

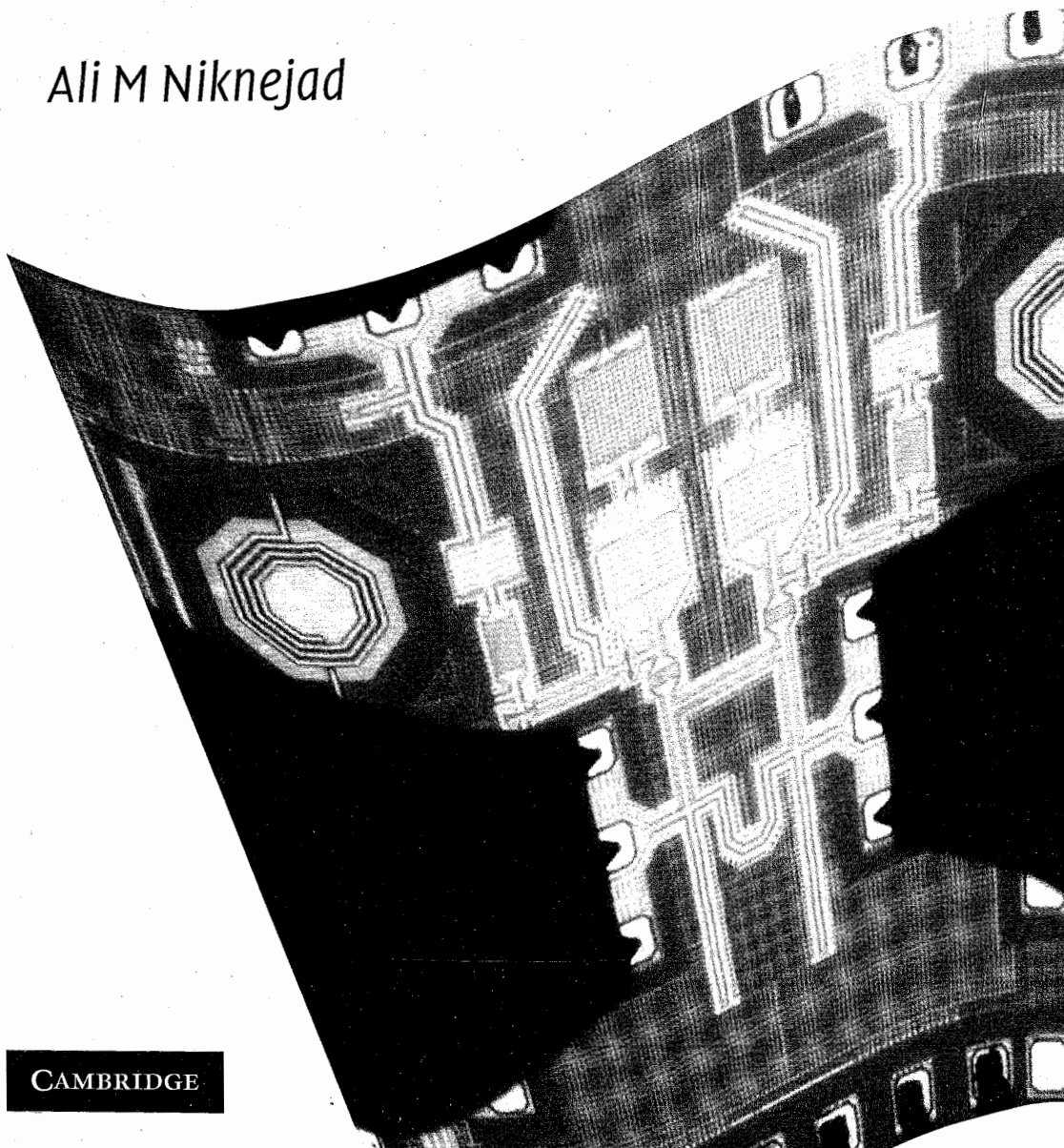
Niknejad

Electromagnetics for High-Speed Analog
and Digital Communication Circuits

Electromagnetics

for High-Speed Analog *and*
Digital Communication Circuits

Ali M Niknejad



CAMBRIDGE

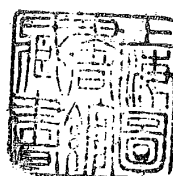
TN03
N692

RIDGE

7/20/04
NB92

Electromagnetics for High-Speed Analog and Digital Communication Circuits

ALI M. NIKNEJAD



7/20/04

782

CAMBRIDGE UNIVERSITY PRESS

Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo

Cambridge University Press

The Edinburgh Building, Cambridge CB2 2RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org

Information on this title: www.cambridge.org/9780521853507

© Cambridge University Press 2007

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 2007

Printed in the United Kingdom at the University Press, Cambridge

A catalog record for this publication is available from the British Library

Library of Congress Cataloging in Publication data

ISBN-13 978-0-521-85350-7 hardback

ISBN-10 0-521-85350-8 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.

Contents

<i>Preface</i>	<i>page</i>	ix
<i>Acknowledgments</i>		xi
1	Introduction	1
1.1	Motivation	1
1.2	System in Package (SiP): chip and package co-design	13
1.3	Future wireless communication systems	13
1.4	Circuits and electromagnetic simulation	15
2	Capacitance	18
2.1	Electrostatics review	18
2.2	Capacitance	32
2.3	Non-linear capacitance	41
2.4	References	52
3	Resistance	53
3.1	Ohm's Law	53
3.2	Conduction in semiconductors	59
3.3	Diffusion	66
3.4	Thermal noise	68
3.5	References	73
4	Ampère, Faraday, and Maxwell	74
4.1	Ampère: static magnetic fields	74
4.2	Magnetic materials	82
4.3	Faraday's big discovery	88
4.4	Maxwell's displacement current	91
4.5	References	95
5	Inductance	96
5.1	Introduction	96
5.2	Inductance	97
5.3	Magnetic energy and inductance	101
5.4	Discussion of inductance	106

5.5	Partial inductance and return currents	119
5.6	Impedance and quality factor	120
5.7	Frequency response of inductors	121
5.8	Quality factor of inductors	130
5.9	Inductors and switching circuits	133
5.10	Preview: how inductors mutate into capacitors	135
5.11	References	136
6	Passive device design and layout	137
6.1	Ring inductor	137
6.2	The classic coil	141
6.3	Spirals	143
6.4	Symmetric inductors	145
6.5	Multilayer inductors	147
6.6	Inductor equivalent circuit models	149
6.7	Integrated capacitors	150
6.8	Calculation by means of the vector potential	153
6.9	References	165
6.10	Appendix: Filamental partial mutual inductance	165
7	Resonance and impedance matching	168
7.1	Resonance	168
7.2	The many faces of Q	180
7.3	Impedance matching	186
7.4	Distributed matching networks	199
7.5	Filters	199
7.6	References	200
8	Small-signal high-speed amplifiers	201
8.1	Broadband amplifiers	202
8.2	Classical two-port amplifier design	220
8.3	Transistor figures of merit	242
8.4	References	244
9	Transmission lines	246
9.1	Distributed properties of a cable	246
9.2	An infinite ladder network	248
9.3	Transmission lines as distributed ladder networks	249
9.4	Transmission line termination	253
9.5	Lossless transmission lines	255
9.6	Lossy transmission lines	260
9.7	Field theory of transmission lines	264
9.8	T-line structures	265
9.9	Transmission line circuits	272

9.10	The Smith Chart	282
9.11	Transmission line-matching networks	287
9.12	References	292
10	Transformers	293
10.1	Ideal transformers	293
10.2	Dot convention	294
10.3	Coupled inductors as transformers	295
10.4	Coupled inductor equivalent circuits	296
10.5	Transformer design and layout	299
10.6	Baluns	301
10.7	Hybrid transformer	302
10.8	Transformer parasitics	305
10.9	Transformer figures of merit	305
10.10	Circuits with transformers	310
10.11	References	319
11	Distributed circuits	320
11.1	Distributed RC circuits	320
11.2	Transmission line transformers	325
11.3	FETs at high frequency	332
11.4	Distributed amplifier	335
11.5	References	342
12	High-speed switching circuits	343
12.1	Transmission lines and high-speed switching circuits	343
12.2	Transients on transmission lines	345
12.3	Step function excitation of an infinite line	346
12.4	Terminated transmission line	348
12.5	Reactive terminations	357
12.6	Transmission line dispersion	360
12.7	References	363
13	Magnetic and electrical coupling and isolation	364
13.1	Electrical coupling	364
13.2	Magnetic coupling	367
13.3	Ground noise coupling	373
13.4	Substrate coupling	378
13.5	Package coupling	383
13.6	References	385
14	Electromagnetic propagation and radiation	386
14.1	Maxwell's equations in source-free regions	386
14.2	Penetration of waves into conductors	390

viii Contents

14.3	Poynting vector	395
14.4	EM power carried by a plane wave	397
14.5	Complex Poynting Theorem	399
14.6	Reflections from a perfect conductor	402
14.7	Normal incidence on a dielectric	404
14.8	References	406
15	Microwave circuits	407
15.1	What are microwave circuits?	407
15.2	Microwave networks	409
15.3	Lorentz reciprocity theorem	409
15.4	The network formulation	412
15.5	Scattering matrix	414
15.6	Properties of three-ports	421
15.7	Properties of four-ports	429
15.8	Two conductor coupler	438
15.9	References	440
	<i>References</i>	441
	<i>Index</i>	445

1 Introduction

1.1 Motivation

The history of electronics has been inextricably linked with the growth of the communications industry. Electronic communication served as a major enabling technology for the industrial revolution. When scientists and engineers learned to control electricity and magnetism, it did not take long for people to realize that the electromagnetic force would enable long-range communication. Even though the basic science of Maxwell's equations was well understood, it took much longer for practical applications to fully exploit all the fantastic possibilities such as radio, television, and personal wireless communication.

At first only crude wires carrying telegraph signals were rolled out sending Morse code,¹ digital signals at speeds limited by human operators. In this regard it is ironic that digital communication predates analog communication. Telegraph wires were laid alongside train tracks, making long-range communication and transportation a practical reality. Sending signals faster and further ignited the imagination of engineers of the time and forced them to study carefully and understand the electromagnetic force of nature. Today we are again re-learning and inventing new digital and analog communication systems that are once again compelling us to return to the very fundamental science of electricity and magnetism.

The topic of this book is the high-frequency electromagnetic properties of passive and active devices. For the most part, passive devices are resistors, capacitors, transformers, and inductors, while active devices are transistors. Most applications we draw from are high-frequency circuits. For example, radio frequency (RF) circuits and high-speed digital circuits both depend on a firm understanding of passive devices and the environment in which they operate.

Circuit theory developed as an abstraction to electromagnetics. Circuit theory is in effect the limit of electromagnetics for a circuit with negligible dimension. This allows spatial variations and time delay to be ignored in the analysis of the circuit. As such, it allowed practicing engineers to forego solving Maxwell's equations and replaced them with simple concepts such as KCL and KVL. Even differential equations were eliminated and replaced with algebraic equations by employing Laplace transforms. The power and popularity of circuit theory was due to its simplicity and abstraction. It allowed generations of engineers to

¹ Or as Paul Nahin suggests in [41] we should more correctly call this "Vail" code.

solve difficult problems with simple and yet powerful tools. In effect, it allowed generations of engineers to forego reading a book such as this one.

So why read another book on electromagnetics? Why bother learning all this seemingly complicated theory when your ultimate goal is to build circuits and systems for communication and information management?

We live today at the intersection of several interesting technologies and applications. Integrated circuit technology has enabled active devices to operate at increasingly higher frequencies, turning low-cost Si technology into a seemingly universal panacea for a wide array of applications. CMOS digital circuits are switching at increasingly higher rates, pushing multi GHz operation. Si CMOS, bipolar, and SiGe technology have also enabled a new class of low-cost RF and microwave devices, with ubiquitous deployment of cellular phones in the 800 MHz–2 GHz spectrum, and high-speed wireless LAN in the 2–5 GHz bands. There seems to be very little in the way of enabling Si technology to exploit the bandwidths up to the limits of the device technology. In a present-day digital 130 nm CMOS process, for instance, circuits are viable up to 60 GHz [55] [13].

At the same time, wired communication is pushing the limits. Gigabit Ethernet and high-speed USB cables are now an everyday reality, and people are already pursuing a 10 Gb/s solution. Optical communication is of course at the forefront, with data rates in the 40 Gb/s range now commercially viable and at relatively low cost.

The simultaneous improvement in active device technology, miniaturization, and a host of new applications are the driving force of today's engineering. As integrated circuits encompass more functionality, many traditionally off-chip components are pushed into the IC or package, blurring the line between active devices and circuits and passive devices and electromagnetics. This is the topic of this book.

