, which is then shaped by the loop, i.e., $-W'$ acts as a coarse approximation of $-W$ but with shaped quantization noise. Figure 12.44 depicts the final architecture.

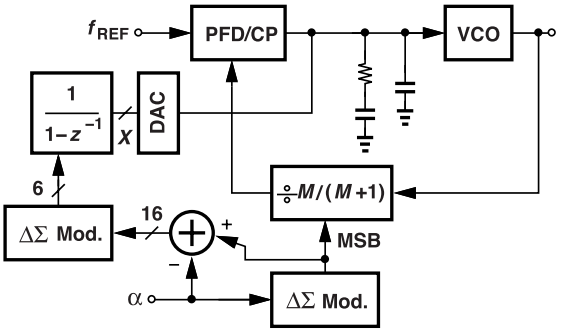


Figure 12.44 Overall fractional-N synthesizer with feedforward cancellation.

While relaxing the resolution required of the DAC, the above approach still demands a high DAC linearity. If, for example, the current sources comprising the DAC suffer from mismatches, its characteristic becomes nonlinear (Fig. 12.45), folding the high-frequency $\Delta\Sigma$ noise [6].

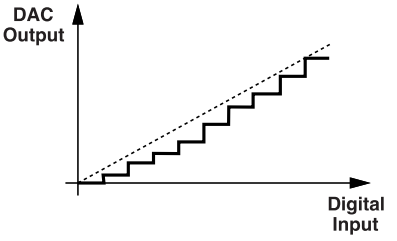


Figure 12.45 DAC nonlinearity example.

Even with a perfectly linear DAC, the $\Delta\Sigma$ noise cancellation in Fig. 12.44 is ultimately limited by the mismatch between the CP and DAC outputs. The problem is particularly difficult because the two outputs have different peak values and different durations. As illustrated in Fig. 12.46, the CP output has a constant height and a width equal to the divider output phase jump, whereas the DAC output has a variable height

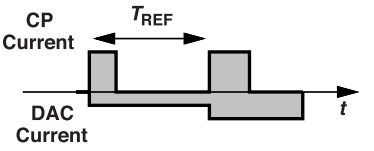


Figure 12.46 Illustration of CP and DAC pulses.

and a width equal to T_{REF}. The areas under these pulses suffer from two mismatches: (1) the CP and DAC currents originate from the same (bandgap) reference but inevitably have random mismatches, and (2) it is difficult to match the CP's channel-length modulation with that of the DAC.

In view of the difficulties associated with DAC feedforward, many designers prefer to simply reduce the synthesizer loop bandwidth instead. In such a case, the VCO phase noise is not suppressed much, requiring a higher power consumption. Nonetheless, the DAC feedforward idea leads to a simpler approach involving digital-to-time converters (DTCs), as described in the next section.

12.5.2 Noise Cancellation by DTC

The DAC feedforward method described in the previous section cancels the $\Delta\Sigma$ noise in the current domain. We might ask whether it is possible to perform the cancellation in the time domain. Specifically, since the phase jumps at the divider output are known ($\pm T_{VCO}$, $\pm 2T_{VCO}$, etc.), can we reproduce these jumps in the *reference* waveform so that the phase error presented to the PFD is free from the $\Delta\Sigma$ noise? Figure 12.47(a) conceptually illustrates the idea. We must determine how the block with a question mark in it is implemented. This block must introduce phase jumps in the reference path equal to those at the divider output.

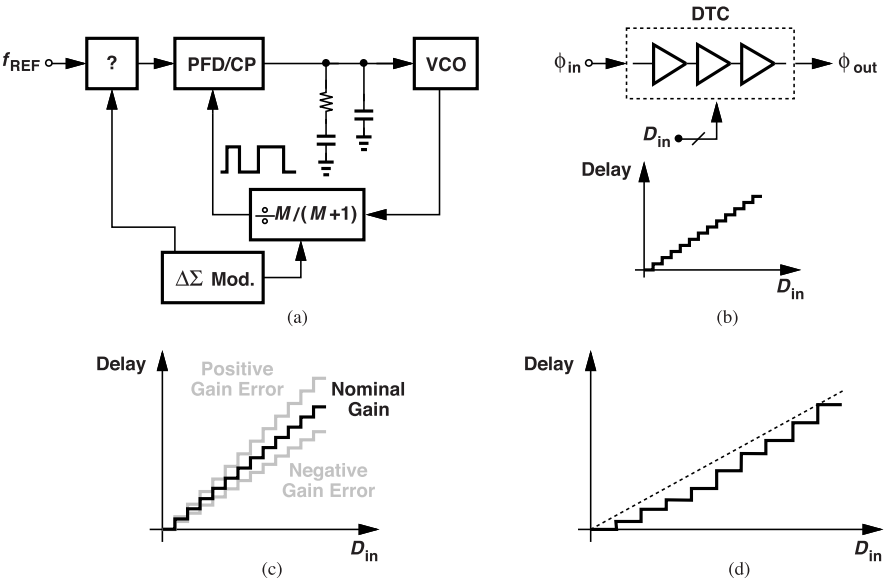


Figure 12.47 (a) Conceptual illustration of $\Delta\Sigma$ noise cancellation in the reference path, (b) high-level diagram of a DTC, (c) problem of DTC gain error, and (d) problem of DTC nonlinearity.

The $\Delta\Sigma$ modulator can provide a digital representation of the phase jumps (similar to X in Fig. 12.44). We must convert this information to the analog domain. To add or subtract phase in the reference path, we can employ a variable-delay line. Driven by a digital control code, such a line is called a "digital-to-time converter" (DTC). As shown in Fig. 12.47(b), the DTC's phase shift from the input to the output can

change in discrete steps according to the value of D_{in}. We recognize that this approach is simpler than DAC feedforward.

The design of the DTC entails two issues. First, the DTC gain error translates to undercompensation or overcompensation of the $\Delta\Sigma$ noise. As exemplified by Fig. 12.47(c), if the gain is lower than the nominal value, the DTC does not produce sufficiently wide phase jumps, yielding some residual noise at the PFD/CP output. Second, nonlinearity in the DTC characteristic [Fig. 12.47(d)] folds the high-frequency $\Delta\Sigma$ noise down. The latter issue is particularly problematic because calibration techniques to reduce the nonlinearity tend to be quite complex.

12.5.3 Reference Frequency Doubling

Our study of PLLs in this book suggests that the performance is generally improved if the reference frequency can be raised.[3] For example, a wider loop bandwidth can be accommodated, reducing the VCO phase. In the case of fractional-N loops, the $\Delta\Sigma$ modulator can be clocked faster so as to lower its in-band noise and shift its noise peak to higher frequencies. The former property is evident from Eq. (12.36). We therefore wish to multiply the reference frequency before it reaches the PFD [3], but without the use of a PLL or a DLL (why?).

In order to double f_{REF}, one can simply delay the waveform and XOR it with itself [Fig. 12.48(a)]. The delay must be long enough to yield an output pulsewidth with which the PFD can operate. However,

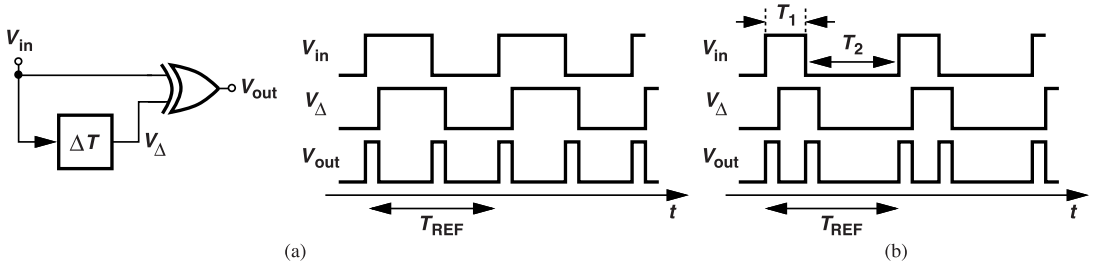


Figure 12.48 (a) Frequency doubler implementation, and (b) effect of duty cycle error.

if V_{in} does not have a 50% duty cycle, the output pulses are not evenly spaced [Fig. 12.48(b)], leading to unwanted frequency components. To formulate this effect, we first recognize that the output waveform in Fig. 12.48(b) in fact repeats every T_{REF} seconds rather than every $T_{REF}/2$ seconds. That is, V_{out} exhibits a first, weak harmonic at f_{REF}, a second, strong harmonic at $2f_{REF}$, etc. We view the first harmonic as a spurious frequency.

Example 12.19

For the waveforms shown in Fig. 12.48(b), determine the relative magnitude of the output component at f_{REF} if $T_1 = (T_{REF}/2) - \epsilon$.

Solution

Let us decompose the output into two half-rate waveforms as shown in Fig. 12.49. For the sake of simplicity, we approximate the pulses by impulses and write:

$$V_{out}(t) = V_{out1}(t) + V_{out2}(t) \tag{12.49}$$

$$= A \sum_{k=-\infty}^{+\infty} [\delta(t - kT_{REF}) + \delta(t - kT_{REF} - T_1)], \tag{12.50}$$

[3] Our assumption is that the reference phase noise is relatively independent of the crystal frequency.

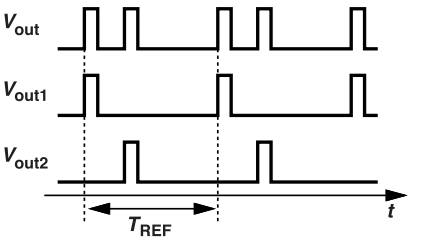


Figure 12.49 Decomposition of V_{out} into two periodic waveforms.

where A denotes the weight of each impulse. The Fourier transform of this signal is given by

$$V_{out}(f) = \frac{A}{T_{REF}} \sum_{k=-\infty}^{+\infty} \delta(f - kf_{REF})(1 + e^{-j2\pi f T_1}) \tag{12.51}$$

$$= \frac{2A}{T_{REF}} e^{-j\pi f T_1} \cos \pi f T_1 \sum_{k=-\infty}^{+\infty} \delta(f - kf_{REF}). \tag{12.52}$$

If $T_1 = T_{REF}/2$, then the envelope, $\cos \pi f T_1$, falls to zero at $f = f_{REF}$ and V_{out} does not contain the fundamental. On the other hand, if $T_1 = (T_{REF}/2) - \epsilon$, we have $\cos \pi f T_1 \approx \cos(\pi f T_{REF}/2) + \pi \epsilon f \sin(\pi f T_{REF}/2)$; i.e., the component at f_{REF} has a normalized amplitude equal to $\pi \epsilon f_{REF}$. This is a good rule of thumb to remember. For example, if $\epsilon = 0.01 T_{REF}$, then the spur has a magnitude of -30 dB with respect to the component at $2f_{REF}$.

The foregoing example suggests that the input duty cycle must be tightly controlled. In reality, the situation is much more difficult: the spur at f_{REF} is "amplified" by the loop's frequency multiplication factor as it reaches the VCO output, but it is also filtered by the loop. This means that the loop bandwidth must be drastically reduced to suppress the spur. As an example, suppose we double a 20-MHz reference and apply the result to a 5-GHz synthesizer. If $\epsilon = 0.01 T_{REF}$ and we wish to have a spur level of -60 dBc at the final output, then the loop must attenuate it by $20 \log(5,000/40) + 60 - 30 = 72$ dB. Approximating the closed-loop frequency response by a first-order roll-off, we conclude that the bandwidth must be decreased to roughly 20 MHz/$10^{72/20} = 5$ kHz.

In practice, our strategy is to add more poles to the loop filter so as to sharpen the roll-off and hence avoid such a dramatic decrease in the bandwidth. For example, a second pole at 1 MHz provides an additional suppression factor of 20 for a 20-MHz spur, allowing the bandwidth to increase to 100 kHz.

This example underscores the importance of duty cycle control in the reference path. The crystal oscillator itself should be designed for a nominal 50% duty cycle, and duty cycle correction (Section 11.10) should be performed on this oscillator's output. In some stringent applications, however, the phase noise contributed by the duty cycle correction circuit becomes critical.

In addition to duty cycle control, we can also insert a narrowband notch filter in the oscillator control path so as to attenuate the spur at f_{REF}. The topologies described in Section 12.3.5 prove useful here.

References

[1]. R. Schreier and G. C. Temes, *Understanding Delta-Sigma Data Converters,* Wiley, 2004.

[2]. T. A. D. Riley, M. A. Copeland, and T. A. Kwasniewski, "Delta-Sigma modulation in fractional-N frequency synthesis," *IEEE J. Solid-State Circuits,* vol. 28, pp. 553-559, May 1993.

[3]. H. Huh et al, "A CMOS dual-band fractional-N synthesizer with reference doubler and compensated charge Pump," *ISSCC Dig. Tech. Papers,* pp. 186-187, Feb. 2004.

[4]. B. Razavi, "An alternative analysis of noise folding in fractional-N synthesizers," *Proc. ISCAS*, May 2018.

[5]. S. E. Meninger and M. H. Perrott, "A 1-MHz bandwidth 3.6-GHz 0.18-μm CMOS fractional-N synthesizer utilizing a hybrid PFD/DAC structure for reduced broadband phase noise," *IEEE J. Solid-State Circuits,* vol. 41, pp.966-981, April 2006.

[6]. S. Pamarti, L. Jansson, and I. Galton, "A wideband 2.4 GHz Delta-Sigma fractional-N PLL with 1 Mb/s in-loop modulation," *IEEE J. Solid State Circuits,* vol. 39, pp. 49-62, Jan. 2004.

Problems

12.1. An RF synthesizer is designed for the 5-GHz WiFi band. If $f_{REF} = 20$ MHz and $K_{VCO} = 300$ MHz/V, determine the loop filter component values. Assume $I_p = 1$ mA.

12.2. Explain what happens to f_{out} if the divide ratio in Fig. 12.6(a) toggles between M and $M + 2$.

12.3. For the $g(t)$ waveform in Fig. 12.9(b), plot the divider phase noise as a function of time from Eq. (12.10).

12.4. If we halve f_{REF} in Fig. 12.7 and change the divide ratios to $2M$ and $2M + 1$, explain what happens to Eqs. (12.10) and (12.13).

12.5. Suppose we add an integrator in the feedback path of Fig. 12.13(b). What happens to Y/Q and Y/X?

12.6. Repeat the previous problem if the integrator is added between the input subtractor and the first integrator.

12.7. Suppose we add a constant value to Q in Fig. 12.15(f). Explain how the feedback loop counteracts this change.

12.8. If Q in Fig. 12.15(f) has a zero average, what are the average values of the subtractor's output and the digital integrator's output?

12.9. If A has a constant value in Fig. 12.16(a), plot B and E as a function of time.

12.10. We have explained intuitively why $1 - z^{-1}$ represents differentiation in the time domain. What can we say about $1 - z^{-2}$?

12.11. If we consider z^{-1} as a one-clock-cycle delay function, we can write its continuous-time equivalent as $\exp(-sT_{CK})$. Does this agree with our derivations leading to Eq. (12.25)?

12.12. Do the spectra in Fig. 12.19(b) change if the delaying integrator in Fig. 12.19(a) is replaced with a nondelaying integrator?

12.13. Plot the output of the subtractor in Fig. 12.22(a) as a function of time for $\alpha = 0.1$ and $\alpha = 0.2$.

12.14. Consider the second-order $\Delta\Sigma$ modulator shown in Fig. 12.24(b). Replacing the quantizer with a subtractor [as in Fig. 12.19(a)] and setting α to zero, determine the noise transfer function, Y/Q.

12.15. What happens in Fig. 12.24(c) if both integrators employ a nondelaying topology?

12.16. Derive the noise transfer functions for both quantizers in Fig. 12.26(d) if the integrators employ a nondelaying topology.

12.17. Modeling the noise of M_1 in Fig. 12.30(a) by a voltage source, $\overline{V_n^2}$, in series with its gate, calculate the resulting noise voltage in V_{cont}. Is the response from V_n to the VCO output phase a low-pass, band-pass, or high-pass function?

12.18. Compute the phase shift from V_{in} to V_{out} in Fig. 12.31(a) if $R_1 = 2R_2 = 2R$ and $C_1 = 2C_2 = 2C$. How much is the phase shift at one-tenth of the notch frequency?

12.19. Determine the output noise voltage of the notch filter in Fig. 12.31(a) at the notch frequency if $R_1 = 2R_2 = 2R$ and $C_1 = 2C_2 = 2C$.

12.20. Suppose τ in Eq. (12.48) is approximately equal to T_{VCO}. Calculate the error in Q with respect to a straight line at $t = T_{VCO}$ and $t = 2T_{VCO}$.

13

Clock and Data Recovery Fundamentals

In addition to the applications described in the previous chapters, phase locking also finds usage in clock and data recovery (CDR). CDR circuits are necessary when we receive asynchronous data, i.e., a stream of random bits or symbols with no accompanying clock. In this chapter, we study the basic concepts in clock and data recovery and introduce various architectures for realizing this function.

We should make a remark about notation. In data communication circuits, the letter Q has four different meanings: the quality factor, the quadrature phase of a waveform, the output of a latch or a flipflop, and the error function (the integral of a Gaussian distribution). The proper meaning should be clear from the context.

13.1 General Considerations

Software companies such as Google and Facebook store and provide data through the use of "servers." Their "data centers" (also known as "server farms") contain thousands of such servers that must be connected to one another for storing and routing data. As illustrated in Fig. 13.1(a), the random data propagates through a "copper" medium (a cable or a trace on a printed-circuit board) or an optical medium (an optical fiber) as it

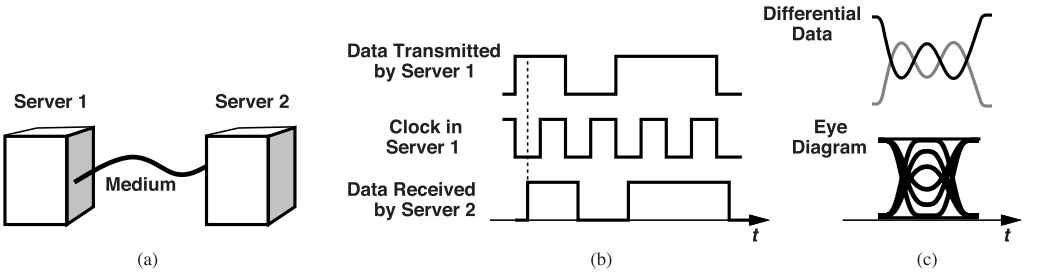


Figure 13.1 (a) Communication link between two servers, (b) transmitted and received waveforms, and (c) actual received eye diagram.

goes from one server to another. Since, in most applications, only the data—and not the clock—is received (why?), the timing information is lost [Fig. 13.1(b)]. Note that the clock contains periodic transitions but the data does not. Owing to the medium imperfections, the actual data also experiences dispersion and jitter [Fig. 13.1(c)]. We must therefore "recover" the clock from the data and then use this clock to retime and "clean up" the data.

The foregoing observations lead to the conceptual CDR function depicted in Fig. 13.2(a). The received data, D_{in}, drives a clock recovery circuit, which generates a periodic clock waveform, CK. A retimer (e.g., a flipflop) then receives this clock along with D_{in}, and recovers the data. In practice, an "equalizer" precedes this arrangement to compensate for the medium's imperfections.

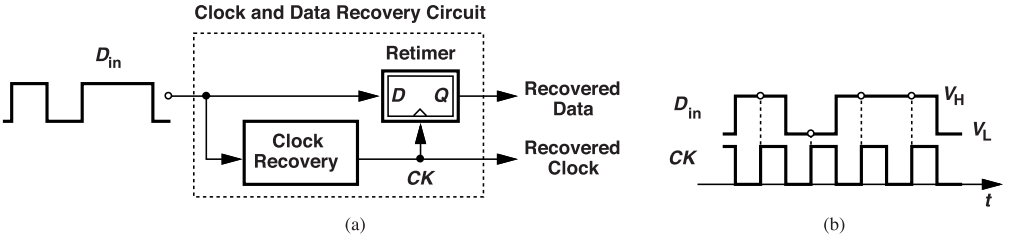


Figure 13.2 (a) Conceptual clock and data recovery system, and (b) its waveforms.

Let us ponder the operation of the clock recovery circuit in Fig. 13.2(a). For the recovered clock to be able to retime the data, it must have (1) a frequency equal to the data rate, e.g., 10 GHz for a 10-Gb/s random binary stream, and (2) a proper phase so that it samples the data at its highest and lowest values [Fig. 13.2(b)], i.e., farthest from the data transitions. We say the clock samples the data in the middle of the data eye.

Example 13.1

A D flipflop acts as a data retimer. Explain the effect of retiming on the data jitter.

Solution

As illustrated in Fig. 13.3, the clock samples D_{in} away from the transitions, thereby masking the jitter. That

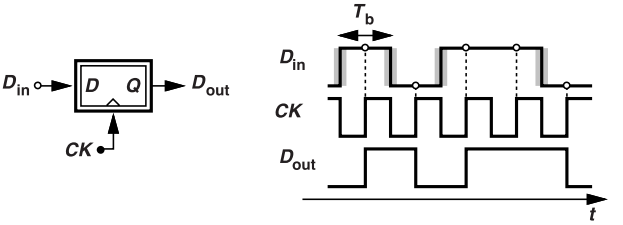


Figure 13.3 Effect of jitter at flipflop input.

is, D_{out} inherits no jitter from D_{in}. In this case, the principal source of jitter in D_{out} is the recovered clock, CK, itself. Note that the clock period is equal to the bit period, T_b.

That the clock must be periodic points to two possibilities for clock generation: (1) manipulate the input random data so as to extract a periodic component from it, or (2) employ an oscillator to produce a periodic waveform and add means to ensure proper frequency and phase alignment. Before delving into these techniques, we must familiarize ourselves with the nature of random data (Section 13.2).

Example 13.2

A certain data format exhibits the spectrum shown in Fig. 13.4. Can we generate a clock at a frequency of $1/T_b$ by means of a narrowband filter centered around this value?

Solution

No, we cannot. For the result to be periodic and have little noise, the filter bandwidth must be very narrow, thereby passing only an extremely small amount of energy (equal to the area under the spectrum from f_1 to f_2). We thus conclude that an *impulse* of energy is necessary at $1/T_b$ for clock recovery.

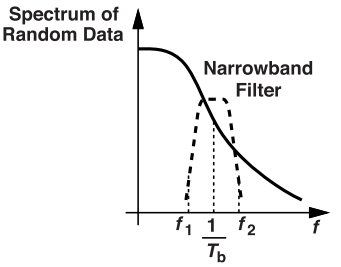
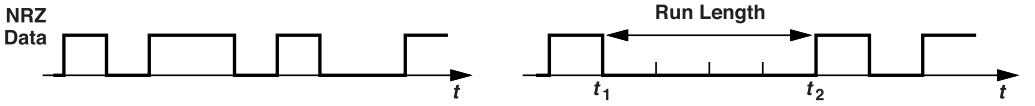


Figure 13.4 Narrowband filter to extract energy in the vicinity of $1/T_b$.

13.2 Properties of Random Binary Data

The most common data format used in digital transmission simply consists of a random sequence of ONEs and ZEROs [Fig. 13.5(a)]. Also called "nonreturn-to-zero" (NRZ) data, this format exhibits interesting prop-



(a) (b)

Figure 13.5 (a) NRZ data, and (b) illustration of run length.

erties in the time and frequency domains. If the data is completely random, it can contain an arbitrarily long sequence of consecutive ONEs or ZEROs, called a "run." We assume in this chapter that the ONEs and ZEROs occur with equal probabilities. Shown in Fig. 13.5(b) is an example having a "run length" of 4 bits. The absence of transitions in this time interval means that no frequency or phase information is available from t_1 to t_2, thus making CDR design difficult. In practice, the data is encoded to ensure an upper bound on the run length, as specified by the communication standard.

13.2.1 Spectrum of NRZ data

The frequency-domain behavior of NRZ data plays a central role in our study of CDR circuits. Since the signal is random, we must consider its power spectral density (spectrum) rather than its Fourier transform. In the first step, we remove the dc value of the waveform shown in Fig. 13.5(a), arriving at the dc-free sequence in Fig. 13.6. Now, we recognize that the waveform consists of shifted and possibly negated copies of a basic

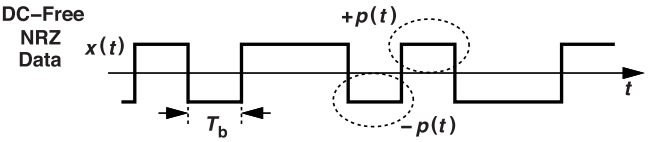


Figure 13.6 DC-free NRZ data.

pulse, $p(t)$. That is,

$$x(t) = \sum_{k=-\infty}^{+\infty} a_k p(t - kT_b), \tag{13.1}$$

where a_k is a random number assuming the values of $+1$ or -1, and T_b is the bit period. We say the bit rate is $1/T_b$ bits per second. As seen throughout this chapter, to simplify our analyses, we assume that the data toggles between 0 and $+1$ in some cases or between -1 and $+1$ in some others.

If the a_k values in (13.1) are independent and the $+1$ and -1 values occur with equal probabilities, it can be shown [11] that the spectrum of the waveform is given by

$$S_x(f) = \frac{1}{T_b}|P(f)|^2, \tag{13.2}$$

where $P(f)$ denotes the Fourier transform of $p(t)$. For example, if $p(t)$ is a rectangular pulse of height V_0 and width T_b seconds, then

$$S_x(f) = \frac{1}{T_b}\left[V_0 T_b \frac{\sin(\pi f T_b)}{\pi f T_b}\right]^2. \tag{13.3}$$

Let us examine this sinc^2 function. Plotted in Fig. 13.7(a), the spectrum exhibits nulls at integer multiples

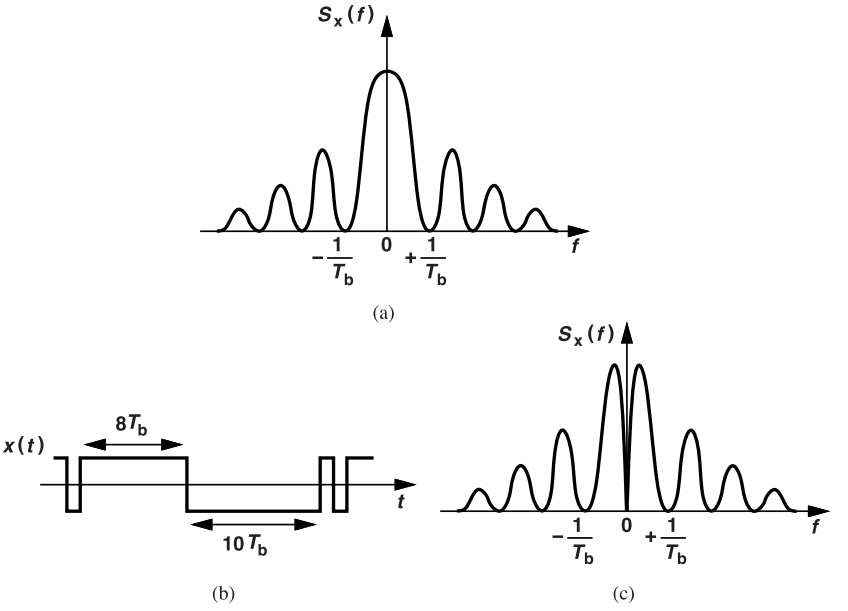


Figure 13.7 (a) Spectrum of NRZ data, (b) long runs in the data stream, and (c) spectrum of encoded data.

of the bit rate $(1/T_b)$, i.e., it carries no energy at these frequencies. Thus, a clock cannot be generated by simply filtering the NRZ data. Note that, if $x(t)$ in Fig. 13.6 is not shifted down and hence has a finite dc value, then an impulse, $\delta(f)$, appears in $S_x(f)$. Another interesting attribute observed in Fig. 13.7(a) is that $S_x(f)$ has a finite energy near the zero frequency; i.e., the data spectrum carries arbitrarily low frequencies. This occurs because the ONEs and ZEROs appear independently, sometimes creating long runs [Fig. 13.7(b)]. For example, a sequence of 100 ZEROs followed by 100 ONEs exhibits energy components roughly in the vicinity of $1/(200T_b)$.[1] As mentioned earlier, communication standards typically encode the data to limit the run length, in which case $S_x(f)$ begins to vanish as f approaches zero [Fig. 13.7(c)].

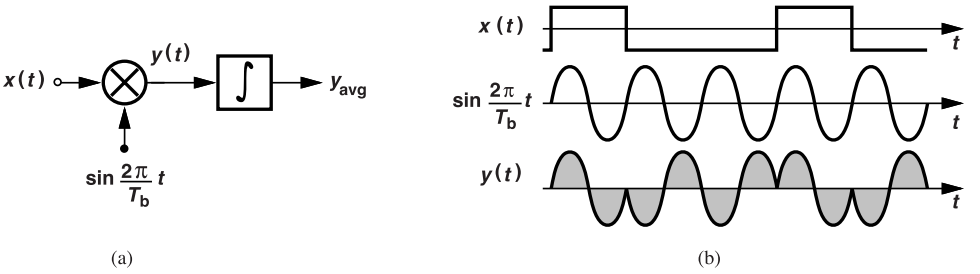
Example 13.3

Using the concept of correlation, prove that $x(t)$ in Fig. 13.6 does not contain a periodic component at $1/T_b$. [This is also evident in Fig. 13.7(a).]

[1] This property should not be confused with the dc content: a nonzero dc value in the time domain translates to an impulse at $f = 0$, while arbitrarily long runs yield $S_X(f = 0) > 0$.

Solution

In order to check the existence of a periodic waveform in a random signal, we can "correlate" the latter with the former, i.e., we can multiply the two and compute the average value of the result. If nonzero, this value signifies the periodicity within the random signal. Figure 13.8(a) conceptually depicts the setup for this test. Correlation in essence reveals the "similarity" between two signals.



(a) (b)

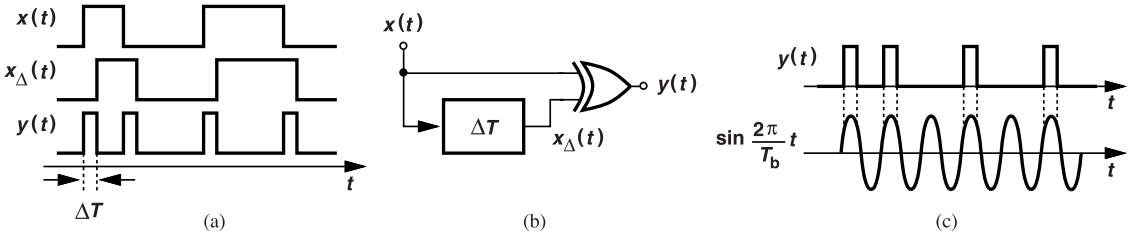
Figure 13.8 (a) Correlating $x(t)$ with $\sin(2\pi t/T_b)$, and (b) graphical illustration.

With the aid of the waveforms shown in Fig. 13.8(b), we observe that the product has a zero average in every bit period, indicating zero correlation. In Problem 13.3, we prove that this holds even if the initial phase of the sinusoid is chosen arbitrarily.

13.3 Clock Recovery by Edge Detection

We have surmised that a periodic clock can be extracted by manipulating random data. In other words, we wish to generate an impulse at $1/T_b$ from the spectrum shown in Fig. 13.7(a), which is possible only if we perform a *nonlinear* operation (why?).

An example of processing NRZ data for clock generation is "edge detection." Illustrated in Fig. 13.9(a), the idea is to produce a pulse (or an impulse) corresponding to each data transition. Figure 13.9(b) shows a



(a) (b) (c)

Figure 13.9 (a) Edge detection of random data, (b) circuit implementation, and (c) correlation between edge-detected data and $\sin(2\pi t/T_b)$.

possible implementation. To confirm that $y(t)$ contains a periodic component having a frequency of $1/T_b$, we follow the procedure in Example 13.3 and note from Fig. 13.9(c) that the product of $y(t)$ and $\sin(2\pi/T_b)t$ indeed has a nonzero average in every bit period. That is, the two waveforms exhibit finite correlation. In practice, we choose $\Delta T \approx T_b/2$.

While containing an impulse at $1/T_b$, the spectrum of $y(t)$ in Fig. 13.9(a) also carries significant "noise" due to the randomness that it has inherited from $x(t)$. Figure 13.10(a) sketches this spectrum, suggesting that a filter must follow the edge detector so as to produce a low-jitter clock. The very narrow bandwidth and the precise center frequency required of such a filter point to a phase-locked loop implementation [Fig. 13.10(b)]. We know from Chapter 7 that a PLL's bandwidth can be a very small fraction of its input frequency, affording an extremely high Q. Thus, $w(t)$ exhibits the spectrum shown in Fig. 13.10(c).

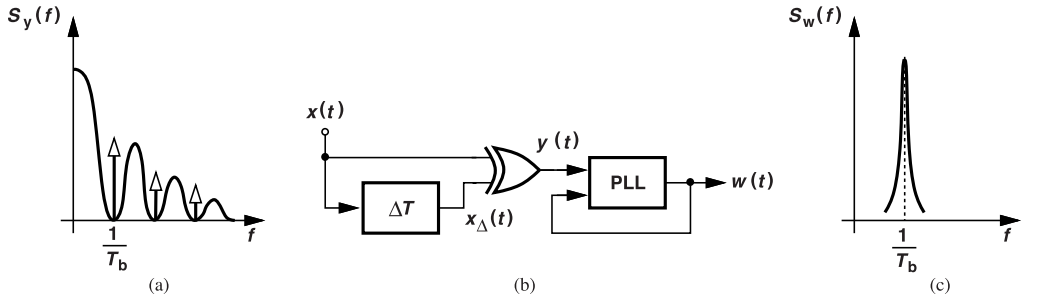


Figure 13.10 (a) Spectrum of edge-detected data, (b) such data driving a PLL, and (c) PLL output spectrum.

In summary, if NRZ data undergoes edge detection, its spectrum exhibits an impulse at $1/T_b$, which can serve as the recovered clock. However, the noise surrounding this component must be removed by a PLL, i.e., the PLL must lock to the impulse at $1/T_b$ while rejecting the noise.

Example 13.4

Explain how the data can be retimed (recovered) in the architecture of Fig. 13.10(b). Assume the retiming FF samples the input data on the falling edge of the clock.

Solution

As shown in Fig. 13.11, the rising edges of the recovered clock, $w(t)$, are aligned with those of $y(t)$ and hence $x(t)$. We therefore use the falling edges of $w(t)$ to sample $x(t)$ by means of a D flipflop.

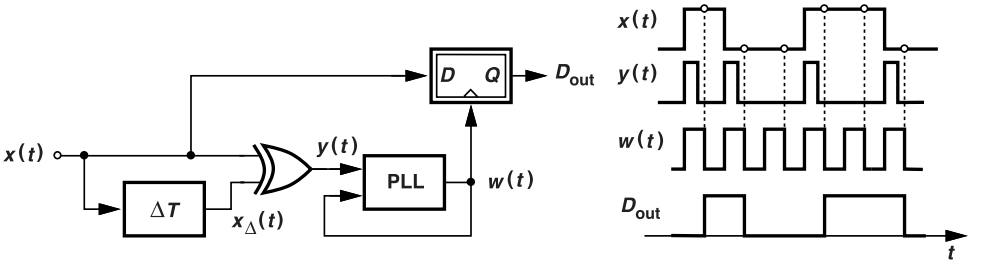


Figure 13.11 Use of edge detection for clock and data recovery.

Example 13.5

A student omits the PLL in Fig. 13.11, chooses $\Delta T \approx T_b/2$, and applies $y(t)$ directly to the retimer [Fig. 13.12(a)]. Explain the flaw in this design.

Solution

As seen in Fig. 13.12(b), this topology indeed retimes the input properly. However, no periodic clock is generated, which proves problematic for subsequent processing of D_{out} (e.g., demultiplexing). Also, D_{out} inherits all of the jitter of $x(t)$ (why?), i.e., D_{out} is indistinguishable from $x(t)$. We conclude that no clock or data recovery has taken place.

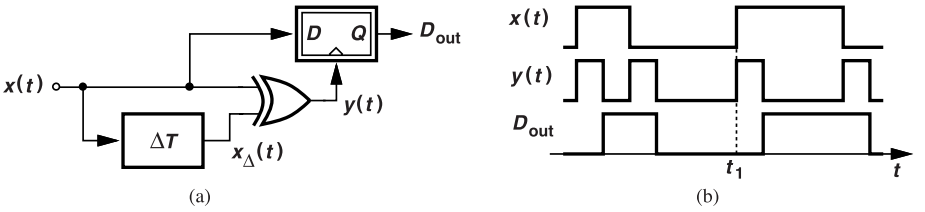


Figure 13.12 (a) Data recovery using edge-detected data, and (b) corresponding waveforms.

Example 13.6

If we consider the finite circuit delays in Fig. 13.11, how do the samples taken of $x(t)$ depart from their ideal points? Assume the falling edges of $w(t)$ sample $x(t)$.

Solution

The XOR has a delay, T_{XOR}, and the data and clock paths within the flipflop generally exhibit unequal delays. For example, if the clock path has a longer delay to the output, we can model the overall circuit as shown in Fig. 13.13(a), where ΔT_{CK} denotes the difference between the flipflop's data path and clock path

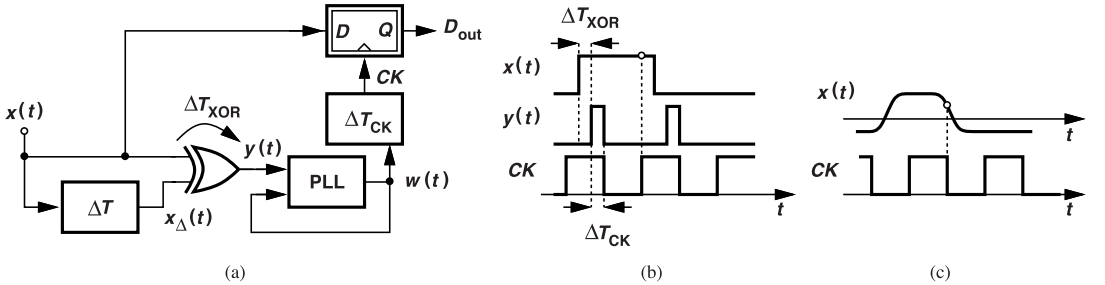


Figure 13.13 (a) Circuit of Fig. 13.11 including skews, (b) resulting loss of timing margin, and (c) more realistic situation.

delays. We observe that the sampling points on $x(t)$ are shifted by $\Delta T_{XOR} + \Delta T_{CK}$ [Fig. 13.13(b)], an issue that becomes critical at high speeds because of (a) the finite transition times of $x(t)$, and (b) the jitter in $x(t)$. As illustrated in Fig. 13.13(c) with more realistic waveforms and in the presence of skew and jitter, the sampling points suffer from degraded levels, causing a high bit error rate.

The difficulties discovered above make the solution shown in Fig. 13.11 less attractive at high speeds. We thus seek other approaches.

13.4 Clock Recovery by Phase-Locking

Let us reexamine the CDR architecture of Fig. 13.11: the final recovered clock is in fact generated by an oscillator that is phase-locked to the data edges. Suppose that we instead apply the input data directly to a PLL [Fig. 13.14(a)]. Does the circuit lock in this case? We recognize that the phase detector must generate a finite dc output in response to the phase difference between D_{in} and CK. We therefore return to the PD and PFD realizations described in Chapter 7 and determine whether they operate properly when one of the inputs is a random binary sequence rather than a periodic signal.

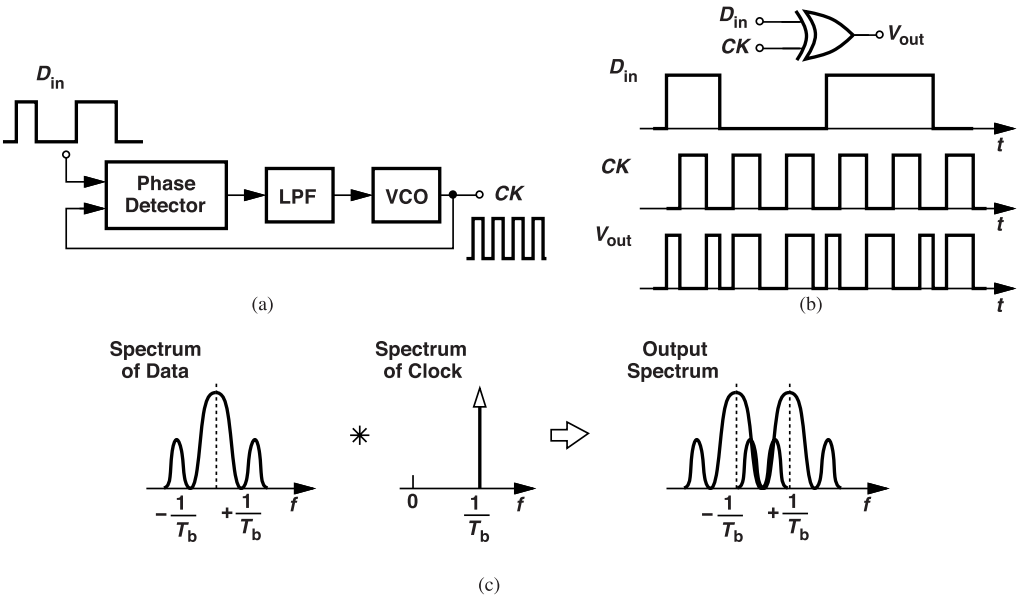


Figure 13.14 (a) Random data driving a PLL, (b) XOR operating as a PD, and (c) spectra at input and output of XOR PD. (For simplicity, the impulses representing the dc values are not shown in the spectra.)

Consider the XOR shown in Fig. 13.14(b). The reader can show that the average value of V_{out} is independent of the phase difference between D_{in} and CK. Thus, an XOR cannot operate as a phase detector in a CDR environment. In Problem 13.8, we prove that the same holds for the PFD developed in Chapter 7.

It is also possible to view the XOR as a mixer (multiplier) and recognize that it convolves the data and clock spectra. Depicted in Fig. 13.14(c), the result confirms that the output contains no dc.

The foregoing examples reveal that the design of PDs that operate with random data is not straightforward. In the next several sections, we develop such PDs.

13.4.1 Bang-Bang Phase Detector

Another circuit element that can act as a PD but has not been studied in the previous chapters is a D flipflop. It is instructive to first study the behavior of this PD with periodic inputs. As depicted in Fig. 13.15(a), if, for example, V_{in2} lags V_{in1} by $+\Delta\phi$, then the positive values of V_{in1} are sampled, producing a positive output average equal to $+V_H$ regardless of the value of $\Delta\phi$. Conversely, if V_{in2} leads V_{in1}, V_{out} assumes a negative average equal to $-V_H$. This analysis yields the "bang-bang" characteristic shown in Fig. 13.15(b), a very nonlinear behavior that nevertheless distinguishes between positive and negative phase errors. We expect a PLL employing such a PD to exhibit interesting and complex properties.

Example 13.7

A PLL with a periodic reference incorporates a DFF PD (Fig. 13.16). Examine the circuit's behavior in the locked condition.

Solution

As evident from Fig. 13.15(a), if $\Delta\phi$ is a constant positive value, the PD output remains equal to $+V_H$, which is not necessarily what the VCO requires. The same holds for a constant negative $\Delta\phi$. Thus, for V_{cont} to have the desired value, the phase difference sensed by the PD must fluctuate between positive and negative values, i.e., ϕ_{out} must toggle between two values. For this fluctuation to be small, the filter's bandwidth is reduced considerably. In other words, the large jumps between $+V_H$ and $-V_H$ generated by the DFF must

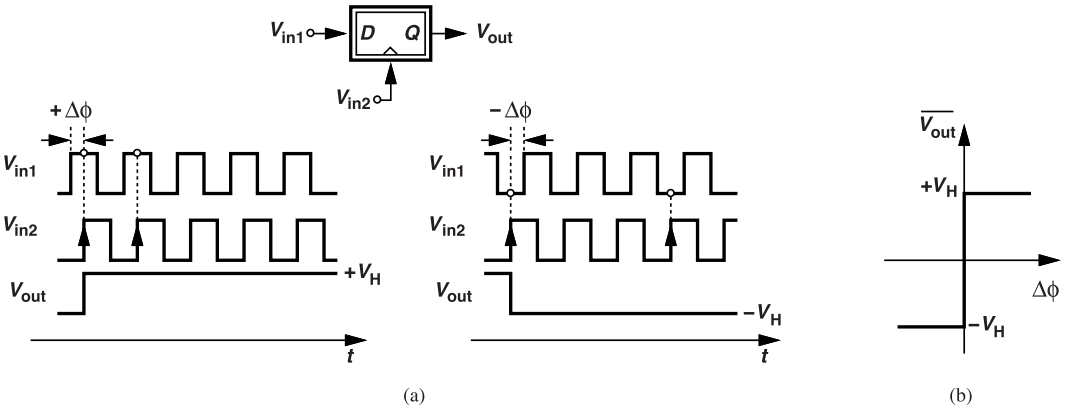


Figure 13.15 (a) D flipflop acting as a PD, and (b) its input-output characteristic.

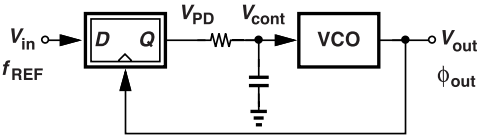


Figure 13.16 Phase-locked loop using a flipflop PD.

be attenuated so much that they create negligible ripple in V_{cont}. The exact analysis of the loop is beyond the scope of this book.

The bang-bang behavior shown in Fig. 13.15(b) suggests an infinite gain at $\Delta\phi = 0$. In practice, two phenomena soften this characteristic. First, the phase noise in V_{in1} and V_{in2} tends to randomize the phase difference sensed by the flipflop, yielding an average $|V_{out}|$ less than V_H for small $\Delta\phi$. To understand this point, let us assume a phase error of $\Delta\phi_1$ such that the rising edges of V_{in2} nominally sample a high level on V_{in1} [Fig. 13.17(a)]. Assuming that $\Delta\phi_1$ is small and V_{in2} has jitter, we recognize that, if this jitter displaces

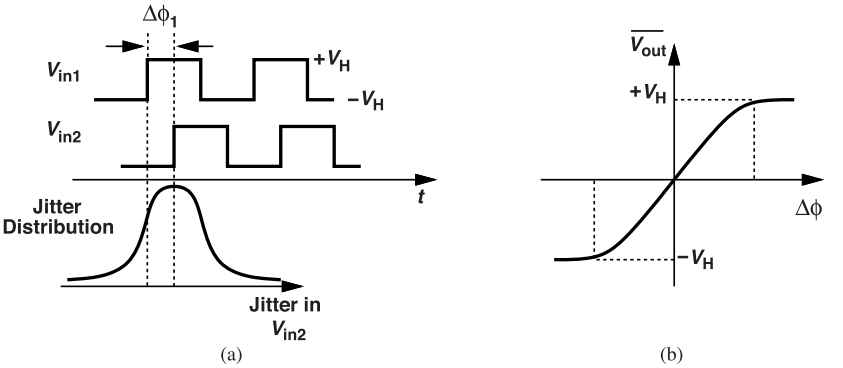


Figure 13.17 (a) Inputs applied to a bang-bang PD, and (b) smoothing of characteristic due to jitter.

the rising edge of V_{in2} to the left by more than $\Delta\phi_1$, then V_{in2} samples $-V_H$ on V_{in1} rather than $+V_H$. That is, statistically, some of the samples produced by the flipflop are negative, leading to an average value less than $+V_H$. Figure 13.17(b) sketches the resulting characteristic, which is in fact obtained as the convolution

of the ideal bang-bang function and the clock jitter distribution (probability density function, PDF) [1]. If the latter is a Gaussian PDF with a standard deviation of σ, then the plot in Fig. 13.17(b) has a relatively linear behavior for approximately $|\Delta\phi| < 3\sigma$ [1].

The second mechanism that lowers the bang-bang PD's gain is flipflop metastability. When the loop in Fig. 13.16 is locked, V_{in} and V_{out} have a small phase difference, the edges of V_{out} sample V_{in} near its zero crossings, and the flipflop becomes metastable (Fig. 13.18). As a result, the output sometimes fails to reach the full swing in one input period, exhibiting an average less than $\pm V_H$. This effect is analyzed in [1].

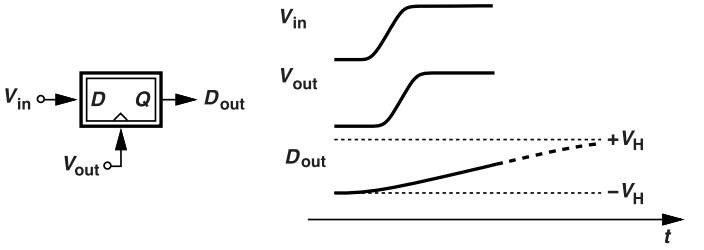


Figure 13.18 Effect of metastability in a bang-bang PD.

Does the PLL of Fig. 13.16 operate properly if V_{in} is replaced with random binary data? We repeat the analysis depicted in Fig. 13.14(b) for a random input (Fig. 13.19), observing that the output exhibits a zero average regardless of the phase error. After all, the DFF simply delays its input data.

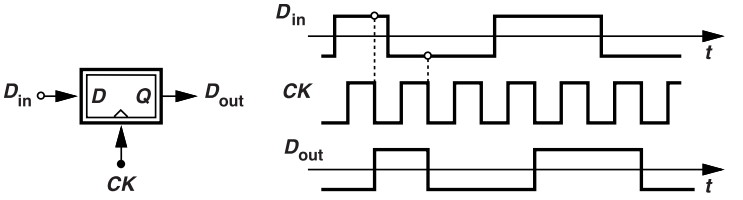


Figure 13.19 A flipflop in which the clock samples the data.

We must try one more configuration: what if the data and clock connections in Fig. 13.19 are swapped? Illustrated in Fig. 13.20, the idea is to sample the *clock* by the rising (or falling) data transitions. Now, the

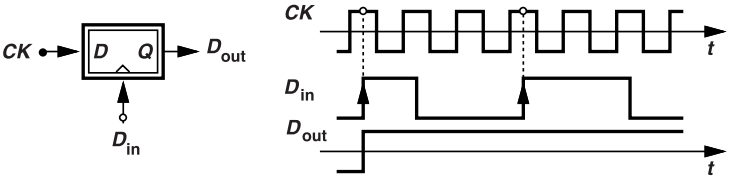


Figure 13.20 A flipflop in which the data samples the clock.

circuit yields a positive or negative output depending on the sign of the phase error, thereby acting as a bang-bang PD.

Why do the arrangements in Figs. 13.19 and 13.20 behave differently? We note that the former samples the data levels on every falling edge of CK even if D_{in} makes no transition. The latter, on the other hand, performs phase comparison only when a data edge occurs. In a sense, the former does not perform edge detection whereas the latter does.

These findings lead to the bang-bang CDR architecture shown in Fig. 13.21(a). It is important to recognize the different data and clock connections to the two flipflops: FF_1 operates as a PD and senses the data at its clock input, whereas FF_2 acts as a retimer and samples D_{in} by the recovered clock (Example 13.4).

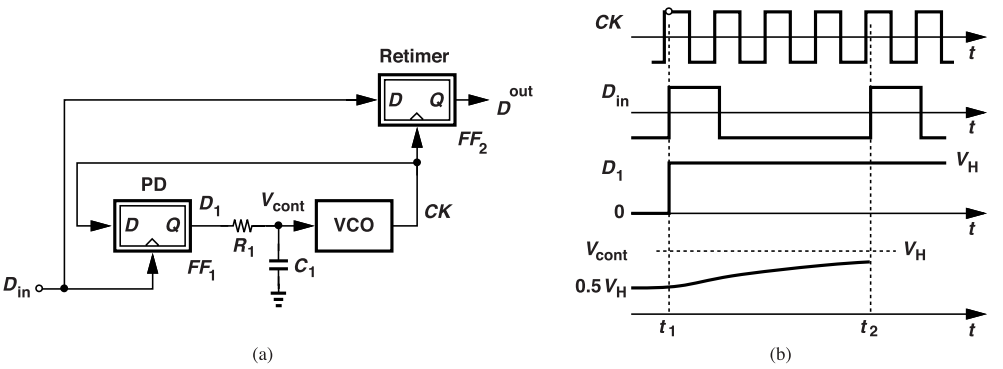


Figure 13.21 (a) Basic bang-bang CDR circuit, and (b) effect of large swings at PD output.

The architecture of Fig. 13.21(a) elicits a number of questions. First, does the loop contain *frequency* acquisition as well (Chapter 7)? No, it does not. As explained in Chapter 7 and also in the next chapter, a frequency detector must be added to the circuit to ensure locking if the initial difference between the data rate and the VCO frequency exceeds the loop's capture range.

Second, how do we choose the loop bandwidth here? This bandwidth is decided by a number of factors related to the jitter. Interestingly, the bang-bang operation in Fig. 13.21(a) can create significant jitter in the recovered clock even if D_{in} has none. This can be seen from the waveforms shown in Fig. 13.21(b), where the absence of data edges between t_1 and t_2 allows the PD output to reach a logical level and push V_{cont} up (or down) considerably. The VCO phase changes according to:

$$\phi(t_2) - \phi(t_1) = K_{VCO} \int_{t1}^{t2} V_{cont} dt, \tag{13.4}$$

potentially a large amount. Called "data-dependent" jitter, the result is more severe for longer runs. To alleviate this issue, R_1C_1 must be much greater than the maximum run length so as to minimize the change in V_{cont}. For this reason, the loop bandwidth is drastically reduced. As an example, if $V_H = 1$ V, $t_2 - t_1 = 10T_b$, and we wish to keep the change in V_{cont} less than 5 mV, we have $R_1C_1 \approx 2000T_b$.

To appreciate another difficulty with this type of PD, recall from Fig. 13.1(a) that high-speed data can be severely attenuated in a medium. Such small data swings do not activate FF_1 properly, failing to sample the clock.

Example 13.8

Repeat Example 13.6 for the architecture of Fig. 13.21(a).

Solution

Modeling the data and clock path delay differences as shown in Fig. 13.22, we observe that, in the locked condition, CK lags D_{in} by ΔT_{CK} seconds due to FF_1. This clock is delayed by another ΔT_{CK} seconds within FF_2. Thus, the actual clock edge sampling the input data in FF_2 is offset from its ideal point by $2\Delta T_{CK}$ seconds, a serious problem at high speeds.

The CDR architecture of Fig. 13.21(a) entails another issue that can raise the jitter. In a typical flipflop design, a capacitive path exists between the clock and data inputs, introducing unwanted coupling. Consider, for example, the latch topology shown in Fig. 13.23 (Chapter 15), where C_{GD5} and $C_{GS1,2}$ allow the transitions to travel from CK to D and D and vice versa. We recognize that both flipflops in Fig. 13.21(a) suffer from this effect and couple the random data transitions to the VCO. To minimize the corruption, a buffer should immediately follow the VCO. This issue arises in all CDR architectures.

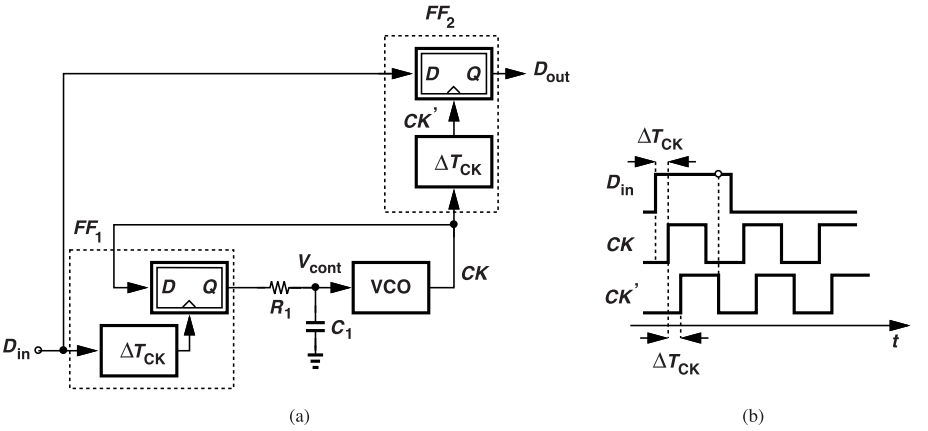


Figure 13.22 Effect of skews in a bang-bang CDR loop.

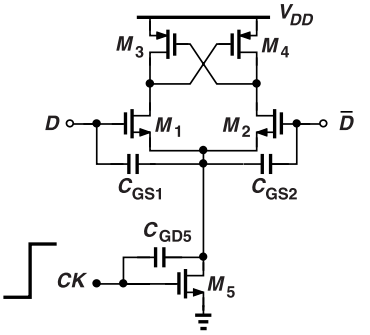


Figure 13.23 Kickback noise in a latch.

Our study of the single-flipflop PD indicates that bang-bang operation produces considerable ripple on the oscillator control line in the presence of long runs. This is because the PD maintains its maximum output level even if no data transition occurs. We thus postulate that jitter can be reduced if the PD output is disconnected from the filter or disabled when the data waveform makes no transition. Conversely, we can say the PD output should be enabled only when there is a data edge.

With the foregoing thoughts, we construct the system shown in Fig. 13.24, where an edge detector allows the PD output to reach the loop filter only after a data transition occurs, and only for a brief period, ΔT. The The control voltage is thus constant in the absence of data edges. This method assumes that the PD samples a

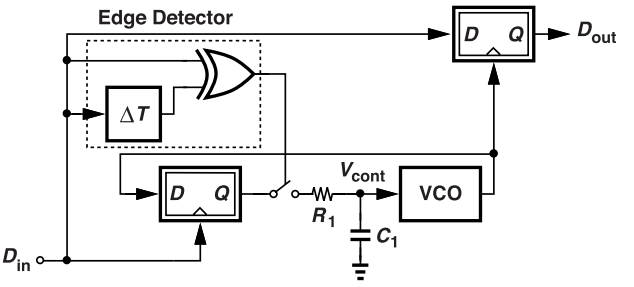


Figure 13.24 Use of an edge detector in a bang-bang CDR circuit.

metastable level and does not have enough time to produce a full swing while the switch is on. Owing to the skews between the XOR and the PD, this is difficult to guarantee. Alternatively, R_1C_1 can be chosen large enough to minimize the ripple on V_{cont}.

Another issue here is that the XOR must produce a rail-to-rail swing so as to activate the switch, a challenging task at high speeds. Note that a complementary switch is necessary here so that V_{cont} can approach zero and V_{DD}.

Example 13.9

Which one of the issues in the architecture of Fig. 13.21(a) are also present in Fig. 13.24?

Solution

While experiencing less data-dependent jitter, this structure must still deal with several issues: (1) it lacks a means for frequency acquisition, (2) it suffers from the phase offset described in Example 13.8, and (3) it entails the unwanted coupling from D_{in} to CK through the flipflop's capacitive paths.

13.4.2 Alexander Phase Detector

Let us continue the thought process in the previous section and seek a phase detector (PD) topology whose output is "disabled" in the absence of data transitions. In this development, we begin with two new principles. (1) Only the clock must sample the data; as explained below, this principle avoids the skews studied in Example 13.8. However, this method does not yield the phase information with a single flipflop (Fig. 13.19), requiring more than one. (2) We utilize both the rising and falling edges of the clock for sampling. Now, consider the situation depicted in Fig. 13.25(a), where CK takes three consecutive samples, S_1-S_3. If the

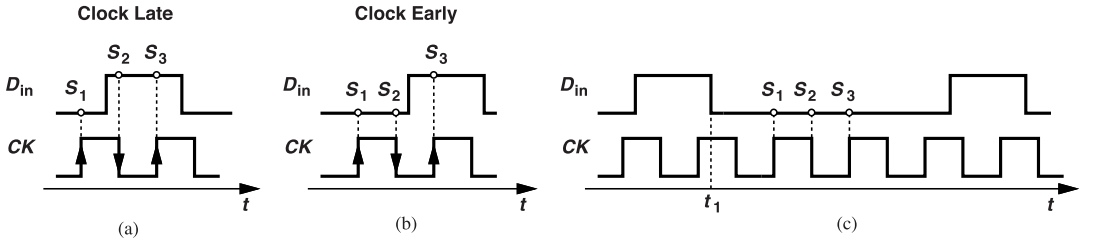


Figure 13.25 (a) Case of late clock, (b) case of early clock, and (c) results of sampling in the absence of data transitions.

falling edge of CK is "late" with respect to the rising edge of D_{in}, then $S_1 \neq S_2 = S_3$. On the other hand, if CK is "early" [Fig. 13.25(b)], we have $S_1 = S_2 \neq S_3$. Thus, S_1, S_2, and S_3 collectively serve two purposes. (1) They provide the sign of the phase error in the form of a bang-bang characteristic: if $S_1 \neq S_2 = S_3$, the PD will be designed to generate a high level and if $S_1 = S_2 \neq S_3$, a low level. (2) They detect when a data edge occurs. For example, we have $S_1 = S_2 = S_3$ after $t = t_1$ in Fig. 13.25(c) since there are no transitions. These are the principles behind the Alexander PD [2].

How do we implement this operation? We expect to use D flipflops for sampling and XOR gates for comparison of S_1-S_3. Also, we envision that S_1 and S_2 must be *stored* before S_3 is available so that the three samples are aligned in time for proper comparison. We begin with two flipflops operating on different clock edges [Fig. 13.26(a)], obtaining S_1 and S_2. Now, we store and align these results by two more flipflops clocked by CK [Fig. 13.26(b)]. That is, Q_3 and Q_4 are updated at the same time and represent S_1 and S_2, respectively. Also, at the time S_1 reaches Q_3, FF_1 samples S_3. Note that FF_2 and FF_4 act as three latches rather than as four (why?). We observe that Q_1, Q_3, and Q_4 provide the three samples at the same time.

Next, we add XOR gates to compare S_1 with S_2 and S_2 with S_3. As illustrated in Fig. 13.26(c), two outputs, A and B, reveal the phase difference polarity: if $A = 1$ and $B = 0$, CK is early, and if $A = 0$ and $B = 1$, CK is late. Exhibiting a bang-bang characteristic, this arrangement is called the Alexander phase detector [2]. With no data transitions, we have $S_1 = S_2 = S_3$ and hence $A = B = 0$.

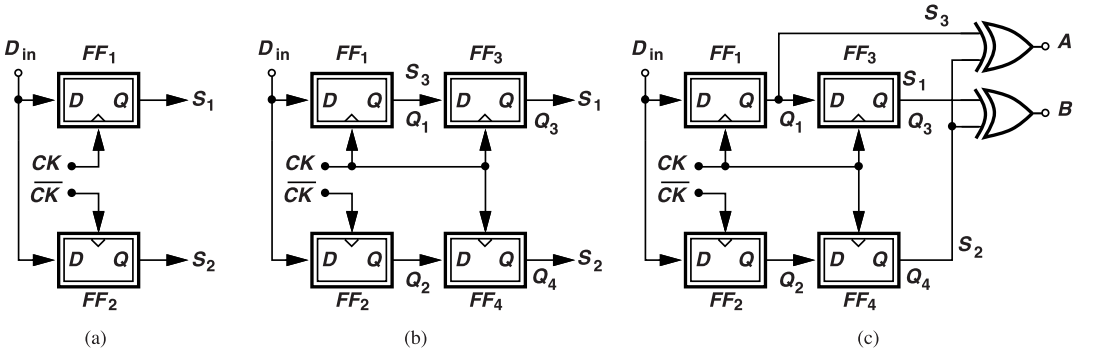


Figure 13.26 (a) Use of two flipflops to take data samples, (b) addition of two flipflops to align the samples, and (c) complete Alexander PD.

It is instructive to draw the input and output waveforms of the Alexander PD carefully, noting that Q_1, Q_3, and Q_4 change on the rising edges of CK, and Q_2 on the falling edges. Depicted in Fig. 13.27 for late and early clocks, the results reveal that (1) the flipflops' outputs are simply shifted replicas of D_{in}, (2) if the clock

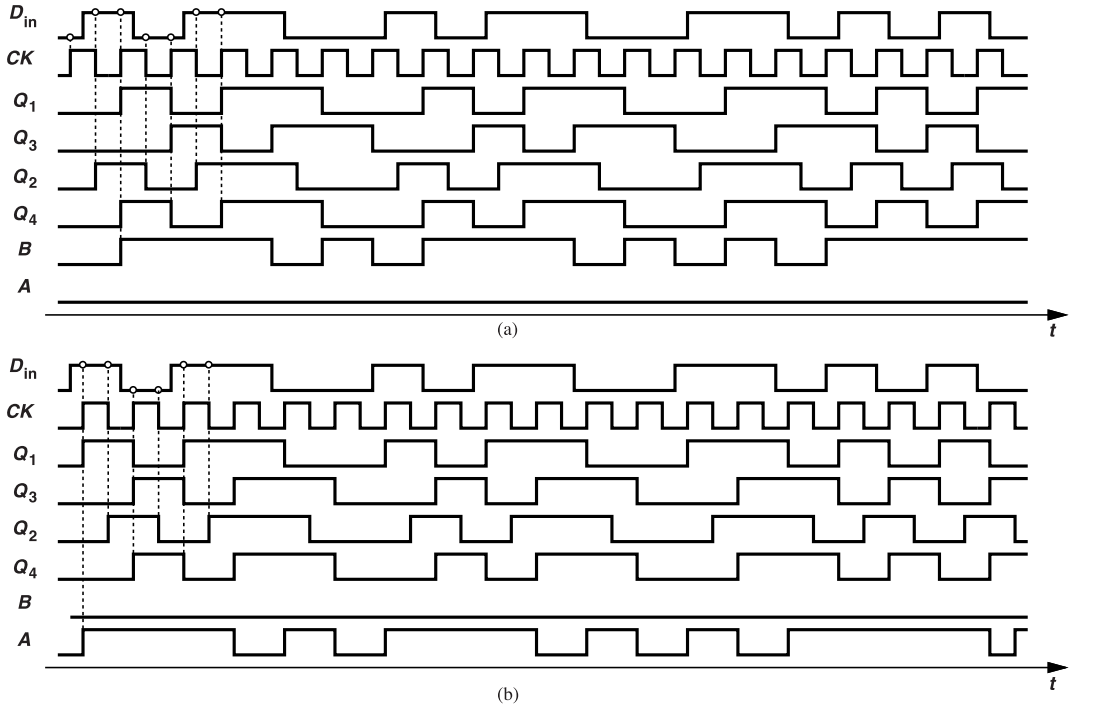


Figure 13.27 Alexander PD waveforms for (a) late, and (b) early clock cases.

is late, then $Q_1 = Q_4$ (i.e., $S_2 = S_3$), and if the clock is early, then $Q_3 = Q_4$ (i.e., $S_1 = S_2$), (3) depending on the polarity of the phase error, only A or B produces pulses, and (4) these pulses are generated only when there is a data transition. For example, in the case of late clock, Q_3 is 1 bit period ahead of Q_4; if a transition is absent in Q_4, it is also absent in Q_3 after 1 UI, yielding $Q_3 = Q_4$ in between and hence $B = 0$.

We must recognize four important properties of the Alexander PD. First, it provides phase error information even though its flipflops sample the data by the clock. [Recall that in the single-FF PD of Fig. 13.21(a), the data samples the clock.] This attribute originates from taking three samples on the consecutive edges of

the clock. We say the Alexander PD uses an "oversampling" factor of 2 because it takes two data samples per clock period.

Second, the structure inherently retimes the data as well, generating the recovered data at the flipflop outputs, again because the clock samples the data. For example, the retimed data is available at Q_3 in Fig. 13.26(c). Third, in contrast to the architecture of Fig. 13.21(a), the flipflops' internal clock and data delay difference does not translate to phase offset (Problem 13.14). Fourth, the circuit produces no output pulses in the absence of data transitions, as if it were "disabled." As explained below, the stage following the PD can be configured such that it does not disturb the loop filter when there are no input edges. Thus, with a long run in the input data, the oscillator control voltage does not change significantly, another point of contrast to the behavior of the loop in Fig. 13.21(a).

In summary, the Alexander PD acts as a bang-bang circuit but it generates much less ripple than the single-FF topology of Fig. 13.21(a). It also retimes the data. Note that the XOR output nodes in Fig. 13.26(c) need not have a wide bandwidth (unless they drive a charge pump) as only the average values of A and B are of interest.

CDR Loop How do we utilize the PD outputs in Fig. 13.26(c)? Since the difference between the average values of A and B signifies the phase error, we apply these two outputs to a voltage-to-current (V/I) converter (a G_m stage) and allow the result to flow through the loop filter (Fig. 13.28). Nodes A and B need not support a wide bandwidth as the V/I converter measures only their average voltage. In fact, we deliberately add some

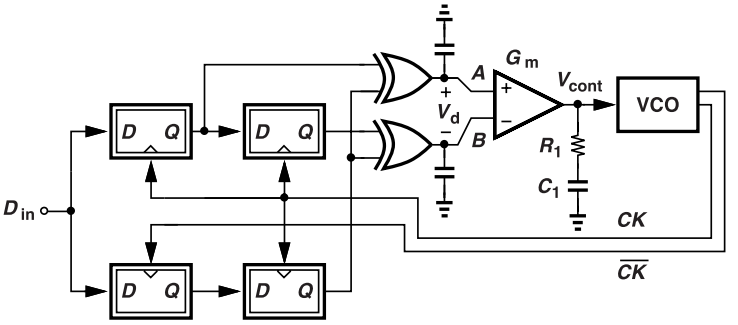


Figure 13.28 CDR loop using an Alexander PD.

capacitance to these outputs so as to limit the high-frequency voltage excursions reaching the V/I converter. Thus, V_A and V_B experience only some ripple.

Example 13.10

Figure 13.29 shows a five-transistor operational transconductance amplifier (OTA) as a candidate for V/I

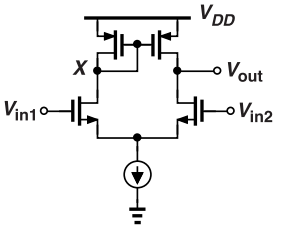


Figure 13.29 Simple V/I converter.

conversion. What design issues does this circuit face?

Solution

The most severe issue relates to the allowable output voltage range. For an input common-mode level equal to V_{CM}, V_{out} cannot fall below $V_{CM} - V_{THN}$ or exceed $V_{DD} - |V_{GSP} - V_{THP}|$. Consequently, the tuning voltage range for the VCO is limited.

The input-referred offset of the circuit is also problematic. Suppose the input data in Fig. 13.28 exhibits a long run, yielding $A = B$. The V/I converter thus produces a current given by $G_m V_{OS}$, where G_m and V_{OS} denote the transconductance and the offset voltage, respectively. Flowing through the filter for the entire length of the run, this current can cause substantial ripple on the oscillator control. In Problem 13.15, we investigate the effect of the OTA's noise. The offset and noise issues are alleviated through the use of long and wide transistors in the OTA design.

XOR Implementation The two inputs of each XOR gate in Fig. 13.28 must be nominally symmetric, i.e., have equal delays in their input paths. Here, we explore XOR topologies with differential or single-ended outputs. Consider the high-speed XOR structure shown in Fig. 13.30(a). We observe that the two inputs see different propagation delays to the output. Also, if the inputs do not have rail-to-rail swings, a level shift

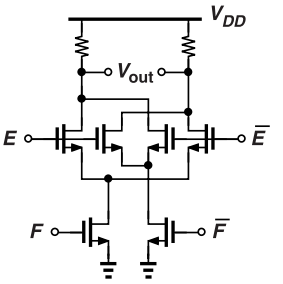


Figure 13.30 Gilbert cell acting as an XOR gate.

circuit is necessary in the F path, causing additional delay. In Problem 13.16, we examine the impact of this skew on the CDR's performance. As a rule of thumb, if the data rate is high enough to require current-mode logic (CML) latches, this skew is problematic.

Example 13.11

How are the above XOR's differential outputs processed in Fig. 13.28?

Solution

The V/I converter has two input terminals whereas two differential XOR gates have four outputs. We therefore have three options: discard one output of the XORs, use a differential-to-single-ended converter after each XOR, or design the V/I converter to have two sets of differential inputs.

We introduce three symmetric XOR topologies here. The first simply places in parallel two instances of a given XOR but with the inputs to one swapped [Fig. 13.31(a)]. The result exhibits symmetric delays.[2] Depicted in Fig. 13.31(b), the second operates with rail-to-rail inputs [3]. Assuming that $V_b \approx V_{DD}/2$, we note that when E is low, only M_4 and M_5 can be active and $V_{out} = \overline{F}$. By symmetry, when F is low, $V_{out} = \overline{E}$. Thus, $V_{out} = \overline{E \oplus F}$. This structure does not require complementary inputs but it draws some static power when E or F is high.

Example 13.12

The single-ended output in Fig. 13.31(b) raises concern about supply noise. Explain whether this is a serious issue.

[2] The same idea applies to the reset NAND gate in the standard PFD topology if the up and down pulsewidths must match precisely.

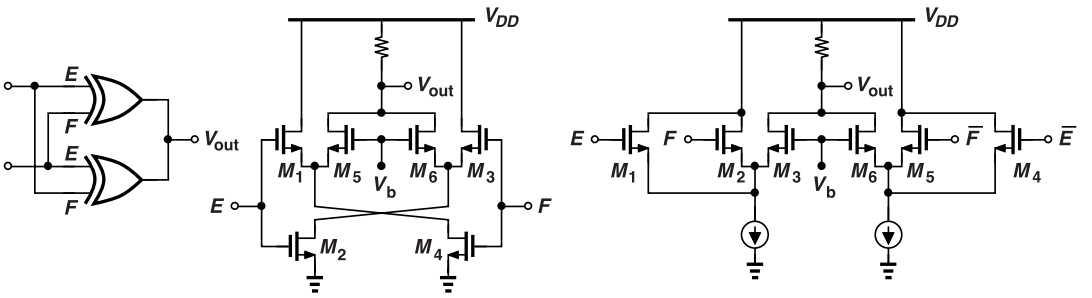


Figure 13.31 Examples of symmetric XOR and XNOR gates.

Solution

In general, it is, but in the Alexander PD environment, it is not. Noting that the V/I converter inputs in Fig. 13.28 arrive from *two* XOR gates, we observe that supply noise becomes a common-mode effect in A and B (Fig. 13.32).

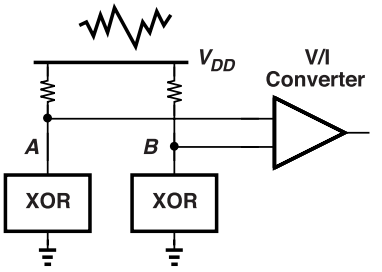


Figure 13.32 Effect of supply noise on XOR outputs.

Figure 13.31(c) shows the third symmetric XOR topology [4], which requires only moderate input swings (a few hundred millivolts). The circuit consists of two current-steering NOR gates (M_1-M_3 and M_4-M_6) sharing an output node. That is, I_{D3} and I_{D6} represent $E + F$ and $\overline{E + F}$, respectively, yielding

$$V_{out} = (E + F + E + F) \tag{13.5}$$

$$= (E + F)(\overline{E} + \overline{F}) \tag{13.6}$$

$$= E\overline{F} + \overline{E}F. \tag{13.7}$$

A complementary output can also be generated by copying the circuit and swapping F and $\overline{F}$, but, as explained in Example 13.12, such an output is not necessary in the Alexander CDR environment. The bias voltage, V_b, is chosen equal to the input common-mode level and must be bypassed to the supply line of the preceding latches by a capacitor if the E and F signals are referenced to that supply voltage (Fig. 13.33). The supply noise thus perturbs V_b and the CM level of the input signals equally.

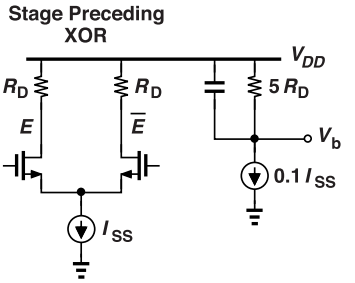


Figure 13.33 Bias generation for XOR gate.

Loop Type and Gain Is the CDR loop of Fig. 13.28 a type-I or a type-II loop? The transfer function from V_d to V_{cont} is given by

$$\frac{V_{cont}}{V_d}(s) = G_m \left[R_{out} || \left(R_1 + \frac{1}{C_1 s} \right) \right] \tag{13.8}$$

$$= \frac{G_m R_{out}(R_1 C_1 s + 1)}{(R_{out} + R_1)C_1 s + 1}. \tag{13.9}$$

Owing to the finite output impedance of the V/I converter, the pole, $\omega_{p1} = 1/[(R_{out} + R_1)C_1]$, has departed from the origin. Thus, in the strict sense, the circuit acts as a type-I loop. The product of this transfer function and K_{VCO}/s behaves as shown in Fig. 13.34: the magnitude falls at -20 dB/dec up to $\omega = \omega_{p1}$, and at -40 dB/dec between ω_{p1} and $\omega_z = 1/(R_1 C_1)$; the phase begins at $-90°$, reaches $-135°$ at ω_{p1} and

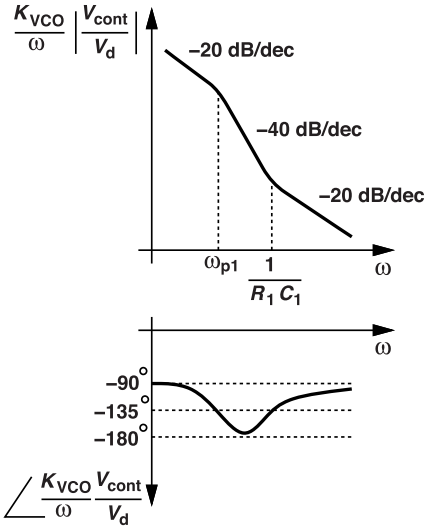


Figure 13.34 Open-loop transfer function of a CDR loop using an Alexander PD (the magnitude must be multiplied by the PD gain).

approximately $-180°$ between ω_{p1} and ω_z, and returns to $-135°$ at ω_z. Note that the magnitude must be multiplied by the PD gain [the slope in Fig. 13.17(b) and the gain of the XOR gates] to yield the total loop gain.

The type-I PLL studied in Chapter 7 suffers from a finite phase offset. Is that the case here as well? We surmise that the bang-bang operation of the Alexander PD would ideally provide infinite gain, allowing no static phase error. However, as explained in Section 13.4.1, the softening of the characteristic due to the clock

and data jitter yields a finite PD gain. In other words, a maximum phase offset of roughly 3σ can appear in the CDR circuit, where σ is the total rms jitter sensed by the PD.

Charge Pump versus V/I Converter Following the PLL developments in previous chapters, we can consider using a charge pump with the Alexander PD, aiming for an infinite gain. Figure 13.35 shows the conceptual realization. Unlike the standard PFD, this PD does not generate up and down pulses whose widths

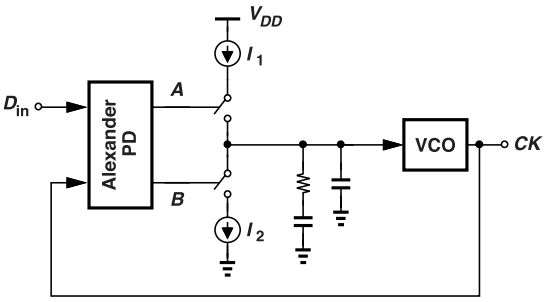


Figure 13.35 Use of an Alexander PD with a charge pump.

are proportional to the phase error. In fact, the pulses appearing on A and B in Fig. 13.27 have a width equal to an integer multiple of the clock period, a property arising from the circuit's bang-bang behavior (Problem 13.13).

In CDR design, charge pumps have a speed disadvantage with respect to V/I converters. While in Fig. 13.28, only the average voltage difference between A and B is of interest, the CP in Fig. 13.35 requires that these nodes be agile and provide rail-to-rail swings so as to turn the switches on and off in one clock period. In other words, the XOR gates in the Alexander PD must have wideband output nodes and, if not providing rail-to-rail swings, must be followed by CML-CMOS level converters.

Alternative Perspective We have observed that, in the locked condition, samples S_1 and S_3 in Fig. 13.25(a) land in the middle of the data bits and sample S_2 on the edge. We thus call the former the "data samples" and the latter the "edge sample." The detailed waveforms in Fig. 13.27 suggest that, if the clock is late, $Q_1 = Q_4$ and hence $A = 0$ while $B = Q_3 \oplus Q_4$ pulsates. Conversely, if the clock is early, $Q_3 = Q_4$ and $B = 0$ while $A = Q_1 \oplus Q_4$ pulsates. It is common to call A and B "up" and "down" pulses, respectively, and verbalize the operation by saying if S_1 and S_2 are unequal (i.e., if the clock is late), the down command is asserted and, similarly, if S_2 and S_3 are unequal, the up command is. That is, the XOR of one data sample and the next edge sample provides the down information, and the XOR of that edge sample and the next data sample yields the up information. Figure 13.36 conceptually illustrates this point. This perspective is consistent with the charge-pump CDR loop shown in Fig. 13.35.

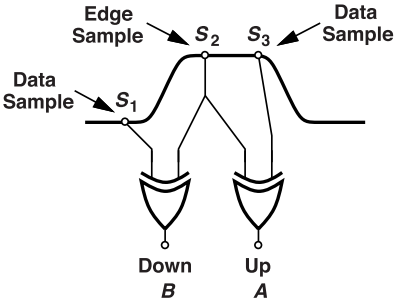


Figure 13.36 Alternative view of Alexander PD outputs.