Technology enhancements

The limitations in frequency and thus speed of operation is usually set by the active device technology. One common figure of merit for a technology is, f_r , the unity-gain frequency f_T , the frequency at which the short-circuit current gain of the device crosses unity. Another important figure of merit is f_{max} the maximum frequency of oscillation, or equivalently the frequency where the maximum power gain of a transistor drops to unity. Since f_{max} is a strong function of layout and parasitics in a process, it is less often employed. In contrast, the f_T depends mostly on the dimensions of the transistor and the transconductance

$$f_T = \frac{1}{2\pi} \frac{g_m}{C_\pi + C_\mu} \quad (1.1)$$

It can be shown [50] that the device f_T is inversely related to the transistor dimensions. For a long-channel MOSFET the key scaling parameter is the channel length L

$$f_T \approx \frac{1}{2\pi} \frac{3\mu(V_{gs} - V_t)}{2L^2} \quad (1.2)$$

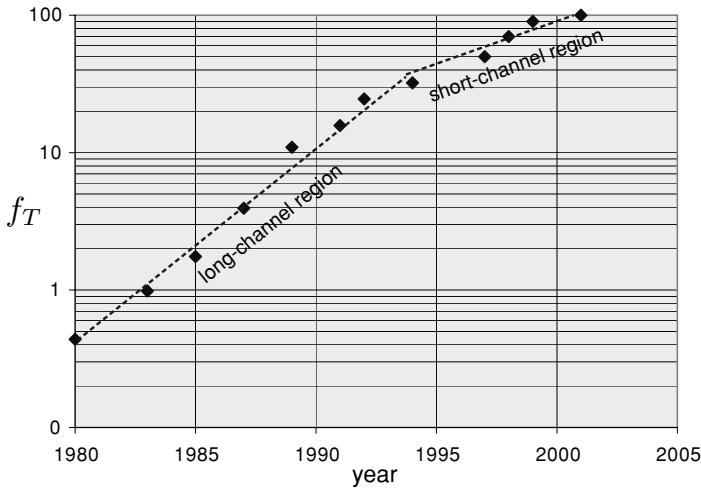


Figure 1.1 The improvements in device unity-gain frequency f_T over the past two decades due to device scaling.

while in the limit for short channel transistors the scaling changes to L^{-1} since the current is limited by velocity saturation

$$I_{ds,sat} = WQ_i v_{sat} = WC_{ox}(V_{gs} - V_t)v_{sat} \quad (1.3)$$

resulting in

$$f_T \propto \frac{v_{sat}}{L} \quad (1.4)$$

For a bipolar junction transistor (BJT) the critical dimension is the base width. In the limit that base transit limits the frequency of operation

$$f_T \approx \frac{1}{2\pi} \frac{1}{\tau_B} \propto \frac{1}{W_B^2} \quad (1.5)$$

As integrated circuit manufacturing technology has improved exponentially in the past three decades, so has the f_T of the device, giving circuit designers increasingly faster devices. A plot of the device f_T over the years for a MOSFET device is shown in Fig. 1.1, and the exponential growth in technological advancements can be seen clearly.

It is important to note that this improvement in performance only applies to the intrinsic device. Early circuits were in fact limited by the intrinsic transistor and not the parasitic routing and off-chip environment. As circuit technology has advanced, though, the situation has reversed and now the limitation is set by the parasitics of the chip and board environment, as well as the performance of the passive devices. This is why the material of this book is now particularly relevant. It can be shown that a good approximation to the CMOS device f_{max} is given by [42]

$$f_{max} \approx \frac{f_T}{2\sqrt{R_g(g_m C_{gd}/C_{gg}) + (R_g + r_{ch} + R_s)g_{ds}}} \quad (1.6)$$

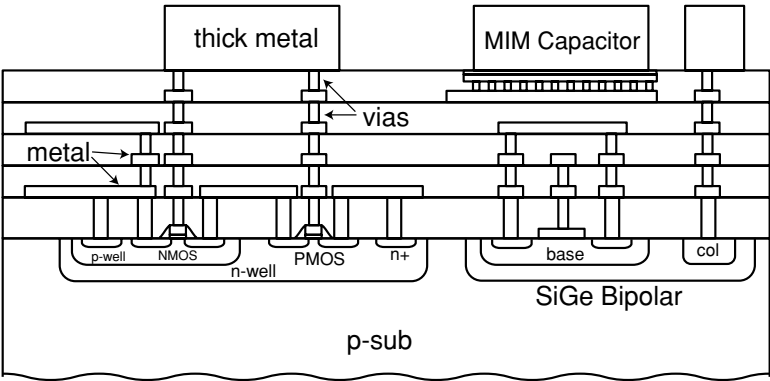


Figure 1.2 Cross section of a SiGe BiCMOS process.

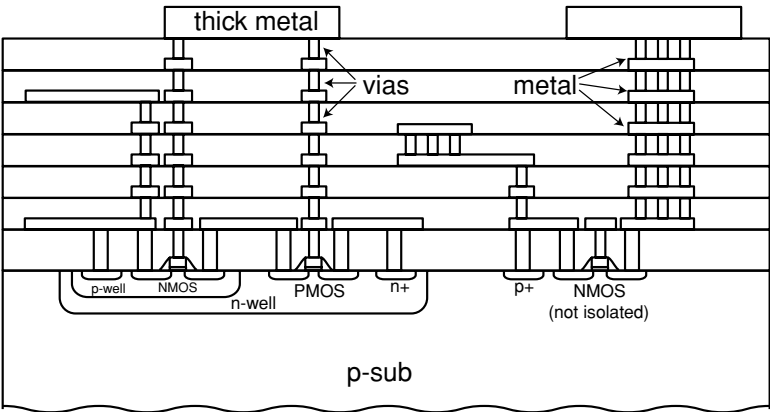


Figure 1.3 Cross section of an advanced CMOS process.

where the device performance is a strong function of the loss, such as the drain/source resistance R_s , R_d , and the gate resistance R_g . These parasitics are in large part determined by layout and the process technology.

While early integrated circuit technologies were limited to a few types of different active devices and a few layers of aluminum interconnect metal, present-day process technology has a rich array of devices and metal routing. In an advanced Si process, shown in Fig. 1.2, high-performance SiGe HBT devices are complemented by MOS and PN-junction varactors, metal-insulator-metal (MIM) high-density and high-quality capacitors, and thick-metal for low-loss interconnect and inductors/transformers. Even a digital CMOS process, as shown in Fig. 1.3, has many advanced capabilities. In addition to several flavors of MOS active devices (fast thin oxide, thick oxide, high/low V_T), there are also enhanced isolation structures and triple-well (deep n-well) devices, and many layers of interconnect that allow construction of high-quality, high-density capacitors and reasonably high-quality inductors.

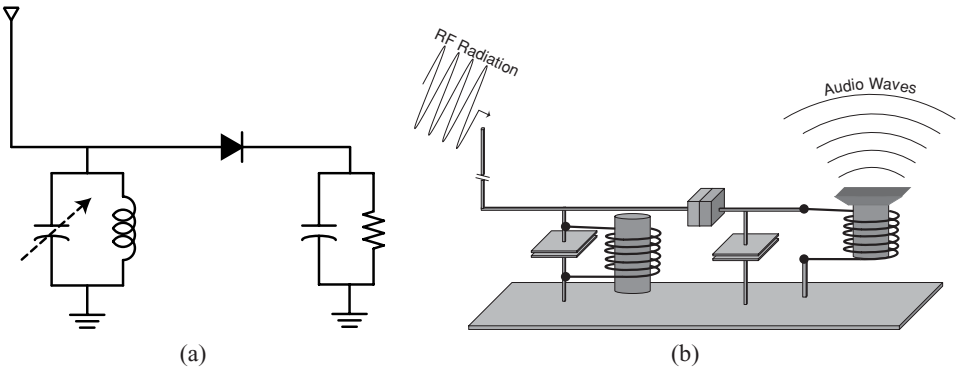


Figure 1.4 (a) A simple AM receiver circuit. The resistor represents a high-input impedance earphone. (b) A physical realization of the simple AM receiver circuit.

Radio and wireless communication

Early radio systems were essentially all passive. To see this look into the back of an old radio where a few active devices (vacuum tubes or transistors) are surrounded by tens to hundreds of passive devices. Consider the circuit diagram of a very simple AM receiver shown in Fig. 1.4a. The antenna drives a resonant tank tuned to the center frequency of the transmitting station. This signal is fed into a peak detector that follows the peak of the RF signal. The low-pass filter time constant is only fast enough to follow the low-frequency audio signal (generically the baseband signal) and yet too slow to follow the RF, thus removing the RF signal and retaining the low-frequency audio. This received signal is usually too weak to drive a speaker but can be heard through a sensitive headphone. A simple audio amplifier can be used to strengthen the signal.

It is interesting to note that this AM receiver can be physically realized by merely using contacts between a few different pieces of metal and semiconductors. This is shown in Fig. 1.4b. The resonant tank is simply a piece of wire wound into a coil which contacts with the capacitor, two metal plates in close proximity. The diode can be realized as the junction of a metal and semiconductor. Finally, to convert electric energy into sound we can use another large inductor coil and use the time-varying magnetic force to move a paper thin cone driven by a magnetic core. Magnetic materials have been known since ancient times and therefore since the metal age we have had the capability to build radio receivers! In fact, it is not surprising that radios often crop up accidentally.²

Most modern radios operate based on an architecture invented by Edwin Armstrong. The block diagram of such a system, called a super-heterodyne receiver, is shown in Fig. 1.7. This receiver incorporates a local oscillator (LO), a block that primarily converts DC power into RF power at the oscillation frequency. A mixer takes the product of this signal and the

² For instance my old answering machine also picked up the radio. Sometimes you could hear it as you were waiting for the tape recorder to rewind. At least this was a desirable parasitic radio.

signal received by the antenna. Recall the following trigonometric identity³

$$2 \cos(\omega_{LO}t) \cos(\omega_{RF}t) = \cos((\omega_{LO} + \omega_{RF})t) + \cos((\omega_{LO} - \omega_{RF})t) \quad (1.7)$$

Note that the product of the received RF signal and the local oscillator signal produces two new signals, one centered at the difference frequency and one centered at the sum frequency. If we put a bandpass filter at one of these frequencies, call it the intermediate frequency, IF, we can electronically tune the radio by simply changing the LO frequency. This is accomplished by using a frequency synthesizer (a PLL or phase locked loop), and thus we avoid building a variable filter common to the early radios. The important point is that the IF is fixed and we can build a very selective filter to pinpoint our desired signal and to reject everything else. Why not simply set LO equal to RF to move everything to DC? This is in fact the direct-conversion or zero-IF architecture. It has some shortcomings such as problems with DC offset,⁴ but its main advantage is that it lowers the complexity of the RF section of a typical radio.

At the heart of the frequency synthesizer is the voltage controlled oscillator (VCO). The VCO is an oscillator where the output frequency is a function of a control voltage or current.⁵ To build a VCO we need a way to change the center frequency of a resonant tank. The resonant tank is simply an inductor in series or in parallel with a capacitor. One typical realization is to use varactor, a variable capacitor. A reversed biased diode serves this purpose nicely, as the depletion region width, and thus the small-signal capacitance, is a function of the reverse bias. It seems that a super-heterodyne receiver has simply moved the variable resonant tank from the antenna front end to a variable resonant tank in the VCO! Have we gained anything? Yes, because the frequency of the VCO can be controlled precisely in a feedback loop (using an accurate frequency reference such as a crystal), eliminating any problems associated with absolute tolerances in components in addition to drift and temperature variation.

The radio has once again emerged as a critical application of passive devices spawned by the growth and popularity of wireless telephones, in particular the cellular phone. By limiting the transmitter powers and taking advantage of spatial diversity (re-using the same frequency band for communication for points far removed – for non-adjacent cell sites), a few hundred radio channels can be used to provide wireless communication to millions of people. Modern cell phones employ complicated radio receivers and transmitters (transceivers) employing hundreds and thousands of passive devices. Early cell phones used simple architectures such as the super-heterodyne receiver but the demand for low-cost and small footprints has prompted a re-investigation of radio architectures.

The layout of a modern 2.4 GHz transceivers for 802.11b wireless LAN (WLAN) is shown in Fig. 1.5 [7]. The IC is implemented in a 0.25 μ CMOS process and employs several integrated passive devices such as spiral inductors, capacitors, and resistors. The spiral inductors comprise a large fraction of the chip area. The next chip shown in Fig. 1.6

³ I recall asking my trig teacher about the practical application of the subject. After scratching her head and pondering the question, her response was that architects use trig to estimate the height of buildings! A much better answer would have been this equation.

⁴ And $1/f$ noise in MOS technology.

⁵ This makes a nice AM to FM modulator, as well.