13.4.3 Hogge Phase Detector

In this section, we develop a linear phase detector that can operate with random binary data. We expect that such a PD delivers an output whose average is linearly proportional to the phase error. Let us return to the edge detector shown in Fig. 13.9(b) and recognize that the delay stage can be instead realized by a flipflop [Fig. 13.37(a)]. As illustrated in Fig. 13.37(b), the delay between D_{in} and D_1 is equal to $T_{CK}/2$ if CK is

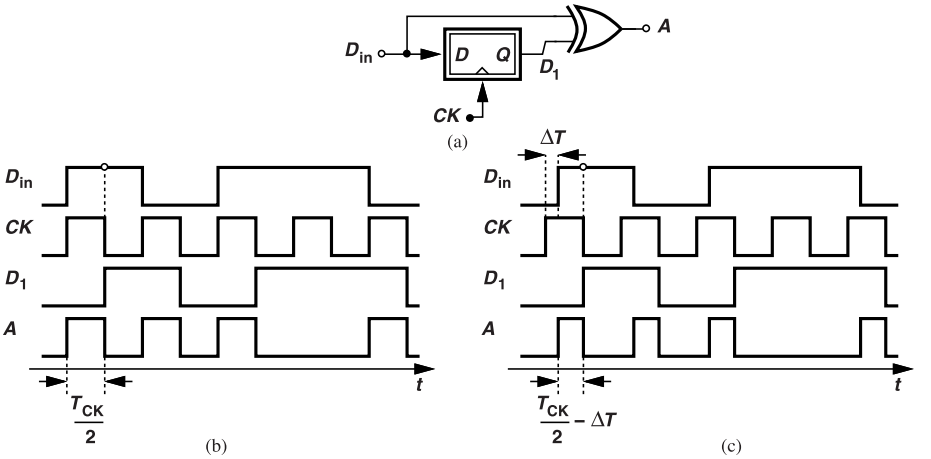


Figure 13.37 (a) Synchronous edge detector, (b) case of aligned clock, and (c) case of early clock.

properly aligned with D_{in}. We assume the falling edge of CK samples the data. The output then produces, for each input data transition, a pulse whose width is $T_{CK}/2$ seconds. Can this edge detector, by itself, serve as a PD? Suppose the rising edge of the clock is early by ΔT seconds [Fig. 13.37(c)]; then, the output pulsewidth, $T_{CK}/2 - \Delta T$, indeed represents the phase difference between D_{in} and CK. The same observation applies if the clock is late. Unlike the bang-bang PDs studied in the previous sections, this structure exhibits a linear characteristic. We say the output, A, contains "proportional pulses."

Example 13.13

Sketch the input-output characteristic of the PD in Fig. 13.37(a).

Solution

Since the PD's output pulsewidth is equal to $T_{CK}/2 - \Delta T$, the average output goes to zero if $\Delta T = T_{CK}/2$, leading to the behavior shown in Fig. 13.38.

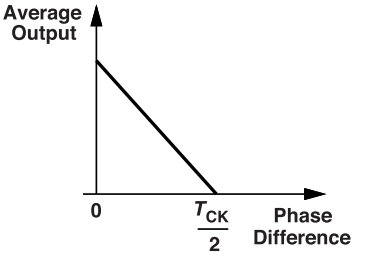


Figure 13.38 Input-output characteristic of PD in Fig. 13.37(a).

How does a CDR circuit lock with such a PD characteristic? Suppose the loop gain is extremely high; it is conceptually helpful to imagine a high-gain amplifier following the loop filter. Owing to the high gain, the loop demands a near-zero dc output for the PD, forcing the input phase error toward $T_{CK}/2$. But this means

that the falling edge of CK (which must sample the data in the middle of the eye) occurs near the data edge, causing a high bit error rate. We must therefore modify the circuit to avoid this behavior.

We surmise that another sequence of output pulses is necessary so that the average *difference* between the two outputs falls to zero when the phase error is zero. Let us then apply D_1 in Fig. 13.37(a) to another flipflop that is clocked by $\overline{CK}$ and XOR the result with D_1 [Fig. 13.39(a)]. Since the phase difference between D_1

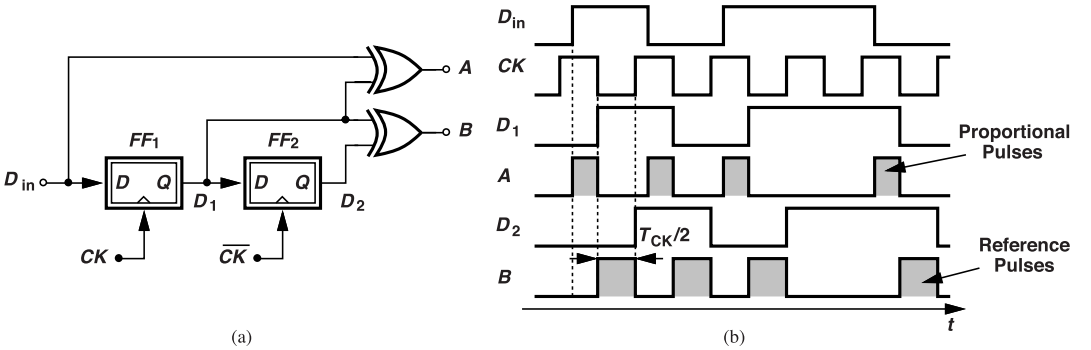


Figure 13.39 (a) Complete Hogge PD, and (b) its waveforms. (The FFs sample their inputs on the falling edges of their respective clocks.)

and D_2 is equal to half of a clock period, B generates pulses of this width for each data transition. We call these the "reference pulses." The width of pulses in A is equal to $T_{CK}/2$ minus the phase error whereas that of the pulses in B is constant and equal to $T_{CK}/2$. Thus, the average difference between A and B is zero only if D_{in} and CK are aligned. This is how a high-gain loop locks. Called the "Hogge PD" [5], this topology also retimes the input data, reproducing it at D_1 and D_2.

It is instructive to review how we have arrived at the Hogge PD. With the FF skew difficulties in the CDR of Fig. 13.21(a), we decided to only sample the data by the clock, realizing that a single FF would not serve as a PD in this mode. The XOR operation in Fig. 13.37(a) happened to allow phase detection, but with a nonzero average output for a zero phase error. For this reason, we added the second FF and XOR gate so that the average difference between A and B is zero if the phase error goes to zero.

Design Issues Operating as a linear PD, the Hogge topology is well-suited to CDR design but it poses several challenges at high speeds. First, the idealized waveforms in Fig. 13.39 do not account for the finite delays within the flipflops. Specifically, as shown in Fig. 13.40, the CK-to-Q delay of FF$_1$, ΔT_1, *widens* the pulses in A because D_1 changes ΔT_1 seconds later than expected. The resulting error in A translates to a static phase error equal to ΔT_1 in a CDR environment.

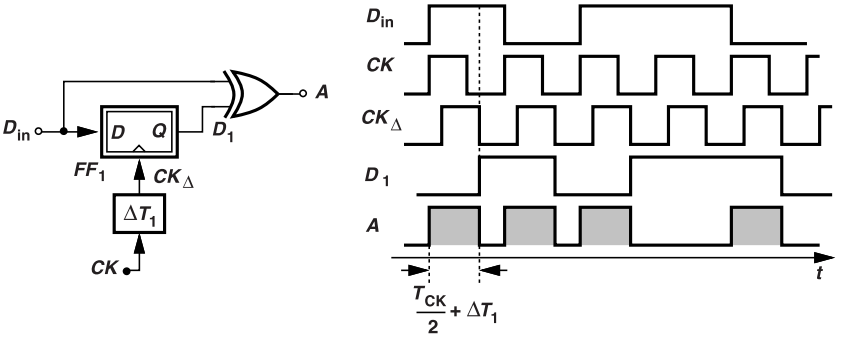


Figure 13.40 Effect of clock skew in Hogge PD.

Example 13.14

What is the effect of the CK-to-Q delay of FF_2 in Fig. 13.39?

Solution

This delay shifts D_2. Since D_1 and D_2 are shifted by the same amount, the pulses in B maintain a width equal to $T_{CK}/2$. Thus, the reference pulses are not affected.

In order to compensate for the flipflop delay in Fig. 13.39, we can make either the proportional pulses narrower or the reference pulses wider. The former method requires inserting a delay between D_{in} and the first XOR input [Fig. 13.41(a)] [5], thereby ensuring that the data and the clock experience equal delays. The

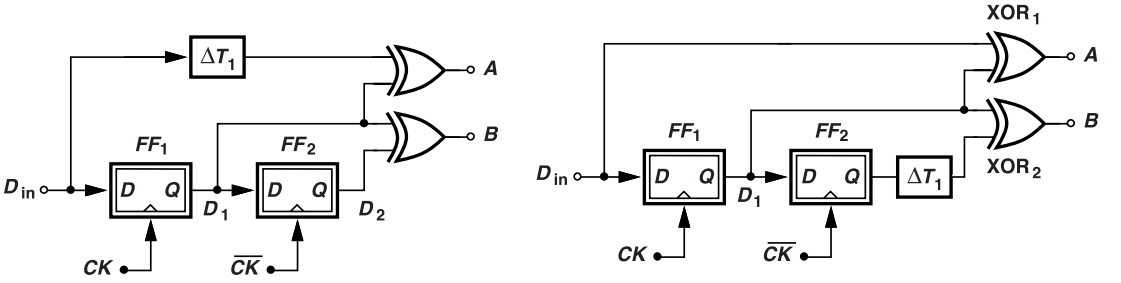


Figure 13.41 Compensation for clock skew by inserting a delay at the input of (a) first XOR, or (b) second XOR.

latter calls for a delay stage at the output of FF_2 [Fig. 13.41(b)] [6] so that the phase difference presented to XOR_2 increases from $T_{CK}/2$ to $T_{CK}/2 + \Delta T_1$. The ΔT_1 stage can be implemented as a replica of the path through the FF.

The second issue in the Hogge PD arises from the misalignment between the proportional and reference pulses. As seen in Fig. 13.42, the pulses in B arrive $T_{CK}/2$ seconds later than those in A. As a result, the

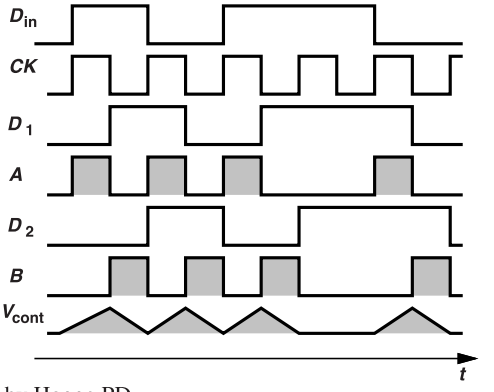


Figure 13.42 Ripple generated by Hogge PD.

loop filter and the control voltage are disturbed twice. In other words, even though it operates as a linear PD, the Hogge topology delivers a large amount of charge to the filter and then retakes it, introducing considerable ripple. This effect can be reduced through the use of staggered reference pulses [7].

The third issue relates to the Hogge PD's finite gain and hence the need for a charge pump so as to force the static phase error to zero. Shown in Fig. 13.43, such a CDR architecture must produce at A and B pulses as narrow as one-half of a bit period in the locked condition. That is, the XOR output nodes must have a wide

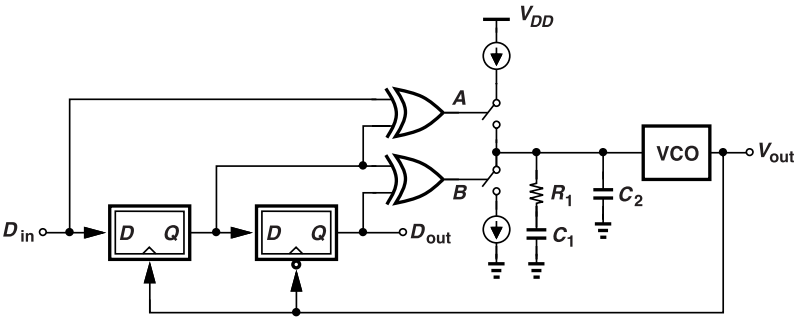


Figure 13.43 CDR loop using Hogge PD and charge pump.

bandwidth and the CP itself must be agile enough. For example, a 25-Gb/s CDR requires that the CP current sources turn on and off in 20 ps. We also observe that, in the locked condition, A and B activate the CP at different times, generating a high ripple.

13.5 Problem of Data Swings

As illustrated in Fig. 13.1(c), actual data suffers from heavy attenuation as it travels through a lossy channel. The small voltage swings thus created pose difficulties with respect to PD and CDR design. We make the following observations for the three PDs studied in this chapter. The single-FF bang-bang PD requires that the data sample the clock, a difficult problem in view of the small data swings. The Alexander PD, on the other hand, allows the clock to sample the data, a more favorable situation as it is simpler to generate large clock swings than large data swings. While sampling the data by the clock, the Hogge PD still demands large data swings for driving its first XOR gate. Thus, the Alexander PD is the most attractive choice at high speeds.

References

[1]. J. Lee and B. Razavi, "Analysis and modeling of bang-bang clock and data recovery circuits," *IEEE J. Solid-State Circuits*, vol. 39, pp. 1571-1580, Sep. 2004.
[2]. J. D. H. Alexander, "Clock recovery from random binary data," *Electronics Letters*, vol. 11, pp. 541-542, Oct. 1975.
[3]. B. Razavi, Y. Ota, and R. G. Swartz, "Design techniques for low-voltage high-speed digital bipolar circuits," *IEEE J. Solid-State Circuits*, vol. 29, pp.332-339, March 1994.
[4]. B. Razavi, K. F. Lee, and R. H. Yan, "Design of high-speed low-power frequency dividers and phase-locked loops in deep submicron CMOS," *IEEE J. Solid-State Circuits*, vol. 30, pp. 101-109, Feb. 1995.
[5]. C. R. Hogge, "A self-correcting clock recovery circuit," *IEEE J. Lightwave Tech.*, vol. 3, pp. 1312-1314, Dec. 1985.
[6]. S. B. Anand and B. Razavi, "A 2.75-Gb/s CMOS clock and data recovery circuit with broad capture range," *ISSCC Dig. Tech. Papers*, pp. 214-215, Feb. 2001.
[7]. L. DeVito et al., "A 52 MHz and 155 MHz clock recovery PLL," *ISSCC Dig. Tech. Papers*, pp. 142-143, Feb. 1991.

Problems

13.1. Repeat the correlation illustrated in Fig. 13.8 if the sine wave has a 45° phase shift: $\sin[(2\pi/T_b)t + \pi/4]$. Is the result still zero?

13.2. In Fig. 13.8, find the correlation between $x(t)$ and the third harmonic, $\sin[3(2\pi/T_b)t]$.

13.3. Consider the waveforms shown in Fig. 13.8(b). Prove that the product, $y(t)$, has a zero average for any phase chosen for the sinusoid.

13.4. Explain how the static phase offset of the PLL in Fig. 13.11 affects the performance.

13.5. Suppose the phase of $x(t)$ in Fig. 13.12 experiences a sudden jump. For example, its rising edge at $t = t_1$ is displaced by a small amount. Explain how this affects $y(t)$ and D_{out}.

13.6. Repeat the previous problem for the topology in Fig. 13.11.

13.7. We wish to compensate for the XOR delay in Fig. 13.13(a). Explain what type of circuit block can be inserted in the path going to the flipflop's D input to perform this compensation.

13.8. Consider the basic NOR-based PFD topology developed in Chapter 7. Assume one input senses an NRZ sequence with a bit period of T_b while the other receives a clock having a frequency of $1/T_b$. Plot the inputs and outputs for a phase difference of 10° and 45° and prove that the average difference between the up and down pulses is independent of the phase error.

13.9. Suppose the random data, $x(t)$, in Fig. 13.9(b) toggles between -1 and $+1$ (rather than between zero and $+1$). Prove that the XOR gate can be replaced with a mixer (an analog multiplier) in this case.

13.10. If FF_1 in Fig. 13.21(a) inverts its D input, does the CDR loop operate properly?

13.11. In the circuit of Fig. 13.26(b), we mistakenly drive FF_4 by $\overline{CK}$. Explain what happens.

13.12. In the Alexander PD of Fig. 13.26(c), we decide to connect Q_2, rather than Q_4, to the XOR gate. Explain what happens to B.

13.13. Let us assume that a CDR loop incorporating the Alexander PD of Fig. 13.26(c) has locked such that the falling edges of CK sample in the middle of the data eye. Plot the circuit's waveforms. What is the minimum pulsewidth observed within this PD?

13.14. Assuming that the flipflops in Fig. 13.26(c) have a slightly longer delay in their clock path than in their data path, repeat the previous problem and show that the clock still samples the data in the middle of the eye.

13.15. The OTA of Fig. 13.29 is used as the V/I converter in Fig. 13.28. First, suppose we insert a small, constant voltage source in series with one of its inputs. Explain what happens to the (average) output phase. Next, determine whether this voltage source experiences a low-pass, band-pass, or high-pass response as it translates to phase noise.

13.16. The propagation delays from E and F in Fig. 13.30 are slightly different. Denoting this skew by ΔT, explain what happens in the CDR loop of Fig. 13.28.

13.17. Suppose we replace the XOR gate in Fig. 13.37(a) with an analog multiplier (a mixer). If all of the circuit's waveforms toggle between -1 and $+1$ (rather than between 0 and 1), draw the output waveform and show that the circuit operates as a synchronous edge detector. This example indicates that a mixer operating with large, differential signals is equivalent to an XOR gate.

13.18. Construct the plot shown in Fig. 13.38 if the D and CK paths of the flipflop have slightly different delays.

13.19. Repeat the previous problem if the two XOR inputs have slightly different delays.

13.20. Does the CDR loop of Fig. 13.43 lock properly if A and B are swapped, i.e., if A serves as the down command and B as the up command?

14

Advanced Clock and Data Recovery Principles

In this chapter, we elevate our study of CDR concepts in several important dimensions. First, we introduce "half-rate" architectures, a variant that proves useful at high speeds. Next, we deal with DLL-based CDR designs and, eventually, arrive at digital CDR implementations. Finally, we analyze the jitter characteristics of CDR circuits.

14.1 Half-Rate Phase Detectors

At very high speeds, it becomes difficult to (1) design a low-noise VCO operating at a frequency equal to the input data rate, (2) distribute such a high-frequency waveform across a chip, and (3) implement flipflops running at the "full rate." We therefore seek PD topologies that operate properly if they sense full-rate data and a slower clock, e.g., one at half of the data rate.

14.1.1 Half-Rate Bang-Bang PDs

In this section, we explore how an Alexander PD can be so modified as to operate with a half-rate clock.

Example 14.1

Can the Alexander PD of Fig. 13.26(c) operate with a half-rate clock?

Solution

Since the Alexander PD relies on both the rising and the falling edges of a full-rate cock, we predict that it fails in the half-rate mode. As illustrated in Fig. 14.1(a), each two consecutive samples taken by the clock,

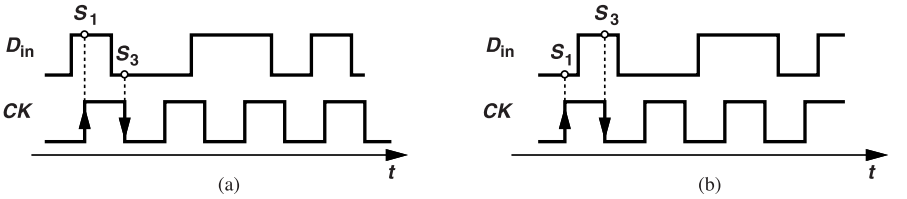


Figure 14.1 Operation of Alexander PD with half-rate clock for (a) late, and (b) early clocks.

e.g., S_1 and S_3, do reveal whether a data transition has occurred, but the "middle" sample (also called the edge sample), S_2, is absent. Consequently, the CDR loop does not know how to position the clock edges. For example, the phase relationship shown in Fig. 14.1(b) yields the same result as does that in Fig. 14.1(a) ($S_1 \oplus S_3 = 1$), implying that the loop cannot find a unique locking point.

The foregoing example suggests that a half-rate Alexander PD would require additional clock edges so as to create the edge sample, S_2. Such a topology can be formed using quadrature phases of the clock, a plausible conjecture because the quadrature phases of the clock together carry as many edges as a full-rate clock. Depicted in Fig. 14.2(a), the idea is to take the three necessary samples around each data transition by

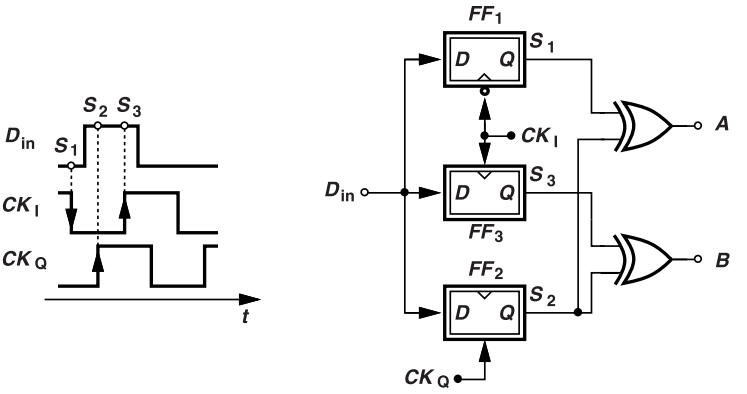


Figure 14.2 (a) Alexander PD operation with quadrature half-rate clocks, and (b) circuit implementation.

means of the falling edge of CK_I, the rising edge of CK_Q, and the rising edge of CK_I. As in the full-rate structure, we compare S_1 with S_2 and S_2 with S_3 so as to determine whether the clock is early or late.

Figure 14.2(b) shows a possible implementation of the half-rate Alexander PD. Driven by CK_I and CK_Q, the three flipflops provide S_1-S_3, which drive the XOR gates as in the full-rate counterpart. When placed in a CDR loop, the circuit positions S_2 around the data transitions, causing FF_2 to become metastable. The other two flipflops then sample D_{in} at optimum points. Note that FF_1 and FF_3 are driven by complementary half-rate clocks, thereby acting as a demultiplexer and producing retimed, *half-rate* data at S_1 and S_3. This PD too exhibits a bang-bang characteristic while producing a zero net output in the absence of data transitions.

Example 14.2

Does the half-rate PD of Fig. 14.2(b) provide information for all input data edges?

Solution

No, it does not. Consider the waveforms shown in Fig. 14.3, noting that the falling edges of CK_Q do not

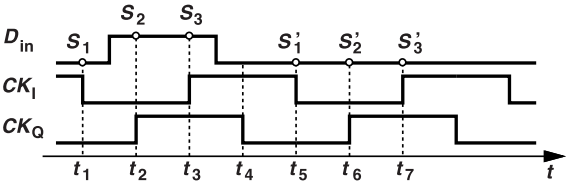


Figure 14.3 Detailed waveforms of half-rate Alexander PD.

sample. Here, S_1-S_3 reveal a transition between t_1 and t_2, and S_1'-S_3' suggest no transition from t_5 to t_7. But D_{in} is not sampled at $t = t_4$ and the data edge between t_3 and t_4 is missed.

In order to detect the missing edge in the foregoing example, we add one more flipflop and drive it by CK_Q [Fig. 14.4(a)]. Here, $A = S_1 \oplus S_2$, $B = S_2 \oplus S_3$, $C = S_3 \oplus S_4$, and $D = S_4 \oplus S_1$, and the results are combined by G_m stages. As shown in Fig. 14.4(b), the average differences between A and B and between C and D represent the phase error.

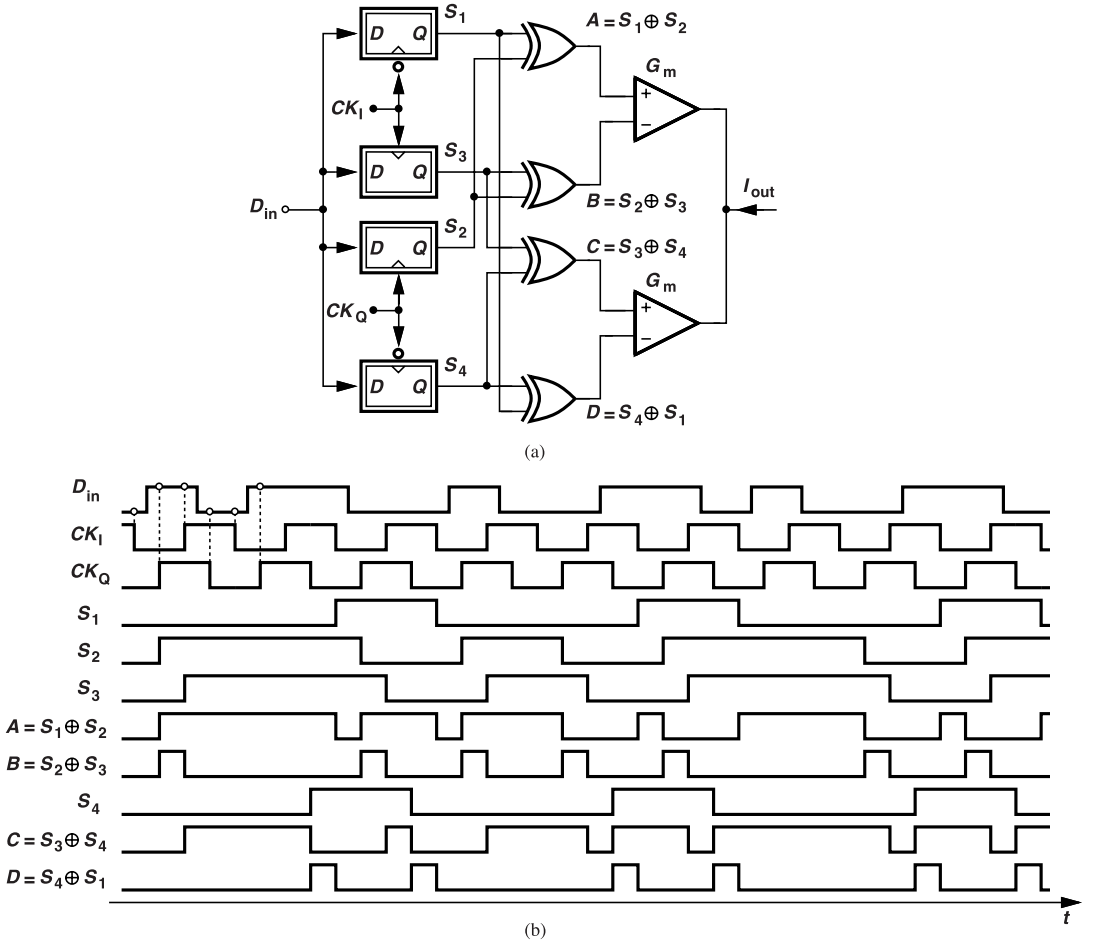


Figure 14.4 (a) Complete half-rate Alexander PD , and (b) its waveforms.

It is interesting to note that, while the full-rate PD waveforms in Fig. 13.27 reveal that A or B remains quiet for an late or early clock, respectively, the half-rate topology in Fig. 14.4(a) still generates fixed-width pulses in B and D. This is because the S_2 and S_3 waveforms suffer from a time offset of $T_{CK}/4$, and so do the S_4 and S_1 waveforms. To ensure that B and D are quiet, these waveforms must be properly aligned by additional flipflops.

The need for quadrature clock phases in the half-rate Alexander PD proves problematic. Since the choice of a half-rate architecture is motivated by speed limitations, it is difficult to operate the VCO at full rate and divide its output frequency by 2 to obtain quadrature phases.[1] One can resort to the quadrature LC oscillator described in Chapter 5, but at the cost of power and area. Also, the input data drives twice as many flipflops in the half-rate PD, experiencing a large capacitive load.

The half-rate Alexander PD becomes attractive in digital CDR circuits, where the quadrature clock phases are derived from the transmitter. It also serves as the foundation for lower clock rates, e.g., $1/4$ or $1/8$ of the data rate. We study these principles in Section 14.2.

[1] In other words, if a $\div 2$ circuit can be designed to operate at the full rate, so can the entire CDR circuit.

14.1.2 Half-Rate Linear PDs

In this section, we study linear half-rate PD circuits. We begin with the Hogge topology.

Example 14.3

Examine the behavior of the Hogge PD if the clock is at half rate.

Solution

As explained in Section 13.4.3, the full-rate Hogge topology relies on both the rising and falling edges of the clock to generate the proportional and reference pulses. With a half-rate clock, on the other hand, half of these edges are absent, causing failure. As exemplified by the waveforms in Fig. 14.5, Q_1 is not a copy of D_{in} and B does not generate a reference pulse for every data transition.

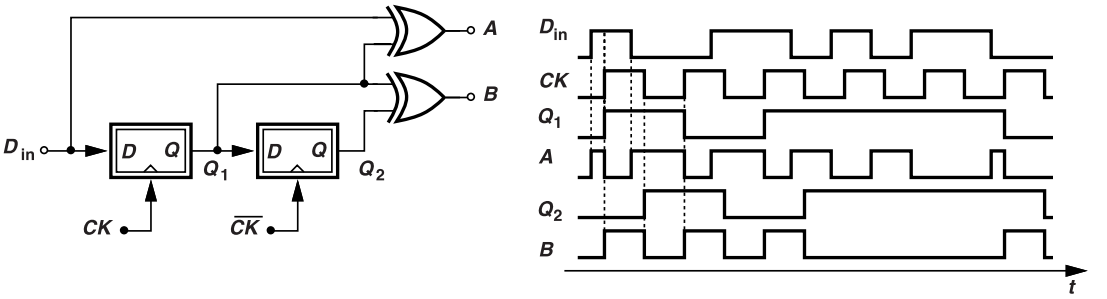


Figure 14.5 Failure of Hogge PD in half-rate operation.

In order to operate the Hogge PD at half rate, we can utilize the quadrature phases of the clock and ensure that all necessary samples of D_{in} are taken. To this end, let us revisit the full-rate topology in Section 13.4.3 and recall from Fig. 13.39 that the rising edges of the full-rate clock sample the data and the falling edges resample the result. Now, consider the waveforms shown in Fig. 14.6(a). To sample every bit of D_{in}, we

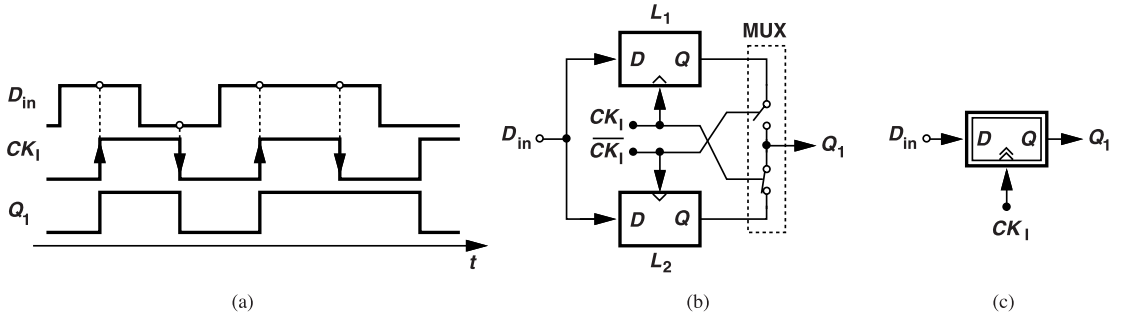


Figure 14.6 (a) Use of half-rate clock to sample data, (b) realization using latches and a MUX, and (c) double-edge-triggered FF symbol.

must utilize both the rising and falling edges of CK_I, and hence ensure Q_1 is a copy of D_{in}. How can this be accomplished? We employ a "double-sampling" or "double-edge-triggered" flipflop. Illustrated in Fig. 14.6(b), such a circuit consists of two latches operating with CK_I and $\overline{CK_I}$ and a multiplexer. When CK_I is high, L_1 is in the sense mode, L_2 is in the hold mode, and the MUX selects the output of the latter. Similarly, when CK_I is high, the output of L_1 is selected. The circuit thus updates Q_1 on both edges of the clock. We denote this FF by the symbol shown in Fig. 14.6(c).

Example 14.4

Using the complementary latch topology in Fig. 13.23, construct a double-edge-triggered flipflop.

Solution

As shown in Fig. 14.7, we incorporate two instances of the latch for sampling the data when CK is high

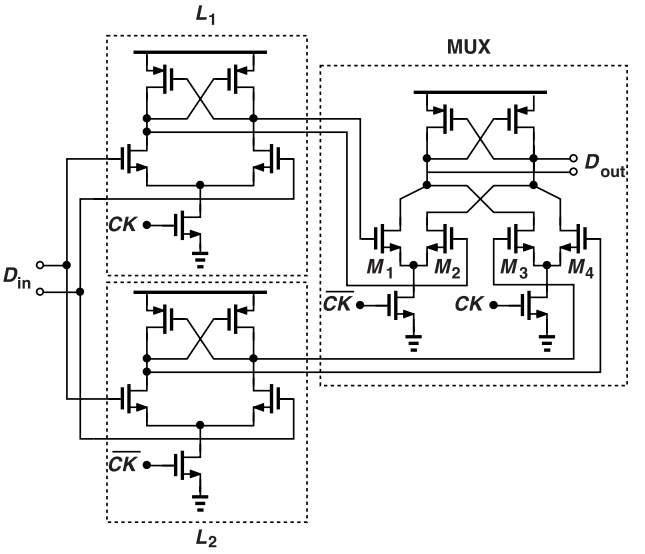


Figure 14.7 Circuit realization of a double-edge-triggered FF.

or low. We also extend the latch topology to form a 2-to-1 multiplexer. Note that L_1 and M_1-M_2 form a master-slave flipflop (why?) and so do L_2 and M_3-M_4.

Must we also use both edges of CK_Q to resample the data? Yes, we must; recall from Fig. 13.39 that the resampling must occur half of a bit period after the first sample. These thoughts lead to the half-rate Hogge PD shown in Fig. 14.8, where FF_1 and FF_2 are double-edge-triggered flipflops. The reader can readily verify

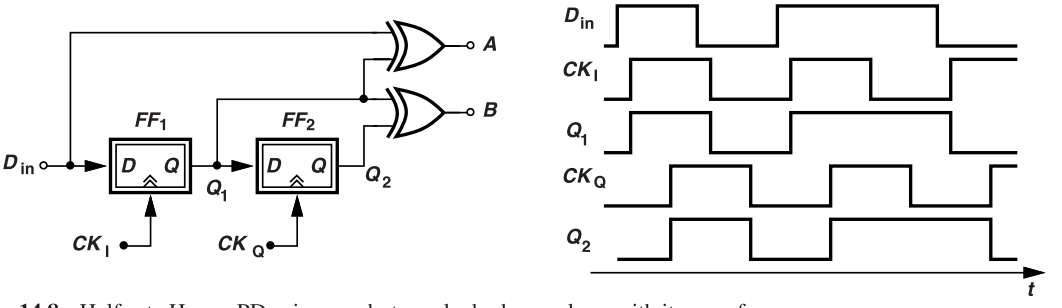


Figure 14.8 Half-rate Hogge PD using quadrature clock phases along with its waveforms.

that FF_1, driven by both edges of CK_I, is equivalent to the first full-rate flipflop in the PD of Fig. 13.39, and similarly, FF_2 is equivalent to the second flipflop in that figure. In other words, both realizations generate the same waveforms for A and the same waveforms for B.

As mentioned earlier, it is difficult to generate quadrature clock phases at high speeds. We now study a half-rate Hogge PD that does not require such phases. We recognize that the circuit must still sample the data on both the rising and falling edges of the clock. Interestingly, if we simply drive two *latches* by complementary half-rate clocks, we can detect the phase error. Depicted in Fig. 14.9(a), the circuit generates two outputs: Q_1, which tracks D_{in} when CK is high and is frozen when CK is low, and Q_2, which does

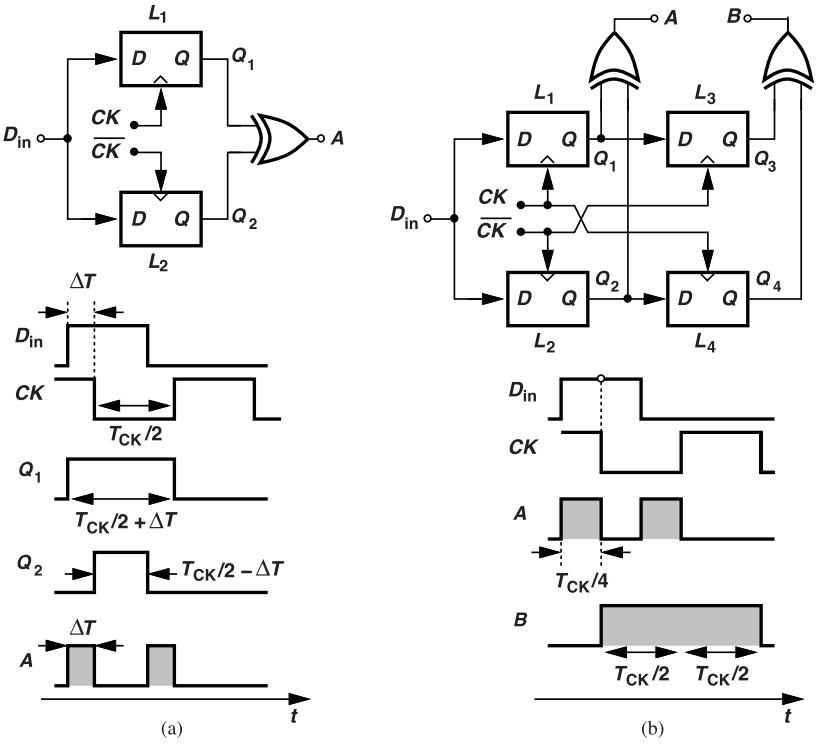


Figure 14.9 (a) Half-rate linear PD without quadrature clocks, (b) complete PD and its waveforms in locked condition.

the same according to $\overline{CK}$. With the phase error example shown here, Q_1 follows D_{in} and rises before CK falls and remains high until the next rising edge of CK, generating a pulse of width $T_{CK}/2 + \Delta T$. On the other hand, Q_2 goes high when CK falls and remains high until D_{in} returns to zero, producing a pulse of width $T_{CK}/2 - \Delta T$. As a result, the XOR gate delivers a pulse of width ΔT for each input data transition; A is therefore the "proportional" output.

In a manner similar to the full-rate counterpart shown in Fig. 13.39, we must also generate a "reference" output. As illustrated in Fig. 14.9(b), this is accomplished by adding two more latches and another XOR gate [1]. Operating as master-slave flipflops while running with a half-rate clock, the L_1-L_3 and L_2-L_4 cascades also demultiplex the data by a factor of two, yielding half-rate data streams at Q_3 and Q_4, respectively. Since Q_3 and Q_4 can change only on opposite edges of the clock, they bear a phase difference of $T_{CK}/2$, which is detected by the second XOR gate. For each data transition, A exhibits a pulse of width $T_{CK}/4$ and B a pulse of width $T_{CK}/2$. The circuit thus acts as both a linear PD and a 1-to-2 DMUX with no need for quadrature clocks. The reader is encouraged to plot the circuit's waveforms for the case where the falling edges of CK land in the middle of the data eye.

Suppose the foregoing half-rate PD operates within a CDR loop. Let us plot the input and output waveforms when the circuit is locked with a zero phase error. As shown in Fig. 14.9(b), for optimum sampling, the falling edge of CK occurs in the middle of the input bits. As a result, the pulses in A and B are $T_{CK}/4$ and $T_{CK}/2$ seconds wide, respectively, yielding unequal averages. This difference can be accommodated if the effect of the reference pulses is scaled down by a factor of two. For example, if the PD is followed by a charge pump, the current source driven by B should be half of that driven by A.

14.2 Oscillatorless CDR Architectures

Some systems incorporate a large number of CDR circuits on one chip. For example, a wireline transceiver supporting 16 I/O "lanes" requires as many CDR loops. In such cases, it is difficult to employ LC oscillators for clock recovery, while ring oscillators may not provide a sufficiently low jitter. In this section, we study several CDR architectures that operate without oscillators of their own.

14.2.1 DLL-Based CDR Circuits

In order to arrive at an alternative CDR implementation, we observe that the *transmitter* in a wireline system also employs a PLL for delivering the clocks needed by the multiplexer. This PLL operates with a crystal reference and achieves low phase noise (Fig. 14.10). We can then ask, is it possible to utilize the TX PLL output, CK_{TX}, in the RX CDR loop so as to eliminate the VCO in the latter? In other words, given this

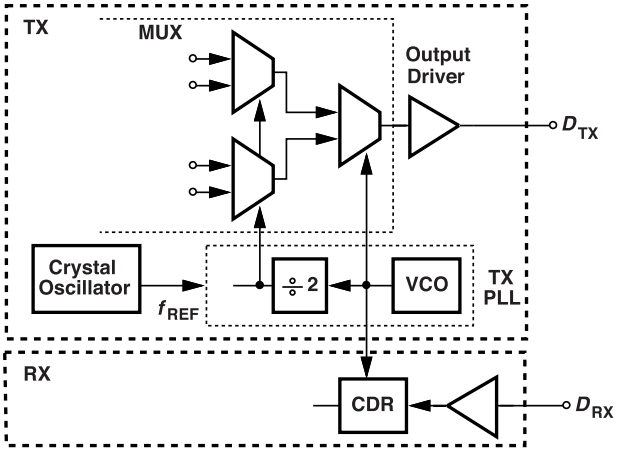


Figure 14.10 Data transceiver architecture.

clock, can we avoid clock recovery and simply retime the received data?

In the first step of our investigation, we consider a solution as simple as one flipflop [Fig. 14.11(a)]. If the TX PLL frequency is equal to the input data rate, then CK_{TX} can retime and clean up D_{in}. However, the

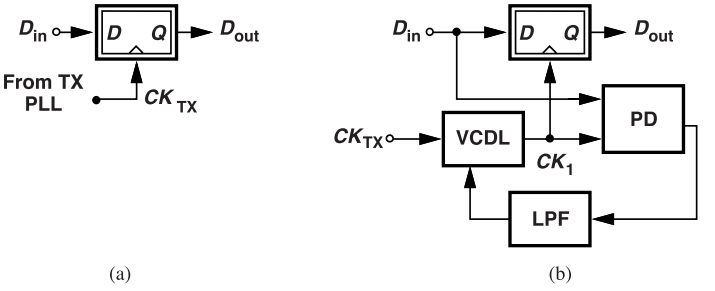


Figure 14.11 (a) Received data retiming using TX clock, and (b) use of a DLL to remove skews.

phase of D_{in} with respect to CK_{TX} depends on the physical length across which D_{in} has traveled (e.g., a few centimeters to tens of centimeters on a backplane) and is unknown. We must therefore adjust the clock (or data) phase to reach the optimum sampling point on D_{in}. This can be accomplished through the use of a phase detector and a variable-delay line, as depicted in Fig. 14.11(b). Here, the flipflop is clocked by CK_1, whose phase is aligned with D_{in} by means of the loop consisting of the PD, the LPF, and the voltage-controlled

delay line. The reader recognizes that these three blocks form a delay-locked loop (Chapter 11), except that our objective is to align CK_1 and D_{in} rather than CK_1 and CK_{TX}. The PD can be based on any of the topologies studied in this and the previous chapters.

Example 14.5

Derive the transfer function from the phase of D_{in} to that of CK_1 in Fig. 14.11(b).

Solution

We note that if D_{in} experiences a phase step, CK_1 does not immediately follow because of the LPF. If the filter has a transfer function given by $1/(1 + s/\omega_{LPF})$, we have from Fig. 14.12

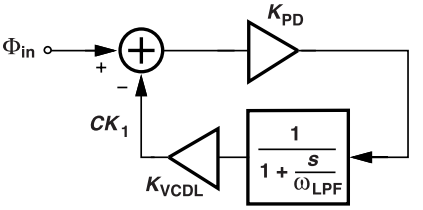


Figure 14.12 DLL model.

$$H(s) = \frac{\dfrac{K_{PD}K_{VCDL}}{1+\dfrac{s}{\omega_{LPF}}}}{1+\dfrac{K_{PD}K_{VCDL}}{1+\dfrac{s}{\omega_{LPF}}}} \tag{14.1}$$

$$= \frac{K_{PD}K_{VCDL}}{\dfrac{s}{\omega_{LPF}} + K_{PD}K_{VCDL} + 1}, \tag{14.2}$$

where K_{VCDL} denotes the gain of the variable delay line. Note the difference between this transfer function and that of the basic DLL in Chapter 11. This low-pass response implies that, if the phase of D_{in} fluctuates slowly, CK_1 can track it, but fast fluctuations are suppressed.

The CDR architecture of Fig. 14.11(b) appears much simpler and more efficient than those developed earlier in this and the previous chapter. In fact, since CK_{TX} is free from data-dependent jitter (why?) and since the DLL produces much less jitter than a PLL, we are tempted to readily adopt this approach over the previous ones. However, a practical issue must be taken into account that calls for considerably higher design complexity. To see this issue, consider the system shown in Fig. 14.13, where two transceivers communicate through a cable. Transceiver 1 employs a PLL with a reference frequency of f_{REF1} to multiplex the data and transmit it to Transceiver 2. The latter utilizes its own TX clock, which is based on f_{REF2}, to perform data recovery according to the implementation in Fig. 14.11(b). The crystal oscillators located in the two transceivers operate at *nominally* equal frequencies, but inevitably exhibit some mismatch (on the order of 100 parts per million). That is, the frequency of CK_{TX2} is not exactly equal to the data rate received from the cable. Since the DLL in Fig. 14.11(b) cannot remove this frequency "offset," the edges of CK_{TX2} drift with respect to the input data transitions, causing sampling errors.

14.2.2 PI-Based CDR Circuits

We wish to reexamine the CDR circuit of Fig. 14.11(b) in the presence of a frequency offset, Δf, between D_{in} and CK_{TX}. This situation is called "plesiochronous operation." A key point that helps understand the operation is described in the following example.

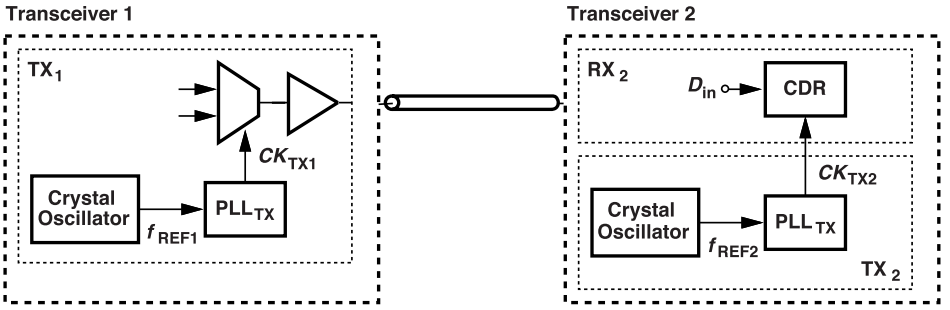


Figure 14.13 Problem of frequency mismatch between two transceivers.

Example 14.6

Can a VCDL generate an output frequency unequal to its input frequency?!

Solution

Yes, it can. Recall that the VCDL output phase is given by $\phi_{out} = \phi_{in} + \phi_0 + K_{VCDL}V_{cont}$, where ϕ_0 corresponds to a minimum, inevitable delay. Suppose the input is periodic and expressed as $V_0 \cos \phi_{in} = V_0 \cos \omega_{in} t$. If V_{cont} is a ramp, αt, then the output emerges as $V_0 \cos(\omega_{in} t + \phi_0 + K_{VCDL}\alpha t) = V_0 \cos[(\omega_{in} t + K_{VCDL}\alpha)t + \phi_0]$. Thus, the output frequency is equal to $\omega_{in} + K_{VCDL}\alpha$. Of course, this would require that V_{cont} either grow linearly toward $+\infty$ or $-\infty$ or be reset every time the phase difference between the input and the output crosses 2π.

We assume that the VCDL provides a delay of ΔT and the PD is followed by a CP so as to obtain an infinite loop gain at dc (Fig. 14.14). Owing to the frequency offset, the phase difference between D_{in} and

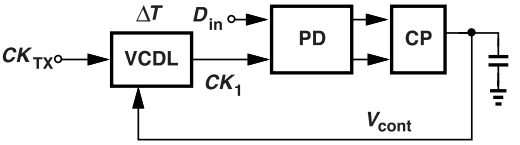


Figure 14.14 A DLL for studying effect of frequency mismatch.

CK_{TX}, $\Delta\phi$, grows linearly with time. But for proper operation, CK_1 must be free from this offset. Thus, the CP provides a ramp to the VCDL so as to keep CK_1 aligned with D_{in}.

The circuit of Fig. 14.14 fails in reality. This is because the CP continues to raise V_{cont} until the delay line "saturates" and then stops—while the phase difference between D_{in} and CK_{TX} grows with no bound. We must therefore seek a delay line implementation that provides an unbounded delay, i.e., one whose value can reach 2π, 4π, etc.

To this end, we make an important observation: a phase interpolator (PI) can also act as a variable delay line but with an unbounded phase shift. We have studied various PI implementations in Chapter 11, but, if we abstract the circuit as shown in Fig. 14.15, we see that (1) CK_1 is a "delayed" copy of $CK_{TX,I}$ (or $CK_{TX,Q}$), and (2) as V_{cont} varies, so does this delay. In other words, with quadrature phases of the TX clock available, we can realize a VCDL in the form of a phase interpolator. The key point here is that the PI can introduce a monotonically increasing phase for CK_1 that grows without saturation. This is possible if CK_1 rotates from $CK_{TX,I}$ to $CK_{TX,Q}$ to $CK_{TX,\overline{I}}$ to $CK_{TX,\overline{Q}}$ to $CK_{TX,I}$ to $CK_{TX,Q}$ and so on. We note that CK_1 must rotate at a rate equal to Δf.

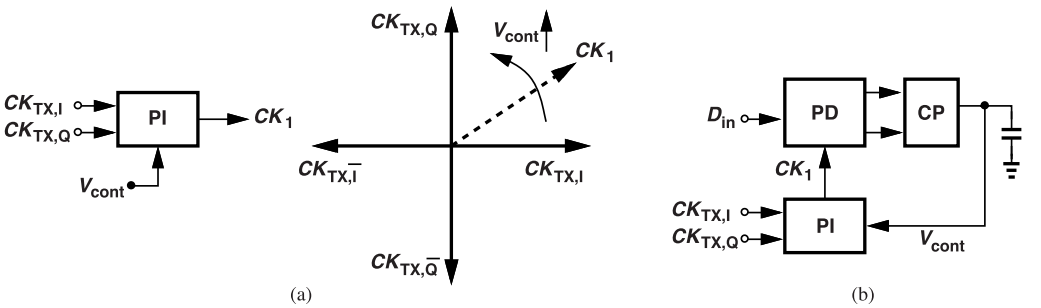


Figure 14.15 (a) Conceptual view of a phase interpolator, and (b) use of a PI in a DLL.

Returning to the architecture of Fig. 14.14, we now replace the VCDL with a PI [Fig. 14.15(b)]. Of course, interpolation between only $CK_{TX,I}$ and $CK_{TX,Q}$ rotates the phase of CK_1 by no more than 90°; we must also interpolate between $\overline{CK_{TX,I}}$ and $CK_{TX,Q}$, etc., so as to obtain a delay as large as one period.

In the presence of a frequency offset between D_{in} and CK_{TX}, the loop in Fig. 14.15(b) must continuously change V_{cont} so as to minimize the phase error between D_{in} and CK_1. But how can this be accomplished while keeping V_{cont} bounded? This point is better understood in the context of digital CDR circuits.

14.2.3 Digital CDR Circuits

Most PI-based (VCO-less) CDR circuits incorporate digital interpolation. The reader is encouraged to review Chapter 11 before starting this section. Consider the PI topology in Fig. 14.16(a), where the differential pairs sensing CK_I and CK_Q switch completely and $V_{cont1} - V_{cont2}$ produces $\alpha CK_I + \beta CK_Q$ at the output. For

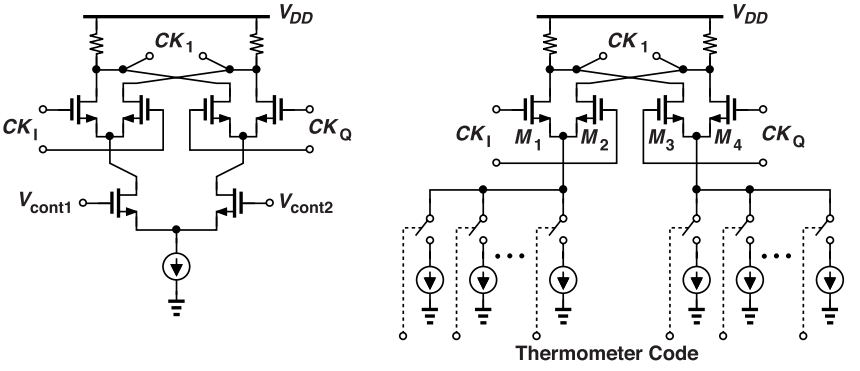


Figure 14.16 PI with (a) analog control, or (b) digital control.

example, if $V_{cont1} - V_{cont2}$ is large and positive, $\alpha = 1$ and $\beta = 0$. We replace the bottom differential pair with digitally-controlled current sources, i.e., a DAC, as shown in Fig. 14.16(b). As more current sources are switched into the tail of M_1 and M_2 and fewer into the tail of M_3 and M_4, α increases and β decreases. As explained in Chapter 11, we wish to keep the sum of the currents drawn from the two differential pairs roughly constant so that the output voltage swing does not change significantly. For guaranteed monotonicity, the current sources are nominally equal and driven by a thermometer code, with the LSB size chosen to provide the necessary phase resolution. We elaboate on these points in Problems 14.10-14.12.

Example 14.7

How is the phase interpolator resolution chosen?

Solution

If the CDR architecture of Fig. 14.15(b) incorporates a digital PI, then the phase of CK_1 can only change in discrete steps. That is, the phase difference between D_{in} and CK_1 cannot be less than the PI's output LSB size. This mechanism produces jitter in CK_1. The PI resolution is thus chosen to achieve sufficiently small phase jumps. For example, for a total peak-to-peak jitter budget of 5 ps, we must ensure a PI resolution well below this value.

It is important to note that the PI of Fig. 14.16(b) covers only the first quadrant, i.e., it can generate an output phase between CK_I and CK_Q. For other quadrants, e.g., for phases between CK_Q and $\overline{CK_I}$, additional differential pairs and DACs are necessary (Example 14.10).

Digital phase interpolation entails two issues. First, due to the low jitter required in most applications, the DAC becomes complex. For example, to obtain a 1-ps resolution at 10 GHz, each tail network in Fig. 14.16(b) must employ more than 100 unit current sources (Problem 14.12). Second, the architecture of Fig. 14.15(b) must be modified such that the PI control input is a *digital* quantity. That is, like the digital PLLs studied in Chapter 10, the design now turns into a "digital CDR" circuit, requiring the PD to act as a time-to-digital converter (TDC) and the LPF to be realized in the digital domain.

Designing a TDC that operates with *random* data appears to be a daunting task, but we can simply utilize a bang-bang PD as a 1-bit TDC. Following this thought, let us return to the idea of a single D flipflop acting as a PD (Section 13.4.1) and first construct the analog bang-bang loop shown in Fig. 14.17(a). Here, the FF output is high if CK_1 is early, causing the charge pump to ramp up V_{cont}. The reverse occurs if the clock is late. The loop therefore places one edge of CK_1 near the transitions of D_{in}.

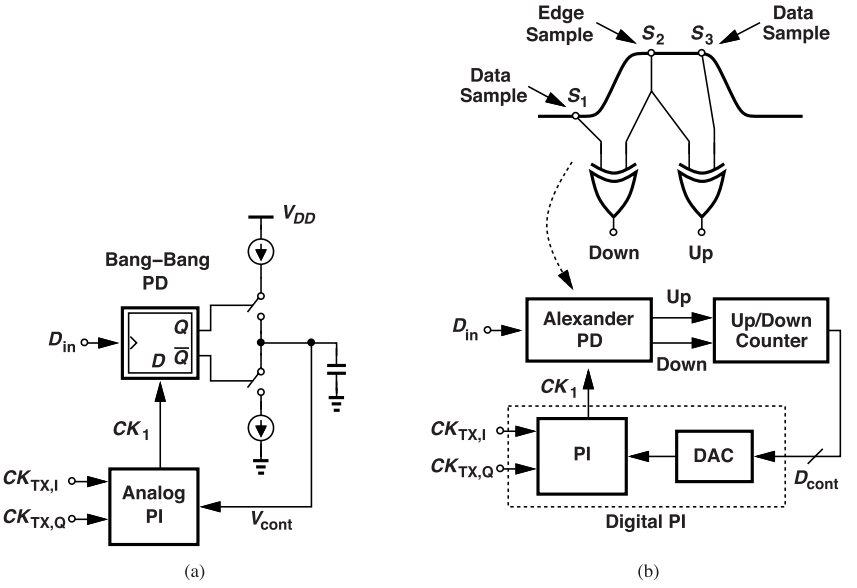


Figure 14.17 CDR loop using (a) an analog PI and a FF PD, or (b) a digital PI and an Alexander PD.

We now emulate the CP up/down behavior in the digital domain: as shown in Fig. 14.17(b), an Alexander PD takes three consecutive samples of D_{in} and generates the "up" and "down" signals according to the perspective illustrated in Fig. 13.36. A binary counter then increments or decrements its output based on the PD's decision, driving the phase interpolator through a DAC. For example, if CK_1 is early, the counter output increases, forcing the PI to rotate the phase toward $CK_{TX,Q}$. We call this architecture a "full-rate" digital CDR loop.

The circuit of Fig. 14.17(b) merits a few remarks. First, the counter behaves as an integrator, accumulating (averaging) the phase error information for the past 2^K phase comparisons, where K is the number of bits in

the counter. Second, K is chosen according to the necessary PI and DAC resolution, typically, in the range of 8 to 10 bits. Third, if we had used a single-flipflop as the bang-bang PD, the loop would act like its analog counterpart in Fig. 14.17(a) and keep increasing or decreasing D_{cont} even in the absence of data transitions. On the other hand, with an Alexander PD, D_{cont} remains constant if there are no input edges.

Example 14.8

How does the topology of Fig. 14.17(b) behave if D_{in} and CK_{TX} have no frequency mismatch?

Solution

At some point in time, the PD decides that CK_1 is early, generates a high level, and commands the counter to count up. As a result, the PI reduces the phase error between D_{in} and CK_1. This continues until CK_1 becomes slightly late, at which point the PD output goes low and the reverse occurs. Thus, the sampling edges of CK_1 fluctuate around the data edges by at least 1 LSB of the phase interpolator. In practice, the input and clock jitter further randomizes this fluctuation.

It is interesting to recognize that the bang-bang PD in Fig. 14.17(b) acts as a 1-bit phase quantizer that, in conjunction with the counter, forms a loop similar to a first-order $\Delta\Sigma$ modulator (Chapter 12). In other words, we can model the circuit as shown in Fig. 14.18(a), where K_{PI} denotes the gain of the phase interpolator (defined as the PI output phase change for a 1-LSB change in its input). Unlike a $\Delta\Sigma$ modulator,

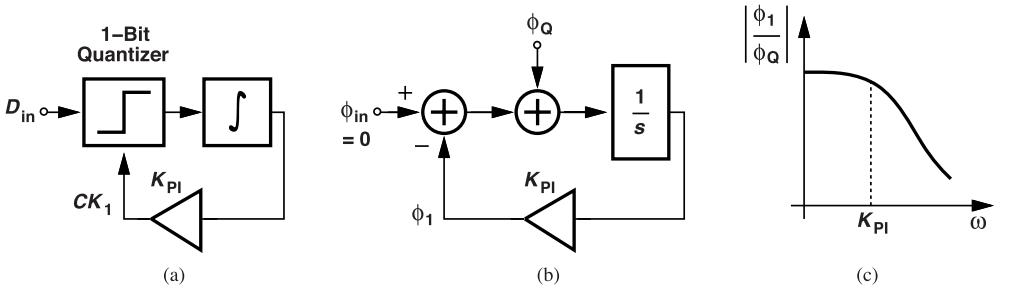


Figure 14.18 (a) Analogy between a bang-bang CDR loop and a $\Delta\Sigma$ modulator, (b) simplified model, and (c) transfer function from quantization noise to recovered clock phase.

this loop places the quantizer before the integrator. Drawing the phase-domain model as in Fig. 14.18(b) and representing the PD quantization noise by ϕ_Q, we obtain the PI output phase as:

$$\phi_1 = \frac{1}{1 + \dfrac{s}{K_{PI}}}\phi_Q. \tag{14.3}$$

That is, ϕ_Q experiences a low-pass response having a corner frequency equal to K_{PI} [Fig. 14.18(c)], as does ϕ_{in}. The result stands in contrast to the high-pass noise shaping of a $\Delta\Sigma$ loop, a difference arising because here the output is taken at the quantizer input rather than at its output.

Example 14.9

Explain intuitively how increasing the PI resolution alters the response in Fig. 14.18(c).

Solution

If the PI resolution and the counter length are increased, then, for a given input phase error, the loop takes *longer* to make the correction because the PI's output phase steps are smaller. Thus, the corner frequency decreases.

We must now address one issue in the loop of Fig. 14.17(b). Which signal clocks the counter? We expect the recovered clock, CK_1, should perform this task. Since it is difficult to design the counter and the DAC for high speeds, we consider a lower clock frequency. However, if driven by a submultiple of the recovered clock frequency, the counter ignores some of the PD output values, in essence losing some information regarding the input transitions.

Let us explore the possibility of using half-rate or quarter-rate clocks in a digital CDR environment. Returning to the half-rate Alexander PD waveforms and the implementation in Fig. 14.4, we recall that the four XOR outputs' average values (in the analog domain) collectively reveal whether the clock is early or late. But in a digital loop, it is difficult to compute these averages. We therefore take a different approach. We first recall that a full-rate Alexander PD can generate "up" and "down" signals by XORing one data sample and the next edge sample, and this edge sample with the next data sample. This principle applies to lower-rate clocks as well if enough clock phases are available. For example, we can replace the full-rate PD in Fig. 14.17(b) with eight flipflops driven by half-quadrature (45°), quarter-rate clock phases [Fig. 14.19(a)]. Here,

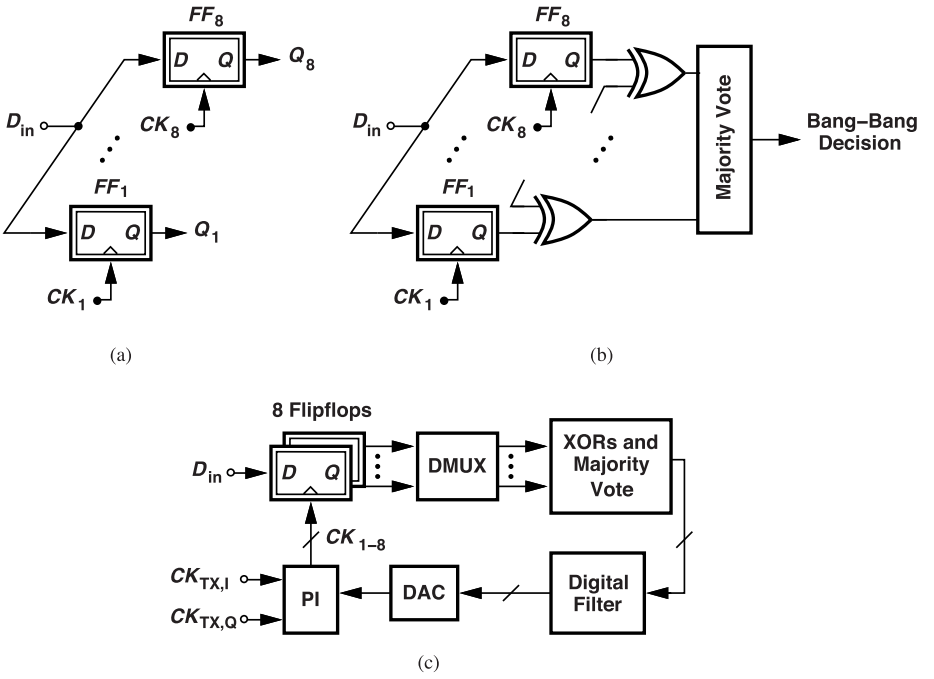


Figure 14.19 (a) Sampling input data by eight clock phases, (b) use of a majority vote block to obtain a decision, and (c) digital CDR loop.

the odd-numbered and even-numbered flipflops generate the data samples and the edge samples, respectively (or vice versa). Thus, $Q_1 \oplus Q_2$, $Q_3 \oplus Q_4$, etc., can act as up commands, and $Q_2 \oplus Q_3$, $Q_4 \oplus Q_5$, etc., as down commands.

The question that remains is how to combine these up and down commands so as to produce a single bang-bang decision for the up/down counter. This can be accomplished through the use of "majority voting," i.e., by simply calculating the number of ups and the number of downs, finding the difference, and incrementing or decrementing the counter accordingly. (The counter does not change state if the two numbers are equal.) Figure 14.19(b) conceptually illustrates this approach. With a simple bang-bang (1-bit) decision, the counter counts up or down by 1 LSB.

In practice, two other modifications are necessary: (1) since most applications replace the counter with a digital loop filter (as in the digital PLLs studied in Chapter 10), the samples must be further demultiplexed so as to ease the speed required of the filter, and (2) the flipflops' outputs must be properly aligned before the XOR operations.

The overall CDR loop is shown in Fig. 14.19(c), where the PI generates CK_1-CK_8 from $CK_{TX,I}$ and $CK_{TX,Q}$. The eight flipflops demultiplex the data by a factor of 4, and the DMUX by another factor of 8 or 16. Note that the digital filter simultaneously adjusts the phases of CK_1-CK_8 based on the majority vote result. As explained below, this architecture must be further modified for plesiochronous operation. We should recognize two drawbacks of this architecture: D_{in} must drive the input capacitance of eight flipflops, and CK_{1-8} suffer from a jitter that is at least equal to the PI's output LSB size.

Plesiochronous Operation Let us return to the issue that originally led us to the use of phase interpolation, namely, the fact that the phase difference between D_{in} and CK_{TX} grows with no bound in the presence of a frequency mismatch. The architectural evolution culminating in the CDR loop of Fig. 14.19(c) has not addressed this effect yet. Specifically, we ask how the PI can continue to rotate all of its output phases. As shown in Fig. 14.20, each PI output must be able to rotate, indefinitely, at a rate equal to the frequency offset.

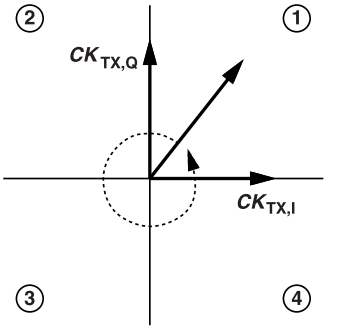


Figure 14.20 Rotation of clock phase through different quadrants.

Considering the PI implementation in Fig. 14.16(b), however, we note that the circuit's output can only go from CK_I (when M_3 and M_4 are off) to CK_Q (when M_1 and M_2 are off). The other three quadrants in Fig. 14.20 are not covered. We call this circuit a "one-quadrant" PI.

In order to deal with this shortcoming, we must incorporate a switching action within the interpolator that changes the *polarity* of each of its clock inputs as the phasor in Fig. 14.20 reaches the boundary of a quadrant. For example, for the phasor to go from the first quadrant to the second, one of the PI inputs must switch from $CK_{TX,I}$ to $CK_{TX,I}$. We thus need "boundary detectors."

Example 14.10

A student reasons that a phase interplator can be configured such that it naturally crosses the quadrant boundaries and hence does not require explicit detectors. Using two Gilbert cells as four-quadrant multipliers, the student draws the circuit as shown in Fig. 14.21. The idea is to interpolate between CK_I and CK_Q when cells A and C are active, between $\overline{CK_I}$ and CK_Q when cells B and C are active, etc. Can this topology operate with no boundary detectors?

Solution

No, it cannot. The difficulty is that, if the digital filter in Fig. 14.19(c) generates an output that only increases or decreases, it cannot directly drive this PI. For example, if the filter simply increments cell A and decrements cell C, we interpolate only between CK_I and CK_Q. We observe that to interpolate between $\overline{CK_I}$ and CK_Q, we must disable A and enable B. Similarly, to interpolate between $\overline{CK_I}$ and $\overline{CK_Q}$, only B and D must remain active. A finite-state machine must follow the filter to keep track of the quadrants and perform the necessary switching.

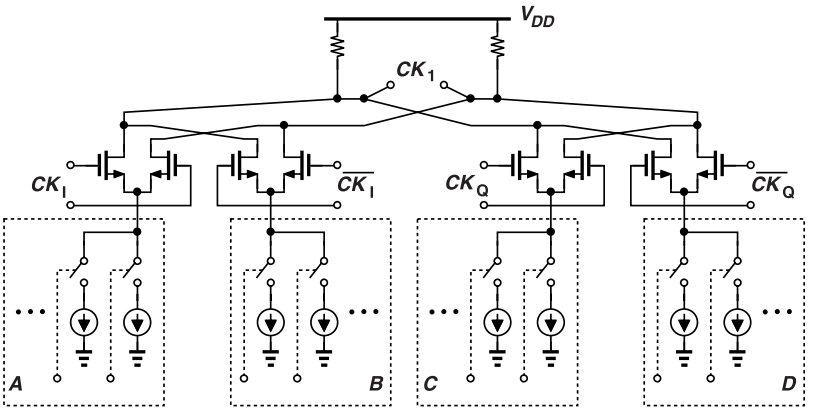
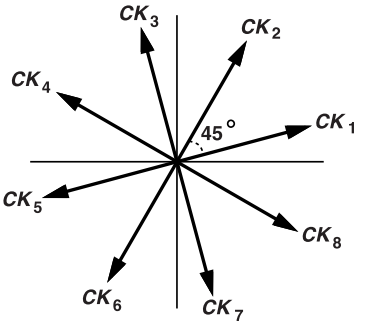
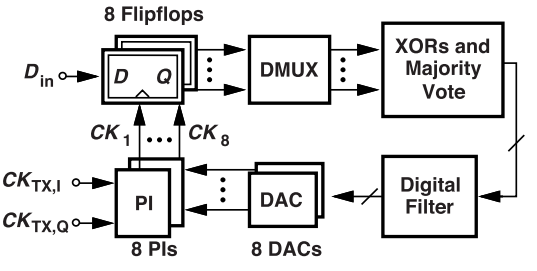


Figure 14.21 Use of two Gilbert cells to implement a four-quadrant PI.

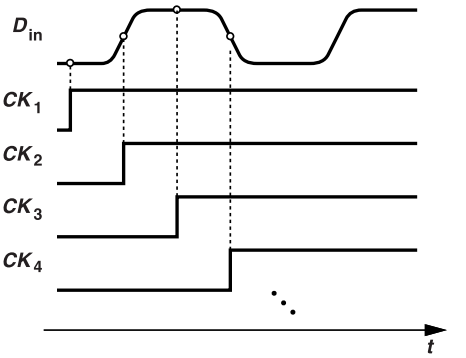
We should now look at the situation in Fig. 14.19(c) more closely: the PI receives two inputs that are $90°$ apart and interpolates between them to achieve sufficient resolution. In plesiochronous operation, the eight PI outputs must continue to rotate. Figure 14.22(a) shows a snapshot at a given moment. The key observation



(a)



(b)



(c)

Figure 14.22 (a) Eight equally-spaced clock phases at some point in time, (b) CDR loop using eight PIs, and (c) input and clock waveforms.

here is that when one clock phase crosses a quadrant boundary, some others do not; thus, a single interpolator does not suffice. These thoughts lead to the arrangement depicted in Fig. 14.22(b), where each sampling flipflop is driven by a dedicated PI. The phase error information provided by the samplers and the majority

vote must therefore control the PIs such that even-numbered (or odd-numbered) flipflops sample near the input data transitions [Fig. 14.22(c)].

14.3 Frequency Acquisition

CDR circuits typically suffer from a relatively narrow bandwidth (equal to a small fraction of the input data rate), because of the communication standard's specifications and/or to suppress the data-dependent ripple on the oscillator's control voltage. Since the phase detectors that operate with random data do not provide frequency error information, additional means are necessary to guarantee lock when the VCO frequency begins far away from the desired value.

Recall from Chapter 7 that a second loop containing a frequency detector (FD) can be added to a phase-locked loop so as to allow frequency acquisition. In a CDR environment, the FD must operate properly with random data, a difficult issue. For example, [2] describes an FD that employs double-edge-triggered flipflops and quadrature clock phases to compare the data rate with the clock frequency. Such "reference-less" CDR circuits prove useful in receivers where a crystal oscillator is not available.

Most systems do have a crystal oscillator, thereby affording a simpler approach to frequency acquisition. Illustrated in Fig. 14.23, the idea is to bring the VCO frequency close to the desired value by means of a

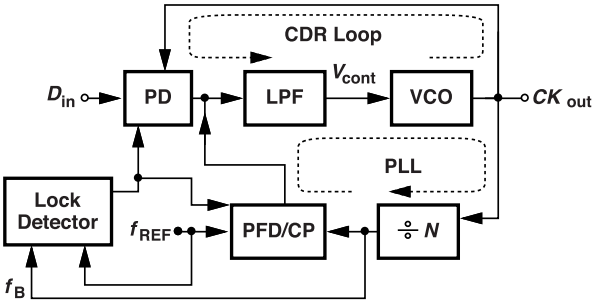


Figure 14.23 Addition of a PLL and lock detector to a CDR loop for frequency acquisition.

simple integer-N PLL while the CDR loop is deactivated, and subsequently disable the former and enable the latter so that the VCO locks to D_{in}. In this example, the two loops meet at the input of the low-pass filter. The operation is as follows. First, the CP is turned on while the PD remains off. The PLL forces the VCO to lock to Nf_{REF}, a condition confirmed by the lock detector (LD). As explained in Section 9.10, the LD measures the difference between f_{REF} and f_B, asserting its output when $|f_{REF} - f_B|$ falls below a certain threshold, e.g., $10^{-4} f_{REF}$. Next, the CP is disabled and the PD is enabled. Now, both the LPF state and the VCO frequency are close to the desired values, i.e., within the acquisition ability of the CDR loop. This loop thus takes over and successfully locks the VCO to D_{in}. If the PD is followed by a charge pump, as in the case of the Hogge topology, then the LD controls both CPs. The design of lock detectors is described in Chapter 9. Note that the two loops are not active simultaneously, avoiding the possibility of a "fight" between them.

Example 14.11

In order to make the VCO in Fig. 14.23 less sensitive to noise on V_{cont}, an engineer decides to decompose the control into fine and coarse branches and include only the latter in the PLL (Fig. 14.24). A large capacitor can be tied from V_{cont2} to ground so as to keep this line quiet. Explain the issue here.

Solution

The difficulty here is that, after the CP is disabled, LPF_2 begins to lose its state due to leakage currents. Consequently, V_{cont2} gradually reaches 0 or V_{DD}, causing the CDR loop to lose lock.

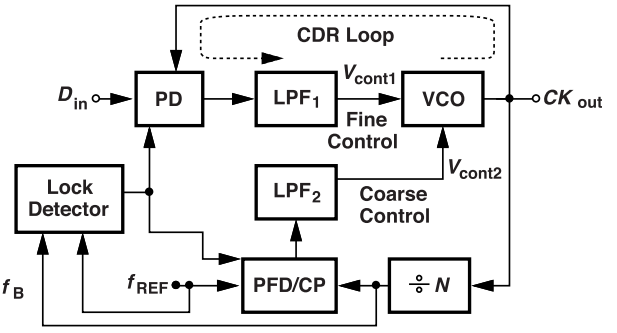


Figure 14.24 Frequency acquisition with fine and coarse VCO controls.

The architecture of Fig. 14.23 merits a few remarks. First, after the switch-over to CDR operation, the lock detector must continue to monitor $f_{REF} - f_B$; if due to, say, a glitch on the supply, the CDR loses lock, the LD must enable the PLL to repeat the frequency acquisition process. Second, the low-pass filter must be designed according to the dynamics and bandwidth required of the CDR loop. Such a design typically suffices for the PLL stability as well; if not, the CP current provides an additional degree of freedom for the PLL's stability.

Example 14.12

Can the value of the capacitor(s) in the LPF of Fig. 14.23 be changed as the loop switch-over occurs?

Solution

While allowing more flexibility in the design, this method faces a serious issue. Suppose, as shown in Fig. 14.25, C_1' is switched into the filter when the CDR loop takes over. Then, due to charge sharing between the

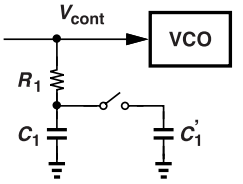


Figure 14.25 Switching a capacitor into the loop filter.

two capacitors, V_{cont} experiences a jump, possibly producing a large frequency shift at the VCO output. The loop may therefore fail to lock.

Another approach to ensuring lock draws upon the discrete VCO calibration techniques described in Chapters 8 and 9. If the VCO frequency is digitally set to be close to the desired value, then the CDR always begins with a small frequency error and, therefore, locks. This is illustrated in Fig. 14.26. Of course, to set the digital control properly, a divider and a lock detector may still be required.

14.4 Jitter Characteristics

Since CDR circuits sense a random, possibly jittery input, their jitter characteristics are more complex than those of frequency-multiplying PLLs. In this section, we study three aspects of CDR jitter behavior. In this

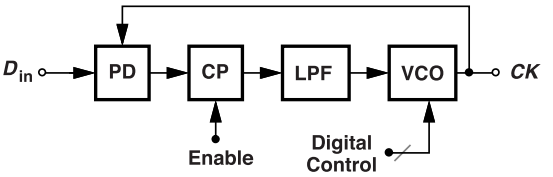


Figure 14.26 Digital setting of VCO frequency to ensure CDR lock.

context, we shall refer to the input data period as the "unit interval" (UI). The reader is encouraged to review the PLL transfer functions derived in Chapters 7 and 8.

14.4.1 Jitter Generation

As with other PLLs, a CDR loop generates output jitter even with a jitterless input . We can identify several sources: the VCO intrinsic phase noise, ripple on its control voltage due to random input data (i.e., pattern-dependent jitter), and supply and substrate noise. The first is shaped according to the loop bandwidth and order (Chapter 8) while the second also depends on the type of phase detector (Section 13.4). For example, a single flipflop acting as a bang-bang PD generates a large amount of ripple. Thus, even if the communication standard allows a wide loop bandwidth, the pattern-dependent jitter may still demand a narrow one.

The difficulty with the jitter generated by a CDR circuit is that it reduces the budget for other imperfections in the receiver. As illustrated in Fig. 14.27, D_{in} itself suffers from slow edges—due to the limited bandwidth

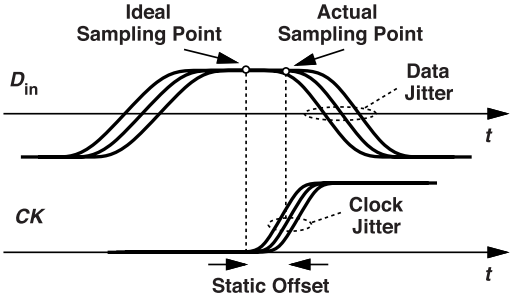


Figure 14.27 Data waveform showing effect of jitter and clock phase offset.

of the receive path—and random jitter. In addition, the clock edge incurs some static offset resulting from the loop nonidealities. The clock jitter exacerbates the situation, especially when the jitter reaches several times its rms value. As a rule of thumb, we aim for a generated rms jitter of about 0.01 UI in the recovered clock.

14.4.2 Jitter Transfer

In addition to aligning the recovered clock with the input data, a CDR circuit should also act as a narrowband filter and remove most of the input jitter. That is, the data jitter is filtered by the closed-loop input-output transfer function, simply called the "jitter transfer." A unique property of phase-locked loops is that they can realize a jitter (phase) filter having an arbitrarily narrow bandwidth for an arbitrarily high input frequency. This high-Q action can be chosen to greatly suppress the input jitter before it reaches the recovered clock output. However, as explained in the next section, jitter "tolerance" requirements work against such a choice.