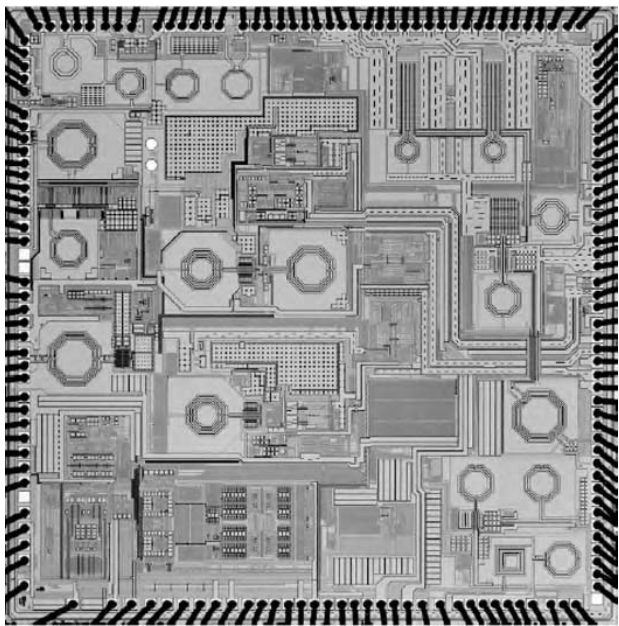


Figure 1.5 A 2.4 GHz CMOS 802.11b Wireless LAN Transceiver [7]. (Copyright 2003, IEEE)

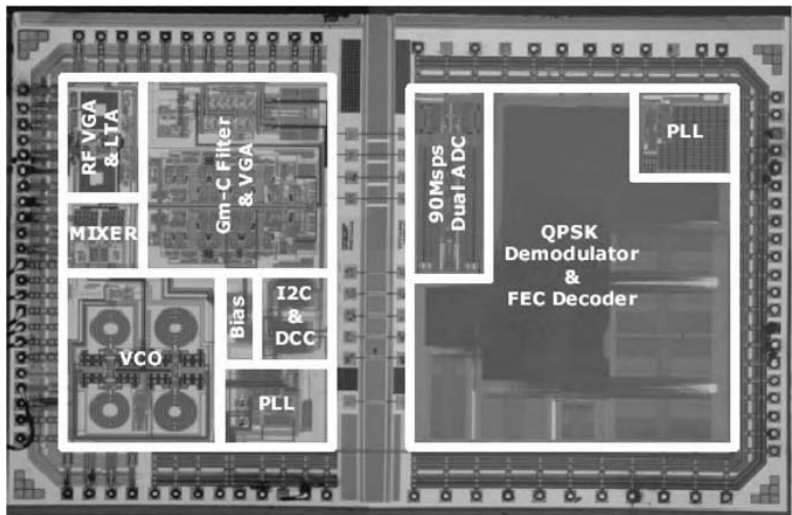


Figure 1.6 A direct-conversion satellite broadband tuner-demodulator SOC [17] operates from 1–2 GHz. (Copyright 2003, IEEE)

Copyright © 2007, Cambridge University Press. All rights reserved.

Niknejad, Ali M.. Electromagnetics for High-Speed Analog and Digital Communication Circuits, Cambridge University Press, 2007.
ProQuest Ebook Central, <http://ebookcentral.proquest.com/lib/sjsu/detail.action?docID=288460>.
Created from sjsu on 2023-03-15 17:14:50.

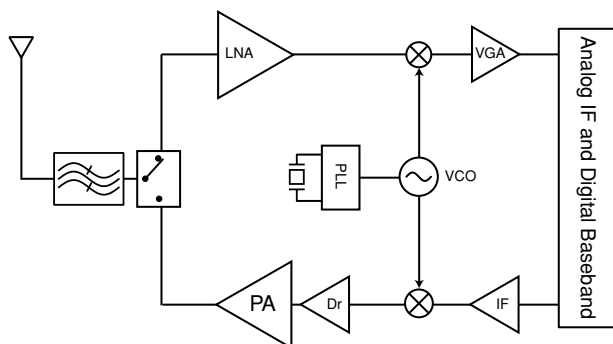


Figure 1.7 The block diagram of an Armstrong super-heterodyne transceiver.

[17] is an integrated direct-conversion satellite broadband tuner-demodulator “system-on-a-chip” (SOC). The chip is implemented in a $0.18\ \mu$ CMOS process and employs MIM capacitors and spiral inductors. It operates in the 1–2 GHz band, requiring broadband operation and high linearity. Notice that the digital baseband has been integrated on to a single chip along with the sensitive analog and RF blocks. This brings about several important challenges in the design due to the parasitic coupling between the various blocks. A triple-well process and lead-less package technology are used to maximize the isolation.

In general, integrating an entire transceiver on to a single chip has many challenges. The power amplifier (PA) or PA driver can injection lock the VCO through the package and substrate, causing a spurious modulation. Digital circuitry can couple seemingly random switching signals into the analog path, effectively increasing the noise floor of the sensitive RF and analog blocks. As the level of integration increases, a single chip or package may contain several systems in operation simultaneously, requiring further understanding and modeling of the coupling mechanisms.

Computers and data communication

Computers and data communication, particularly the Internet, have given rise to a new tidal wave in the information revolution. The speed of computers has improved drastically due to technological improvements in transistor, microprocessor, memory, and system bus architectures. Computer circuits move and process discrete time signals at a frequency determined by the system clock. For instance, in the current generation of computers the clock speed inside the microprocessor is several GHz, while the speed of the system bus and memory lag behind by a factor of 2–3. This is because inside the microprocessor everything is small and dense and signals travel short distances in the presence of small parasitics (mainly capacitance). Off-chip, though, the system bus environment is characterized by much longer distances and much larger parasitics, such as non-ideal dispersive transmission lines along the board traces. Modern computer networks, like gigahertz Ethernet LAN, also

operate at high frequencies over wires, necessitating a complete understanding of distributed transmission line effects. These topics are covered in Chapters 9 and 12.

High-speed wireless data communication is the focus of much research and development. The next and future generations of cellular technology will bring the Internet from our homes and offices into virtually every location on earth. Wireless LAN systems enable short-range high-speed data communication without the expensive network infrastructure. A physical network infrastructure requires time-consuming distribution of cables to every office in a building. A wireless system can be up and running in minutes or hours as opposed to days or months.⁶

In such systems cost and size will force many external passive components on to the chip environment, where knowledge of parasitic coupling and loss is critical in a successful low-cost implementation. In this book we spend a great deal of time discussing inductors, capacitors, transformers, and other key passive elements realized in the on-chip environment.

Microwave systems

Microwave systems employ higher frequencies where the wavelength $\lambda = c/f$ is of the order of centimeters or millimeters. Thus the lumped circuit approach fails since these structures are a significant fraction of a wavelength and spatial variation begins to play as important a role as time variation. Such systems were first employed in World War II for radar systems.⁷ In a radar system, the small wavelength allows us to construct a highly directional antenna to focus a beam of radiation in a given direction. By observing the reflection, we can compute the time-of-flight and hence the distance to an object. By also observing the Doppler frequency shift, we can compute the speed of the object.

Perhaps the greatest difficulty in designing microwave systems below 10 GHz is that the operating frequency is in an intermediate band where lumped element circuit techniques do not strictly apply and microwave methodology results in prohibitively large circuits. At 3 GHz, the wavelength is 10 cm in air and about 5 cm in silicon dioxide, while an integrated circuit has dimensions of the order of millimeters, thus precluding distributed elements such as quarter-wave transmission lines. But using advances pseudo-lumped passive devices such as inductors, transformers, and capacitors, microwave ICs can be realized with minimal off-chip components.

Many early microwave systems were designed for military applications where size and cost were of less concern in comparison to the quality and reliability. This led to many experimental and trial-and-error design approaches. Difficult system specifications were met by using the best available technology, and often expensive and exotic processes were employed to fabricate high-speed transistors. New microwave systems, in contrast, need to be mass produced and cost and size are the main concerns. Fortunately high-volume process

⁶ I seem to recall that it took a year for a network upgrade to occur in Cory Hall at Berkeley!

⁷ It is ironic that the EEs of the time lacked the necessary skills to build such systems and the project was handed off to the physicists at the MIT Radiation Lab.

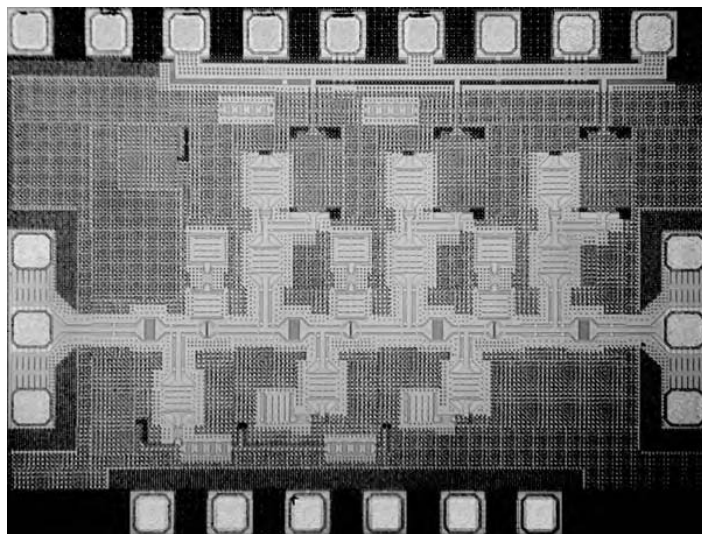


Figure 1.8 A three-stage 60 GHz CMOS LNA implemented in a digital 130 nm process.

technology using silicon is now readily available. The speed is now sufficient to displace many specialized technologies. Since high-volume microwave systems are primarily being designed by circuit engineers as opposed to microwave engineers, the lack of knowledge of electromagnetics and distributed circuits can be an impediment to successful integration and implementation.

Higher-frequency bands offer new opportunities to exploit sparsely used spectrum. The 60 GHz “oxygen absorption” band is a prime example, providing 7 GHz of unlicensed bandwidth in the US. An example of a 60 GHz multi-stage low-noise amplifier (LNA) is shown in Fig. 1.8. Here transmission lines play a key role as inductors, interconnectors, and resonators. A 60 GHz single-transistor mixer, shown in Fig. 1.9, employs a hybrid coupler (see Section 15.7) to combine the RF and LO signal. Spiral inductors are also employed in the IF stages. Both of these chips were fabricated in a digital 130 nm CMOS process. Another CMOS microwave circuit is shown in Fig. 1.10. This is a circular standing-wave 10 GHz oscillator, employing integrated transmission lines in the resonator [11]. There is a beautiful connection between this oscillator and the orbit of an electron in a hydrogen atom. Similar to the wave function of an electron, the electromagnetic mode must satisfy the periodic boundary condition, and this determines the possible resonant modes of the structure.

Optical communication

Fiber-optic communication systems allow large amounts of data to be transmitted great distances with relatively little attenuation. At optical frequencies, metals are too lossy for long-haul communication without amplification, and so the energy is confined inside a thin fiber of glass by total internal reflection. The flexibility and low cost of this material has displaced more traditional waveguides made of rigid or semi-rigid and expensive materials.

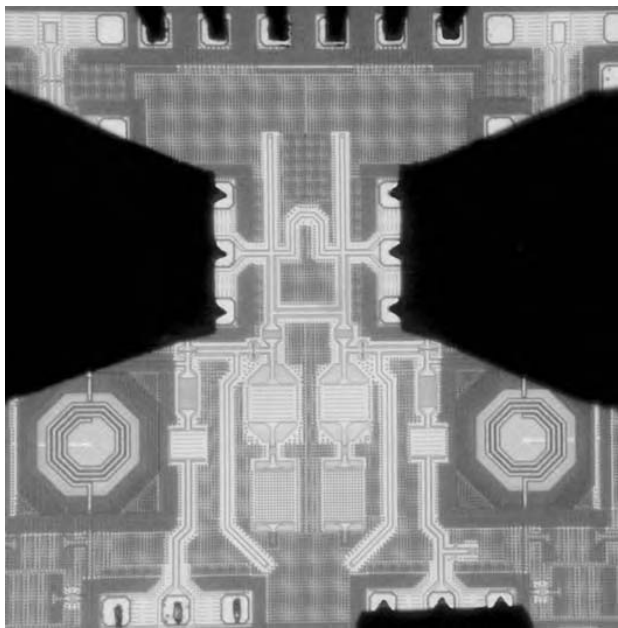


Figure 1.9 A single-transistor 60 GHz CMOS mixer implemented in a digital 130 nm process.

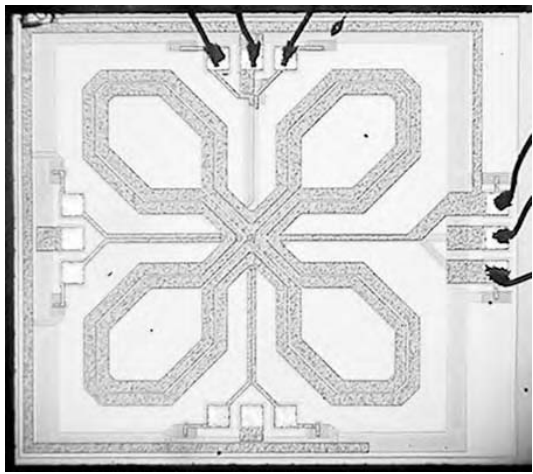


Figure 1.10 A circular standing-wave oscillator operates at 10 GHz [11]. (Copyright 2003, IEEE)

The main bottleneck in optical systems, though, is the electronics. While new optical amplifiers may displace electronic regeneration techniques, routing digital signals on fiber lines requires electronic circuits. Much research effort is dedicated to completely displacing electronic routing through mechanical means. This may seem like a backward step

Copyright © 2007, Cambridge University Press. All rights reserved.