For a type-II CDR circuit, the closed-loop jitter transfer is expressed as

$$H(s) = \frac{2\zeta\omega_n s + \omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2}. \tag{14.4}$$

This transfer function quantifies how jitter in the input data propagates to the recovered clock and the retimed data. We have studied this response in Chapter 7 but will analyze it further here in the context of clock and data recovery. It is important to note that both ζ and ω_n depend on the data "transition density," D_T, here. We define D_T (a dimensionless quantity) as the number of data transitions in unit time divided by the bit rate. For example, if ONEs and ZEROs occur with equal probabilities and the PD detects both the rising and the falling edges, then $D_T = 0.5$. The dependence arises because the phase detectors studied in this chapter produce a net output only when data transitions occur, exhibiting a gain proportional to D_T. This can be taken into account by multiplying the PD or the VCO gain by D_T:

$$\zeta = \frac{R_1}{2}\sqrt{\frac{I_p D_T K_{VCO} C_1}{2\pi}} \tag{14.5}$$

$$\omega_n = \sqrt{\frac{I_p D_T K_{VCO}}{2\pi C_1}}. \tag{14.6}$$

Let us reexamine the 3-dB bandwidth and the problem of jitter peaking. Recall from Chapter 7 that

$$\omega_{-3dB}^2 = [2\zeta^2 + 1 + \sqrt{(2\zeta^2+1)^2+1}]\omega_n^2. \tag{14.7}$$

In practice, we wish to choose a narrow CDR bandwidth so as to minimize the pattern-dependent jitter. How do we reduce ω_{-3dB} without degrading the stability? If ζ is greater than unity, we have

$$\omega_{-3dB} \approx 2\zeta\omega_n \tag{14.8}$$

$$\approx \frac{R_1 I_p D_T K_{VCO}}{2\pi}. \tag{14.9}$$

Thus, R_1 can be lowered to reduce ω_{-3dB} while the loop filter capacitor is increased to ensure a proper value for ζ. For this reason, CDRs often require large capacitors in their filters.

Since, for a given ζ, the bandwidth can be reduced only by decreasing ω_n, we can return to the transfer function in Eq. (14.4), neglect the term ω_n^2 for some range of s, and write

$$H(s) \approx \frac{2\zeta\omega_n}{s + 2\zeta\omega_n}. \tag{14.10}$$

That is, the loop can be approximated by a *one-pole* system having a 3-dB bandwidth of $2\zeta\omega_n$. This is expected because the zero and the first pole almost coincide for a large ζ (Chapter 7). We surmise that this approximation holds more accurately at higher jitter frequencies (why?).

Example 14.13

Determine the range of ω across which the one-pole behavior is observed.

Solution

We must examine the magnitude of the transfer function in Eq. (14.4). We have

$$|H(j\omega)|^2 = \frac{4\zeta^2\omega_n^2\omega^2 + \omega_n^4}{(\omega_n^2 - \omega^2)^2 + 4\zeta^2\omega_n^2\omega^2}. \tag{14.11}$$

For this to approximate the magnitude squared of (14.10), $4\zeta^2\omega_n^2/(\omega^2 + 4\zeta^2\omega_n^2)$, we must make ω_n^4 negligible in both the numerator and the denominator. A sufficient condition for this is $4\zeta^2\omega_n^2\omega^2 \gg \omega_n^4$ and hence

$$\omega^2 \gg \frac{\omega_n^2}{4\zeta^2}. \tag{14.12}$$

In other words, for ω greater than roughly $3\omega_n/(2\zeta)$, the transfer function assumes a first-order shape. Note that this value is typically below the 3-dB bandwidth, $2\zeta\omega_n$.

Jitter Peaking Some systems process the data in a cascade of CDR circuits before delivering it to the final destination. Long-haul optical transceivers, for example, incorporate "repeaters" every hundreds of kilometers so as to retime and clean up the data. In such cases, the transfer functions of the CDR circuits in the data path are multiplied, making jitter peaking more significant.

Let us return to the jitter transfer in Eq. (14.4) and note that:

$$\omega_z = -\frac{\omega_n}{2\zeta} \tag{14.13}$$

$$\omega_{p1,2} = (-\zeta \pm \sqrt{\zeta^2 - 1})\omega_n \tag{14.14}$$

$$= \left(-1 \pm \sqrt{1 - \frac{1}{\zeta^2}}\right)\zeta\omega_n. \tag{14.15}$$

This expression can be simplified if $\zeta^2 > 2$. A good approximation is to write $\sqrt{1-\epsilon} \approx 1 - \epsilon/2 - \epsilon^2/8$ for $\epsilon < 0.5$. It follows that

$$\omega_{p1,2} \approx \left[-1 \pm \left(1 - \frac{1}{2\zeta^2} - \frac{1}{8\zeta^4}\right)\right]\zeta\omega_n, \tag{14.16}$$

and hence

$$\omega_{p1} \approx -\frac{\omega_n}{2\zeta} - \frac{\omega_n}{8\zeta^3} \tag{14.17}$$

$$\omega_{p2} \approx -2\zeta\omega_n + \frac{\omega_n}{2\zeta} + \frac{\omega_n}{8\zeta^3}. \tag{14.18}$$

These results provide a great deal of insight. As shown in Fig. 14.28, we observe that (a) the zero appears

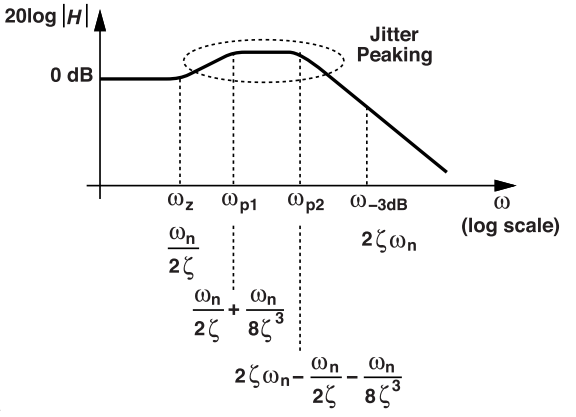


Figure 14.28 Jitter peaking.

before the first pole, inevitably causing jitter peaking; (b) ω_z and ω_{p1} almost coincide if $8\zeta^3 \gg 2\zeta$, i.e., if $4\zeta^2 \gg 1$; this is the only parameter that controls jitter peaking; (c) ω_{p2} and ω_{-3dB} have a difference equal to ω_{p1} and approach each other as ζ increases.

For ζ values greater than approximately 2, ω_{p2} is much greater than ω_{p1}, negligibly affecting the response between ω_z and ω_{p1}. Thus, $|H(j\omega_{p1})|/|H(j\omega_z)| \approx \omega_{p1}/\omega_z$, which is the amount of jitter peaking, J_p. From

(14.13) and (14.17),

$$J_p \approx 1 + \frac{1}{4\zeta^2}. \tag{14.19}$$

One can argue that only the term $1/(4\zeta^2)$ should be called jitter peaking, but the expression $1 + 1/(4\zeta^2)$ better lends itself to decibel calculations. For example, $\zeta = 2$ yields 0.53 dB of peaking.

Another phenomenon that can introduce jitter peaking relates to "excess delay" in the loop. For example, the frequency divider in a simple PLL exhibits a finite delay, T_{div}, causing the loop gain to be multiplied by $\exp(-T_{div}s)$. This factor simplifies to $1 - T_{div}s$ for loop dynamics that are much longer than T_{div} (Chapter 15). The resulting right-half-plane zero degrades the phase margin. Similarly, the delay and latency associated with the TDC and the loop filter in digital PLLs can create jitter peaking. In a digital CDR circuit, the demultiplexer, the counter, and the PI contribute excess delay.

14.4.3 Jitter Tolerance

Our analysis of jitter transfer in the previous section was based on the general understanding that a CDR circuit should have a narrow bandwidth so as to suppress the input data jitter and generate a "clean" clock. In reality, however, another interesting effect enters the picture that requires a *wide* bandwidth.

We know that random data incurs jitter as it begins from a transmitter, travels through a (lossy) medium, and arrives at a receiver. The jitter contains various frequency components, i.e., some components produce slow phase fluctuations and some, fast. For a given CDR bandwidth, very low jitter frequencies reach the recovered clock with little attenuation. Let us examine the loop's behavior under this condition. As shown in Fig. 14.29(a), even though the bit period fluctuates and the data transitions deviate from their ideal points in

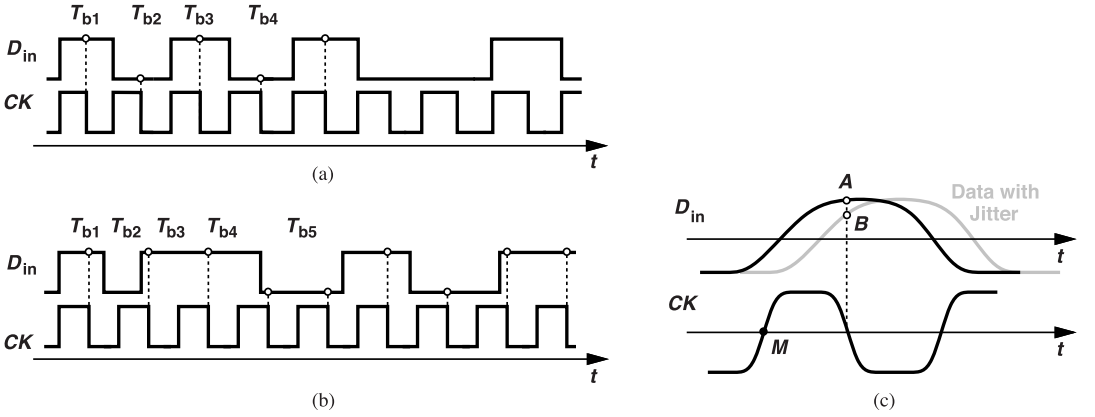


Figure 14.29 (a) Phase tracking in the case of slow jitter, (b) phase tracking in the case of fast jitter, and (c) effect of data jitter on the sampled value.

time, the recovered clock "keeps up" with the data, adjusting its phase to ensure sampling in the middle of the eye. In other words, it is *fortunate* that the input jitter is inherited by the clock!

As another scenario, suppose the input contains jitter frequencies beyond the CDR bandwidth. Filtered by the circuit, such components do not affect the recovered clock [Fig. 14.29(b)]. That is, the clock edges continue to sample at evenly-spaced points in time, occasionally coming dangerously close to the data edges. As depicted in the more realistic waveforms in Fig. 14.29(c), the clock samples at point B rather than point A in the presence of data jitter. The clock jitter further exacerbates this effect. We conclude that, for sufficiently fast input phase fluctuations, the CDR suffers from a high error rate because the clock edge sometimes lands near the data zero crossings. In such a case, the instantaneous phase error between the data edge and the proper edge of the clock grows to as much as 0.5 UI.

The foregoing behavior of CDR loops is quantified by the "jitter tolerance," J_{tol}, i.e., the maximum amount of input jitter that the circuit can tolerate before it begins to incur errors. This quantity represents the ultimate

performance as it embodies (1) the loop's agility, (2) the input jitter, (3) the input rise and fall times, (4) the clock jitter, (5) the clock rise and fall times, and (6) the static phase offset. We expect J_{tol} to be large for slow jitter components— the CDR can maintain a small phase error—and small for fast phase fluctuations. To derive the corresponding response, we recognize from Fig. 14.29(c) that the loop frequently samples incorrectly as the phase difference between D_{in} and the proper edge of the clock approaches 0.5 UI. The maximum tolerable phase error is therefore expressed as

$$\phi_{in} - \phi_{out} < \pm \frac{1}{2} \text{ UI}, \tag{14.20}$$

where ϕ_{out} denotes the recovered clock phase corresponding to that edge which is nominally aligned with the data transitions, namely, point M in Fig. 14.29(c). Since ϕ_{out}/ϕ_{in} is given by the jitter transfer function, $H(s)$, we have $\phi_{in}|1 - H(s)| < 0.5$ UI. It follows that

$$\phi_{in} < \frac{0.5 \text{ UI}}{|1 - H(s)|}. \tag{14.21}$$

This inequality means that the input phase fluctuation, ϕ_{in}, must remain less than 0.5 UI/$|1 - H(s)|$ to avoid errors.[2] For fast jitter components, the right-hand side of (14.21) decreases in magnitude, imposing a smaller tolerable peak for ϕ_{in}. In fact, replacing for $H(s)$ from (14.4) yields

$$\phi_{in} < \frac{1}{2} \frac{s^2 + 2\zeta\omega_n s + \omega_n^2}{s^2}. \tag{14.22}$$

We denote the transfer function on the right hand side by $G_{JT}(s)$ and call it the jitter tolerance:

$$G_{JT}(s) = \frac{1}{2} \frac{s^2 + 2\zeta\omega_n s + \omega_n^2}{s^2} \text{ UI}, \tag{14.23}$$

which contains two zeros equal to the poles of the jitter transfer and two poles at the origin. In other words, for a given jitter frequency, $|G_{JT}|$ gives the maximum input jitter amplitude that can be tolerated. Plotted in Fig. 14.30, this function falls at -40 dB/dec up to ω_{p1} and at -20 dB/dec between ω_{p1} and ω_{p2}. Beyond ω_{p2}, $|G_{JT}|$ asymptotically approaches 0.5.

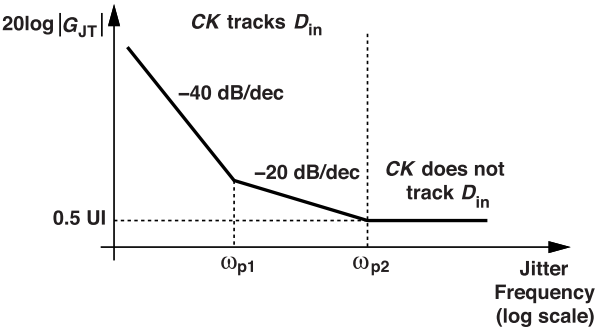


Figure 14.30 Jitter tolerance plot.

Example 14.14

How is the plot of Fig. 14.30 measured in the laboratory?

[2] In practice, due to the other imperfections mentioned above, the maximum tolerable ϕ_{in} is smaller.

Solution

In the jitter tolerance test, the random data contains *sinusoidal* jitter, with a jitter frequency corresponding to each point on the horizontal axis of Fig. 14.30. This is accomplished by modulating the phase of the reference frequency used by the data generator. As illustrated in Fig. 14.31, the data is provided by a pseudo-random binary sequence (PRBS) generator, which itself is driven by a periodic clock, CK_1. The RF signal generator

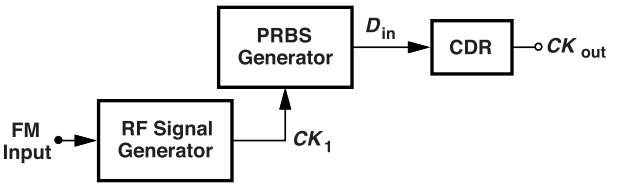


Figure 14.31 Jitter tolerance test setup.

producing this clock typically has an FM input, which can sense a sinusoid to modulate the clock phase and hence the data phase. We can express CK_1 as $A_0 \cos[\omega_0 t + k \int V_m(t)]$. For example, an FM input of the form $A \sin \omega_{p1} t$ modulates the phase of D_{in} at a rate equal to ω_{p1} and tests the CDR tolerance in Fig. 14.30 at $\omega = \omega_{p1}$.

From this study, we surmise that CDR circuits must be designed for great agility, i.e., a wide bandwidth, so that the clock edge can track the data edge in the presence of fast input jitter. However, such a remedy entails other issues that are discussed below.

It is instructive to estimate $|G_{JT}(s)|$ at $\omega = \omega_{p1}$. From (14.23), we have

$$|G_{JT}(j\omega_{p1})|^2 = \frac{1}{4} \frac{(\omega_n^2 - \omega_{p1}^2)^2 + 4\zeta^2 \omega_n^2 \omega_{p1}^2}{\omega_{p1}^4}. \tag{14.24}$$

As explained in Section 14.4.2, $|\omega_{p1}| \approx \omega_n/(2\zeta) + \omega_n/(8\zeta^3)$, which reduces to $\omega_n/(2\zeta)$ if ζ is greater than roughly 1.5. It follows that $|G_{JT}(j\omega_{p1})|^2 \approx 8\zeta^4$ and hence

$$|G_{JT}(j\omega_{p1})| \approx 2\sqrt{2}\zeta^2. \tag{14.25}$$

For example, if $\zeta = 1.5$, we have $|G_{JT}(j\omega_{p1})| \approx 6.36$ UI. That is, the input can contain at most 6.36 UI of peak jitter at this jitter frequency. The reader can also show that the same approximations lead to $|G_{JT}(j\omega_{p2})| \approx 0.5$ UI, as predicted by the plot in Fig. 14.30.

The jitter tolerance profile expressed by (14.23) and plotted in Fig. 14.30 places a fundamental limit on the performance of CDR circuits, revealing that the clock jitter, the static phase offset, and the finite rise and fall times must be small enough to allow a peak input jitter of 0.5 UI for $\omega > \omega_{p2}$. Indeed, it is this frequency range in which typical CDR loops perform poorly and suffer from a low tolerance.

In order to improve the jitter tolerance, we can increase the loop bandwidth. From Fig. 14.28, we have $\omega_{p1} \approx \omega_n/(2\zeta)$ and $\omega_{p2} \approx 2\zeta\omega_n$, recognizing that increasing ω_n pushes both ω_{p1} and ω_{p2} to higher values, whereas raising ζ lowers ω_{p1} and increases ω_{p2}. Since $G_{JT}(s) \approx \omega_n^2/(2s^2)$ for small s values, we arrive at the behavior shown in Fig. 14.32(a) if ω_n increases and ζ remains constant. In this case, the tolerance improves for all frequencies below ω'_{p2}. On the other hand, if ζ rises and ω_n is constant, we obtain a greater tolerance for $\omega_{p1} < \omega < \omega'_{p2}$ [Fig. 14.32(b)].

The scenario depicted in Fig. 14.32(a) appears particularly promising: we can simply increase ω_n (by decreasing the main loop filter capacitor) and keep ζ constant (by increasing the loop filter resistor). This method, however, faces two issues: (1) a higher ripple on the control voltage and hence greater pattern-dependent jitter, and (2) a wider jitter transfer bandwidth, which is objectionable in some applications, e.g., in optical communication repeaters. In other words, the bandwidth must be optimized with respect to all of

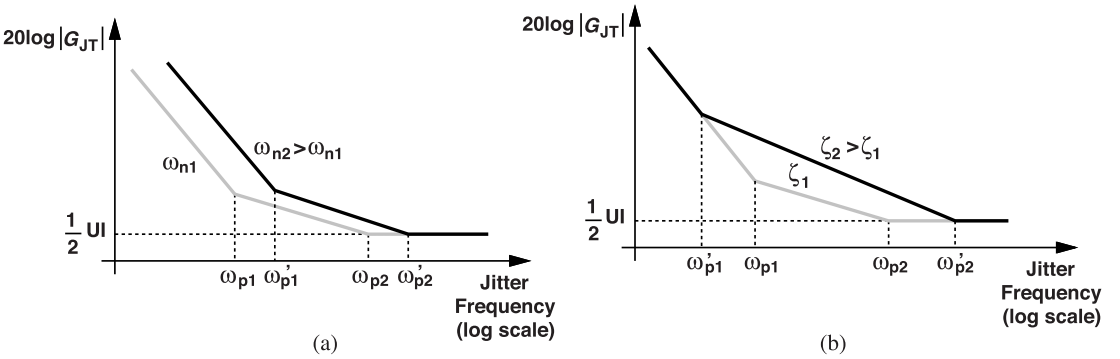


Figure 14.32 Effect of changing (a) ω_n, or (b) ζ on jitter tolerance.

these constraints. We should also point out that jitter peaking due to excess delay (Section 14.4.2) can cause a dip in the jitter tolerance response.

References

[1]. J. Savoj and B. Razavi, "A 10-Gb/s CMOS clock and data recovery circuit with a half-rate linear phase detector," *IEEE J. of Solid-State Circuits,* vol. 36, pp. 761-768, May 2001.

[2]. A. Pottbacker, U. Langmann, and H. U. Schreiber, "A Si bipolar phase and frequency detector for clock extraction up to 8 Gb/s," *IEEE J. Solid-State Circuits,* vol. 27, pp. 1747-1751, Dec. 1992.

Problems

14.1. As explained in Chapter 13, a D flipflop can act as a PD with random data. Can it also operate properly with a half-rate clock?

14.2. Prove that an XOR gate sensing random data and a half-rate clock does not provide an output average proportional to the input phase error.

14.3. If the rising edge of CK_Q in Fig. 14.2(a) is offset from its point in time by ΔT, determine how a CDR loop employing the PD of Fig. 14.2(b) locks and retimes the data.

14.4. Does the analysis in Example 14.2 imply that the half-rate outputs, S_1 and S_3 in Fig. 14.2(b), also miss some data samples?

14.5. Compare the total capacitance driven by the clock(s) in the half-rate PDs of Fig. 14.2(b) and Fig. 14.8.

14.6. Suppose the supply voltage of the VCDL in Fig. 14.11 fluctuates with time. Explain what happens to the phase of CK_1 and D_{out} in the presence of slow or fast fluctuations.

14.7. Suppose the phase of D_{in} in Fig. 14.11 fluctuates slightly with time. Explain what happens to the phase of CK_1 and D_{out} in the presence of slow or fast fluctuations.

14.8. Can the VCDL in Fig. 14.14(a) (along with the rest of the loop) be considered a circuit that changes the frequency of CK_{TX} by Δf as it delivers CK_1? Can you think of any other circuit that can perform this operation?

14.9. We have assumed that the loop in Fig. 14.14(a) nulls the phase difference between D_{in} and CK_1. Does this assumption hold for large frequency offsets? Estimate the maximimum frequency offset at which the phase error remains small.

14.10. In the PI of Fig. 14.16(b), each tail node is tied to 16 unit current sources of value I_u. We begin with $16I_u$ on the left and $0 \times I_u$ on the right so that $CK_1 = CK_I$. Now, we turn off one I_u on the left and turn on one I_u on the right. Determine the phase change in CK_I.

14.11. Repeat the previous problem if we begin with $8I_u$ on the left and $8I_u$ on the right. Is the phase change in this case equal to that found in the previous problem?

14.12. The PI of Fig. 14.16(b) must provide an output phase resolution of 1 ps for a 10-Gb/s CDR circuit. Determine the minimum number of unit current sources. Note that this number must be multiplied by 4 to cover all quadrants (Example 14.10). (Hint: first, solve the previous two problems.)

14.13. We double K_{VCO} in a CDR design. Does the approximate transfer function expressed by Eq. (14.10) become more accurate or less?

14.14. Repeat the previous problem if we double the main capacitor in the loop filter.

14.15. Consider the jitter transfer plotted in Fig. 14.28. Explain in detail which values change if K_{VCO} is doubled.

14.16. In Fig. 14.28, is it possible to increase ω_z while keeping the amount of jitter peaking constant?

14.17. In Fig. 14.28, is it possible to increase ω_{p1} without changing the amount of jitter peaking?

14.18. Plot the jitter transfer response of Fig. 14.28 for $\zeta = 4$ and $\zeta = 6$. Assume only the loop filter resistor is adjusted to change ζ.

14.19. Explain what happens to the jitter tolerance behavior depicted in Fig. 14.30 if K_{VCO} is doubled.

14.20. Repeat the previous problem if the main capacitor in the loop filter is doubled.

15

Frequency Dividers

We have seen in the previous chapters that many phase-locked systems incorporate frequency dividers. Such dividers scale the VCO frequency down so that it becomes equal to the reference frequency. In some applications, the feedback divider has a constant modulus while in others, e.g., in RF synthesizers, the modulus must change in steps of unity. This chapter deals with the analysis and design of various divider topologies. We study static and dynamic latches, divide-by-2 circuits, and dual-modulus prescalers. We then present divider design for RF synthesis and introduce the Miller and injection-locked topologies.

15.1 General Considerations

The performance of frequency dividers is generally specified in terms of four parameters: (1) the divide ratio, (2) the maximum allowable input frequency, f_{max}, (3) the power dissipation, and (4) the minimum allowable input voltage swing (also called the "sensitivity"). While the phase noise of dividers also matters, in most cases it is negligible.[1] Dividers are often conservatively designed so as not to pose a risk upon the overall system, i.e., their f_{max} is well above the necessary value. The supply noise of dividers can cause spurs and must be managed carefully.

Let us sketch the input sensitivity of a general divider as a function of the input frequency. We expect that higher frequencies require greater input swings. This trend is conceptually shown in Fig. 15.1(a), where the sensitivity, V_p, is sketched as a function of the input frequency. Each point on this plot represents an f_{max}

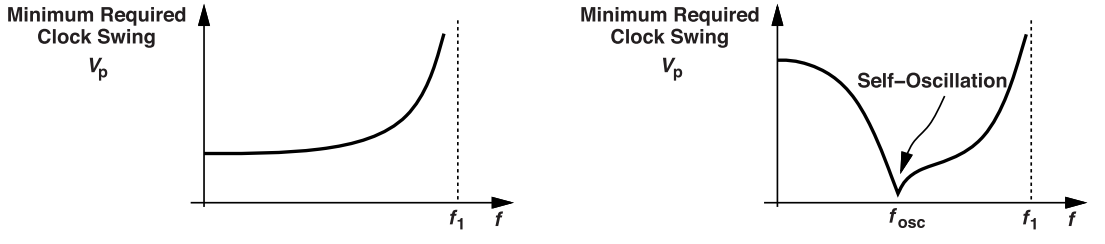


Figure 15.1 (a) Divider sensitivity as a function of input frequency, and (b) behavior of self-oscillating dividers.

corresponding to a certain input swing. For $f > f_1$, the circuit fails regardless of how large the swing is. Some divider loops can act as an *oscillator* when their input is zero, exhibiting the behavior depicted in Fig. 15.1(b). In this case, the circuit oscillates at f_{osc} with a zero input amplitude.

[1] Most divider topologies do not accumulate phase noise around their feedback loop, exhibiting a phase noise profile similar to that of delay lines.

Most dividers operate with rail-to-rail input swings, approaching failure if these swings drop by, say, more than 10%. In some cases, the failure occurs even for frequencies well below f_1 in Fig. 15.1. We must distinguish between two types of failure, "static" or "dynamic." The former occurs even at low frequencies, indicating insufficient voltage swings, transconductances, or dc voltage gain. The latter arises from the finite delay of the stages. Thus, a divider failing at high speeds must first be tested at lower frequencies to ensure proper static conditions. For example, a divider targetting an input frequency of 10 GHz is first tested at a few gigahertz to check its static behavior. Dividers may also fail at very low frequencies if the input edges are not sufficiently sharp. Moreover, the interface between the VCO and the divider must be designed carefully. The following example elaborates on this point.

Example 15.1

Shown in Fig. 15.2(a), an LC oscillator employs a cross-coupled PMOS pair with a single-ended peak-to-

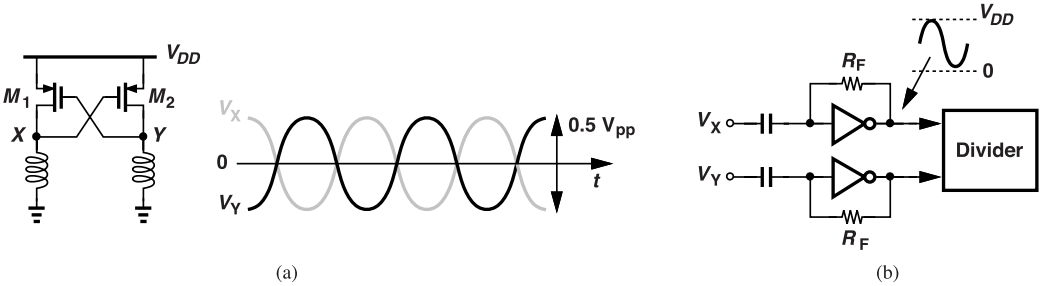


Figure 15.2 (a) LC oscillator with zero output CM level, and (b) use of self-biased inverters as buffers.

peak voltage swing of only 0.5 V to reduce flicker noise upconversion (Chapter 3). How can the oscillator drive a frequency divider that requires rail-to-rail swings?

Solution

We note from the waveforms in Fig. 15.2(a) that V_X and V_Y fall *below* zero for half of the cycle. That is, they cannot directly drive the divider even if their swings were as large as 1 V_{pp}.

In such cases, we can interpose capacitively-coupled, self-biased inverters between the VCO and the divider [Fig. 15.2(b)]. By virtue of their voltage gain, the inverters provide rail-to-rail swings. The coupling capacitors are chosen about five to ten times the input capacitance of the inverters, and the feedback resistors must be much greater than the inverters' output resistance (why?). Also, the high-pass corner frequency of this network must be chosen well below the minimum input frequency. We examine this circuit's supply coupling issues in Problems 15.1-15.3.

The reader can prove that the small-signal input resistance of the self-biased inverters is approximately equal to $1/(g_{mN} + g_{mP})$ if R_F is large. This resistance degrades the oscillator Q to some extent.

One concern in the arrangement of Fig. 15.2(b) is the supply noise: this noise modulates the delay of the inverters, thereby adding phase noise to the oscillator signal. Note that this effect arises even though the circuit is quasi-differential. For this reason, the inverters (and the VCO) are often supplied from on-chip, low-noise regulators.

15.2 Latch Design Styles

Most frequency dividers employ D latches. In this section, we study various static and dynamic latch design styles and describe their pros and cons. Of interest to us are such parameters as the maximum speed, the power consumption, and the number of clocked transistors (and hence the load presented at the input).

Example 15.2

Construct a $\div 2$ circuit using two noninverting D latches.

Solution

If two noninverting D latches are placed in a loop and clocked by CK and $\overline{CK}$, they assume a certain state, e.g., $Q_1Q_2 = 11$, and remain therein indefinitely. For the loop to toggle, a net inversion is necessary (i.e., the feedback must be negative). We thus arrive at the arrangement shown in Fig. 15.3, where the two latches form a master-slave FF.

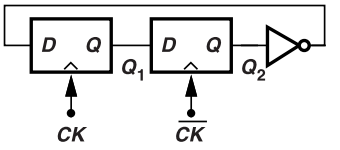


Figure 15.3 Simple divide-by-2 circuit using two latches in a negative-feedback loop.

15.2.1 Static Latches

Some frequency dividers must operate at *low* frequencies reliably. For operation below roughly 100 MHz, we prefer static latches as they are not prone to failure due to the subthreshold and junction leakage of the transistors. As for the upper end, static latches can reach speeds of 5 to 10 GHz in today's CMOS technologies. Figure 15.4(a) depicts an example where D_{in} is sensed when CK is high and its value is stored when CK

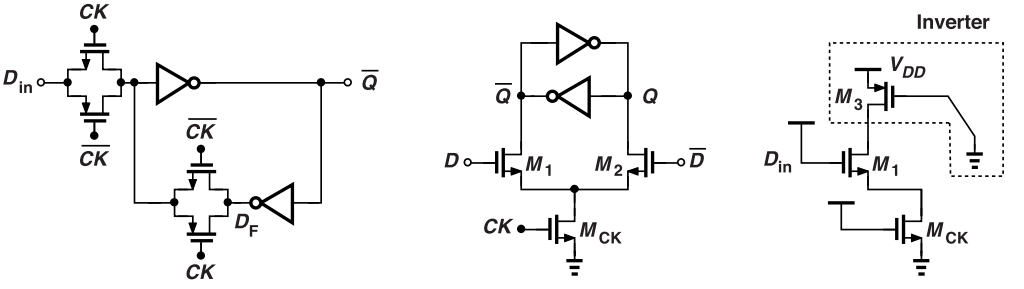


Figure 15.4 (a) Static latch with single-ended data path, (b) static latch with complementary data paths, and (c) simplified circuit of (b) during a write operation.

goes low and the feedback loop consisting of the two back-to-back inverters is enabled. This single-ended topology consists of eight transistors and presents two to each phase of the clock. Note that the clock transition times must be short enough to minimize the fight between D_{in} and D_F when the circuit goes from the sense mode to the latch mode (Problem 15.4). This latch exemplifies an "unratioed" configuration, requiring no particular ratio between transistor widths for proper static operation. Of course, we may still choose wider PMOS devices to optimize the speed.

Figure 15.4(b) shows a D latch with complementary data inputs and outputs. Here, when CK goes high, M_1 or M_2 turns on, overwriting the previous state held by the two inverters. When CK goes low, the new state is retained indefinitely. Employing seven transistors with only one clocked device, this topology tends to be more efficient than that in Fig. 15.4(a). However, the design requires proper *ratioing* of the devices. Specifically, as illustrated in Fig. 15.4(c), for the input data to overwrite the previous state, the series combination of M_{CK} and M_1 (or M_2) must be strong enough to overcome the PMOS transistor in one of the

inverters, a condition that must be satisfied even in the SF corner of the process. For this reason, we choose $M_{1,2}$ and M_{CK} at least as wide as the PMOS devices.

Example 15.3

As seen later in this chapter, some frequency dividers must precede a latch with an AND gate. How can this be done in Fig. 15.4(b)?

Solution

We insert an NMOS NAND branch on one side and an NMOS NOR branch on the other, arriving at the structure shown in Fig. 15.5. Note that the AND branch makes the left side weaker unless it incorporates wider transistors.

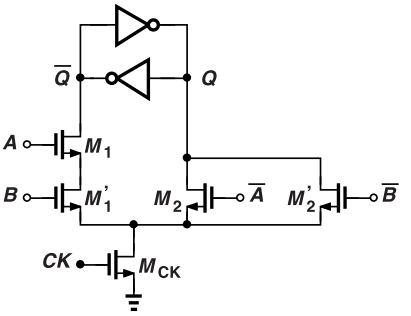


Figure 15.5 Latch including an AND function.

In both topologies of Fig. 15.4, the state is stored by means of two back-to-back inverters, or back-to-back "inverting amplifiers." This is a signature of static latches: when the clocked device is deactivated, the state is held by a *low-impedance* path to ground or V_{DD}. The next arrangement employs the same idea.

A third static latch incorporates current steering and operates with non-rail-to-rail input and output swings. Called the "current-mode logic" (CML) topology and shown in Fig. 15.6(a), this structure draws static power and is used only at very high frequencies. In the sense mode, CK is high, M_5 is on, M_6 is off, and the circuit reduces to that in Fig. 15.6(b). The input is thus amplified by M_1 and M_2 and impressed on nodes X and Y. Next, CK goes low, M_5 turns off, M_6 turns on, and the regenerative pair, M_3-M_4, inherits, amplifies, and stores the differential voltage $V_X - V_Y$ [Fig. 15.6(c)]. We recognize this pair and its load resistors as two back-to-back inverting (common-source) amplifiers. The single-ended voltage swing is equal to $R_D I_{SS}$ if the transistors switch completely.

The speed advantage of CML circuits originates from two properties: (a) the use of moderate voltage swings, e.g., $I_{SS}R_D \approx$ 300-400 mV, so that the transitions can occur faster, and (b) the use of only NMOS devices in the data and clock paths. The static power penalty, however, limits their use to only cases where other wideband divider topologies fail.

The positive-feedback amplification provided by M_3 and M_4 merits further analysis. If the initial value of $V_X - V_Y$ sensed by this pair is relatively small, we can construct the small-signal model shown in Fig. 15.6(d). Writing KCLs at X and Y gives

$$\frac{V_X}{R_D} + C_X \frac{dV_X}{dt} = -g_{m3,4} V_Y \tag{15.1}$$

$$\frac{V_Y}{R_D} + C_Y \frac{dV_Y}{dt} = -g_{m3,4} V_X . \tag{15.2}$$

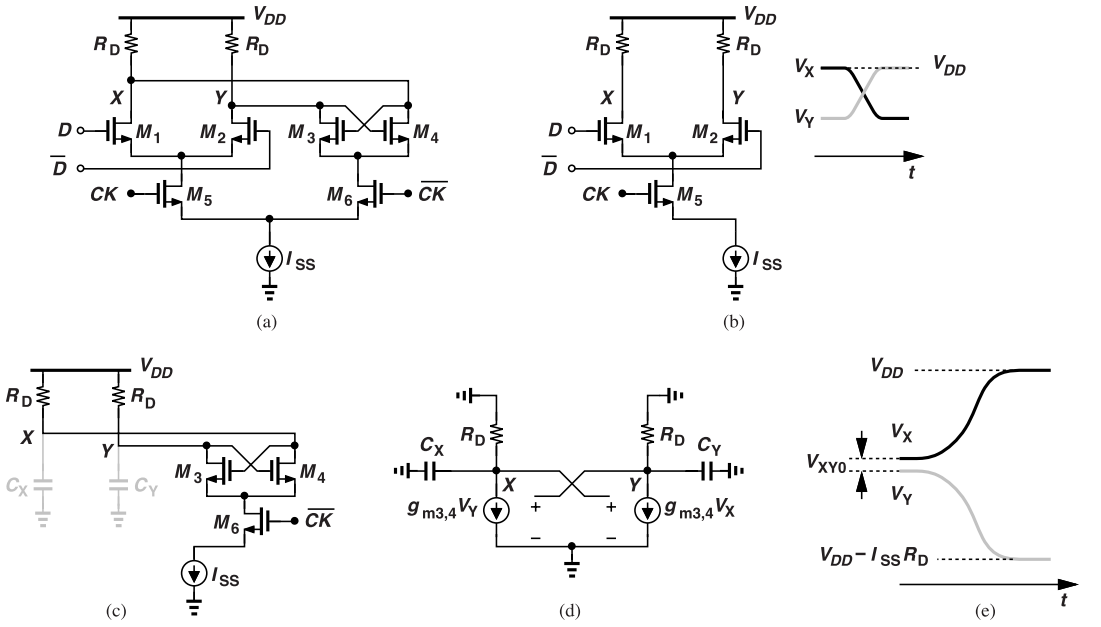


Figure 15.6 (a) CML latch, (b) circuit in the sense mode, (c) circuit in the regenerative (latch) mode, (d) equivalent circuit in the regeneration mode, and (e) output wavforms.

Subtracting each term of (15.2) from the corresponding term in (15.1) and rearranging the result, we have

$$R_D C_D \frac{dV_{XY}}{dt} = (g_{m3,4} R_D - 1) V_{XY}, \tag{15.3}$$

where $C_D = C_X = C_Y$ and $V_{XY} = V_X - V_Y$. If V_{XY} begins with an initial value of V_{XY0}, then (15.3) yields

$$V_{XY} = V_{XY0} \exp \frac{t}{\tau_{reg}}, \tag{15.4}$$

where $\tau_{reg} = R_D C_D / (g_{m3,4} R_D - 1)$ denotes the regeneration time constant. The exponential growth is caused by positive feedback, more fundamentally by the circuit's poles in the right half plane. As illustrated in Fig. 15.6(e), V_{XY} follows (15.4) until the circuit enters the large-signal regime, one transistor's transconductance falls, and the tail current is completely steered to one side. We note that the exponential amplification can occur only if $\tau_{reg} > 0$ and hence if $g_{m3,4} R_D > 1$.

As mentioned earlier, the single-ended voltage swing, $I_D R_D$, is typically chosen around 300 to 400 mV. However, the stacking of M_1-M_2 and M_3-M_4 pairs on top of the M_5-M_6 pair in Fig. 15.6(a) prohibits the CML latch from low-voltage operation. As a remedy, we remove the tail current source and bias the clocked transistors by means of a current-mirror arrangement (Fig. 15.7). Here, the bias current is defined by I_{REF}, and the coupling capacitors isolate the gate bias of M_5 and M_6 from the CK common-mode level. The resistor and capacitor values are large enough to yield a high-pass corner far below the clock frequency.

Example 15.4

How should the value of C_1 and C_2 be chosen in Fig. 15.7?

Solution

We can select this value to be five to ten times the capacitance seen at the gates of M_5 and M_6 so as to

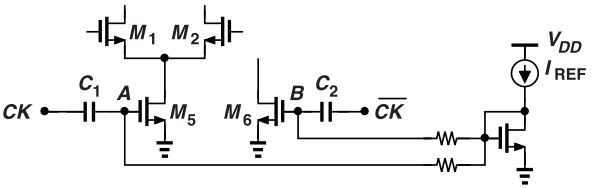


Figure 15.7 Use of capacitive coupling and current-mirror biasing in the clock path.

minimize the attenuation of the clock swings. However, if the clock swings are rail-to-rail, we can allow some attenuation, e.g., by a factor of 2 or 3, for the amplitude reaching A and B because M_5 and M_6 can operate with moderate swings. In fact, it is desirable to avoid rail-to-rail swings at A and B so that M_5 and M_6 operate in saturation and the circuit maintains some common-mode rejection for M_1 and M_2. To understand this point, suppose V_A rises to V_{DD}, thereby driving M_5 into the deep triode region. Then, the current flowing through M_1 and M_2 has a strong dependence on the CM level at their gates.[2] In such a case, the CM gain of the circuit (through the data path) may exceed unity, causing significant CM errors in a cascade of latches.

Owing to their limited input swings, CML latches cannot easily employ a NAND gate. The reader can observe that the NAND principle illustrated in Fig. 15.5 fails here because M_1 and M_1' cannot be turned off with only moderate input swings. A CML latch including proper NAND operation is shown in Fig. 15.8, where V_X is low only if both A and B are high in the sense mode. This circuit requires a DC level shift for the B input to ensure that M_1 and M_2 do not enter the triode region. This structure is rarely used in low-voltage design.