Niknejad, Ali M.. Electromagnetics for High-Speed Analog and Digital Communication Circuits, Cambridge University Press, 2007.
ProQuest Ebook Central, <http://ebookcentral.proquest.com/lib/sjsu/detail.action?docID=288460>.
Created from sjsu on 2023-03-15 17:14:50.

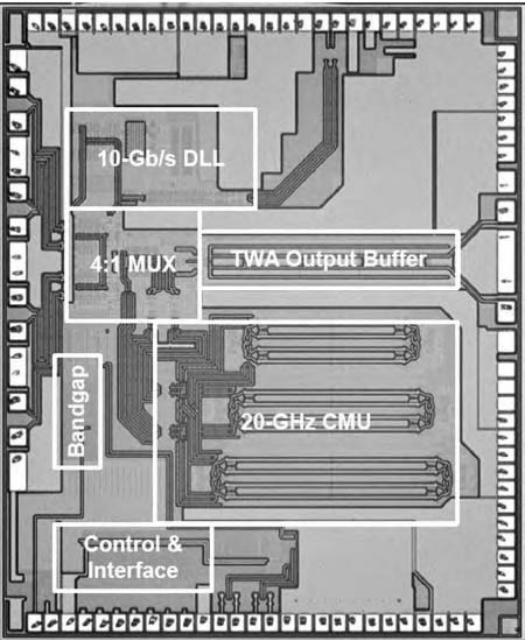


Figure 1.11 A 40 Gb/s SONET MUX/CMU implemented in SiGe technology. The integrated 20 GHz CMU employs coupled oscillators and transmission line resonators [56]. (Copyright 2003, IEEE)

since mechanical systems are inherently much slower than electronic systems. But it is the throughput of electronic systems, not the switching rate, that is limiting optical communication systems. Transistors can only process signals up to a certain limit due to intrinsic f_T and extrinsic parasitics. Micro-electrical-mechanical systems (MEMS) technology allow micromechanical systems to reside on the same substrate as low-cost electronic circuits. A simple MEMS-based hinged mirror switch does not limit the throughput of an optical system, whereas an electronic system requires high-speed digital signals to travel through transistors and parasitics. The co-design of MEMS and electronic systems in a chip requires an understanding of electromagnetics and transduction.

Even if routing of optical signals is done without electronic means, the generation of high-speed optical signals requires electronic systems to pool and un-pool dozens of lower-speed signals on to a fast common serial line. This requires accurate timing and synchronization, such as voltage-controlled oscillators phase locked to reference signals. Optical and high-speed digital circuits require routing of high-speed signals in the presence of board and transmission line parasitics, as well as the design of passive elements for frequency synthesis. An example optical transceiver chip incorporating the multiplexers (MUX) and clock-multiplication unit (CMU) is shown in Fig. 1.11, where transmission lines are used in a coupled oscillator to generate a low-phase noise high-frequency oscillation.

Copyright © 2007, Cambridge University Press. All rights reserved.

1.2 System in Package (SiP): chip and package co-design

Present design methodology draws a clear line between the chip and the package. This clean interface between the chip and the package simplifies design considerably, since the two can be analyzed independently. As already noted, the package is a major source of parasitic energy coupling in current transceivers and the careful evaluation of the package is a critical step in the design. At higher frequencies, though, the line between the package and chip will blur even further. In fact, the package could be a crucial and beneficial part of the system, as many key passive components could reside in the package. Package and chip co-design will require new design techniques and a more intimate understanding of the electromagnetic effects.

The simultaneous simulation of the interaction of the board environment, the package, and the on-chip passives is presently beyond the scope of present-day simulation tools. One problem is the vastly different scales employed in board environment (centimeters), package (sub-cm), and chip environment (sub-micron geometries). While future tools may overcome this limitation, present-day engineers must have a firm grasp of fundamental engineering to analyze and design such structures. For instance, instead of simulating the entire system, key coupling issues can be dealt with only by simulating critical sub-portions of the system. Identification of key elements and analysis of such systems are two of the goals of this book.

1.3 Future wireless communication systems

The conceptual transceiver shown in Fig. 1.12 naturally partitions into several sections, most notably the RF front-end section, the analog core, and the digital signal processing (DSP) unit. The arrows represent the flow and transfer of information among the various sections. The RF receive section amplifies the weak incoming RF signal and converts it into an intermediate frequency. The choice of IF frequency mostly determines the architecture of the system. The analog section performs further amplification and filtering and converts the signal into digital form. From here, the signal is usually buffered and driven off-chip into the DSP core where the final demodulation and detection are performed. Similarly, on the transmit side, information is encoded and modulated in the DSP core. The analog section converts the signal from digital discrete time to continuous time and up-converts the center frequency to the transmit band. In this transceiver each section communicates with every other section in a well-defined manner. Isolation between the various sections is achieved by physical separation and proper frequency planning to avoid interference.

There are several things to note about this system. Note that the entire system is physically much smaller than the wavelength. Thus, the quasi-static approximation is applicable and much of the internal circuitry can be considered as lumped elements. Also, due to the small size, there is no significant coupling between the external electromagnetic fields propagating in free space and the on-chip signals. The transduction of RF energy from free space to the volume of the conductors is achieved externally through an antenna element. This naturally partitions the radiation incident on the system from the transceiver. The RF energy flows

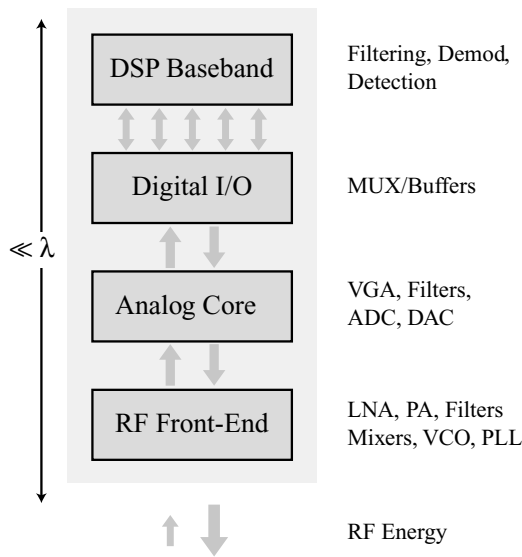


Figure 1.12 A block diagram of a present-day advanced wireless transceiver.

into the system along a well-defined path from the package interface between the board and the chip. Due to the physically large size and requirement for high quality, many passive devices in this system are realized externally, such as surface acoustic wave (SAW) filters, wire-wound high-quality inductors, and low-insertion loss diode switches. To save cost and power, the trend has been to integrate as many of the passive devices on-chip as possible. At higher frequencies this is increasingly a necessity due to the large parasitic impedances present in the board environment. MEMS technology may allow many off-chip components to be miniaturized and moved on chip.

To improve isolation in present-day transceivers, many options such as deep trenches, triple-well devices, or special technologies, such as Silicon-on-Insulator (SoI), are effective. The proper grounding and regulation of supplies are also important. While these techniques are effective at mitigating the coupling at lower microwave frequencies, they are much less effective at higher frequencies.

A hypothetical transceiver at a much higher frequency is shown in Fig. 1.13. Due to the large physical size of the chip and package relative to the wavelength, an array of antenna elements is integrated directly into the package. The DSP core is the heart of the system and there is constant feedback between the baseband and RF sections to optimize the power and throughput in the system. For instance, the phased antenna array can be used to boost the directivity of the receiver or transmitter to improve the SNR or to lower the power. Spatial diversity in the antenna elements can also boost the data throughput per fixed bandwidth beyond Shannon's limits.

We see that the package now plays a much larger role in this hypothetical transceiver. The package is physically much larger than the chip and contains an ever-increasing density

Copyright © 2007. Cambridge University Press. All rights reserved.

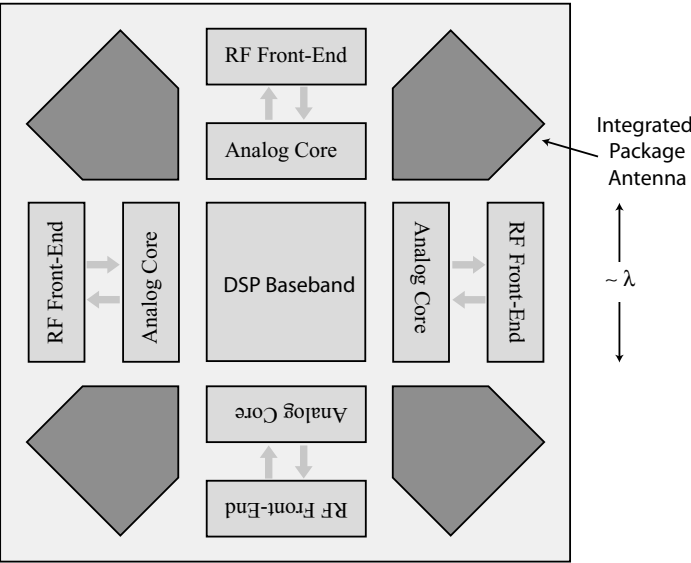


Figure 1.13 A block diagram of a hypothetical future wireless transceiver.

of signal lines in close proximity. If the package design is not performed simultaneously as an integral part of the system, the electromagnetic coupling can serve as a parasitic feedback path that can compromise the stability and lower the sensitivity of the system. These problems are naturally more severe at higher frequencies. An example of a high-frequency system with large levels of microwave integration is shown in Fig. 1.14 [24], where a bank of eight receivers operate in parallel to down-convert signals from an antenna array. The receiver can select the appropriate phase shift in each path to form a high-directivity antenna, or to improve the spatial diversity of the receiver. Proper routing of signals and isolation are key elements in the successful design of such a system.

1.4 Circuits and electromagnetic simulation

Circuits are really a part of electromagnetics. As we alluded to earlier, circuit theory can adequately replace electromagnetics when circuit dimensions are much smaller than the wavelength of electromagnetic fields. This allows all time variation to be discarded and difficult partial differential equations (PDE) resulting from Maxwell's equations to be reduced to ordinary differential equations (ODE).

Many tools are available for solving a system of non-linear ODEs resulting from applying KVL and KCL to nodes and loops in a circuit. Since Maxwell's equations are linear, the non-linearity is introduced primarily through the transistors and junctions in the circuit. The most common tool for solving ODEs in the circuit designer's arsenal is SPICE and its many variants. Frequency domain solvers (harmonic balance) and period steady-state (PSS)

Copyright © 2007. Cambridge University Press. All rights reserved.

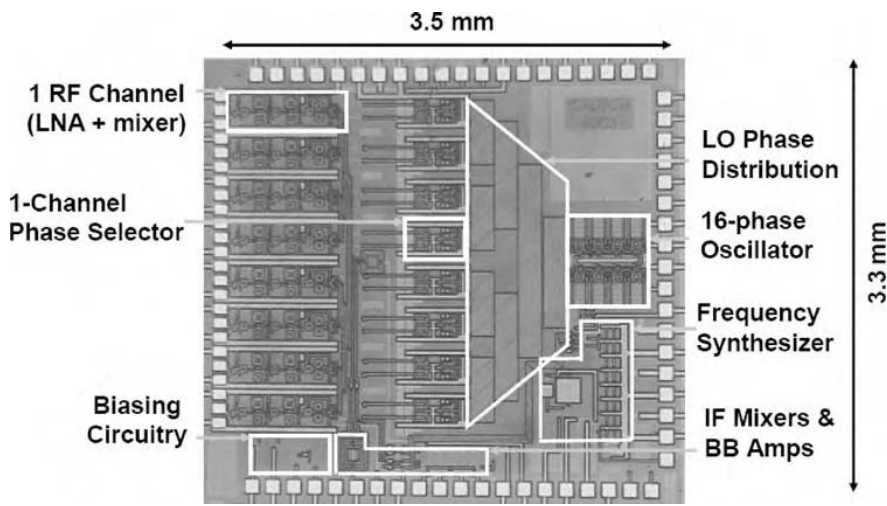


Figure 1.14 A fully integrated SiGe phased array front-end operating at 24 GHz [24]. (Copyright 2003, IEEE)

solvers are also available and are helpful when analyzing periodic time-variant circuits in the presence of noise.