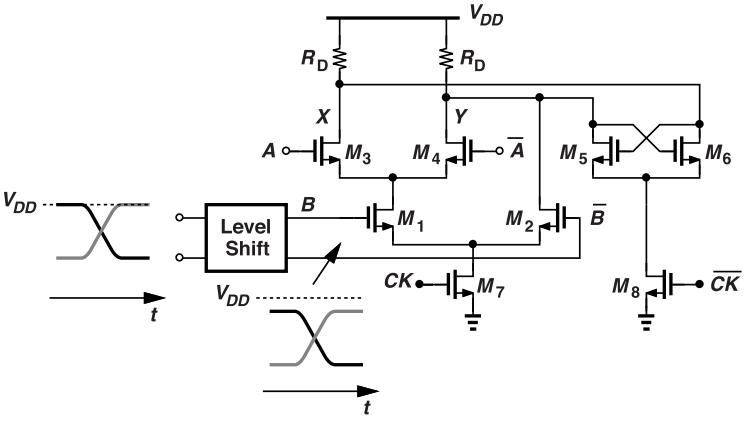


Figure 15.8 Need for level shift in a CML NAND circuit.

It is possible for a CML latch to incorporate an OR gate, as depicted in Fig. 15.9(a), where M_1 and M_1' can impress a low level at X and a high level at Y. In this case, M_2 is simply driven by a constant bias voltage, V_b, chosen equal to the common-mode level of A and B. The single-ended nature of the inputs makes this circuit somewhat less robust than fully-differential CML topologies, necessitating close attention to the inputs. Specifically, the A and B voltage swings and the widths of M_1 and M_1' must be large enough to guarantee complete steering of the tail current. Also, V_b must be so generated as to track the CM level of A and B. As depicted in Fig. 15.9(b), R_1 and I_1 (or, more generally, $I_1 R_1$) are chosen accordingly. For

[2]The CM level is defined as the point at which differential signals cross each other.

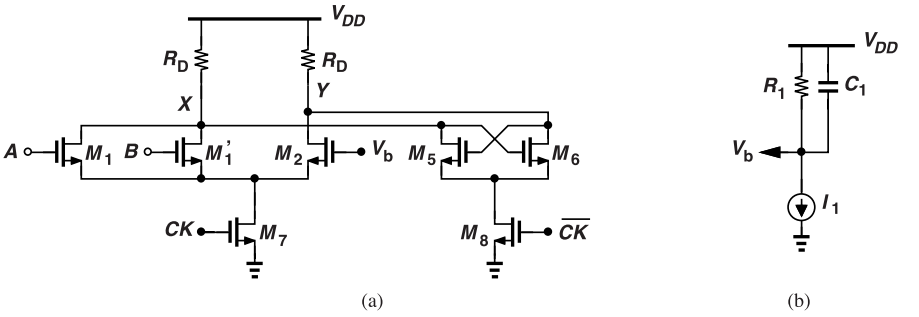


Figure 15.9 (a) CML NOR gate and latch, and (b) circuit for generating V_b.

example, if the preceding stages have an average tail current of I_{SS} and a load resistance of R_D, then their output CM level is equal to $V_{DD} - R_D I_{SS}/2$, requiring that $R_1 I_1 = R_D I_{SS}/2$.

CML circuits draw static power and should be used only if the required speed is so high that other topologies cannot meet it. Among the three static latches studied in this section, the first two draw no static power but suffer from a limited speed whereas the third—the CML topology—can run at very high frequencies and consumes high power. Dynamic latches, on the other hand, generally operate faster than the first two topologies, offering an attractive alternative.

15.2.2 Dynamic Latches

In dynamic latches, the states are stored on device capacitances rather than by back-to-back amplifying stages. Such latches contain fewer transistors than their static counterparts and generally offer a more favorable speed-power trade-off. But, owing to device leakage, they can lose their states if the clock frequency is not sufficiently high. That is, dynamic latches impose a lower bound on the operation frequency.

C^{2}MOS Logic Figure 15.10 shows a "clocked CMOS" (C^{2}MOS) dynamic latch. Here, when CK is high,

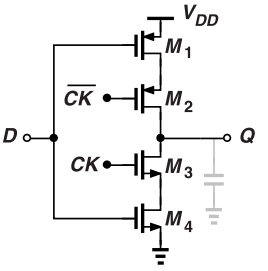


Figure 15.10 C^{2}MOS latch.

the circuit acts as an inverter, sensing the input. When CK goes low, the P and N branches are disabled and the state is stored on the output capacitance. In comparison to the static structures studied in the previous section, this stage requires only four transistors—with no ratioing necessary for proper operation. In Problem 15.7, we show that, if M_1 and M_4 are clocked and D is applied to M_2 and M_3, then the circuit suffers from charge sharing and a degraded output level.

Example 15.5

Construct a $\div 2$ circuit using C^{2}MOS latches.

Solution

Since each latch inverts, two C^2MOS latches do not suffice (Example 15.2). We must therefore insert an inverter in the loop, as shown in Fig. 15.11, to ensure negative feedback.

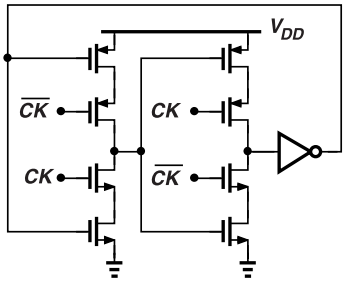


Figure 15.11 C^2MOS divide-by-2 circuit.

How fast can this divider run? For a given input frequency, f_{in}, we require that the loop support a waveform having a frequency of $f_{in}/2$. If we roughly view the circuit as three inverters, and hence a ring oscillator, we can estimate the maximum output frequency to be about $1/(6T_D)$, where T_D denotes the average delay of the stages. We consider $1/(6T_D)$ as the self-oscillation frequency of the circuit. In this case, $f_{in} \approx 2/(6T_D) = 1/(3T_D)$. (The presence of the clocked transistors within the inverters increases their delay.) From Fig. 15.1(b), we note that the loop can divide properly even at higher frequencies. We may even say that this circuit is a three-stage ring oscillator that is "injection-locked" to the input clock. Note that this $\div 2$ circuit does not provide quadrature outputs.

The principal drawback of C^2MOS latches is their "transparency" during clock transitions. To understand this point, let us consider the master-slave FF shown in Fig. 15.12 with realistic clock waveforms. We recog-

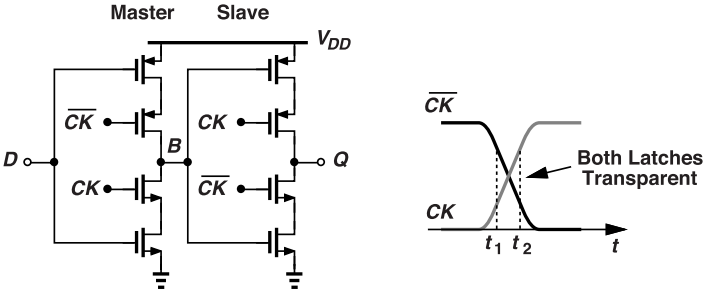


Figure 15.12 Problem of transparency in C^2MOS logic.

nize that, while CK rises to place the slave in the store mode and the master in the sense mode, both stages become transparent for some time. Consequently, D can change B and hence Q. Stored on the parasitic capacitance at Q, the output state can degrade substantially due to this "feedthrough" if the new value of D disagrees with Q. As a conservative measure, the two latches should be driven by nonoverlapping clock phases: i.e., two nonoverlapping clocks and their complements. But for frequency divider implementations, fast clock transitions suffice.

True Single-Phase Clocking A simple, elegant dynamic logic style is based on "true single-phase clocking" (TSPC) [1]. Originally conceived to avoid the use of complementary clocks, TSPC circuits exhibit a higher speed and lower power consumption than their C^2MOS counterparts. Let us begin with the C^2MOS

stage of Fig. 15.10 and remove the clocked PMOS device [Fig. 15.13(a)]. Here, when CK is high, the circuit operates as an inverter and, when CK is low, the output state is retained if Q is high or if D does not fall.

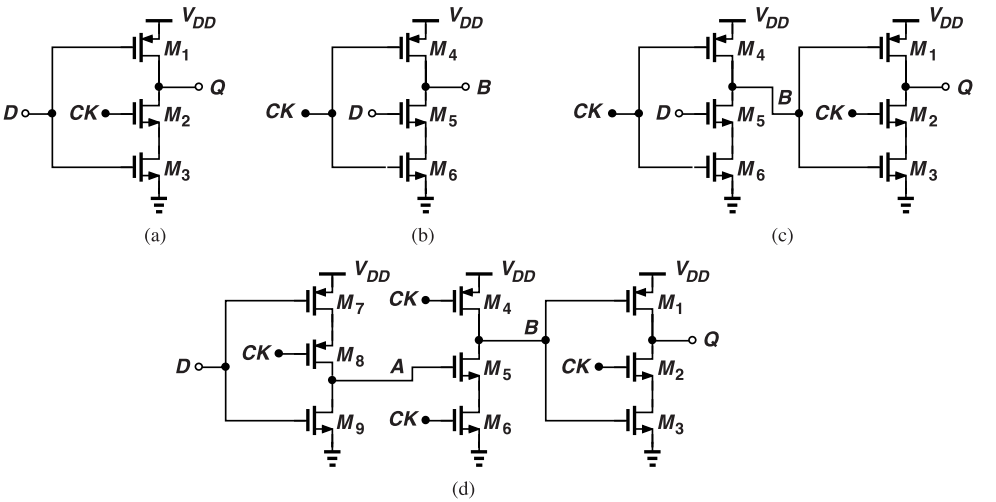


Figure 15.13 (a) Inverter using one clocked device to act as a latch, (b) topology of (a) with CK and D swapped, (c) cascade of circuits in (b) and (a), and (d) addition of one more stage to create a TSPC flipflop.

That is, to guarantee that the state is stored properly, we must ensure D remains high when CK is low.

How can this be accomplished? Let us precede the circuit by another similar structure whose output satisfies this condition. As shown in Fig. 15.13(b), an inverter controlled by CK allows the D input to write a low level when CK is high. Combining the two stages, we arrive at the structure in Fig. 15.13(c). This configuration operates as follows. When CK is low, B is precharged to V_{DD} and Q retains its state. When CK goes high, Q assumes a logical value equal to that of D: if D is high, B falls and Q rises, and, if D is low, B remains high and Q falls. The two-stage topology thus acts as a D latch. In Problems 15.8 and 15.9, we examine other variants of these circuits.

The circuit of Fig. 15.13(c) incorporates six transistors, *more* than the C^2MOS design in Fig. 15.10, but it does not require multiple clock phases. Moreover, to construct a master-slave configuration, we need not duplicate the entire cascade: as illustrated in Fig. 15.13(d), we precede the circuit with a single TSPC stage. Here, when CK is low, the first stage acts as an inverter and is in the sense mode while B is held high and Q retains its state. When CK goes high, the second stage creates $B = A$, and the third stage, $Q = B = A$. This flipflop does require some hold time: suppose $A = 1$; from the time CK goes high, D must be stable for some time until A writes to B. After this hold time, D can change without affecting the state present at Q.

The foregoing TSPC flipflop can operate as a divide-by-2 circuit if D is tied to Q. However, the circuit provides neither quadrature outputs nor an output with a 50% duty cycle.

Example 15.6

A student changes the clocked device in the first stage of Fig. 15.13(d) to an NMOS transistor (Fig. 15.14). Does this circuit operate as a master-slave flipflop?

Solution

No, it does not. When CK is low, the first stage is not completely in the sense mode as it cannot write a logical ZERO onto A.

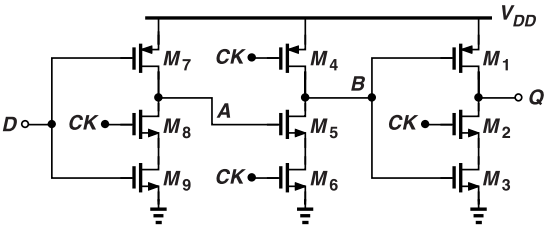


Figure 15.14 TSPC flipflop with clocked NMOS device in the first stage.

TSPC flipflops can readily incorporate logic as well. Figure 15.15 shows an example containing a NAND gate at the input.

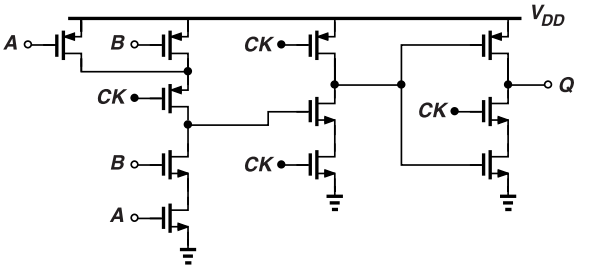


Figure 15.15 TSPC circuit including a NAND function.

The master-slave flipflop of Fig. 15.13(d) employs nine transistors, four of which are clocked. Applications requiring a large number of flipflops can use another TSPC realization that contains six devices, with only two clocked. Depicted in Fig. 15.16 [1], this topology exploits the "split" levels available at the drain and

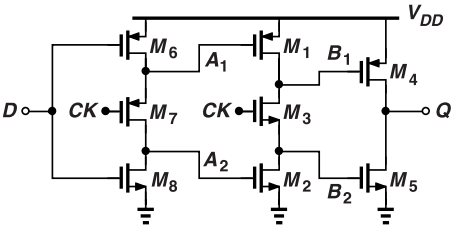


Figure 15.16 TSPC circuit with split signal paths.

source of the clocked transistors while eliminating the clock from the third stage. The reader is encouraged to analyze the circuit's operation, but we should remark that the logical levels at nodes A_1 and B_2 are degraded because they are created at the *sources* of transistors. As a result, the pull-up and pull-down actions of M_1 and M_5, respectively, are slowed down.

It is possible to improve the speed considerably by changing the clocked transistors in Fig. 15.16 to complementary pairs [2]. Shown in Fig. 15.17, the result runs two to three times faster, but it requires complementary clocks. The addition of M_7' and M_3' ensures that the logical levels at A_1 and B_2 are not degraded, thereby providing a high overdrive voltage for M_2 and M_5, respectively.

Ratioed TSPC The foregoing TSPC implementations are unratioed and consume no static power (except for device leakage). A variant of TSPC logic incorporates ratioed devices to achieve a higher speed but at the cost of static power. Shown in Fig. 15.18 [3], the structure omits the stacked transistors in the last two stages. Here, when CK is high, the first stage is in the sense mode, the second stage holds B low, and Q retains its state. In this case, M_5 must be strong enough to overcome M_4 if both are on. When CK falls, $B = \overline{A}$ and $Q = \overline{B} = A$ if M_7 is stronger than M_6. We observe that the second and third stages draw a static current in

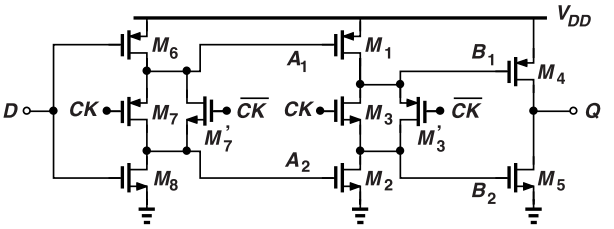


Figure 15.17 Flipflop using complementary switches for faster operation.

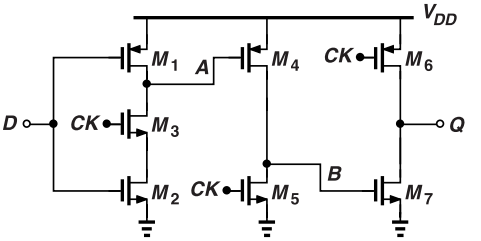


Figure 15.18 Ratioed TSPC circuit.

the first and second modes, respectively. We typically choose the same width for all of the transistors, except for M_5, which should be two to three times wider for optimum speed.

Differential Logic The dynamic latches studied thus far operate with single-ended inputs and outputs. In applications requiring complementary signals and/or quadrature outputs, we must turn to differential (or, more precisely, complementary) implementations. Shown in Fig. 15.19(a) is a simple differential C^2MOS D latch, employing twice as many transistors. We note that, if D and $\overline{D}$ are guaranteed to be complementary, then the circuit provides both an inverted output, $\overline{Q}$, and a noninverted output, Q. Can we use this stage to construct a $\div 2$ circuit with quadrature outputs? We consider a master-slave cascade with negative feedback, depicted in Fig. 15.19(b), as a candidate. Examining the circuit closely, however, we recognize that it in fact consists of four cascaded latches [Fig. 15.19(c)], with *no* net inversion around the loop! Thus, the loop simply latches up, holding a constant 1010 sequence at its four nodes and failing to divide.

To resolve this issue, we must create some interaction between the two paths in Fig. 15.19(a) so that each input influences *both* outputs. To this end, we attach a cross-coupled pair between Q and $\overline{Q}$ [Fig. 15.20(a)]. The idea is to discourage these nodes from assuming equal values in a $\div 2$ loop. If both Q and $\overline{Q}$ tend toward zero, then M_9 and M_{10} act as two diode-connected devices (why?), resisting such a state. On the other hand, if Q and $\overline{Q}$ change in opposite directions, the positive feedback around M_9 and M_{10} encourages the transition. Interestingly, the cross-coupled pair also helps retain the logical levels at Q and $\overline{Q}$ even if the latch is transparent during clock transitions. In practice, a cross-coupled NMOS pair can also be added, leading to the arrangement shown in Fig. 15.20(b). The strength of the cross-coupled devices is typically chosen about $1/4$ to $1/2$ of that of the main path. The circuit now acts as a static latch because the inverters can hold the state indefinitely.

Another differential dynamic latch can be derived from the structure in Fig. 15.4(b) by removing the NMOS devices within the inverters. Depicted in Fig. 15.21(a), this circuit contains only five transistors but requires proper ratioing. In fact, the initial fight between the NMOS and PMOS devices degrades the speed to some extent. This arrangement behaves as a dynamic latch because the low state stored on one side can disappear while the corresponding input transistor is off.

The latch of Fig. 15.21(a) exhibits an output low level that can go *below* zero. Illustrated in Fig. 15.21(b), this phenomenon arises on the falling edge of CK. First, suppose the clock is high and so is D. Thus, M_2 is on and Q is at zero while M_4 is off. When CK falls, it couples to node P through C_{GD5}, drawing a current from the capacitance at Q. The voltage at this node therefore drops below zero, a benign effect.

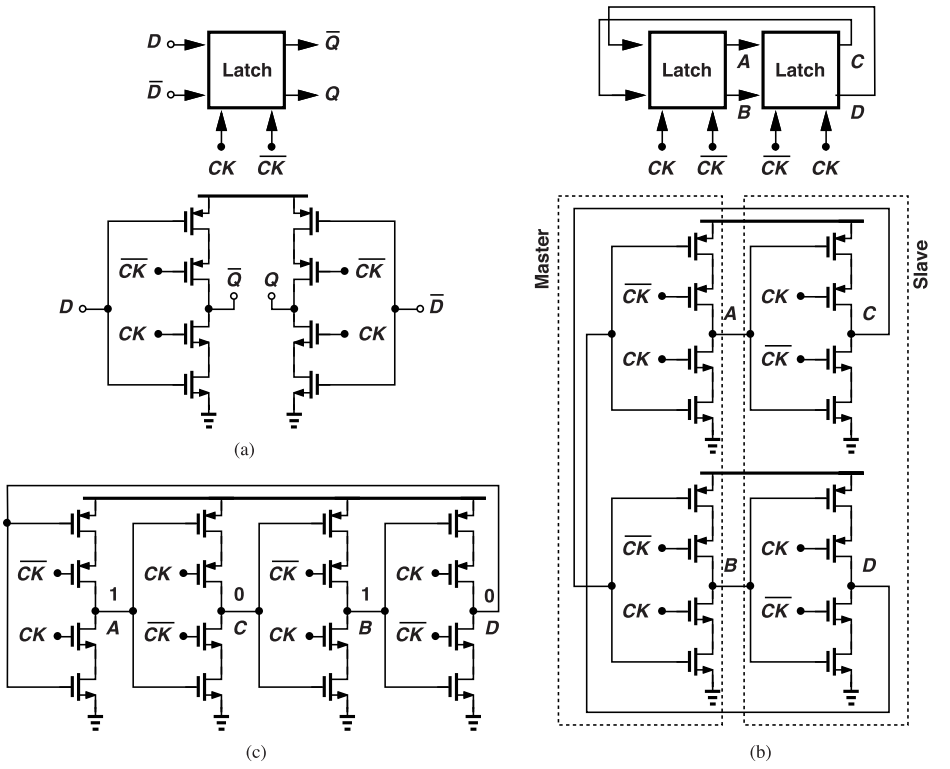


Figure 15.19 (a) Latch with complementary signals based on two C^2MOS stages, (b) divide-by-2 circuit using two instances of (a) as master and slave latches, and (c) unwrapped circuit illustrating the problem of latch-up.

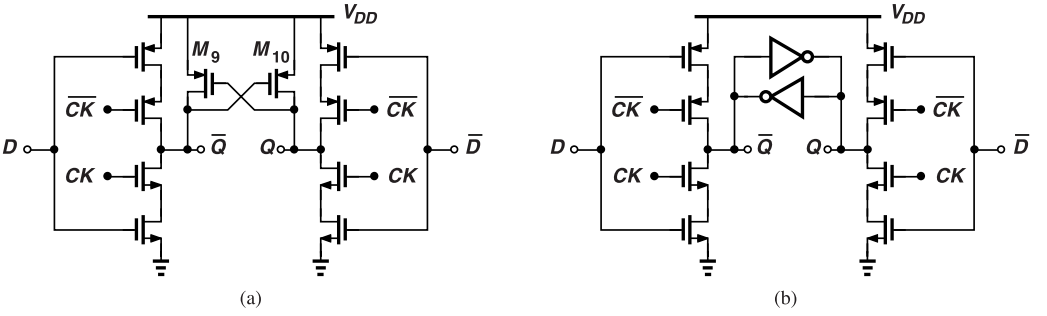


Figure 15.20 Addition of cross-coupled (a) PMOS devices, or (b) inverters to avoid latch-up.

15.3 Divide-by-2 Circuit Design

Divide-by-2 circuits are primarily sought in two applications: (1) to halve the frequency of a very high-speed VCO so that the subsequent divider (e.g., a pulse-swallow counter) can operate properly, and (2) to generate quadrature phases. In both cases, we prefer to use a master-slave flipflop for its robustness, although the former need not be so configured, i.e., a $\div 2$ circuit does not necessarily generate quadrature phases.

The various latches studied in this chapter can serve in flipflop-based divide-by-2 circuits. For example, the CML latch of Fig. 15.6(a) can operate at very high speeds, but at the cost of static power. The differential latch introduced in Fig. 15.21(a) draws only dynamic power while suffering from a limited speed. In a $\div 2$ environment, this structure can be modified to run faster by noting that the PMOS transistors are the bottle-

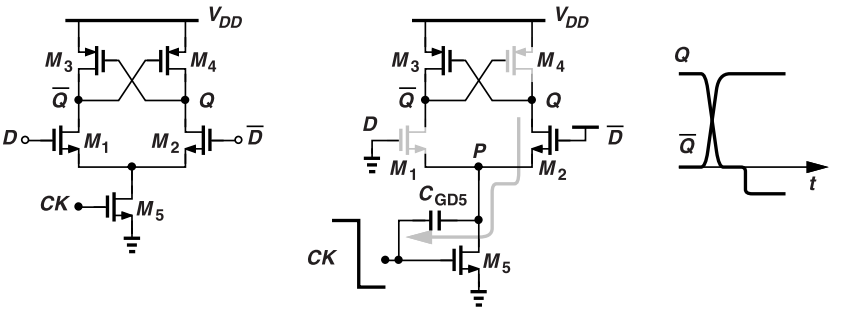


Figure 15.21 (a) Differential dynamic latch, and (b) simplified circuit during write operation.

neck. We can help the pull-up process by adding NMOS source followers that impress a rising input at an output that must also rise. As shown in Fig. 15.22(a) [4], M_6 or M_7 pulls up its source node when its gate goes up, thus shortening the delay (while increasing the input capacitance).

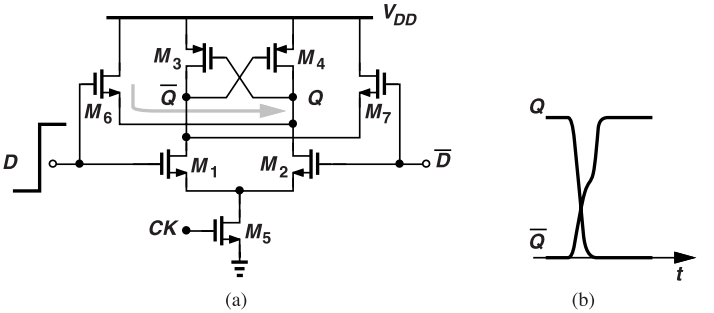


Figure 15.22 (a) Addition of source followers to dynamic latch to improve the speed in a divider environment, (b) appearance of a kink in the waveforms at lower speeds.

The source followers in Fig. 15.22 create another mechanism for improving the speed as well: they couple the rising transition of D or $\overline{D}$ to the output even if CK is low. That is, M_6 and M_7 provide a "feedforward" path that "prepares" the output before CK goes up. Of course, if the input and clock edges are far apart in time, the output signals experience a "kink" [Fig. 15.22(b)] and, at sufficiently low clock frequencies, the divider fails. From another perspective, M_6 and M_7 act as an asynchronous path, bypassing the clock path and causing failure at low frequencies.

In some applications, it is desirable for a $\div 2$ circuit to generate quadrature outputs with a 25% duty cycle. Shown in Fig. 15.23(a) is an FF-based divider accomplishing this task [5]. Assuming that the NMOS devices are stronger than the PMOS transistors, we show that only one of the four outputs can be high when CK is high or low. We can readily show that all four outputs cannot stay low at the same time. For example, if CK is low and $\overline{CK}$ is high, we can assume that $V_{out90} = V_{out270} = 0$, keeping M_3 and M_4 off, but the regenerative action around M_1 and M_2 forces either V_{out0} or V_{out180} to be high.

As a more detailed analysis, suppose, at $t = t_1$ in Fig. 15.23(b), CK is low and V_{out0} is high. This means that M_2 is on, V_{out180} is low, and M_1 is off. In addition, M_9 is on, pulling V_{out90} to zero (because M_{11} is off). At this point, M_8, M_{10}, and M_{12} are off, allowing V_{out270} to be low or high. But V_{out270} cannot be high because it would then pull V_{out0} low through M_3 and violate our initial assumption. When CK goes high, at $t = t_2$, M_{12} turns on, raising V_{out270} to V_{DD} because both M_8 and M_{10} are off. The other three outputs remain low (why?). When CK goes low again, V_{out180} rises because both M_2 and M_4 are off. By the same token, V_{out90} rises on the next rising edge of CK. In summary, when CK is low, V_{out0} and V_{out180} assume opposite logical values while V_{out90} and V_{out270} remain at zero, and vice versa.

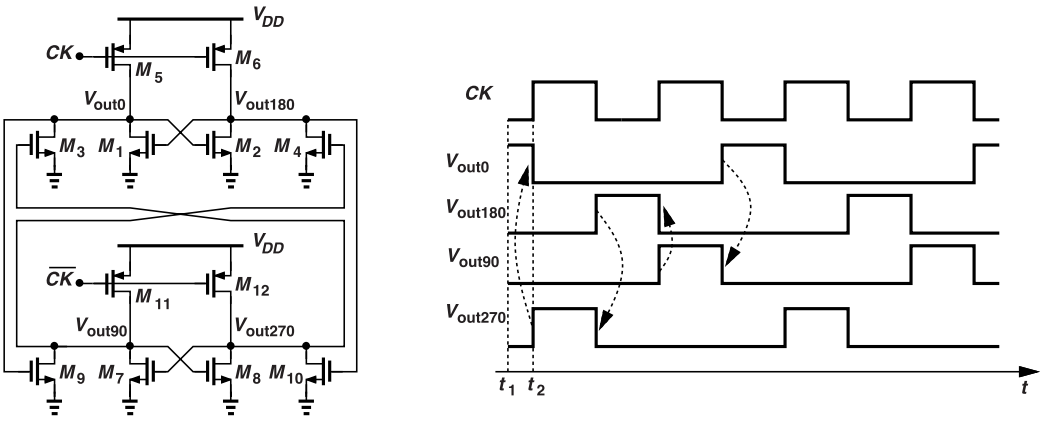


Figure 15.23 (a) Divide-by-2 circuit generating outputs with 25% duty cycle, and (b) circuit's waveforms.

Example 15.7

Upon closer examination, we observe that the output waveforms of the divider shown in Fig. 15.23(a) exhibit a degraded ZERO level [Fig. 15.24(a)]. Explain how this occurs.

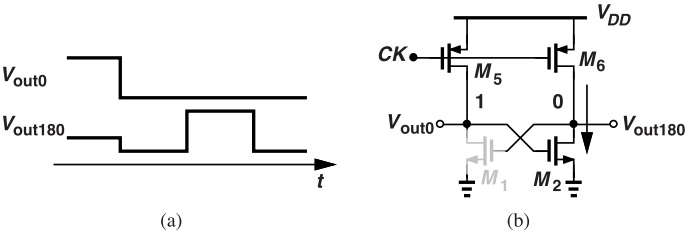


Figure 15.24 (a) Degradation of low output level in the circuit of Fig. 15.23(a), and (b) static current path showing the cause of level degradation.

Solution

When the PMOS devices are on, one of them fights an NMOS transistor that is also on. For example, if V_{out0} is high [Fig. 15.24(b)], then M_2 draws a static current from M_6, producing $V_{out180} > 0$. Thus, the PMOS devices must be weak enough so as to not degrade the logical ZERO level significantly.

It is possible to avoid the degraded level and the static current incurred by the foregoing circuit if some speed penalty is acceptable. Figure 15.25 depicts such a modification: a cross-coupled PMOS pair, M_5'-M_6', is placed in series with the clocked transistors so as to cut off the current path on the side that holds a low level. For example, if CK is low, $V_{out0} = \text{ONE}$, and $V_{out180} = \text{ZERO}$, then M_6' is off.

Example 15.8

Explain why the PMOS cross-coupled pair in Fig. 15.25 reduces the speed.

Solution

Suppose, at $t = 0^-$, CK is high, $V_{out0} = V_{out180} = \text{ZERO}$, and one input is high (Fig. 15.26). When CK falls, M_6 and M_6' appear in series and their gate voltages are at zero. Thus, V_{out180} charges up more slowly.

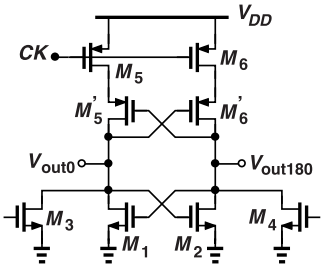


Figure 15.25 Modified divider avoiding level degradation.

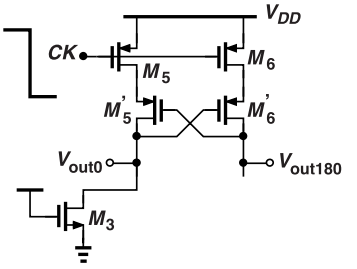


Figure 15.26 Degradation of speed in Fig. 15.25.

Note that M_5' and M_6' do not provide regenerative amplification because V_{out0} is held relatively constant and near zero by M_3.

15.4 Dual-Modulus Prescalers

Recall from Chapter 12 that RF synthesizers incorporate a feedback divider whose modulus must vary in unity steps, e.g., from N to $N + 1$. This operation requires the use of dual-modulus "prescalers," i.e., high-speed dividers whose modulus can switch between 2 and 3 or between 3 and 4, etc. We call such arrangements "$\div 2/3$" and "$\div 3/4$" circuits, respectively.

Let us begin with a $\div 3$ circuit. Shown in Fig. 15.27(a) is an example where two D flipflops cycle through the three necessary states. Note that the next value of Q_2 is equal to $\overline{Q_1 \cdot Q_2}$. With the aid of the waveforms plotted in Fig. 15.27(b) and assuming that the FFs change their output on the rising edge of the clock, we

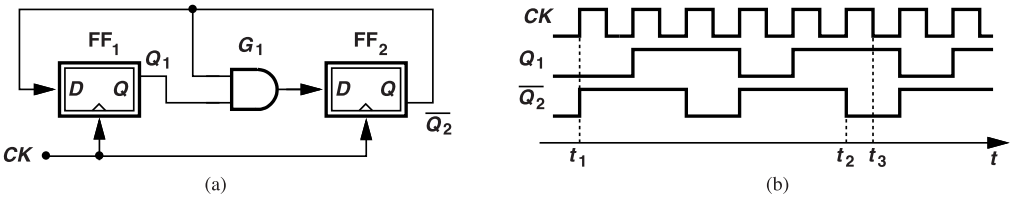


Figure 15.27 (a) Divide-by-3 circuit, and (b) its waveforms.

note that, if $Q_1\overline{Q_2} = 00$ and CK goes high (at t_1), then Q_1 remains at ZERO while $\overline{Q_2}$ rises to ONE. In the next clock cycle, Q_1 assumes the value of $\overline{Q_2}$ and goes high and $\overline{Q_2}$ is unchanged. This state, $Q_1\overline{Q_2} = 11$,

generates a ONE at the AND output, requiring that the next value of $\overline{Q_2}$ be ZERO. In the third clock cycle, therefore, Q_1 stays high and $\overline{Q_2}$ falls. The state $Q_1\overline{Q_2} = 00$ actually does not occur again. The two outputs exhibit a duty cycle of 1/3, a useful property in some applications and undesired in some others. It is helpful to remember the 110 or 011 output pattern for $\div 3$ circuits.

Example 15.9

Explain why the $\div 3$ circuit of Fig. 15.27(a) is generally slower than $\div 2$ topologies.

Solution

The feedback around the second flipflop must deal with the additional delay presented by the AND gate. Specifically, in Fig. 15.27(b), as the clock rises at $t = t_2$ and $\overline{Q_2}$ falls, this transition must propagate through the AND gate and be established within the master latch of FF_2 before the clock falls. Figure 15.28 illustrates this timing constraint, revealing that the total delay from CK to the output of the master latch (B) must be less than half of the clock period. As a rule of thumb, we say the maximum speed of $\div 3$ circuits is about half of that of their $\div 2$ counterparts.

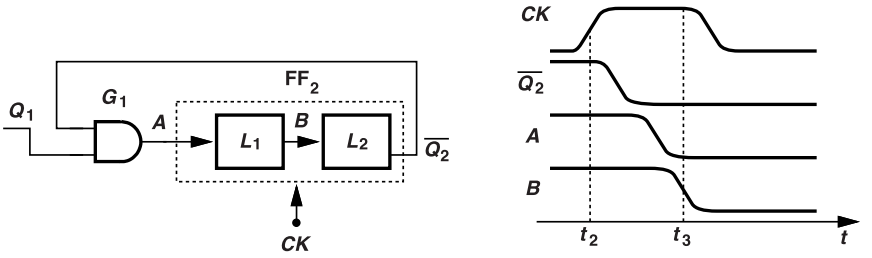


Figure 15.28 Critical path in a divide-by-3 circuit illustrating the speed limitation.

The AND gate in Fig. 15.28 is preferably merged with the following latch (Section 15.2) so as to reduce the delay. Since in CML designs, we prefer NOR inputs, we can readily modify the divider structure to accommodate a NOR implementation. To this end, we sense the Q output of FF_2 and carry the missing inversion to the inputs of FF_1 and G_1 [Fig. 15.29(a)]. Next, we allow the inversion at the input of FF_1 to

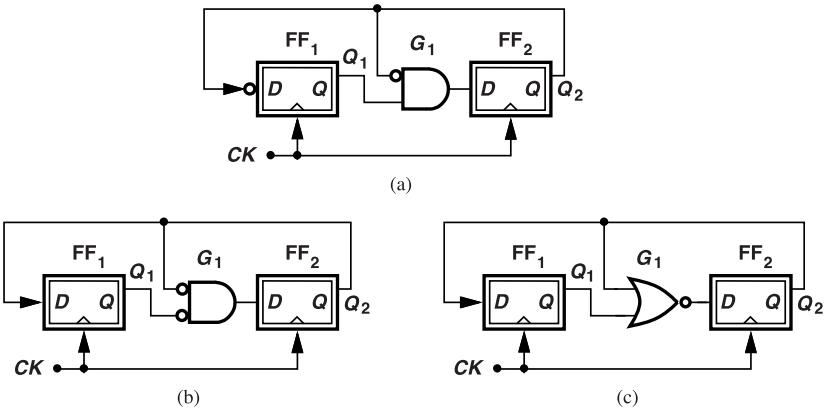


Figure 15.29 (a) Divide-by-3 circuit with inversion moved to inputs of G_1 and FF_1, (b) modified circuit, and (c) divide-by-3 circuit using a NOR gate.

propagate through the flipflop and reach the input of the AND gate [Fig. 15.29(b)]. The result is equivalent

to a NOR realization [Fig. 15.29(c)]. An interesting byproduct of this modification is that the FFs now need not invert, a property that proves useful for C^2MOS flipflops but not for TSPC designs (why?).

The $\div 3$ topologies of Fig. 15.27(a) and Fig. 15.29(c) can also accommodate a $\div 2$ mode. Illustrated in Fig. 15.30, the idea is to reduce the circuit to one flipflop when the circuit must divide by 2. If the "modulus

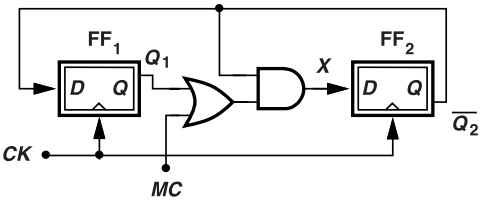


Figure 15.30 Dual-modulus $\div 2/3$ circuit.

control" input, MC is low, Q_1 simply travels through the OR gate, and the circuit divides by 3. On the other hand, if MC is high, Q_1 has no effect, the AND gate passes Q_2 to X, and FF_2 acts as a $\div 2$ circuit. Note that the stages following this divider must sense the signal at $\overline{Q_2}$ (or any other point within the loop around FF_2) but not the signal at Q_1 (why?).

One can similarly contemplate $\div 3/4$ implementations. Such arrangements still require only two flipflops and can evolve from the $\div 3$ circuits of Fig. 15.27(a) or Fig. 15.29(c). For example, we wish to modify the former so that, when MC is high, the structure reduces to a $\div 4$ circuit, i.e., to two FFs in a loop. Depicted in Fig. 15.31, this occurs if an OR gate is inserted in the loop around FF_2. If $MC = 0$, the OR gate is out of

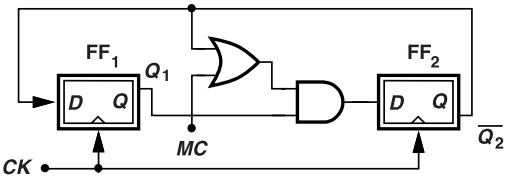


Figure 15.31 Dual-modulus $\div 3/4$ circuit.

the picture and the circuit divides by 3; if $MC = 1$, the structure acts as two flipflops in a loop and divides by 4. Of course, the longer delay due to the two gates lowers the maximum speed.

Example 15.10

Using the ratioed TSPC flipflop of Fig. 15.18, design a $\div 3/4$ circuit.

Solution

Since this FF inverts the signal, we need an extra inversion to satisfy the structure in Fig. 15.31. Merging the OR and AND gates with the first stage of FF_2, we arrive at the design shown in Fig. 15.32.