PDE solvers, by contrast, are much slower and harder to use than ODE solvers such as SPICE. This is because a space of three to four dimensions (three spatial dimensions and one dimension for time) is involved. This has a large impact on traditional matrix solvers that require $O[n^{2-3}]$ time to solve a system of n unknowns. Time variation can be removed for a harmonic steady-state solution, but this even further complicates incorporating PDE solutions in an ODE solver. For instance, if you simulate an inductor with an electromagnetic (EM) solver, how do you take the frequency domain solution into a time-domain solver and efficiently compute the solution?

Electromagnetic simulation and analysis is a critical component in successfully designing active and passive components. This is missing from the circuit curriculum, since circuit designers usually do not deal with the innards of components such as transistors and inductors. Instead, they employ a lumped circuit model that captures the details. For instance, the BSIM models incorporated in Berkeley SPICE and commercial simulators model transistor device physics with hundreds of elements and parameters to describe a sub-micron transistor accurately. A transistor, though, is a very small element and most of the physics fit well into a lumped element model. A typical inductor, by contrast, is relatively simple, containing perhaps three elements to model the inductance, the resistive loss, and capacitance of the structure. This model serves as long as the inductor structure is very small relative to the wavelength. But as the frequency of operation increases, several shortcomings of the model are present. For example, the loss is a complicated function of frequency. Even if we only model skin effect, the resistive loss varies like $\sqrt{f}$. How do we put such a variation in resistance in a time-domain SPICE simulation? There is no easy solution.

Copyright © 2007, Cambridge University Press. All rights reserved.

In addition to SPICE, then, the circuit designer of today and of the future must be able to understand and model electromagnetic effects. In the past each block was partitioned into a lumped model (designed by device and EM engineers) and the circuit designer could ignore the complicated details. For modern circuits this is no longer possible as even short wires on chip have sufficient parasitic inductance, loss, and capacitance as to impact the performance of circuits. The situation is exacerbated by the increasingly higher densities at which circuits are packed on a single die, and electromagnetic coupling becomes the bottleneck in the achievable performance. As EM solution techniques expand with improvements in computational speed and capacity,⁸ the use of an EM solver will become an integrated part of the design of circuits much as using SPICE is standard practice today.

⁸ It is interesting to note that each generation of advance is fueled by new computation tools running on the previous generation of computers. This self-feeding cycle of advancement means that the current generation of tools is never good enough for the next generation. That's why we need engineers who understand the fundamentals and can partition difficult problems into digestible pieces. That's why engineers will never be replaced by computers!

5 Inductance

5.1 Introduction

Inductance is an intrinsic part of every physical circuit in which currents flow. Even the simplest circuit shown in Fig. 5.1a cannot be fully understood unless the inductance is taken into account. Even though no inductor is drawn in the figure explicitly, there is a net electromagnetic force

$$V_L = L \frac{dI}{dt} \quad (5.1)$$

induced around the loop that tends to oppose any change in the current magnitude. The inductance L is a property of the loop related to the physical geometry of the circuit. We can represent this induced voltage by an element that has the voltage–current relationship embodied in Eq. (5.1), as shown schematically in Fig. 5.1b. In this case, the inductance is a property of the loop and it does not stem from a separate coil (as suggested by the inductor’s symbol) introduced into the circuit. While we can calculate the resistance of the loop by breaking the loop into three separate conductors, as suggested by $R_1 - R_3$, this cannot be done easily with the inductance because the currents in the conductors interact magnetically.

But is this intrinsic inductance really important? Using techniques of this chapter we can determine the magnitude of L for a given circuit. From the application at hand, we can estimate dI/dt or ωL in a steady-state periodic circuit and from Eq. (5.1) we can estimate the magnitude of the induced voltage. If L is sufficiently small, then we can ignore the magnetic effects of the circuit. Clearly then, for circuits operating at higher frequencies or at high switching rates dI/dt , we cannot ignore the inductive behavior. But even at lower frequencies, if the circuit is physically large then the L term is significant and the induced voltages are important.

It is interesting to note that in this circuit there is no one point where this voltage is induced but rather it is induced around the entire loop. In fact, voltages are induced in the entire space and not just along the circuit path. This gives rise to *eddy currents*, or magnetically induced currents that flow in other conducting bodies. Furthermore, the voltage is not induced uniformly along the path of the circuit, but in a distributed fashion where the induction is stronger at points where more magnetic energy is stored.

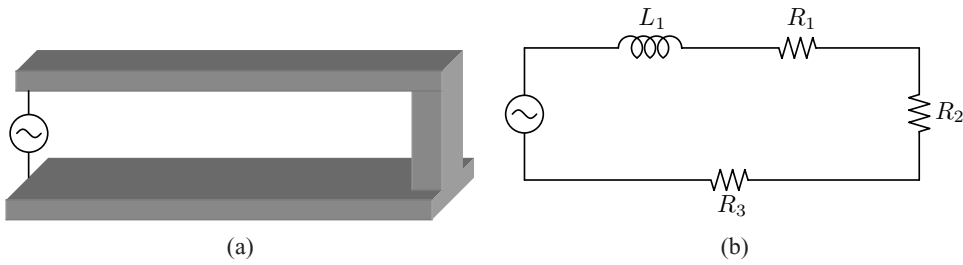


Figure 5.1 (a) A simple circuit with inductance. (b) Schematic representation of circuit.

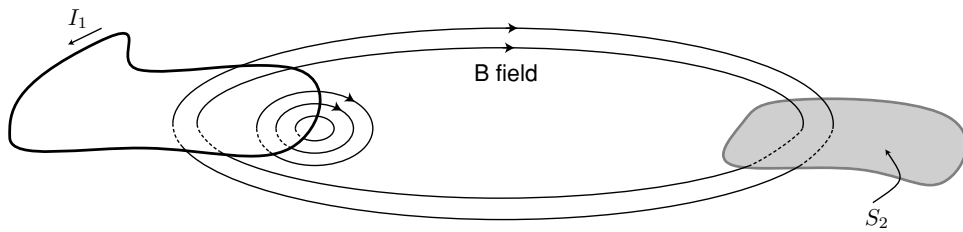


Figure 5.2 The flux linkage between two current loops.

Is the distributed nature of inductance important? Consider the input impedance of the above circuit seen from the terminals of the voltage source. Isn't it simply $j\omega L + R$? Surprisingly, no! In later chapters we will compute the input impedance of this distributed circuit and find the conditions under which the value deviates from the lumped circuit approximation.

5.2 Inductance

Magnetic flux

Magnetic flux plays an important role in many EM problems (in analogy with electric charge)

$$\Psi = \int_S \mathbf{B} \cdot d\mathbf{S} \quad (5.2)$$

Due to the absence of magnetic charge

$$\Psi = \oint_S \mathbf{B} \cdot d\mathbf{S} \equiv 0 \quad (5.3)$$

but net flux can certainly cross an open surface. The magnetic flux crossing a loop is shown in Fig. 5.2 by the integral of B crossing the cross-hashed surface. Magnetic flux is independent of the surface but only depends on the curve bounding the surface. This is easy to show

Copyright © 2007, Cambridge University Press. All rights reserved.

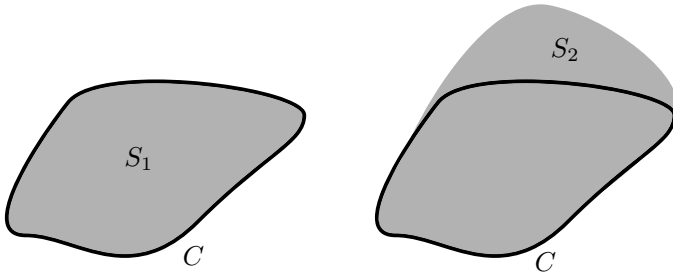


Figure 5.3 Any open surface bounded by the circuit can be used to calculate the magnetic flux.

since

$$\Psi = \int_S \mathbf{B} \cdot d\mathbf{S} = \int_S \nabla \times \mathbf{A} \cdot d\mathbf{S} = \oint_C \mathbf{A} \cdot d\mathbf{l} \quad (5.4)$$

$$\Psi = \int_{S_1} \mathbf{B} \cdot d\mathbf{S} = \int_{S_2} \mathbf{B} \cdot d\mathbf{S} \quad (5.5)$$

So in Fig. 5.2, we can select either S_1 or S_2 to calculate the flux.

Flux linkage

In Fig. 5.3, consider the flux-crossing surface S_2 due to a current I_1 flowing in loop S_1

$$\Psi_{21} = \int_{S_2} \mathbf{B}_1 \cdot d\mathbf{S} \quad (5.6)$$

Likewise, the “self-flux” of a loop is defined by the flux crossing the surface of a path when a current is flowing in the path

$$\Psi_{11} = \int_{S_1} \mathbf{B}_1 \cdot d\mathbf{S} \quad (5.7)$$

The flux is linearly proportional to the current and otherwise only a function of the geometry of the path. To see this, let us calculate Ψ_{21} for filamental loops

$$\Psi_{21} = \oint_{C_2} \mathbf{A}_1 \cdot d\mathbf{l}_2 \quad (5.8)$$

but

$$\mathbf{A}_1 = \frac{1}{4\pi\mu_0^{-1}} \oint_{C_1} \frac{I_1 d\mathbf{l}_1}{R - R_1} \quad (5.9)$$

Substituting, we have a double integral

$$\Psi_{21} = \frac{I_1}{4\pi\mu_0^{-1}} \oint_{C_2} \left(\oint_{C_1} \frac{d\mathbf{l}_1}{R - R_1} \right) \cdot d\mathbf{l}_2 \quad (5.10)$$

The geometry of this double integral is shown in Fig. 5.4. We thus have a simple formula that only involves the magnitude of the current and the average distance between every two

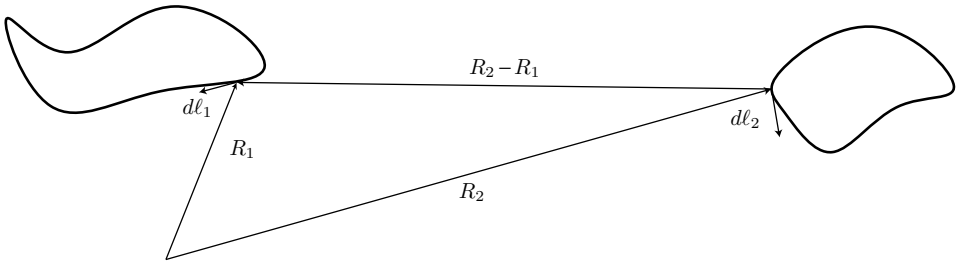


Figure 5.4 The distance between two arbitrary points in the loop.

points on the loops

$$\Psi_{21} = I_1 \times \frac{1}{4\pi\mu_0^{-1}} \oint_{C_2} \oint_{C_1} \frac{d\ell_1 \cdot d\ell_2}{R_2 - R_1} \quad (5.11)$$

Mutual and self inductance

Since the flux is proportional to the current by a geometric factor, we may write

$$\Psi_{21} = M_{21} I_1 \quad (5.12)$$

We call the factor M_{21} the mutual inductance

$$M_{21} = \frac{\Psi_{21}}{I_1} = \frac{1}{4\pi\mu_0^{-1}} \oint_{C_2} \oint_{C_1} \frac{d\ell_1 \cdot d\ell_2}{R_2 - R_1} \quad (5.13)$$

The above expression is known as Neumann's equation. It illustrates compactly the geometric nature of mutual inductance. In the above computation R takes on the values of the distance between every pair of points on the loops. In other words, mutual inductance is the weighted average value of $1/R$ between every pair of points on the loops weighted by $\cos(\theta)$, where θ is the local angle between the points in the loop. Thus, points lying on segments mutually perpendicular do not contribute to the mutual inductance.

An interesting observation is that mutual coupling is reciprocal. In other words

$$M_{12} = \frac{\mu}{4\pi} \oint_{C_1} \oint_{C_2} \frac{1}{R} d\ell_2 \cdot d\ell_1 = M_{21} \quad (5.14)$$

since the weighted “averaging” may be done in any order. This is a profound result that has many implications in circuit theory. Consider the example shown in Fig. 5.5 where a small loop couples to a larger loop. Even though the magnetic flux of loop B is small, loop A “captures” a large portion of the flux. Similarly, even though loop A has a large magnetic flux, only a small portion of it is captured by loop B and thus $M_{AB} = M_{BA}$.

The units of M are H since $[\mu] = \text{H/m}$. The “self-flux” mutual inductance is simply called the self inductance and denoted by $L_1 = M_{11}$.

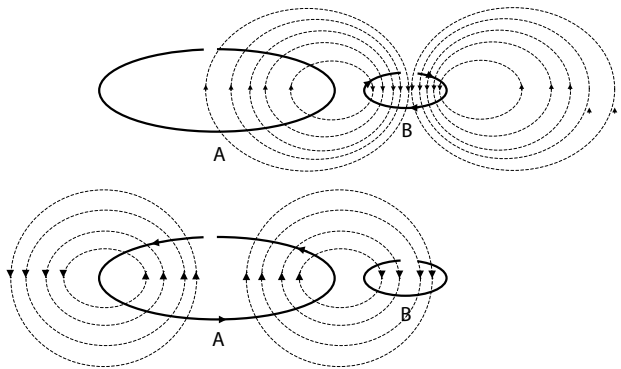


Figure 5.5 Illustration of reciprocity through the *forward* and *reverse* coupling between two vastly different loops.

System of mutual inductance equations

If we generalize to a system of current loops we have a system of equations

$$\begin{aligned}\Psi_1 &= L_1 I_1 + M_{12} I_2 + \dots M_{1N} I_N \\ &\vdots \\ \Psi_N &= M_{N1} I_1 + M_{N2} I_2 + \dots L_N I_N\end{aligned}$$

Or in matrix form $\psi = M\mathbf{i}$, where M is the inductance matrix. This equation resembles $\mathbf{q} = C\mathbf{v}$, where C is the capacitance matrix.

Example 9

Coaxial conductor

In transmission line problems, we need to compute inductance/unit length. Consider the shaded area in Fig. 5.6, from $r = a$ to $r = b$. The magnetic field in the region between conductors is easily computed

$$\oint \mathbf{B} \cdot d\mathbf{l} = B_\phi 2\pi r = \mu_0 I \tag{5.15}$$

The external flux (excluding the volume of the ideal conductors) is given by

$$\psi' = \int_a^b B_\phi dr = \frac{\mu_0 I}{2\pi} \int_a^b \frac{dr}{r} = \frac{\mu_0 I}{2\pi} \ln\left(\frac{b}{a}\right) \tag{5.16}$$

The inductance per unit length is therefore

$$L' = \frac{\mu_0}{2\pi} \ln\left(\frac{b}{a}\right) \text{ [H/m]} \tag{5.17}$$

Copyright © 2007. Cambridge University Press. All rights reserved.

Niknejad, Ali M.. Electromagnetics for High-Speed Analog and Digital Communication Circuits, Cambridge University Press, 2007.
ProQuest Ebook Central, <http://ebookcentral.proquest.com/lib/sjsu/detail.action?docID=288460>.
Created from sjsu on 2023-03-15 17:16:07.

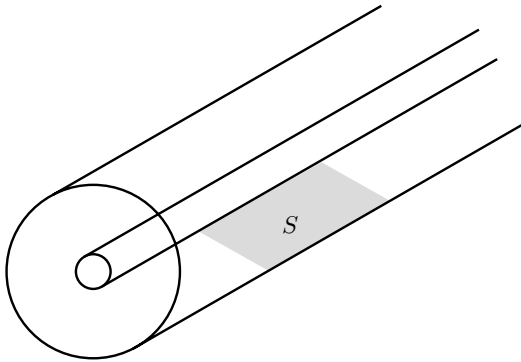


Figure 5.6 The partial inductance of a coaxial conductor is computed by calculating the flux in the shaded area.

Recall that the product of inductance and capacitance per unit length for a transmission line is a constant (see Ch. 9)

$$L'C' = \frac{1}{c^2} \quad (5.18)$$

where c is the speed of light in the medium. Thus we have calculated the capacitance per unit length without any extra work.

5.3 Magnetic energy and inductance

Energy for a system of current loops

We would like to calculate the magnetic energy associated with a system of current loops, shown in Fig. 5.7. In the electrostatic case, we assembled our charge distribution one point charge at a time and used electric potential to calculate the energy. This can be done for the magnetostatic case but there are some complications.

Energy for two loops

As we move in our second loop with current I_2 , we would be cutting across flux from loop 1, as shown in Fig. 5.8a, and therefore an induced voltage around loop 2 would change the current. When we bring the loop to rest, the induced voltage would drop to zero. To maintain a constant current, therefore, we would have to supply a voltage source in series to cancel the induced voltage. The work done by this voltage source represents the magnetostatic energy in the system.

A simpler approach is to bring in the two loops with zero current and then increase the current in each loop one at a time, shown in Fig. 5.8b. First, let us increase the current in

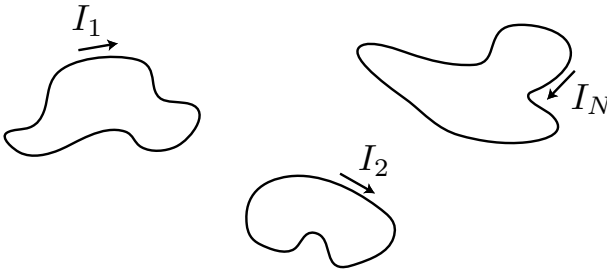


Figure 5.7 A system of current loops store magnetic energy.

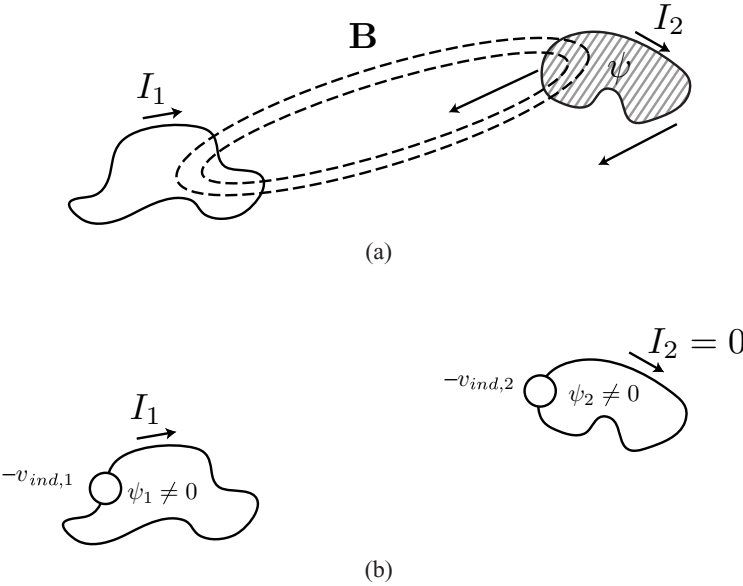


Figure 5.8 (a) When moving loop 2 into our system, the rate of change of flux in loop 2 generates a current in loop 2, which changes the field configuration. (b) Instead, we bring in loops 1 and 2 without any current, and turn on the current in each loop one at a time.

loop 1 from zero to I_1 in some time t_1 . Note that at any instant of time, a voltage is induced around loop number 1 due to its changing flux

$$v_{ind,1} = -\frac{d\psi}{dt} = -L_1 \frac{di_1}{dt} \tag{5.19}$$

where i_1 represents the instantaneous current.

Note that this induced voltage will tend to decrease the current in loop 1. This is a statement of Lenz's Law. In other words, the induced voltage in loop 1 tends to create a magnetic field to oppose the field of the original current! To keep the current constant in loop 1, we must connect a voltage source to cancel the induced voltage. The work done by

this voltage source is given by

$$w_1 = \int_0^{t_1} p(\tau) d\tau = - \int_0^{t_1} v_{ind,1} i_1(\tau) d\tau \quad (5.20)$$

where $p(t) = -v_{ind,1} i_1(t) = L_1 i_1 \frac{di_1}{dt}$. The net work done by the source is simply

$$w_1 = L_1 \int_0^{t_1} i_1 \frac{di_1}{d\tau} d\tau = L_1 \int_0^{I_1} i_1 di_1 = \frac{1}{2} L_1 I_1^2 \quad (5.21)$$

Note that to keep the current in loop 2 equal to zero, we must also provide a voltage source to counter the induced voltage

$$v_{ind,2}(t) = -M_{21} \frac{di_1}{dt} \quad (5.22)$$

This voltage source does not do any work since $i_2(t) = 0$ during this time.

By the same argument, if we increase the current in loop 2 from 0 to I_2 in time t_2 , we need to do work equal to $\frac{1}{2} L_2 I_2^2$. But is that all? No, since to keep the current in loop 1 constant at I_1 we must connect a voltage source to cancel the induced voltage

$$-v_{ind,1} = \frac{d\psi_1}{dt} = M_{12} \frac{di_2}{dt} \quad (5.23)$$

The additional work done is therefore

$$w'_1 = \int_0^{t_2} M_{12} \frac{di_2}{d\tau} I_1 d\tau = M_{12} I_1 I_2 \quad (5.24)$$

The total work to bring the current in loop 1 and loop 2 to I_1 and I_2 is therefore

$$W = \frac{1}{2} L_1 I_1^2 + \frac{1}{2} L_2 I_2^2 + M_{12} I_1 I_2 \quad (5.25)$$

But the energy should not depend on the order we turn on the currents. Thus we can immediately conclude that $M_{12} = M_{21}$. We already saw this when we derived an expression for M_{12} using the Neumann equation.

Generalize to N loops

We can now pretty easily generalize our argument for two loops to N loops

$$W = \frac{1}{2} \sum_i L_i I_i^2 + \sum_{i>j} \sum_j M_{ij} I_i I_j \quad (5.26)$$

The first term represents the “self” energy for each loop and the second term represents the interaction terms.

$$W = \frac{1}{2} \sum_{i=j} \sum_j M_{ij} I_i I_j + \frac{1}{2} \sum_{i \neq j} \sum_j M_{ij} I_i I_j \quad (5.27)$$

$$W = \frac{1}{2} \sum_i \sum_j M_{ij} I_i I_j \quad (5.28)$$

The factor $1/2$ accounts for the double counting in the double summation. Let's rewrite this equation and combine terms

Energy in terms of vector potential

We derived the mutual inductance between two filamentary loops using Neumann's equation

$$M_{ij} = \frac{\mu_0}{4\pi} \oint_{C_i} \oint_{C_j} \frac{d\ell_i \cdot d\ell_j}{R_{ij}} \quad (5.29)$$

Let's substitute the above relation into the expression for energy

$$W = \frac{1}{2} \sum_i I_i \left[\sum_j M_{ij} I_j \right] \quad (5.30)$$

$$W = \frac{1}{2} \sum_i I_i \left[\sum_j I_j \frac{\mu_0}{4\pi} \oint_{C_i} \oint_{C_j} \frac{d\ell_i \cdot d\ell_j}{R_{ij}} \right] \quad (5.31)$$

Let's change the order of integration and summation

$$W = \frac{1}{2} \sum_i I_i \oint_{C_i} \left[\sum_j I_j \frac{\mu_0}{4\pi} \oint_{C_j} \frac{d\ell_j}{R_{ij}} \right] \cdot d\ell_i \quad (5.32)$$

Each term of the bracketed expression represents the vector potential due to loop j evaluated at a position on loop i . By superposition, the sum represents the total voltage potential due to all loops

$$W = \frac{1}{2} \oint_{C_i} I_i A \cdot d\ell_i \quad (5.33)$$

We derived this for filamentary loops. Generalize to an arbitrary current distribution and we have

$$W_m = \frac{1}{2} \int_V \mathbf{J} \cdot \mathbf{A} dV \quad (5.34)$$

Compare this to the expression for electrostatic energy

$$W_e = \frac{1}{2} \int_V \rho \phi dV \quad (5.35)$$

Thus the vector potential $\mathbf{A}$ really does represent the magnetic potential due to a current distribution in an analogous fashion as ϕ represents the electric potential.

Energy in terms of fields

Let's replace $\mathbf{J}$ by Ampère's Law $\nabla \times \mathbf{H} = \mathbf{J}$

$$W_m = \frac{1}{2} \int_V (\nabla \times \mathbf{H}) \cdot \mathbf{A} dV \quad (5.36)$$

Using the identity

$$\nabla \cdot (\mathbf{H} \times \mathbf{A}) = (\nabla \times \mathbf{H}) \cdot \mathbf{A} + \mathbf{H} \cdot (\nabla \times \mathbf{A}) \quad (5.37)$$

$$W_m = \frac{1}{2} \int_V \nabla \cdot (\mathbf{H} \times \mathbf{A}) dV + \frac{1}{2} \int_V (\nabla \times \mathbf{A}) \cdot \mathbf{H} dV \quad (5.38)$$

Apply the Divergence Theorem to the first term to give

$$W_m = \frac{1}{2} \int_S \mathbf{H} \times \mathbf{A} \cdot d\mathbf{S} + \frac{1}{2} \int_V (\nabla \times \mathbf{A}) \cdot \mathbf{H} dV \quad (5.39)$$

We would like to show that the first term is zero. To do this, consider the energy in all of space $V \rightarrow \infty$. Use a large sphere of radius r and take the radius to infinity. We know that if we are sufficiently far from the current loops, the potential and field behave like $A \sim r^{-1}$ and $H \sim r^{-2}$. The surface area of the sphere goes like r^2 . The surface integral, therefore, gets smaller and smaller as the sphere approaches infinity.