The foregoing dual-modulus dividers are based on synchronous operation, i.e., their FFs are clocked simultaneously. For greater divide ratios, we can combine one of these prescalers with additional asynchronous dividers. As an example, Fig. 15.33(a) depicts how a $\div 2/3$ circuit (Div23) is followed by an asynchronous $\div 2$ stage to form a $\div 5$ counter. To study the operation, let us assume that the flipflops change their outputs on the rising edge of their respective clocks. Note that Q_1 always follows $\overline{Q_2}$ after one clock delay. We begin with $Q_1\overline{Q_2}\,\overline{Q_3} = 000$ [Fig. 15.33(b)] and recognize that $MC = \overline{Q_3} = 0$, $X = 0$, and the $\div 2/3$ stage divides by 3. As CK goes high at $t = t_1$, so does $\overline{Q_2}$, but Q_1 remains low (why?). In response to the rising edge on Q_2, the second (asynchronous) stage changes Q_3 to 1 as well, commanding Div23 to divide by 2 in the next cycle. Thus, $\overline{Q_2}$ and Q_1 simply toggle between 0 and 1 on the subsequent rising edges of CK. This continues until t_2, at which time $\overline{Q_2}$ rises and hence $\overline{Q_3}$ falls, driving Div23 into the $\div 3$ mode. At $t = t_3$, Q_1 goes high

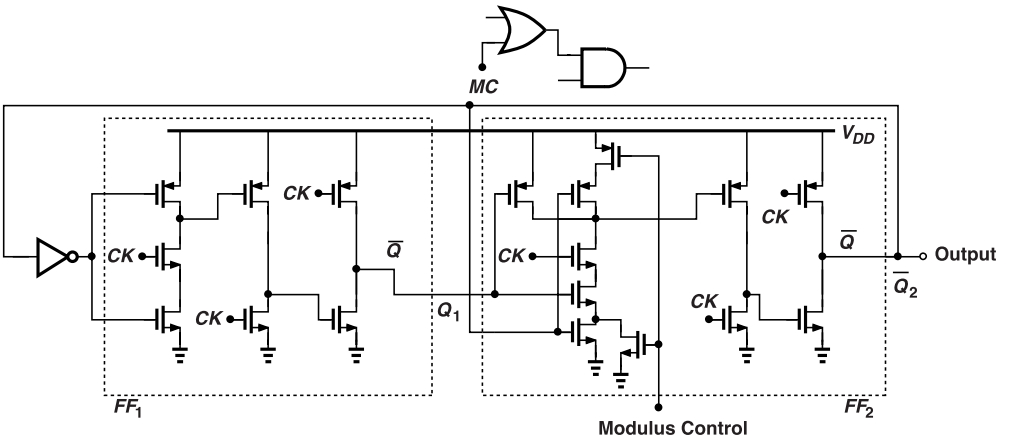


Figure 15.32 TSPC $\div 3/4$ circuit.

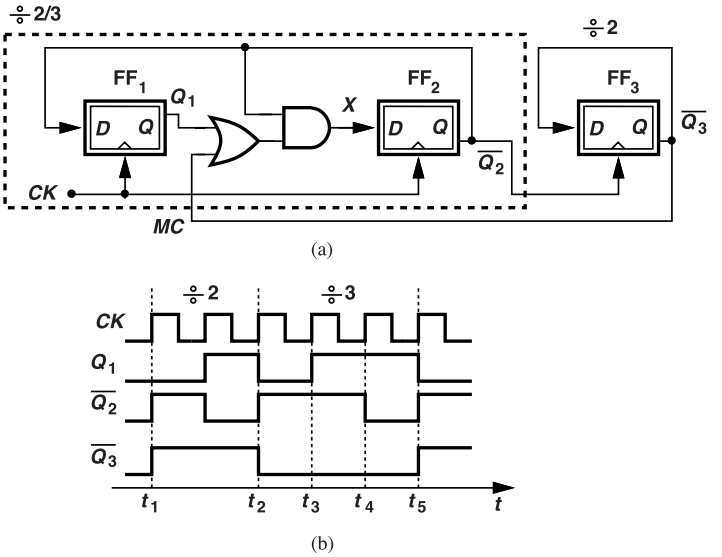
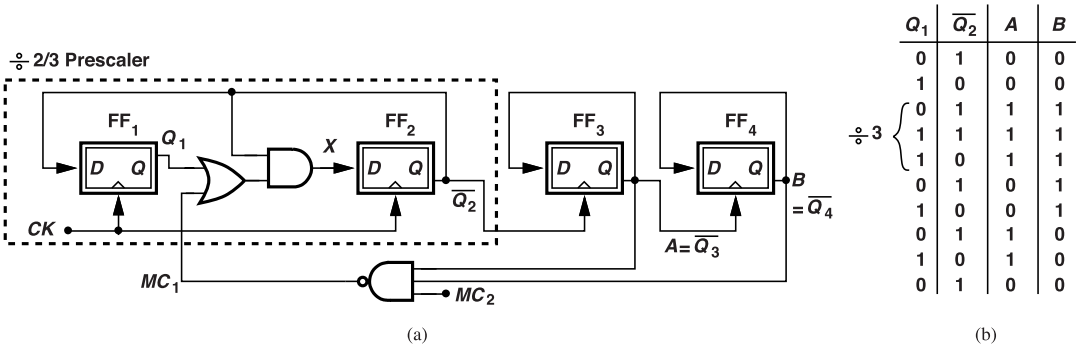


Figure 15.33 (a) Divide-by-5 circuit, and (b) its waveforms.

but $\overline{Q_2}$ does not change (why?), and the present state of $Q_1\overline{Q_2} = 11$ leads to $\overline{Q_2} = 0$ at $t = t_4$. At $t = t_5$, Q_1 falls, $\overline{Q_2}$ rises, and so does $\overline{Q_3}$. The circuit has now reached the same state as that at $t = t_1^+$, revealing its periodic behavior. We observe that (a) Div23 divides by 2 for two input cycles (from t_1 to t_2) and by 3 for three input cycles (from t_2 to t_5), and (b) only $\overline{Q_3}$ carries a periodic output with the proper period and can be sensed by the next stage.

In analyzing dividers, we can also consider the number of input *pulses*—rather than input edges—that result in one output pulse. In Fig. 15.33, for example, Div23 receives two clock pulses between t_1 and t_2, and three between t_2 and t_5. Similarly, the asynchronous stage receives two pulses between t_1 and t_5. This perspective will prove useful later.

We now study the $\div 8/9$ counter shown in Fig. 15.34(a). To understand the principle, we first recognize that, if the prescaler divides by 2, then the overall circuit behaves as three cascaded divide-by-2 stages and hence divides by 8. This occurs when MC_2 is low and hence MC_1 is high. For $\div 9$ operation, we keep MC_2 high, allowing MC_1 to go low only when $AB = 11$. Figure 15.34(b) shows the sequence of the states,



Q_1	$\overline{Q_2}$	A	B
0	1	0	0
1	0	0	0
0	1	1	1
1	1	1	1
1	0	1	1
0	1	0	1
1	0	0	1
0	1	1	0
1	0	1	0
0	1	0	0

Figure 15.34 (a) Dual-modulus $\div 8/9$ circuit, and (b) its state table.

assuming the flipflops change their output on the rising edge of their clock input. We observe that the $\div 2/3$ circuit divides by 3 for three input cycles (while $AB = 11$) (recall the pattern 110 or 011 for $\div 3$ operation) and by 2 for six input cycles. Note that only B carries a periodic waveform with the desired period and can serve as the counter's output.

As another example, we study the $\div 15/16$ circuit depicted in Fig. 15.35(a). This arrangement incorporates

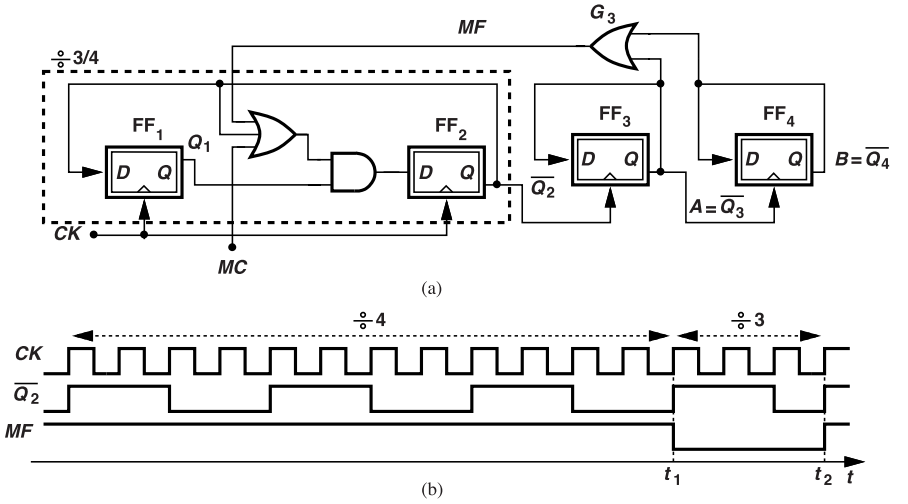


Figure 15.35 (a) Dual-modulus $\div 15/16$ circuit, and (b) its waveforms.

the $\div 3/4$ prescaler of Fig. 15.31 and two asynchronous $\div 2$ stages. If MF or MC is high, the prescaler divides by 4, and the overall circuit by 16. Based on our analysis of the topologies in Figs. 15.33 and 15.31, we can predict the operation of this counter in the $\div 15$ mode. We surmise that, with $MC = 0$, the circuit divides by 4 for 12 input cycles and by 3 for three input cycles [Fig. 15.35(b)]. The $\div 3/4$ stage thus generates a total of four pulses at $\overline{Q_2}$, which, upon driving the asynchronous stages, produce one pulse at MF. This pulse lasts from t_1 to t_2, causing the $\div 3$ mode.

The use of both synchronous and asynchronous logic in Fig. 15.35(a) can lead to a race condition and hence failure in extreme PVT corners. Suppose FF_3 and FF_4 change their output states on the *rising* edge of their respective clock inputs. If MC is low, the circuit initially divides by 16, with $Q_1\overline{Q_2}$ cycling through 01, 11, 10, and 00 until both A and B are low. As shown in Fig. 15.36(a), Q_1Q_2 then skips the state 00 after the state 10. From the time $\overline{Q_3}$ goes low until the time $Q_1\overline{Q_2}$ skips one state, three CK cycles have passed, allowing a relatively relaxed delay constraint through FF_3 and G_3.

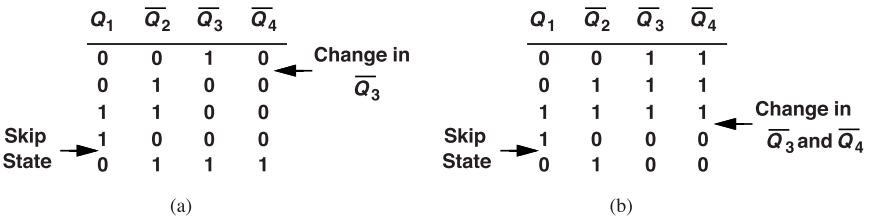


Figure 15.36 (a) Proper, and (b) improper choice of edges in Fig. 15.35(a).

On the other hand, if FF_3 and FF_4 in Fig. 15.35(a) change their states on the *falling* edge of their clock inputs, the delay through FF_3 and G_3 must be less than half of the input clock period. This can be seen from Fig. 15.36(b), where, immediately after $Q_3\,Q_4$ falls to 00, the prescaler must skip the state 00. We therefore prefer the former choice and must select the signals such that FF_3 and FF_4 change their states on the rising edge of their clock inputs.

15.5 Divider Design for RF Synthesis

As explained in Section 15.1, RF synthesizers employ a feedback divider whose modulus must change in unity steps. Moreover, the divider must operate reliably at the maximum VCO frequency and/or with the minimum VCO output swing. This is a critical point as the VCO frequency may well exceed the desired value during the synthesizer's lock transient, possibly causing the divider to fail.

RF synthesizers typically cover a narrow frequency band, e.g., from 2.400 GHz to 2.480 GHz for Bluetooth. We thus surmise that the output frequency can be expressed as $(K + S)f_{REF}$, where K is a large, fixed integer and S is one whose value can change in unity steps. In Bluetooth, for example, we can choose $K = 2400, S = 1, ..., 80$, and $f_{REF} = 1$ MHz. We seek a divider architecture that provides a modulus of the form $K + S$, and study two different realizations in this section.

15.5.1 Pulse Swallow Divider

Figure 15.37 shows an implementation called the "pulse swallow divider." The circuit consists of a dual-

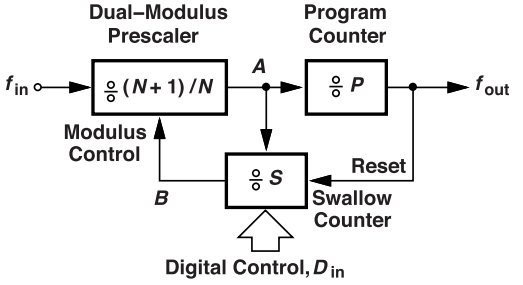


Figure 15.37 Pulse-swallow divider.

modulus prescaler (or, more generally, a dual-modulus divider), a "program counter" (a fixed-modulus divider), and a "swallow counter" (a resettable divider whose modulus can be changed, by the digital control, from 1 to a maximum value in unity steps). For simplicity, we call these circuits the N, P, and S counters, respectively. As explained below, the architecture requires that $P > S$. Note that the S counter controls the modulus of the prescaler, and the P counter can reset the S counter. We will prove that $f_{out} = f_{in}/(NP+S)$.

To study the divider's operation, let us assume that the prescaler begins in the $\div(N + 1)$ mode, and the P and S counters are reset. The prescaler generates a pulse at A after $N + 1$ input clock cycles. Thus, a total of $(N + 1)S$ input cycles are necessary for the S counter to change B. The prescaler then begins to divide by N. By now, the P counter has received S pulses, and it needs another $P - S$ input pulses to fill up and

reset the swallow counter. The prescaler must therefore count another $(P - S)N$ input clock cycles (hence P must be greater than S). We observe that the P counter produces one output pulse in response to every $(N + 1)S + (P - S)N = NP + S$ input cycles. It follows that $f_{out}/f_{in} = 1/(NP + S)$.

The pulse swallow divider is a clever invention: it provides an overall divide ratio, $NP + S$, that can be varied in unity steps by simply adjusting the value of S. In the previous Bluetooth example, we can choose $NP = 2400$ and $S = 1, ..., 80$; e.g., if $N = 8$, P is chosen equal to 300.

Example 15.11

Can the waveforms at nodes A or B in Fig. 15.37 be used as an output?

Solution

No, they cannot. Since the prescaler modulus changes from $N + 1$ to N after $(N + 1)S$ cycles, the waveform at A first has a frequency of $f_{in}/(N + 1)$ hertz and then f_{in}/N hertz (Fig. 15.38). Similarly, the frequency of the signal at B toggles between $f_{in}/[(N + 1)S]$ and $f_{in}/(NS)$. These nodes thus do not have a proper period.

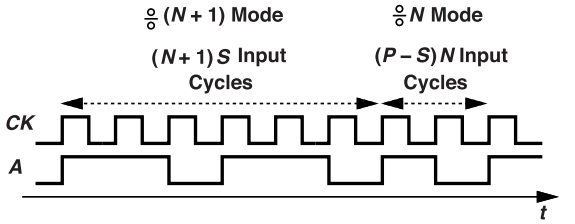


Figure 15.38 Pulse-swallow divider's waveforms.

Let us turn to the implementation of the swallow counter in Fig. 15.37. This counter must provide a programmable modulus ranging from 1 to a maximum. Denoting the desired modulus value by $D_{in} = D_n...D_1$, we observe that the circuit begins from $0...0$ and is clocked by the prescaler until it counts up to $D_n...D_1$. It then changes the prescaler modulus and keeps this output level until it is reset by the P counter. Shown in Fig. 15.39 is an implementation example, where n asynchronous resettable $\div 2$ stages begin from $00...0$ and count until their outputs match the digital input, $D_n...D_1$. At this point, all of the XNOR gates generate high logical levels, raising the RS latch output, changing the prescaler modulus, and disabling the $\div 2$ stages. This state remains until the P counter resets the RS latch.

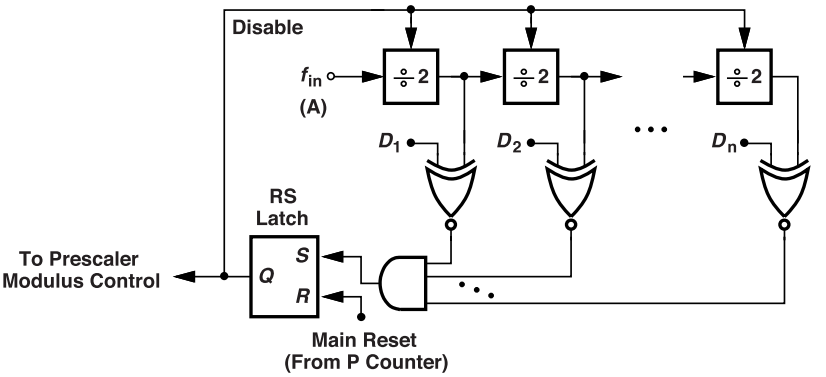


Figure 15.39 Swallow counter implementation.

As the waveforms in Fig. 15.38 suggest, the pulse swallow divider operation requires that P be greater than S. This limits the synthesizer output and/or reference frequency range, as illustrated by the following example.

Example 15.12

A 2.4-GHz fractional-N synthesizer must operate with crystal frequencies ranging from 19 MHz to 25 MHz. If we choose $N = 3$ for the prescaler, what values of P and S are necessary?

Solution

To obtain $f_{out} = 2400$ MHz, we require $NP + S = 127$ for $f_{REF} = 19$ MHz and $NP + S = 96$ for $f_{REF} = 25$ MHz. With $N = 3$, we have $3P + S_1 = 127$ and $3P + S_2 = 96$. Can we choose P and S to satisfy these two conditions? Note that $S_1 - S_2 = 31$. If S is programmable from 1 to 32 and P is chosen greater than S, then $3P + S_2$ exceeds 96. In other words, no values of P and S exist.

The above difficulty can be avoided if the program counter has a variable modulus as well, i.e., if it is based on the structure shown in Fig. 15.39. Let us return to Example 15.12 and assume $P = 30$ while S varies from 6 to 20, obtaining $NP + S = 96, ..., 110$. For greater divide ratios, we change P to 33 and allow S to go from 12 to 28, creating $NP + S = 111, ..., 127$. In other words, with programmable designs for both the S and P counters, $NP + S$ can assume a wide range.

15.5.2 Vaucher Divider

The Vaucher divider [9] is a modular topology incorporating a cascade of $\div 2/3$ stages and providing a programmable divide ratio. It is used as an alternative to the pulse swallow divider. In order to understand the Vaucher divider's operation, we begin with the simple structure of Fig. 15.33, repeated in Fig. 15.40(a). Recall that the first stage divides by 2 (when $MC = 1$) for two input cycles and by 3 (when $MC = 0$) for three input cycles, with $\overline{Q_3}$ carrying the desired output.

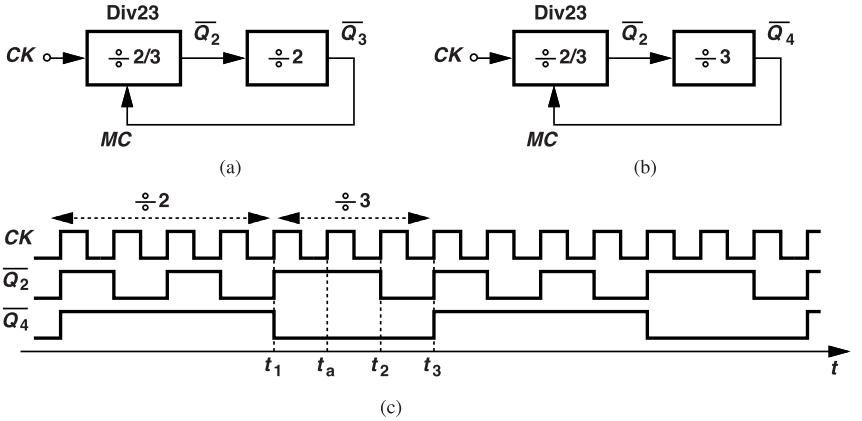


Figure 15.40 (a) A $\div 2/3$ stage and a $\div 2$ stage in a loop, (b) circuit of (a) with the second divider changed to a $\div 3$ stage, and (c) waveforms showing divide-by-7 operation.

Let us now replace the second stage with a $\div 3$ circuit [Fig. 15.40(b)] and determine the new divide ratio. Suppose the first stage divides by 2 when $MC = 1$ and by 3 when $MC = 0$. We return to the $\div 3$ topology of Fig. 15.27(a) and note that, if starting from reset, its output, $\overline{Q_2}$, rises on the first rising edge of the clock and remains high for two input cycles. Thus, assuming all nodes in Fig. 15.40(b) are reset to zero, we obseve that the first rising edge of CK forces both Q_2 and Q_4 to go high [Fig. 15.40(c)]. As a result, the first stage begins to divide by 2. This stage must produce *two* pulses on $\overline{Q_2}$ for the second stage to generate one pulse, as shown at $t = t_1$. Now, the first stage operates in the $\div 3$ mode and keeps $\overline{Q_2}$ high for two cycles of CK,

until $t = t_2$, and low for one cycle of CK, until $t = t_3$. This behavior repeats periodically thereafter. That is, the overall circuit divides by 7.

The operation depicted in Fig. 15.40(c) can be viewed from two perspectives: (1) Div23 divides by 3 from t_1 to t_3, or (2) Div23 continues to divide by 2 until $t = t_a$, at which point and up to $t = t_2$ it "swallows" an input pulse, i.e., it generates no transitions. Both perspectives prove useful in our analysis.

In summary, the topologies in Figs. 15.40(a) and 15.40(b) act as $\div 5$ and $\div 7$ circuits, respectively. Let us consider the structure shown in Fig. 15.41, where each stage divides by 2 when its modulus control is high

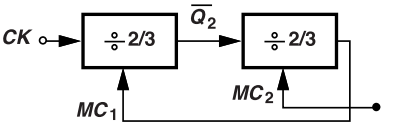


Figure 15.41 Two $\div 2/3$ stages in a loop.

and by 3 when it is low. We observe that the overall circuit divides by 5 if $MC_2 = 1$ [as in Fig. 15.40(a)] and by 7 if $MC_2 = 0$ [as in Fig. 15.40(b)]. Also, suppose we can override MC_1 externally. Then, the structure divides by 4 if MC_1 is forced to 1 and $MC_2 = 1$ and by 6 if MC_1 is forced to 1 and $MC_2 = 0$. That is, this arrangement can provide a divide ratio ranging from 4 to 7.

For modular design, we wish to cascade a number of $\div 2/3$ stages as in Fig. 15.41. To this end, we revisit the $\div 3$ circuit in Fig. 15.27(a) and draw its internal structure as shown in Fig. 15.42(a). Next, we return the output from Q_2 to the input of the first latch and choose $\overline{Q_1}$ to drive the AND gate [Fig. 15.42(b)]. The

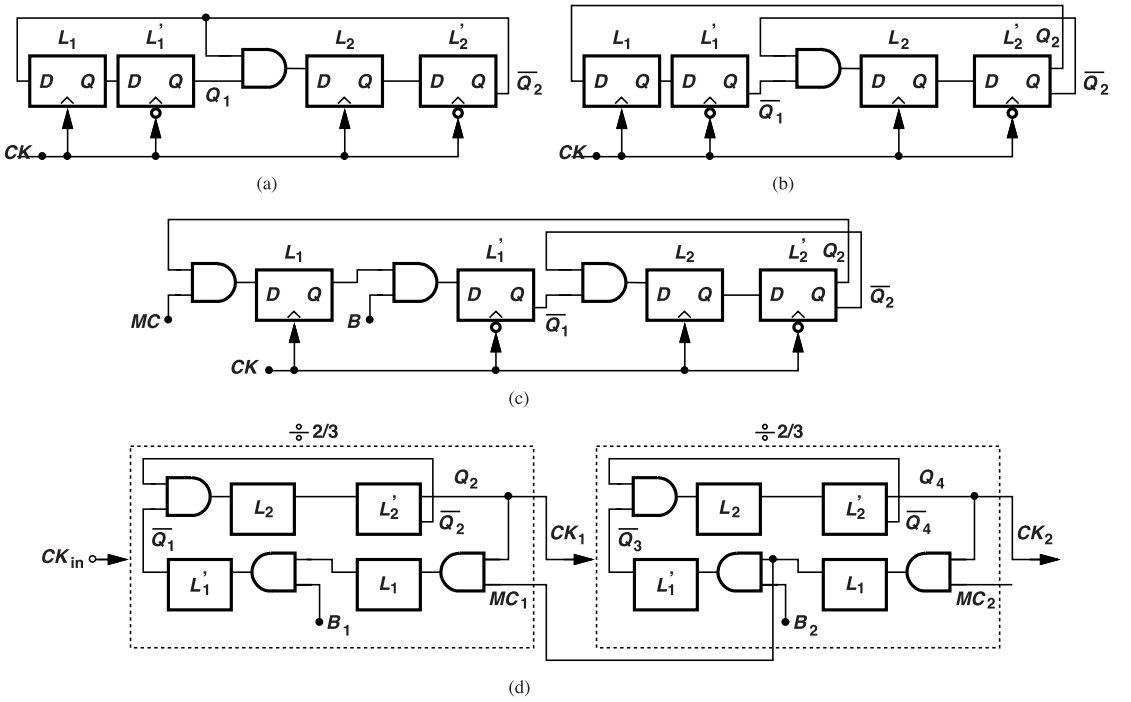


Figure 15.42 (a) A $\div 3$ circuit shown at the latch level, (b) separation of feedback signals, (c) addition of controls MC and B, and (d) modular extension of $\div 2/3$ stages.

circuit's function remains unchanged. Now, we insert two AND gates before L_1 and L_1' [Fig. 15.42(c)], noting that, if $MC = 1$ and $B = 1$, then the circuit divides by 3, but if either MC or B is low, it divides by 2. In the last step of our development, we redraw this block and employ two instances to construct a loop similar to

that in Fig. 15.41 [Fig. 15.42(d)] [9]. This result merits some remarks. First, B_1, B_2, and MC_2 set the divide ratio. If $MC_2 = 1$, then B_1B_2 can set this ratio in the range of 4 to 7. Second, MC_1 is obtained from Q_4 but with half a clock delay due to L_1.

It can be shown that, if the cascade in Fig. 15.42(d) is extended from 2 to n stages, then the output period is given by

$$T_{out} = (2^n + 2^{n-1}B_n + \cdots + 2^0 B_1)T_{in}, \tag{15.5}$$

where MC_n is assumed to be high [9]. That is, the divide ratio ranges from 2^n to $2^{n+1} - 1$ in steps of unity.

As studied in Problem 15.17, the additional AND gates in Fig. 15.42(c) degrade the speed. Moreover, the use of the Vaucher divider in fractional-N synthesizers requires additional measures: when B_1 and/or B_2 in Fig. 15.42(d) are changed by the $\Delta\Sigma$ modulator, some of the divider's internal states do not immediately change, corrupting the divider output sequence.

15.6 Miller Divider

The Miller divider generally achieves a higher speed than do FF-based topologies and proves useful if CML designs fail. Shown in Fig. 15.43(a), Miller's structure consists of a mixer (a multiplier) and a low-pass filter

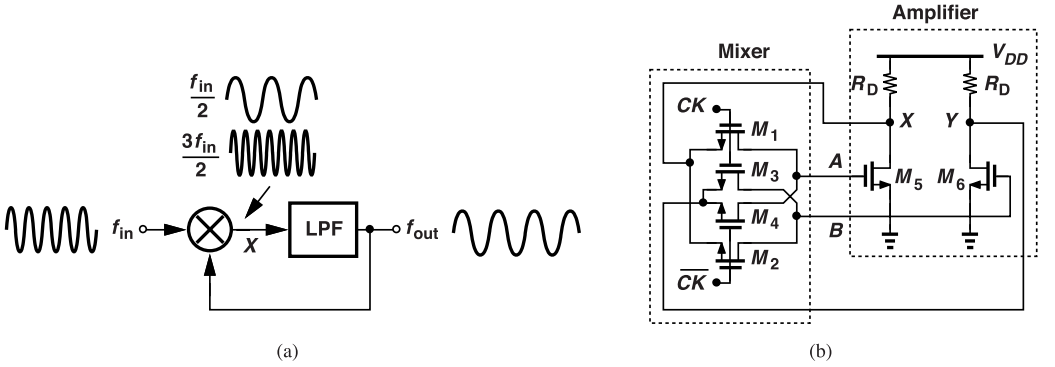


Figure 15.43 (a) Miller divider, and (b) an implementation example.

that form a feedback loop [8]. The mixer generates both $f_{in} + f_{out}$ and $f_{in} - f_{out}$ at node X, with the former suppressed by the filter. Thus, $f_{out} = f_{in} - f_{out}$ and hence $f_{out} = f_{in}/2$. Unlike FF-based configurations, however, the circuit does not produce quadrature outputs. We return to this point later.

The Miller divider achieves a high speed because it need not satisfy the timing constraints governing FF-based topologies. In essence, the loop in Fig. 15.43(a) is an "analog" path that allows the $f_{in}/2$ component to circulate unattenuated. This means that the loop gain is sufficiently large at $f_{in}/2$ and sufficiently small at $f_{out} + f_{in} = 3f_{in}/2$.

Figure 15.43(b) depicts an implementation example. Here, M_1-M_4 form a passive mixer and M_5-M_6 an amplifier. The low-pass filtering action occurs at the output nodes of the mixer and the amplifier. This topology requires rail-to-rail clock inputs. Alternatively, an active mixer can be used so as to allow smaller input swings.

While a single Miller divider does not provide quadrature outputs, two such structures can. This point becomes clear if we recognize that a divide-by-2 circuit receives an input of the form $\cos(\omega_{in}t + \theta)$ and generates an output given by $\cos[(\omega_{in}t + \theta)/2]$. Let us then utilize two Miller dividers, one sensing $\cos\omega_{in}t$ and the other, $\cos(\omega_{in}t + 180^\circ)$ (Fig. 15.44). The outputs are therefore equal to $\cos[(\omega_{in}t)/2]$ and $\cos[(\omega_{in}t + 180^\circ)/2]$. This principle can be applied to any divide-by-2 circuit, including the TSPC topology.

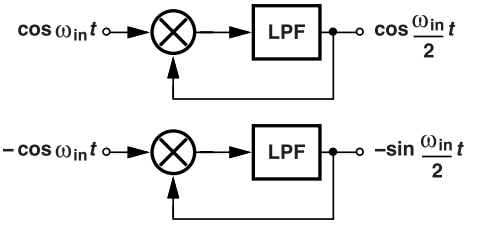


Figure 15.44 Use of two Miller dividers to generate quadrature outputs.

Example 15.13

Prove by contradiction that the output of the bottom Miller divider in Fig. 15.44 cannot be of the form $-\cos(\omega_{in}t/2)$.

Solution

The mixer output is then equal to $\cos\omega_{in}t\cos(\omega_{in}t/2) = (1/2)\cos(3\omega_{in}t/2) + (1/2)\cos(\omega_{in}t/2)$. The second component travels through the low-pass filter, contradicting our assumption that the output is equal to $-\cos(\omega_{in}t/2)$.

The Miller divider fails if the component at $f_{in}+f_{out} = 3f_{in}/2$ is not attenuated sufficiently. This appears counter-intuitive as $3f_{in}/2$ is simply the third harmonic of $f_{in}/2$ and should be *helpful* (by sharpening the edges) or at least benign. But a closer look reveals otherwise. Let us open the loop and examine the waveforms emerging from the mixer (Fig. 15.45). With equal amplitudes, the two components create a composite waveform exhibiting multiple zero crossings. The key point is that the low-frequency part of the sum waveform cannot be sustained around the loop unless the third harmonic is attenuated by at least a factor of 3 [10]; hence the need for the LPF. We also note that the circuit begins to fail if f_{in} is less than two-thirds of the corner frequency of the LPF because the LPF provides little attenuation for the third harmonic in such a case.

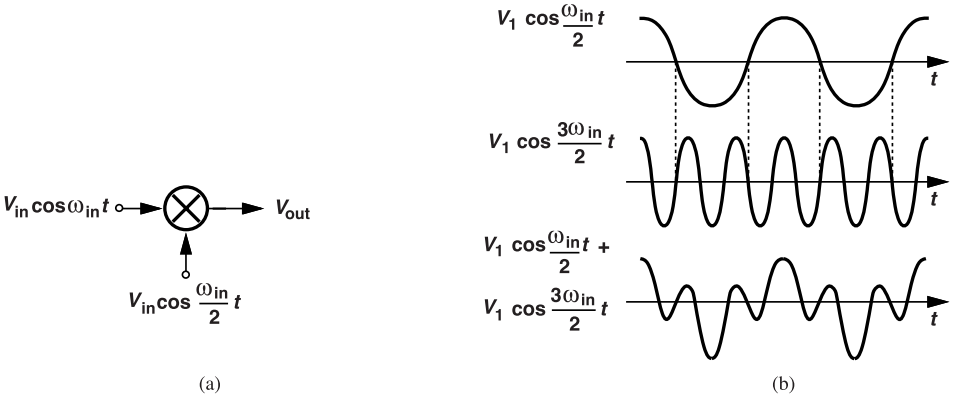


Figure 15.45 (a) Open-loop Miller divider without an LPF, and (b) its waveforms.

It is possible to replace the resistive loads in Fig. 15.43(b) with LC tanks so as to achieve higher speeds. Designed for a resonance frequency equal to the middle of the output range of interest, the tanks also suppress the $3f_{in}/2$ component. In this case, however, the frequency range has a lower bound that is higher than that of the resistively-loaded counterpart.

15.7 Injection-Locked Dividers

An oscillator can act as a divider if it is injection-locked to a harmonic of its oscillation frequency. For a good understanding of injection pulling and locking in oscillators, the reader is referred to the literature [6, 12]. We provide a brief description here.

Recall from Chapter 6 that an external signal can be coupled (injected) into an oscillator. For example, Fig. 15.46 shows how a periodic input at f_{in} is unilaterally injected into an LC topology. Even if the oscillator's

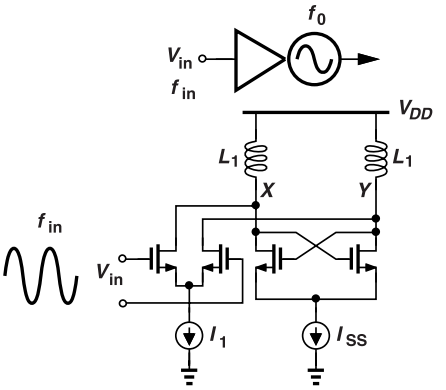


Figure 15.46 Injection of a periodic signal into an LC oscillator.

free-running frequency, i.e., the resonance frequency of the tanks, is equal to $f_0 \neq f_{in}$, it is possible to ensure that the circuit oscillates at f_{in} rather than at f_0. This "injection locking" phenomenon occurs if $|f_{in} - f_0|$ is sufficiently small and the input injection strength (determined by I_1/I_{SS} if all four transistors switch completely) is sufficiently large. The maximum $|f_{in} - f_0|$ for which the circuit is locked is approximately given by $[f_0/(2Q)](I_1/I_{SS})$, where Q denotes the quality factor of the tanks. For example, with $I_1 \approx 0.25 I_{SS}$ and $Q = 5$, the "lock range" is about $\pm 2.5\%$.

What happens if f_{in} is around $2f_0$ or $3f_0$? In this case, the circuit has little tendency to lock. Thus, for injection locking an oscillator to a higher harmonic than the fundamental, a mixing action is necessary that translates that harmonic to the vicinity of f_0. We call this action "downconversion." As illustrated conceptually in Fig. 15.47(a), the input and output are mixed, generating $f_{in} \pm f_{out}$. If $f_{in} - f_{out}$ falls near f_0 and is strong enough, the loop can operate as a frequency divider. For example, with $f_{in} = 2f_0 + \Delta f$, the mixer generates a component near $2f_0 + \Delta f - f_0 = f_0 + \Delta f$, which can injection lock the oscillator. The loop then settles so that $f_{out} = f_0 + \Delta f/2$.

Let us apply the principle illustrated in Fig. 15.47(a) to an LC oscillator. We recognize that a differential pair such as that in Fig. 15.47(b) can act as a mixer if M_1 and M_2 turn on and off. Since the mixer output is available in the current domain, it can be directly injected into the oscillator [Fig. 15.47(c)]. We must now decide how to connect the gates of M_1 and M_2 to the oscillator output. If they are respectively tied to X and Y, then these two transistors are configured as diode-connected devices, failing to perform mixing. On the other hand, if the gate of M_1 attaches to Y and that of M_2 to X, the differential pair is properly switched by the oscillator output. This leads us to the arrangement shown in Fig. 15.47(d). A closer look reveals that M_1 and M_2 in fact form a cross-coupled pair—just like M_3 and M_4. We thus simplify the design by merging these two transistors with M_3 and M_4 and applying the input to the tail current source of the oscillator [Fig. 15.47(e)]. In this case, M_3 and M_4 act as both a negative-G_m pair and a mixer. This topology is a candidate for $\div 2$ operation.

In summary, the injection-locked divider of Fig. 15.47(e) operates as follows: the current waveform generated by M_5 at f_{in} is mixed with the switching frequency of M_3 and M_4, f_{out}, yielding in their drain currents components at $f_{in} \pm f_{out}$. The sum is attenuated by the tanks and the difference is now similar to the injection in Fig. 15.46. The circuit thus locks such that $f_{in} - f_{out} = f_{out}$.

The mixing action provided by M_1 and M_2 in Fig. 15.47(b) or by M_3 and M_4 in Fig. 15.47(e) entails a

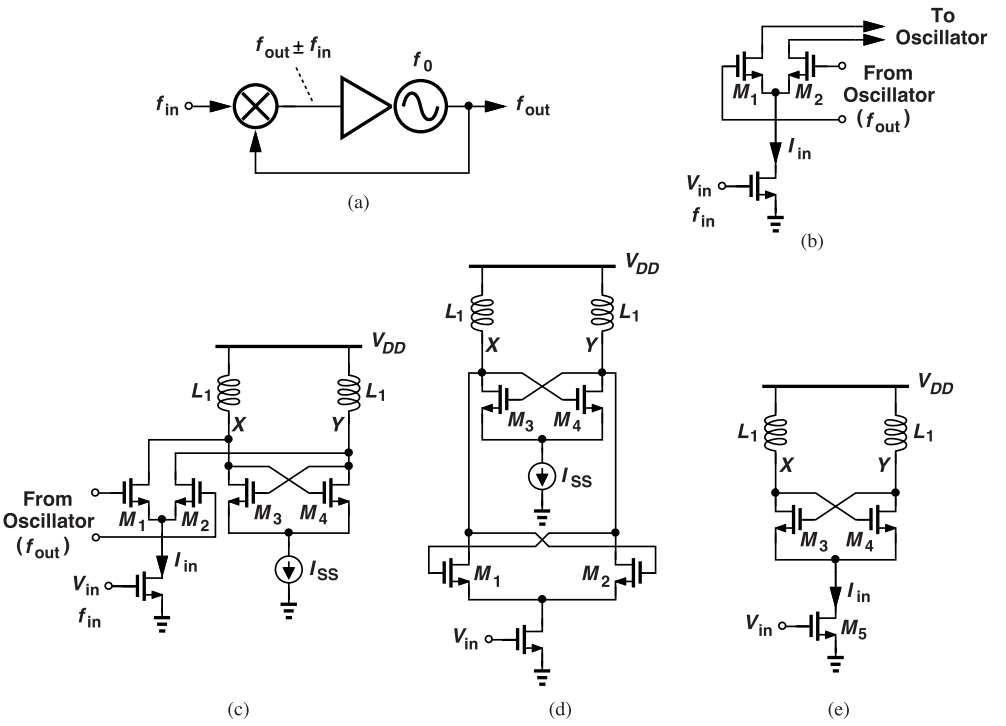


Figure 15.47 (a) Injection-locked divider model, (b) simple mixer, (c) injection and mixing results applied to LC oscillator, (d) circuit of (d) redrawn, and (e) injection through the tail path.

gain less than unity, thereby weakening the injection. Defined as the output differential current amplitude at the frequency of interest divided by the amplitude of I_{in}, this gain is equal to $2/\pi$. Since $I_{in} = g_{m5}V_{in}$, the (one-sided) lock range is approximately given by

$$|f_{in} - 2f_0| = \frac{f_0}{2Q}\frac{(2/\pi)g_{m5}V_{in}}{I_{SS}}, \tag{15.6}$$

where I_{SS} denotes the bias current of M_5. In practice, it is difficult to achieve a lock range greater than about $\pm 10\%$.

Injection locking can reduce the phase noise of an oscillator considerably [12]. However, the phase noise of injection-locked oscillators begins to approach their free-running profile as the circuit operates closer to the *edge* of the lock range. For example, if the topology in Fig. 15.47(e) provides a lock range from 55 GHz to 60 GHz, then its phase noise is negligibly reduced when it runs near one of these two extremes.

Ring oscillators, too, can be injection-locked to an input so as to act as frequency dividers. Consider, for example, the two-stage ring shown in Fig. 15.48(a), where the cross-coupled pairs alter the transfer function to allow oscillation (Chapter 5). To perform mixing according to Fig. 15.47(e), we can simply add two tail transistors that are controlled by complementary inputs [Fig. 15.48(b)]. This configuration achieves a much wider (fractional) lock range than the LC topology of Fig. 15.47(e). To understand this point, we recognize that each stage in Fig. 15.48(b) is in fact a latch similar to that in Fig. 15.21(a). In other words, this circuit can also be viewed as a simple FF-based frequency divider. The drawback, however, is that the maximum operation frequency of this structure is much lower than that of the LC arrangement.

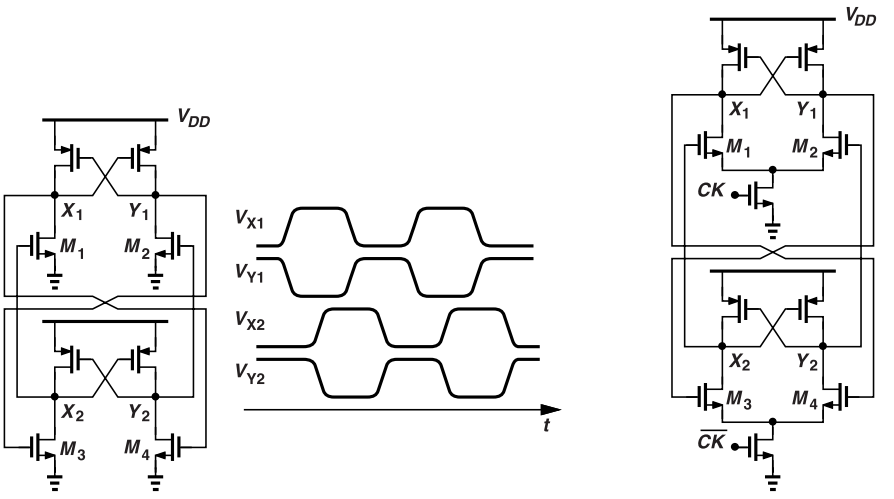


Figure 15.48 (a) Two-stage ring oscillator and its waveforms, and (b) use of tail devices for injection.

15.8 Fractional Dividers

The divider configurations studied thus far provide an integer modulus. In some applications, however, a fractional divide ratio proves useful. For example, consider an RF transmitter delivering a high amount of power to an antenna at a carrier frequency of f_0 (Fig. 15.49). Owing to the power amplifier (PA) nonlinearity,

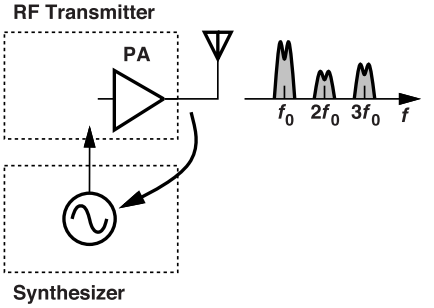


Figure 15.49 Problem of oscillator pulling in an RF transmitter.

the transmitter produces power around the higher harmonics as well. Since the PA output inevitably leaks to various parts of the system, it can potentially couple to the synthesizer's VCO and corrupt its phase. The difficulty here is that, even if it runs at $2f_0$ or $3f_0$, the VCO is not immune to this effect because the PA's higher harmonics may not be negligible. On the other hand, if the VCO oscillates at, say, $1.25f_0$, the corruption is much less.

Recall that the dividers studied previously clock their flipflops only on one edge. What happens if we utilize a double-edge-triggered flipflop (DETFF) such as those introduced in Chapter 14? Since a DETFF changes its output state on both clock edges, we expect the divide ratios to be *halved*. For example, let us return to the $\div 3$ circuit of Fig. 15.27(a) and analyze its behavior if it incorporates DETFFs. As depicted in Fig. 15.50, we begin with $Q_1\overline{Q_2} = 00$, noting that the rising edge of CK at $t = t_1$ causes $\overline{Q_2}$ to go high (why?). The falling edge of CK at $t = t_2$ also causes the FFs to change their states, with FF_1 and FF_2 reading $\overline{Q_2}$ and $Q_1 \cdot \overline{Q_2}$, respectively. Continuing the switching actions, we observe that one period of Q_1 or $\overline{Q_2}$ corresponds to 1.5 periods of CK. The circuit thus divides by 1.5. Note that these outputs exhibit a

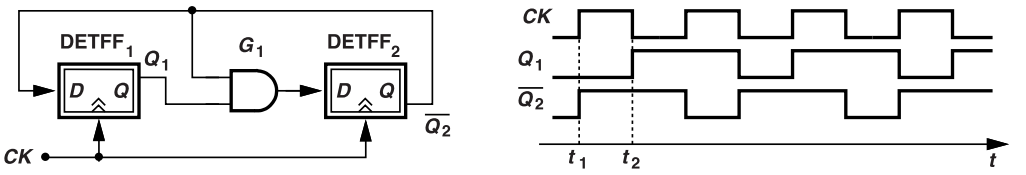


Figure 15.50 Use of a $\div 3$ circuit with double-edge-triggered flipflops to realize $\div 1.5$ operation.

duty cycle equal to 2/3. Similarly, if applied to the $\div 5$ topology of Fig. 15.33(a), this principle yields a $\div 2.5$ circuit.

The foregoing $\div 1.5$ structure must deal with two issues. First, the reader can show that, if the input duty cycle departs from 50%, the output contains spurious frequency components. Second, since the Q_1 or $\overline{Q_2}$ outputs do not have a 50% duty cycle, they cannot directly drive another similar divider. Such a case arises if we wish to obtain an overall divide ratio of, say, 2.25. Thus, duty cycle correction stages (Chapter 11) must precede (and follow) this circuit.

Another approach to obtaining fractional divide ratios is through the use of the Miller topology. Consider the arrangement shown in Fig. 15.51, where a $\div N$ stage is inserted in the feedback path. At node X, we

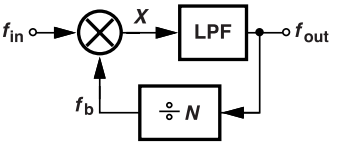


Figure 15.51 Fractional Miller divider.

have two frequency components: $f_{in} \pm f_b = f_{in} \pm f_{out}/N$. The sum is suppressed by the LPF, yielding $f_{out} = f_{in} - f_{out}/N$ and hence

$$f_{out} = \frac{N}{N+1} f_{in}. \tag{15.7}$$

The divider output can also prove useful:

$$f_b = \frac{1}{N+1} f_{in}. \tag{15.8}$$

For example, with $N = 2$, we have $f_{out} = f_{in}/1.5$ and $f_b = f_{in}/3$.

The structure shown in Fig. 15.51 entails two issues. First, the LPF output swing must be large enough to ensure proper operation of the $\div N$ circuit. This point stands in contrast to the behavior of the original Miller divider [Fig. 15.43(a)], where only a sufficiently high loop gain is necessary. Second, since the sum and difference components generated by the mixer are now closer to each other, the LPF must provide a steeper roll-off so as to attenuate the former.

In order to alleviate the second issue, the loop can incorporate a single-sideband (SSB) mixer, an arrangement that, nominally, generates only the difference (or only the sum) component. Based on the relation $\cos(\omega_1 - \omega_2)t = \cos\omega_1 t \cos\omega_2 t + \sin\omega_1 t \sin\omega_2 t$, an SSB mixer requires the quadrature phases of both ω_1 and ω_2, leading to the topology shown in Fig. 15.52 [11]. Here, the input signal is available in quadrature form and so is the feedback signal (by virtue of the $\div 2$ operation). At node X, we obtain $f_{in} - f_{out}$, which must be equal to $2f_{out}$. Thus, $f_{out} = f_{in}/3$. Also, node X provides a frequency equal to $f_{in}/1.5$. This structure achieves a wider frequency range than does the circuit in Fig. 15.51 for $N = 2$.

The principal drawback of the foregoing topology is that it requires quadrature input phases, a serious issue in terms of oscillator design. Moreover, the fractional output available at X suffers from spurious components due to mixer imperfections [7].

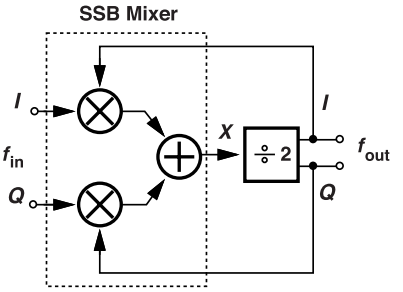


Figure 15.52 Fractional Miller divider using quadrature phases.

15.9 Divider Delay and Phase Noise

Divider circuits suffer from both delay and phase noise, thus affecting PLLs in which they are utilized. In this section, we study these phenomena.

Dividers can potentially exhibit a long delay from their input to their output. For example, the pulse-swallow counter described in Section 15.5.1 must allow the input edge to propagate through the prescaler and the program counter before it causes an output edge. We wish to analyze the effect of this delay on the PLL dynamics. For a stage having a delay of ΔT, the transfer function is given by $\exp(-\Delta T \cdot s)$, which can be approximated as $1 - \Delta T \cdot s$ if ΔT is much smaller than the time scales of interest. This condition holds in a PLL environment as the loop settling time is at least tens of input cycles whereas the divider delay is typically well below one input cycle. For a type-II PLL, we therefore have

$$H_{open}(s) = \frac{I_p}{2\pi}\left(R_1 + \frac{1}{C_1 s}\right)\frac{K_{VCO}}{Ms}(1 - \Delta T \cdot s). \tag{15.9}$$

Let us examine the magnitude and phase of this response, noting that the new zero at $1/\Delta T$ leads to (1) an asymptotic value of $I_p R_1 K_{VCO}\Delta T/(2\pi M)$ for $|H_{open}(j\omega = \infty)|$ and (2) an additional phase contribution of $\tan^{-1}(-\Delta T \cdot \omega)$. The response is plotted in Fig. 15.53, revealing that the unity-gain frequency has risen and the phase has become more negative, both of which degrade the phase margin.[3] To ensure negligible

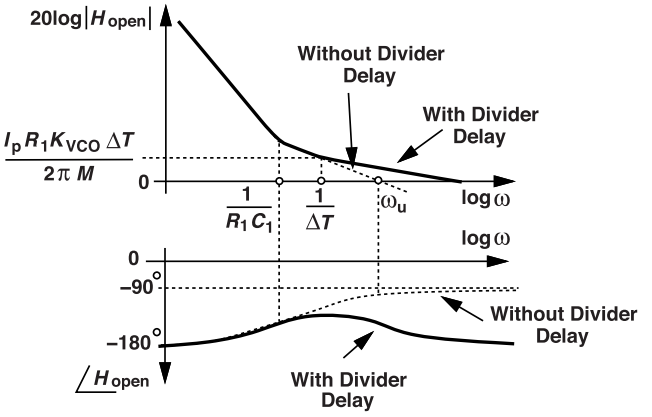


Figure 15.53 Effect of divider delay on PLL stability.

[3]Owing to other poles in the loop, $|H_{open}|$ eventually falls below the asymptotic value shown here.

degradation, we must have $1/\Delta T \gg \omega_u$, for example,

$$\frac{1}{\Delta T} = 5\omega_u \tag{15.10}$$

$$= 5\sqrt{2\zeta^2 + \sqrt{4\zeta^4 + 1}}\omega_n. \tag{15.11}$$

This condition is typically satisfied—unless ΔT exceeds the input period.

The phase noise of dividers corrupts the feedback waveform that reaches the phase detector. As illustrated in Fig. 15.54, this phase noise, $\phi_{n,div}$, is indistinguishable from the input and hence experiences the low-pass

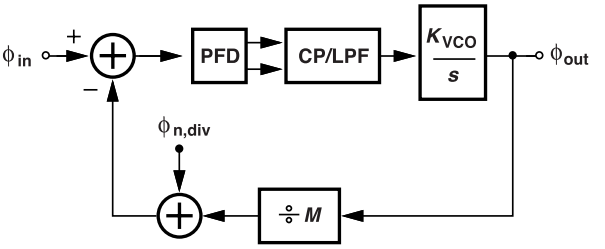


Figure 15.54 Effect of divider phase noise.

transfer function of the loop as it appears in ϕ_{out}. If employing a large number of asynchronous stages, the divider can thus introduce significant in-band noise.

It is possible to remove the feedback divider phase noise by means of a flipflop. Depicted in Fig. 15.55(a), the idea is to retime the divider output using the VCO signal as the clock. Figure 15.55(b) shows that jitter in V_A does not appear in V_B because the latter changes only on the falling edges of V_{out}.

This method, however, faces an issue related to with PVT variations. Since the divider delay is PVT-dependent, the falling edges of V_{out} in Fig. 15.55(b) can occur very close to the transitions of V_A [Fig. 15.55(c)], making the FF metastable. In this case, the FF output is ill-defined for a long time, which can confuse the PFD and also accumulate substantial phase noise. In other words, if the total divider delay varies by half a VCO period with PVT, then it is difficult to avoid this phenomenon. The retiming technique should therefore be used only in very demanding applications.

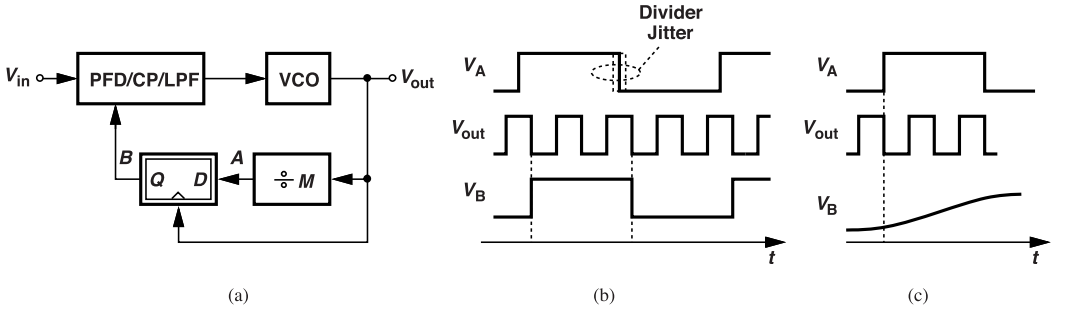


Figure 15.55 (a) Removal of divider phase noise by a retimer, (b) circuits' waveforms, and (c) problem of retimer metastability.

References

[1]. J. Yuan and C. Svensson, "High-speed CMOS circuit technique," *IEEE J. Solid-State Circuits,* vol. 24, pp. 62-70, Feb. 1989.

[2]. L. Kong, Y. Chang and B. Razavi, "A 14 μm x 26 μm 20-Gb/s 3-mW CDR circuit with high jitter tolerance" *Symp. VLSI Circuits Dig. Of Tech. Papers*, pp. 271-272, June 2018.

[3]. B. Chang, J. Park, and W. Kim, "A 1.2-GHz CMOS dual-modulus prescaler using new dynamic D-type flip-flops," *IEEE J. Solid-State Circuits,* vol. 31, pp. 749-754, May 1996.

[4]. J. W. Jung and B. Razavi, "A 25-Gb/s 5-mW CMOS CDR/deserializer," *IEEE J. Solid-State Circuits*, vol. 48, pp. 684-697, Mar. 2013.

[5]. B. Razavi, K. F. Lee, and R. H. Yan, "Design of high-speed low-power frequency dividers and phase- locked loops in deep submicron CMOS," *IEEE J. Solid-State Circuits*, vol. 30, pp. 101-109, Feb. 1995.

[6]. R. Adler, "A study of locking phenomena in oscillators," *Proc. IEEE*, vol. 61, pp. 1380-1385, Oct. 1973.

[7]. B. Razavi, "Cognitive radio design challenges and techniques," *IEEE J. Solid-State Circuits*, vol. 45, pp. 1542-1553, Aug. 2010.

[8]. R. L. Miller, "Fractional-frequency generators utilizing regenerative modulation, " *Proc. IRE,* vol. 27, pp. 446-456, July 1939.

[9]. C. S. Vaucher, et al, "A Family of low-power truly modular programmable dividers in standard 0.35-μm CMOS technology," *IEEE J. Solid-State Circuits,* vol. 35, pp. 1039-1045, July 2000.

[10]. J. Lee and B. Razavi, "A 40-GHz frequency divider in 0.18-μm CMOS technology," *IEEE J. Solid-State Circuits,* vol. 39, pp. 594-601, Apr. 2004.

[11]. C.-C. Lin and C.-K. Wang, "A regenerative semi-dynamic frequency divider for mode-1 MB-OFDM UWB hopping carrier generation," *ISSCC Dig. Tech. Papers,* pp. 206-207, Feb. 2005.

[12]. B. Razavi, "A study of injection locking and pulling in oscillators," *IEEE J. Solid-State Circuits,* vol. 39, pp. 1415-1424, Sep. 2004.

Problems

15.1. In the circuit of Fig. 15.2, a fraction of the oscillator's output signal leaks to the supply of the self-biased inverters, thereby modulating their delay. Examine the output spectrum of the inverters.

15.2. In the circuit of Fig. 15.2(a), the oscillator draws from V_{DD} a large transient current at twice the oscillation frequency. If this component leaks to the supply of the self-biased inverters, examine their output spectrum.

15.3. In the circuit of Fig. 15.2(b), the divider divides by M and draws a transient current from its supply at a frequency of f_{in}/M, where f_{in} is the input frequency. If this component leaks to the supply of the self-biased inverters, examine their output spectrum.

15.4. Suppose CK and CK in Fig. 15.4(a) have very slow transitions, causing the input and feedback switches to remain on for a long time. Explain how the fight between D_{in} and D_F can introduce an error here.

15.5. Suppose CK and $\overline{CK}$ in Fig. 15.6(a) have rail-to-rail swings with a transition time of t_1. Sketch the waveform at their common sourc node and, assuming square-law devices, find the minimum voltage at this node. Assume CK and $\overline{CK}$ cross at $V_{DD}/2$ and, at this point, M_5 and M_6 operate in saturation.

15.6. Repeat the previous problem if M_5 and M_6 operate in the deep triode region when CK and CK cross.

15.7. In the circuit of Fig. 15.10, we apply D to M_2 and M_3 and the clocks to M_1 and M_4. Considering the parasitic capacitances at the drains of M_1 and M_4, explain what happens if Q is high, CK is low, and D goes high.

15.8. Examine the operation of the stages in Figs. 15.13(a) and (b) if transistors M_2 and M_5 are realized as PMOS devices.

15.9. Repeat the previous problem for the circuit of Fig. 15.13(c).

15.10. Does the circuit of Fig. 15.13(c) operate properly if the two stages are swapped?

15.11. The TSPC FF of Fig. 15.13(d) is configured as a $\div 2$ circuit by tying its output to its D input. Plot the waveforms at A, B, and Q. Do you observe any charge sharing between the capacitance at node B and the parasitic capacitance at the drain of M_6?

15.12. We wish to implement the $\div 3$ circuit of Fig. 15.27 using the TSPC FF of Fig. 15.13(d). Noting that this FF always inverts, explain whether FF_1, G_1, and FF_2 in Fig. 15.27 can be so realized to avoid speed penalty due to additional inverters in the signal path. (Gate G_1 can be merged with FF_2.)

15.13. Repeat the previous problem using the C^2MOS latch of Fig. 15.10.

15.14. Plot the waveforms at the outputs of FF_1, G_1, and FF_2 in Fig. 15.29(c). Do all three outputs provide the desired waveform?

15.15. Using the C^2MOS latch of Fig. 15.10, design the $\div 2/3$ circuit of Fig. 15.30. Can the OR and AND gates be merged with FF_2 as in Example 15.10?

15.16. Plot the waveforms in Fig. 15.41 if MC_2 is high or if it's low.

15.17. Consider the $\div 2/3$ circuits shown in Fig. 15.30 and Fig. 15.42(c). Examine and compare their speed limitations. (Hint: start with Example 15.9.)

15.18. In the Miller divider of Fig. 15.43(b), how should the common-mode level of CK and $\overline{CK}$ be chosen to improve the speed?

15.19. Estimate the supply current of the Miller divider in Fig. 15.43(b). Assume square-law devices.

15.20. What is the minimum allowable supply voltage for the circuit of Fig. 15.43(b)?

Index

$1/f$ noise, 36
$\Delta\Sigma$ fractional-N synthesizer, 382
$\Delta\Sigma$ modulator, 379, 380, 382, 439
$\Delta\Sigma$ modulator noise peaking, 388
$\Delta\Sigma$ quantization noise, 383

Absolute jitter, 49, 52, 53
Active notch filter, 389
Additive noise, 144
Alexander CDR loop, 418
Alexander PD, 416
Alexander PD properties, 417
All-digital PLL, 305
AM and PM components, 145
AM/PM conversion, 159
Amplifier noise in type-I PLL, 227
Amplifier noise in type-II PLL, 233
Anti-parallel inverters, 127
Anti-phase coupling, 196
Approximation by impulses, 39
Approximation of noise spectrum, 39
Approximation of transconductance, 152
Array filling, 340
Asymmetry between rise and fall times, 111
Asynchronous divider, 469
Asynchronous stages, 483
Autocorrelation, 49
Average loop gain, 11
Average power of noise, 38

Back-to-back inverters, 456
Backplane, 434
Band-pass tail noise, 156
Bang-bang CDR, 413
Bang-bang characteristic, 411
Bang-bang PD with jitter, 412
Bang-bang phase detector, 411
Barkhausen's criteria, 3
Base station, 370
Baseband signal, 368
Basic inductor model, 142
Basic TDC topology, 310
Blind zone, 85, 95
Blind zone in DCO, 333
Bluetooth, 369, 370
Bode plots of PLL, 221
Bode plots of third-order PLL, 240
Boltzmann constant, 35
Boundary detector, 441
Boundary errors in DLL, 353
Bubble in thermometer code, 308

C^2MOS divide-by-2 circuit, 460
C^2MOS latch, 459
Capacitive coupling, 457
Capacitive coupling of varactors, 191
Capacitor DAC, 332
Carrier, 43
Carrier frequency, 40
Carrier power, 46
Cascaded $\Delta\Sigma$ modulators, 386
CDR bandwidth, 448
CDR using analog PI, 438
CDR using digital PI, 438
Channel spacing, 369
Channel-length modulation in charge pump, 252
Charge injection in charge pump, 255
Charge pump, 227, 422
Charge pump design, 278
Charge pump flicker noise, 272
Charge pump leakage, 256
Charge pump noise, 269, 280
Charge pump nonlinearity, 393
Charge pump output impedance, 255
Charge pump settling, 395
Charge pump switching behavior, 280
Charge pump topologies, 242
Charge pump turn-on, 281
Charge pump voltage compliance, 252
Charge pump white noise, 270
Charge-pump PLL, 227
Charge-pump PLL stability, 231
Charge-pump PLL transfer function, 229
Class-C oscillator, 192
Class-C oscillator design, 193
Clock feedthrough in charge pump, 255
Clock recovery, 404
Clock recovery by phase-locking, 410
Clock skew, 346
CM level of oscillator with complementary pair, 186
CML NOR gate, 458
CMOS-CML level converter, 422
Coarse DAC, 333
Coarse VCO control, 443
Colpitts oscillator, 27
Column thermometer code, 332
Common-mode level modulation, 159
Common-mode rejection, 458
Comparator, 307
Comparator offset, 307
Complementary oscillator design, 190
Complete switching of cross-coupled pair, 161
Continuous tuning, 84, 169
Continuous-time approximation, 250
Control voltage, 84
Control voltage waveform, 286
Convolution, 182

Correlation, 407
Counters acting as a lock detector, 302
Coupled oscillators, 126
Coupling coefficient, 201
Critical path in divider, 468
Cross-coupled oscillator, 23
Cross-coupled pair, 478
Cross-coupled pair transconductance, 151
Cross-coupled pairs, 120
Crystal oscillator, 215, 370
Current regulation, 97
Current-limited phase noise, 186
Current-mode logic, 456
Current-steering DAC, 329
Cycle-to-cycle jitter, 52, 53
Cyclostationary noise, 146

D latch, 455
DAC feedforward, 396, 397
DAC glitch, 330
DAC segmentation, 331
Damping factor, 220
Data center, 404
Data recovery using edge-detected data, 409
Data sample, 422
Data-dependent jitter, 414
DCO noise spectrum, 327
DCO nonlinearity, 337
DCO topologies, 333
DCO tuning by tail current, 339
DCO tuning by variable load resistance, 339
DCO with array filling, 340
DCO with matrix architecture, 332
DCO with noise shaping, 340
DCO with resistor-ladder DAC, 328
Decision threshold, 307
Delay element, 316
Delay line, 436
Delay line with supply noise, 315
Delay mismatch in DLL, 357
Delay-locked loop, 346
Departure of oscillation from resonance, 199
Desired frequency, 72
Deterministic, 50
Deterministic disturbance, 42
Deterministic jitter, 41, 43, 52, 53
Device leakage, 459
Device noise, 35
Differential feedback oscillators, 21
Differential logic, 463
Differential oscillator waveforms, 109
Differential PI, 362
Differential ring design improvement, 111
Differential ring oscillator, 103
Digital CDR, 440
Digital filter, 305
Digital loop filter, 305
Digital PLL, 305
Digital PLL dynamics, 341
Digital PLL transfer function, 343
Digital-to-analog converter (DAC), 328
Digital-to-time converter, 399
Digitally-controlled oscillator (DCO), 305, 326
Discrete frequency tuning, 93
Discrete tuning, 84, 95, 170
Dithering, 384
Divide ratio, 372
Divide-by-1.5 circuit, 480
Divide-by-15/16 circuit, 471
Divide-by-2 circuit, 455, 464
Divide-by-2 circuit with 25% duty cycle, 465
Divide-by-2/3 circuit, 469
Divide-by-2/3 stages in a loop, 474
Divide-by-3 circuit, 299, 467
Divide-by-3/4 circuit, 469
Divide-by-5 circuit, 300, 469
Divide-by-8/9 circuit, 470
Divider behavioral model, 285
Divider delay, 482
Divider failure, 454
Divider modulus, 372
Divider output phase noise, 294, 386
Divider phase noise, 482
Divider sensitivity, 453
Divider with no static current, 466
DLL dynamic behavior, 347
DLL dynamics, 347
DLL generating multiple phases, 352
DLL linear model, 348
DLL nonidealities, 349
DLL with edge combining, 354
DLL with no false lock, 353
DLL with supply noise, 350
DLL-based CDR, 434
DLL/PLL hybrids, 358
Double-edge triggered flipflop, 480
Downconversion, 60
Drain-switched CP, 227
Drawbacks of type-I PLL, 223
DTC gain error, 400
DTC nonlinearity, 400
Dual-modulus divider, 467
Duty cycle correction, 364
Duty cycle error, 400
Duty cycle error in DLL, 346
Dynamic latches, 459
Dynamic PLL behavior, 215

Early clock, 416
Edge detection, 408
Edge detector, 423
Edge detector in bang-bang CDR, 415

Edge misalignment in hybrid DLL/PLL, 359
Edge sample, 422
Effect of duty cycle error on doubler output, 297
Effect of fanout on phase interpolation, 361
Effect of ripple on DLL, 349
Effect of ripple on PLL output, 287
Effect of unequal rise and fall times, 185
Excess delay, 448
Excess phase, 29
Eye diagram, 42

False lock in DLL, 353
Fast and slow paths, 124
Feedback divider design, 299
FF skew, 424
Figure of merit, 79, 80, 82
Fine DAC, 333
Finite state machine, 441
Flash ADC, 307
Flat zone, 88
Flicker noise, 36, 49, 57, 76, 79
Flicker noise corner, 107
Flicker noise corner frequency, 38, 77
Flicker noise of cross-coupled pair, 160
Flicker-noise-induced phase noise, 69
Flipflop metastability, 413, 483
Four-quadrant multiplier, 441
Fourier expansion of transconductance, 151
Fractional divider, 480
Fractional Miller divider, 481
Fractional spur, 373
Fractional-N synthesizer, 368, 371
Free-running, 48
Frequency acquisition, 443
Frequency detector, 443
Frequency divider, 215
Frequency division, 75
Frequency doubler design, 296
Frequency doubling, 400
Frequency doubling issues, 296
Frequency glitch, 330
Frequency mismatch, 435
Frequency multiplication, 215
Frequency multiplication by addition, 355
Frequency multiplication for avoiding false lock, 358
Frequency noise, 55, 67, 375
Frequency offset, 45, 47, 48
Frequency tuning, 83, 166
Frequency-multiplying charge-pump PLL, 232
Frequency-multiplying DLL, 354
Frequency-multiplying PLL, 215
Frequency-multiplying PLL transfer function, 222
Full-scale range, 307

Gain error in DCO, 337
Gate leakage, 261
Gate-switched charge pump, 243
Gaussian jitter, 50
Gaussian PDF, 413
Gilbert cell, 419, 441
Growing amplitude, 117
Growing oscillatory output, 2

Half-Rate Bang-Bang PD, 428
Half-rate Hogge PD, 432
Half-rate linear PD, 433
Half-rate PD with quadrature phases, 429
Half-rate phase detector, 428
Hand-off, 369
High-order $\Delta\Sigma$ modulator, 384
High-speed PFD, 363
Higher-order PLLs, 238
Hogge CDR loop, 425
Hogge PD, 423

Ideal integrator, 217
Ideal tank, 144
Ideal waveforms, 12
Imbalance, 1
Impulse of energy, 405
Impulse sensitivity, 180
Impulse sensitivity function (ISF), 181
In-phase coupling, 196
Inductor, 138
Inductor coupling to substrate, 140
Inductor modeling, 138
Inductor parasitic capacitance, 138
Inductor trace resistance, 140
Injection at zero crossing, 179
Injection locking, 127, 196, 478
Injection pulling, 196
Injection-locked divider, 478
Injection-locked oscillators, 104
Instantaneous frequency, 213
Integer-N synthesizer, 368, 370
Integrated phase noise, 165
integrator, 438
Interconnect capacitance, 114
Interpolation between I and Q, 324
Interpolation in flash ADC, 308
Interpolation network, 130
Inverter-based differential ring, 126
Inverter-based ring with quadrature outputs, 130
Inverter-based rings, 12
ISF and Cyclostationary noise, 186
ISF and Flicker Noise, 184

Jitter, 34, 41, 44, 50–53, 56, 58, 59
Jitter accumulation in ring oscillator, 358
Jitter characteristics, 444
Jitter concepts, 40
Jitter generation, 445

Jitter in analog-to-digital converter, 265
Jitter peaking, 246, 446, 447
Jitter reduction, 294
Jitter tolerance, 448
Jitter transfer, 445
Junction leakage, 172
Junction turn-on, 172

Kickback noise, 414
King in interpolation, 325

Latch design styles, 454
Latch transparency, 460
Latch with merged NAND gate, 299
Latch with source followers, 465
Latch-up, 17
Late clock, 416
Layout, 131
Layout extraction tool, 140
LC circuit, 13
LC DCO, 334
LC DCO floor plan, 328
LC DCO with coarse and fine tuning, 333
LC oscillator, 138
LC oscillator figure of merit, 154
LC oscillators, 13, 58
LDO, 126
Leakage in PLL, 260
Level shift, 458
Linear PD, 425
Linear PLL model, 220
Linear scaling, 121
Linearity, 84
Local oscillator, 368
Lock detector, 301, 443
Loop bandwidth, 220
Loop capacitor reduction, 261
Loop filter, 212
Loop filter noise, 272, 295
Lorentzian spectrum, 49
Lossy, 13
Low-dropout regulator (LDO), 96, 191

Majority voting, 440
Master-slave flipflop, 195
Matrix DAC architecture, 331
Measurement of duty cycle error, 365
Metastable, 59
Miller divider, 476
Miller divider with quadrature outputs, 476
Mixer, 478
Mixer as phase detector, 210
Mixing of tail noise, 156
Modulus randomization, 373
MOS varactor, 167
Multi-path TDC, 321
Multiphase DLL, 352
Multiplexer, 434
Mutual coupling, 196

Narrowband filter, 405
Narrowband FM approximation, 42, 44, 49
Natural frequency, 220
Negative capacitance, 338
Negative resistance, 24, 116
Negative-G_m pair, 478
Negative-resistance circuits, 24
Node capacitances, 73, 113
Noise accumulation, 39, 52
Noise folding, 392, 395
Noise in frequency domain, 34
Noise in time domain, 34
Noise power, 46
Noise shaping, 376
Noise shaping by feedback, 377
Noise transfer function, 37
Noise-shaping function, 378
Nonlinearity in fractional-N loop, 392
Normalized output spectrum, 46
Norton noise equivalent, 149
Notch filter, 389
NRZ data, 406
Number of stages, 74
Nyquist's stability criterion, 30

One-pole PLL approximation, 289
One-pole system, 446
One-port oscillator, 24
One-port systems, 23
One-sided spectrum, 145
Open-loop Miller divider, 477
Open-loop VCO calibration, 303
Oscillation amplitude, 11
Oscillation conditions, 4
Oscillation growth, 6
Oscillator analysis methods, 4
Oscillator design procedure, 173
Oscillator noise-power trade-off, 154
Oscillator topologies, 9
Oscillator tuning, 122
Oscillator with complementary pairs, 186
Oscillator with large transistors, 114
Oscillatorless CDR, 434
Oscillators coupled by diode-connected devices, 204
Oscillators coupled by inverters, 127
Oscillatory feedback system, 2
Oscillatory loop, 3
Output noise spectrum, 38
Output voltage spectrum, 47
Oversampling, 418

Parallel inductor model, 142

Parallel resistance, 13
Parseval's theorem, 38, 51
Pendulum, 1
Periodically-switched noise, 147
Periodically-switched resistance, 149
PFD implementation, 225
PFD operation, 224
PFD phase noise, 251
PFD/CP interface, 283
Phase detector, 209, 483
Phase detector characteristic, 210
Phase detector gain, 209
Phase diffusion constant, 49
Phase error, 209, 448
Phase error due to CP mismatch, 252
Phase fluctuation, 449
Phase fluctuations, 219, 220, 231
Phase impulse response, 178
Phase impulse response of LC tank, 181
Phase interpolation, 124, 130, 360
Phase interpolator, 436
Phase jumps in fractional-N loop, 374
Phase mismatch, 131
Phase noise, 34, 44, 50, 53, 54, 56–60, 66, 67, 75, 77, 80, 81, 128, 143, 153, 163, 201, 369, 375
Phase noise concepts, 40
Phase noise equation, 154
Phase noise in DLL, 351
Phase noise in PLLs, 263
Phase noise of complementary oscillator, 187
Phase noise of delay lines, 64
Phase noise of differential ring, 105
Phase noise of LC tank, 183
Phase noise of ring oscillators, 64
Phase noise reduction by frequency division, 193
Phase noise spectrum, 45, 47
Phase noise suppression bandwidth, 278
Phase noise trade-off with power dissipation, 75
Phase offset, 213
Phase resolution, 437
Phase shift, 66
Phase slope, 66, 67
Phase-locked loop, 211
Phase/frequency detector (PFD), 224
PI using inverters, 360
PI-based CDR, 435
Plesiochronous operation, 441
PLL as a low-pass filter, 216
PLL behavioral model, 284
PLL corner frequency, 249
PLL design procedure, 277
PLL dynamics, 482
PLL jitter estimate, 268
PLL loop gain, 217
PLL noise optimization, 269
PLL output integrated jitter, 263
PLL phase margin, 221
PLL poles, 246
PLL settling time, 244
PLL simplification, 283
PLL static behavior, 213
PLL transfer function, 219, 289
PLL transfer functions, 246
PLL transient behavior, 286
PLL unity-gain bandwidth, 247
PLL using an integrator, 254
PLL with frequency acquisition, 224
PLL with input phase modulation, 291
PLL with input phase noise, 263
PLL with reference frequency doubler, 298
PLL with VCO phase noise, 266
PLLs with discrete tuning, 259
Polarity of feedback in PLL, 232
Pole at origin, 221
Positive feedback, 8, 456
Power, 53, 54
power spectral density, 35
PRBS data, 450
Preliminary design ideas of ring oscillators, 70
Prescaler, 467
Probability density function, 35
Process corner, 280
Program counter, 472
Propagation of noise, 37
Properties of random data, 406
Proportional and integral paths, 341
Proportional pulse, 425
Pull-up resistors in oscillator, 172
Pulse swallow divider, 472
PVT variations, 483

Quadrants, 441
Quadrature generation, 129, 195
Quadrature generation by frequency division, 195
Quadrature LC oscillator, 196
Quadrature oscillator design example, 202
Quadrature oscillator design issues, 201
Quality factor, 16
Quantization, 306
Quantization noise, 306, 374, 382, 439
Quantization noise power, 306
Quantization noise reduction, 396
Quantization noise shaping, 386
Quantization noise spectrum, 306
Quantizer, 305, 306
Quasi-differential delay line, 315

Random, 50, 51
Random data, 404
Random disturbance, 42
Random mismatch in charge pump, 254
Ratioed TSPC, 462

Reciprocal mixing, 60
Reference doubling, 296
Reference spurs, 286
Regeneration, 456
Regeneration time constant, 457
Resistive tuning, 122
Resistor-ladder DAC, 328
Resolution bandwidth, 47
Resonator, 13
Retimer, 404
RF transmitter, 480
Right half plane, 117
Ring DCOs, 339
Ring oscillator, 434, 479
Ring oscillator design, 71
Ring oscillators, 9, 56
Rings with LC loads, 134
Ripple, 214, 238, 425
Row thermometer code, 332
RS latch, 318, 473
Run length, 406

Sampling loop filter, 260
Sawtooth ripple waveform, 287
Second harmonic, 106, 111
Segmented DAC, 331
Self-biased inverter, 104, 454
Self-oscillation frequency, 460
Series resistance, 13
Shaped quantization noise, 397
Shaped VCO phase noise, 266
Sidebands in PLL, 274
Sinusoidal jitter, 450
Skew between clock and VCO, 211
Skew compensation, 425
Skew in bang-bang CDR, 414
Skin effect, 140
Slew rate in oscillator, 152
Slow or fast phase changes, 219
Soft triode region, 162
Source-switched charge pump, 242
spectrum, 35
Spectrum of NRZ data, 406
Spurs, 43
SSB mixer, 481
Startup condition, 3
Static latches, 455
Static offset, 445
StrongArm latch, 317
StrongArm latch waveforms, 317
Supply dependence, 103
Supply noise, 55, 96
Supply noise in PLL, 222, 274
Supply sensitivity, 55, 103, 111
Swallow counter, 472
Switch parasitic capacitance, 170
Symmetric inductor, 138
Symmetric XOR, 419
Synthesizer, 368, 480

Tail capacitance, 111, 160
Tail current, 104
Tail current doublet, 160
Tail current flicker noise, 108
Tail flicker noise, 158
Tail noise, 155, 166
Tail noise equation, 157
Tail noise in quadrature oscillator, 201
Tail noise reduction, 169
Tail thermal noise, 155
TDC characteristic, 309
TDC design, 316
TDC dynamic range, 314
TDC gain variation, 315
TDC imperfections, 315
TDC quantization noise, 312
TDC with delay mismatches, 315
TDC with fanout imbalance, 311
TDC with quadrature clocks, 323
TDC with zero-gain region, 311
TDC/oscillator combination, 323
Temperature corner, 280
Thermal noise, 36, 47, 58, 76, 79, 113
Thermal noise floor, 77
Thermal noise suppression, 166
Thermometer code, 94, 307
Three-point oscillator, 25
Three-stage ring oscillator, 103
Time-to-digital conversion, 309
Time-to-digital converter (TDC), 305
Time-variant system, 183
Tones in $\Delta\Sigma$ modulator, 383
Top-biased oscillator, 189
Transceiver, 368, 434
Transformation of noise current to voltage, 146
Transistor lengths, 74
Transistor noise simulations, 76
Transition density, 446
Transmitter, 368
Triode region, 160, 164
TSPC divide-by-3/4 circuit, 469
TSPC logic, 460
TSPC logic with split paths, 462
TSPC PFD, 363
Tuned amplification, 16
Tuned load, 16
Tuning, 90
Tuning characteristic, 86, 88, 93
Tuning range, 84, 86, 166, 169
Twin-T filter, 389
Two-stage ring design, 120
Two-stage ring oscillator, 116

TX PLL, 434
Type-I PLL, 220
Type-I PLL lock acquisition, 223
Type-I PLL stability, 223
Type-II PLL, 230
Type-II PLL phase margin, 236
Type-II PLL transfer function, 231

Underdamped loop, 215
Unilateral coupling, 196
Unit interval, 58
Unratioed logic, 455
Up and down misalignment, 283
Up and down mismatch, 251, 278
Up and down skew, 251
Upconversion, 368
Upconverter, 368

V/I converter, 418
Varactor, 90, 91
Varactor C/V characteristic, 124
Varactor channel resistance, 167
Varactor nonlinearity, 159
Varactor quality factor, 168
Varactor tuning, 123
Variable capacitance, 90
Variable resistance, 85
Variable-delay stage, 346
Vaucher divider, 474
VCDL static behavior, 347
VCO behavioral model, 284
VCO calibration, 301, 444
VCO gain, 446
VCO models, 30
VCO phase noise, 249
VCO phase noise model, 48, 293
VCO phase noise suppression, 293
Vernier TDC, 319
Voltage regulation, 96
Voltage swing of oscillator with complementary pairs, 187
Voltage-controlled delay line (VCDL), 346
Voltage-controlled oscillators, 27
Voltage-limited phase noise, 186

White noise, 48
White spectrum, 36
White-noise-induced jitter, 48
WiFi, 166

XOR gate, 410
XOR gate as phase detector, 210
XOR implementation, 419