The remaining volume integral represents the total magnetic energy of a system of currents

$$W_m = \frac{1}{2} \int_V (\nabla \times \mathbf{A}) \cdot \mathbf{H} dV \quad (5.40)$$

But $\nabla \times \mathbf{A} = \mathbf{B}$

$$W_m = \frac{1}{2} \int_V \mathbf{B} \cdot \mathbf{H} dV \quad (5.41)$$

And the energy density of the field is seen to be

$$w_m = \mathbf{B} \cdot \mathbf{H} \quad (5.42)$$

Recall that $w_e = \mathbf{D} \cdot \mathbf{E}$

Another formula for inductance

The self inductance of a loop is given by

$$L = \frac{1}{I} \int_S \mathbf{B} \cdot d\mathbf{S} \quad (5.43)$$

Since the total magnetic energy for a loop is $\frac{1}{2}LI^2$, we have an alternate expression for the inductance

$$\frac{1}{2}LI^2 = \frac{1}{2} \int_V \mathbf{B} \cdot \mathbf{H} dV \quad (5.44)$$

$$L = \frac{1}{I^2} \int_V \mathbf{B} \cdot \mathbf{H} dV \quad (5.45)$$

This alternative expression is sometimes easier to calculate. We will see this when we compute the *internal* inductance of a round wire.

Self inductance of filamentary loops

We have tacitly assumed that the inductance of a loop is a well-defined quantity. But for a filamentary loop, we can expect trouble. By definition

$$L = \frac{1}{I} \int_S \mathbf{B} \cdot d\mathbf{S} = \frac{1}{I} \oint_C \mathbf{A} \cdot d\mathbf{l} \quad (5.46)$$

$$L = \frac{\mu}{4\pi} \oint_C \oint_C \frac{d\ell' \cdot d\ell}{R} \quad (5.47)$$

This is just Neumann's equation with $C_1 = C_2$. But for a filamental loop, $R = 0$ when both loops traverse the same point. The integral is thus not defined for a filamental loop!

Magnetic energy of a circuit

In the previous section we derived that the total magnetic energy in a circuit is given by

$$E_m = \frac{1}{2}L_1 I_1^2 + \frac{1}{2}L_2 I_2^2 + M I_1 I_2 \quad (5.48)$$

We would like to show that this implies that $M \leq \sqrt{L_1 L_2}$. Let's re-write the above into the following positive definite form

$$E_m = \frac{1}{2}L_1 \left(I_1 + \frac{M}{L_1} I_2 \right)^2 + \frac{1}{2} \left(L_2 - \frac{M^2}{L_1} \right) I_2^2 \quad (5.49)$$

An important observation is that regardless of the current I_1 or I_2 , the magnetic energy is non-negative, so $E_m \geq 0$.

Consider the current $I_2 = \frac{-L_1}{M} I_1$, which cancels the first term in E_m

$$E_m = \frac{1}{2} \left(L_2 - \frac{M^2}{L_1} \right) I_2^2 \geq 0 \quad (5.50)$$

Since $I_2^2 \geq 0$, we have

$$L_2 - \frac{M^2}{L_1} \geq 0 \quad (5.51)$$

Therefore it is now clear that this implies

$$L_1 L_2 \geq M^2 \quad (5.52)$$

Coupling coefficient

Usually we express this inequality as

$$M = k \sqrt{L_1 L_2} \quad (5.53)$$

where k is the *coupling coefficient* between two circuits, with $|k| \leq 1$. If two circuits are perfectly coupled (all flux from circuit 1 crosses circuit 2), $k = 1$ (ideal transformer). Note that $M < 0$ implies that $k < 0$, which is totally reasonable as long as k lies on the unit interval $-1 \leq k \leq 1$. Negative coupling just means that the flux gets inverted before crossing the second circuit. This is easily achieved by winding the circuits with opposite orientation.

5.4 Discussion of inductance

As we have seen in the previous section, there are many equivalent ways to define inductance. All of the following forms are equally important as they shed new insights into the problem. In this section we discuss these definitions in further detail.

Magnetostatics and quasistatics

Magnetostatics is the study of magnetic fields under time-invariant conditions. Clearly, with no time variation, Maxwell's equations take on very simple forms. While these conditions seem oversimplified for practical applications, especially when disallowing any time variation, we shall see later that many seemingly difficult non-static problems admit solutions when seen in this light. For instance, when studying transmission lines, structures where electric and magnetic energy is in constant dynamic exchange, we shall find that the cross-sectional waveforms take on static field configurations. In other problems, the static behavior will dominate up to very high frequencies and thus we can solve such problems using the *quasistatic* assumption. At a given frequency, or equivalently for a given wavelength of operation, we shall show that this is especially true for structures that are small relative to the wavelength. For instance, even for signals varying tens of billions of times per second, the wavelength is of the order of centimeters, an order of magnitude larger than most integrated circuit structures.

Under the magnetostatic conditions, currents flow uniformly in the volume of the problem space. Thus the net current flow into any volume element V is zero. If it were non-zero, then net charge would continuously accumulate in this volume element, violating the steady-state assumption.

The net current flowing into the volume element V is given by integrating the component of current density $\mathbf{J}$ perpendicular to the bounding surface of the volume S . By the Divergence Theorem we have

$$0 = \oint_S \mathbf{J} \cdot d\mathbf{S} = \int_V \nabla \cdot \mathbf{J} dV \quad (5.54)$$

Assuming Ohm's Law to be valid in the conductive regions and further assuming that the material is isotropic implies the current density is proportional and parallel to the local electric field $\mathbf{J} = \sigma \mathbf{E}$. Substituting in the above equation we have

$$0 = \int_V \nabla \cdot \mathbf{J} dV = \int_V \sigma \nabla \cdot \mathbf{E} dV \quad (5.55)$$

where for simplicity we assume that the material conductivity σ is independent of position. But from Maxwell's equations, in particular Gauss' Law, the divergence of the electric field can only arise from charge. We therefore have

$$0 = \int_V \frac{\sigma}{\epsilon} \rho dV \quad (5.56)$$

Again, let us assume that the permittivity does not vary with position. Take an arbitrary small conductive volume ΔV . Since $\sigma \neq 0$, this forces the net charge to equal zero, or $\rho = 0$. Since the volume element we chose was arbitrary, this implies that all conductors in the problem must have a net charge of zero.

But current flow is a result of charge in motion and from the outset we assume that $\mathbf{J} \neq 0$. Is $\rho = 0$ a contradiction? No, because at any given moment we assume that all moving charges are neutralized by charges of opposite sign. These charges can also be in motion, such as holes in a semiconductor, or they can be immobile ionized host atoms. In

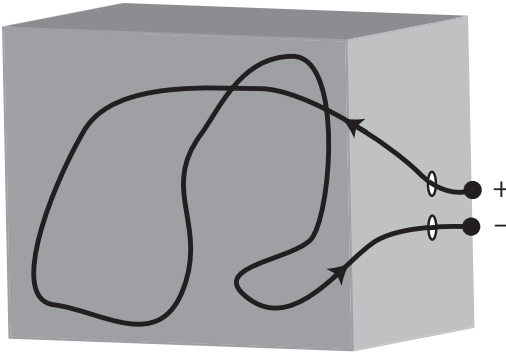


Figure 5.9 A conductive box containing an inductor. The inductor wires are insulated from the box. The inductor may or may not self-intersect.

reality, this is an approximation for finite values of σ since the atoms are in constant thermal agitation causing a net time-varying multi-pole charge moment.

But doesn't $\rho = 0$ imply also that $\mathbf{E} = 0$? No, only that $\nabla \cdot \mathbf{E} = 0$. Since we have assumed steady-state conditions, the electric field has only a curl-free component

$$\mathbf{E} = -\nabla\phi \tag{5.57}$$

The magnetic field does not interact with the electric field since this can only occur under time-varying situations. Maxwell's equations lead to a source-free Poisson's equation

$$\nabla^2\phi = -\frac{\rho}{\epsilon} = 0 \tag{5.58}$$

Since there are no sources in the above equation, the scalar potential can be non-zero only if the boundary conditions are non-zero. If we expand the volume space of the problem to infinity where $\phi = 0$, then we do indeed have a contraction. Since in this case we have $\phi \equiv 0$ and thus $\mathbf{E} \equiv 0$, hence no currents can flow. This would also be the case for a grounded closed cavity with walls constructed from perfect conductors. Without any time variation, current cannot flow in the interior of the cavity. Thus, we have to assume that on the boundary of our problem space there are points where $\phi \neq 0$.

As shown in Fig. 5.9, we must break the circuit loop under consideration and apply a field between the terminals of the loop.

Magnetic flux

Inductance defined in terms of magnetic flux is the most common approach, but it requires some subtle observations. As we have seen, the voltage induced in a loop can be obtained by integrating the electric field along the path of the circuit. By Stoke's Theorem we have

$$V_L = -\oint_C \mathbf{E} \cdot d\mathbf{l} = -\int_S (\nabla \times \mathbf{E}) \cdot d\mathbf{S} \tag{5.59}$$

Copyright © 2007. Cambridge University Press. All rights reserved.

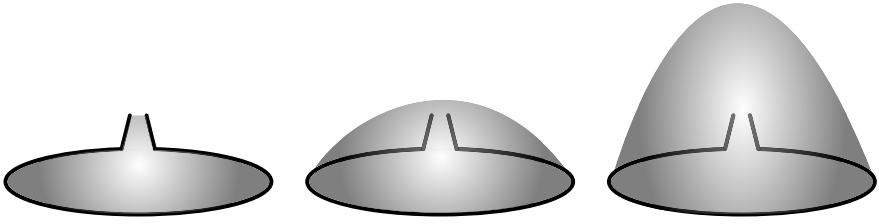


Figure 5.10 Several candidate surfaces bounded by the same curve C .

Note that this is true for any surface S bounded by the circuit path C . From Maxwell's equations we have

$$\nabla \times \mathbf{E} = -\frac{d\mathbf{B}}{dt} \quad (5.60)$$

so that

$$V_L = \frac{d}{dt} \int_S \mathbf{B} \cdot d\mathbf{S} \quad (5.61)$$

The right-hand side is the magnetic flux through the surface bounded by the current. We defined the magnetic flux ψ by

$$\psi \equiv \int_S \mathbf{B} \cdot d\mathbf{S} \quad (5.62)$$

and it follows that

$$V_L = \frac{d\psi}{dt} \quad (5.63)$$

comparing with Eq. (5.1) we have

$$L \equiv \frac{\psi}{I} \quad (5.64)$$

In computing the flux, which surface do we use in the above integral? In Fig. 5.10 we show several candidate surfaces. Clearly, there are an infinite number of surfaces bounded by the closed curve. Is there a unique surface we can attribute to the path? We might try by limiting the surface to the minimal surface (the so-called “soap film” surface generated by dipping the circuit in a soap bath). But the first step of the above manipulation is justified by Stoke's Theorem which admits *any* surface as long as it is enclosed by the path of the circuit.

The magnetic field $\mathbf{B}$ is computed under magnetostatic conditions described above. For a closed circuit path under magnetostatic conditions, this implies a zero electric field in the volume of perfect conductors. Since the voltage drop around a loop of perfectly conducting material is zero, no voltages can be induced in the loop and so the rate of change flux $\dot{\psi}$ must equal zero. A realistic example of this occurs in a superconductor where the static state is literally frozen into space.

Most practical inductors are constructed from materials of finite non-zero resistivity. We break the loop and introduce an electric field at the terminals. Since the electric field is

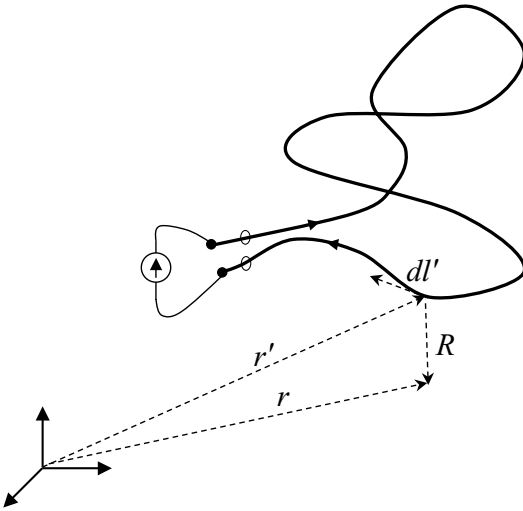


Figure 5.11 Computing the magnetic field using the Biot-Savart equations.

zero or nearly zero along the path of the conductive portion of the loop, the entire induced voltage calculated above appears at the terminals of the circuit.

Under magnetostatic conditions, we can calculate $\mathbf{B}$ from the so-called Biot-Savart Law [52]. This result is easily derived from Ampère’s Law in Maxwell’s equations

$$\mathbf{B}(r) = \mu \oint_C \frac{I(\mathbf{r}')d\mathbf{l}' \times \mathbf{R}}{4\pi R^3} = \frac{1}{4\pi\mu^{-1}} \oint_C \frac{I(\mathbf{r}')d\mathbf{l}' \times \hat{\mathbf{R}}}{R^2} \quad (5.65)$$

where, as shown in Fig. 5.11, the vector $\mathbf{R}$ is drawn from the “source” point r' along the path of integration on the circuit to the “field” point r . The convention of designating the source point with primed coordinates is an arbitrary but common practice.

The above equation is insightful as it states that the influence of current elements on magnetic field follows an inverse-square law dependence similar to the influence of charge on the electric field.

Now the second subtlety enters in the definition of inductance. In computing the surface integral of Eq. (5.62), where does the boundary of the surface lie? Eq. (5.65) clearly shows that the magnetic field is unbounded as we approach a filamental wire so that the integral is undefined for this case. Therefore, we conclude that the inductance of an infinitesimally thin wire is infinite. Should the boundary be taken as only the inside edge of the wire or should the boundary also include the wire itself?

Internal and external inductance

Let us partition the flux as follows

$$\psi = \psi_i + \psi_o \quad (5.66)$$

Copyright © 2007. Cambridge University Press. All rights reserved.

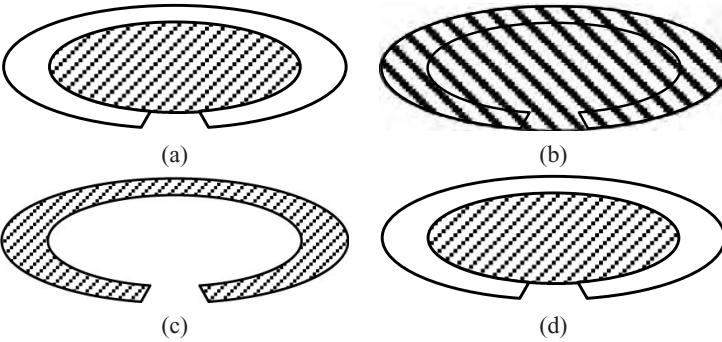


Figure 5.12 (a) The *inner* flux of a circuit. (b) The *outer* flux of a circuit. (c) The *internal* flux of a circuit. (d) The *external* flux of a circuit (same as (a)).

where ψ_i is the inner flux shown in Fig. 5.12a and is computed from the surface integral when the boundary is taken at the internal edge of the loop. The outer flux ψ_o is shown in Fig. 5.12b, where the area of the flux is bounded by the outer edge of the wire forming the loop.

The voltage induced due to $d\psi_i/dt$ is smaller since we are integrating over a smaller area and the flux is non-zero in the volume of the conductors. Comparatively the voltage induced due to $d\psi_o/dt$ is larger. The difference between the voltage induced on the inner edge and outer edge is thus greater the larger the width of the wire. For finite σ we can see then that the current distribution could be non-uniform along the width of the conductor. For good conductors under quasistatic conditions, the current flow is uniform through the width of conductors and the flux is therefore non-zero in the volume of conductors. The voltage induced, therefore, includes both the inner flux ψ_i and the outer flux ψ_o .

The external flux ψ_{ext} is identical with the “inner” flux ψ_i we defined above. The “internal” flux is given by

$$\psi_{int} = \psi_o - \psi_i \quad (5.67)$$

and contains the flux bounded by the inner and outer edges of the conductor. The total inductance is derived by including both terms. This partitioning is convenient once we consider the frequency variation of inductance. As we shall see, at high frequencies the current flow in the volume of conductors decreases (due to induced eddy currents) and current density peaks at the surface of conductors. Then the internal flux ψ_{int} vanishes and the inductance is given by the external flux ψ_{ext} . The external flux is approximately independent of frequency, whereas the internal flux tends to diminish at increasing frequency.

Internal inductance of a round wire

Usually if the wire radius is small relative to the loop area, $L_{int} \ll L_{ext}$. We shall see that at high frequencies, the magnetic field decays rapidly in the volume of conductors and thus the $\psi_{int} \rightarrow 0$ and $L(f \rightarrow \infty) = L_{ext}$. Consider a round wire carrying uniform current. We

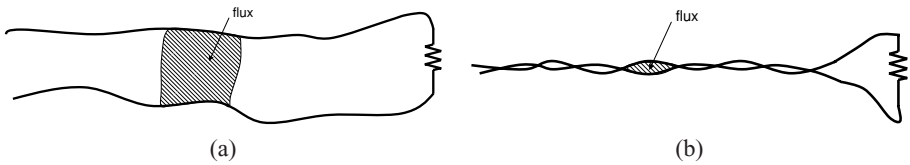


Figure 5.13 (a) Flux of a loop inductor. (b) The magnetic flux of a twisted pair circuit.

can easily derive the magnetic field through Ampère's Law

$$B_{inside} = \frac{\mu_0 I r}{2\pi a^2} \quad (5.68)$$

Using this expression, we can find the internal inductance

$$W_m = \frac{1}{2} \int_V \mathbf{B} \cdot \mathbf{H} dV = \frac{1}{2} \int_{V_{inside}} \mathbf{B} \cdot \mathbf{H} dV + \frac{1}{2} \int_{V_{outside}} \mathbf{B} \cdot \mathbf{H} dV \quad (5.69)$$

$$W_m = \frac{1}{2} L_{int} I^2 + \frac{1}{2} L_{ext} I^2 \quad (5.70)$$

The “inside” term is easily evaluated

$$W_{m,int} = \frac{1}{2} \frac{\mu_0 I^2}{(2\pi a^2)^2} \int_0^a r^2 2\pi r dr = \frac{1}{2} \frac{\mu_0 I^2}{8\pi} \quad (5.71)$$

The internal inductance per unit length is thus

$$L_{int} = \frac{\mu_0}{8\pi} \quad (5.72)$$

Numerically, this is 50pH/mm, a pretty small inductance. Recall that this is only the inductance due to energy stored inside of the wires. The external inductance is likely to be much larger.

Inductance for a general structure

We are now in a position to comment on the inductance of an arbitrary circuit. For instance, to increase the inductance of a circuit we should increase the flux of the circuit. Since the current is divided out of flux, it does not come into play for linear materials where the magnetic permeability μ is independent of current. Hence, the inductance is a purely geometric property of the circuit and we may speak of the inductance in the same way that we speak of other geometric properties, such as the perimeter or surface area of the circuit.

Since the inductance is an integral over the surface of the circuit, it seems as if a circuit that spans a larger area should have higher inductance. For instance, compare the flux area of the two loops shown in Fig. 5.13. In part (a), we show a circuit that spans a large area since the “return current” flows in a wire physically far from the incoming current. In Fig. 5.13b, we keep the return path in close proximity to the incoming current by twisting the wires together. In the limit that the wires are just touching (with some insulation in between to prevent a short), the inductance is driven to arbitrarily low values.

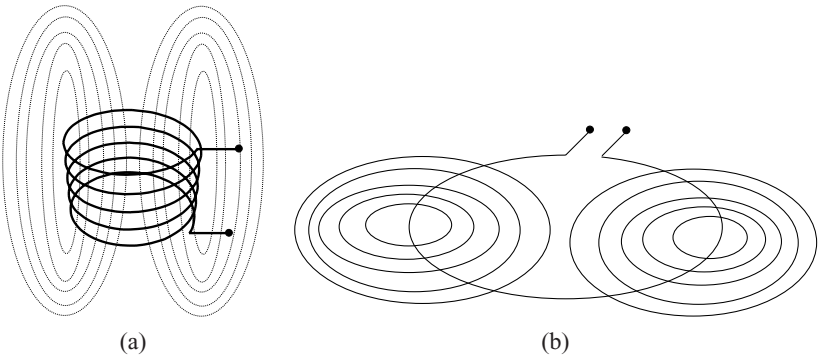


Figure 5.14 (a) A simple loop inductor with area equal to (b) a coil inductor.

Twisting the wires together has another benefit, besides low inductance. The magnetic pickup from a twisted pair is much smaller than a two-wire line with the equivalent inductance. This is because any flux linking with the two-wire line generates an emf, whereas in a twisted pair, for a uniform incoming B field, the flux tends to cancel out since the wire is twisted. In one section of the wire, the flux pickup is positive, but in the next it is negative. In this way, incoming magnetic fields generate a net flux of nearly zero on the twisted pair.

Let us compare two circuits with equal surface area but very different values of inductance. In Fig. 5.14a we see a traditional coil inductor. To help visualize the surface of integration, imagine the portion in the coil as a spiraling ramp resembling a multi-level parking lot structure (Fig. 5.15). For comparison we have also shown a large single loop inductor designed so that the area of the loop is equal to the surface area of the coil. The coil has larger inductance, though, since the magnetic field at any given point inside the coil has additive contributions from several adjacent “loops.” Since the surface area of the single loop is equal to the surface area of the coil, the flux will be larger if the average magnetic field is larger in the coil. Intuitively, it seems like this is the case due to the additive nature of the field from one loop to adjacent loops.

Mutual inductance

The vector nature of the magnetic field leads to an interesting property of magnetic coupling. If the magnetic flux generated by one loop impinges perpendicularly on another loop, it does not generate an induced voltage and hence the mutual coupling $M_{ij} = 0$ for such a case. This principle can be used to isolate structures in a circuit where coupling is undesired.

The net results of mutual inductance can lead to many interesting effects. For instance, loops wound in the opposite orientation add negatively and thus tend to cancel the flux and lower the induced voltage. This is illustrated in Fig. 5.16, where we use the right-hand convention to orient the magnetic fields. Clearly, loop B tends to cancel the self flux of loop A, whereas loop C adds to the flux of loop A.

A nice application of this concept is the ground-fault interrupt circuit (GFI) shown in Fig. 5.17. This circuit is designed to detect a *ground fault*, a condition where an electric

Copyright © 2007, Cambridge University Press. All rights reserved.

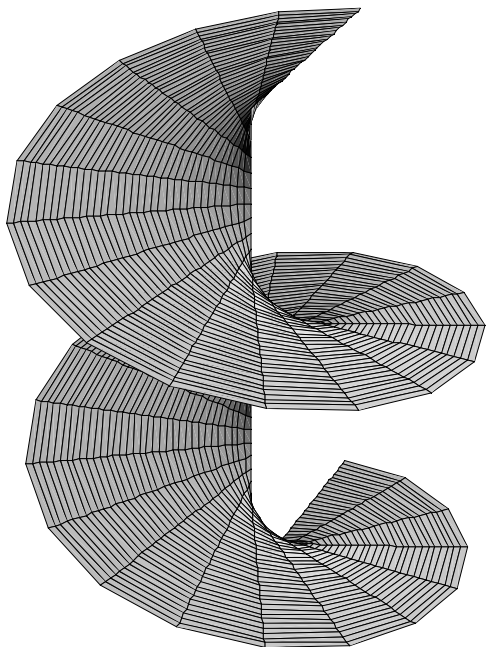


Figure 5.15 A surface bounded by the helix coil inductor.

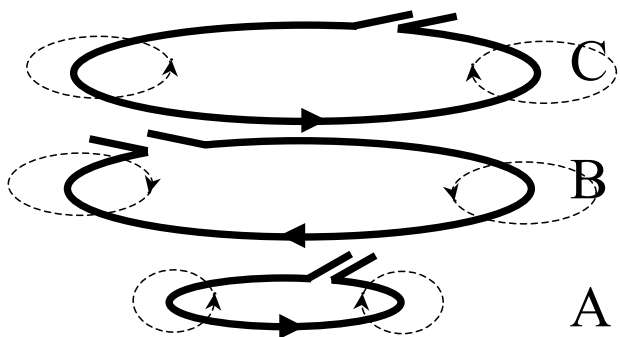


Figure 5.16 Three mutually coupled loops of current. The flux of loop C adds to the flux of loop A, whereas the flux of loop B abstracts from A.

appliance faults by shorting a signal line to ground. Under normal operation all the current that enters the appliance should return to ground via the ground line. If there is any imbalance between the signal line current and ground current, this signifies that a parasitic path to ground has been established, and possibly via conduction current through an unfortunate victim (say a person grounded accidentally touches the high voltage signal line). Under normal operation the voltage on the GFI loop is zero since the flux of the incoming

Copyright © 2007. Cambridge University Press. All rights reserved.

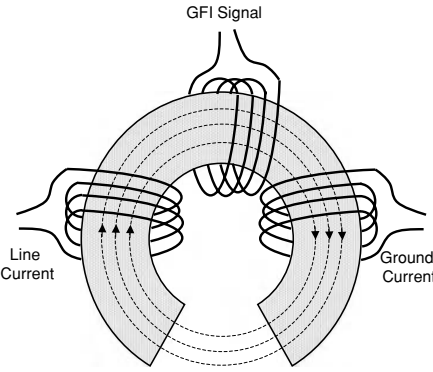


Figure 5.17 Schematic representation of a ground-fault interrupt (GFI) circuit.

current is canceled by the ground current. If any imbalance occurs, a net voltage will be induced at the GFI node and this condition can be detected to quickly shut off power to the appliance.

Magnetic energy perspective

From electromagnetic theory we know that the power stored due to an electromagnetic field is given by [52]

$$P = \frac{1}{2} \frac{d}{dt} \int_V \left(\frac{1}{\mu} |\mathbf{B}|^2 + \epsilon |\mathbf{E}|^2 + \mathbf{E} \cdot \mathbf{J} \right) dV \quad (5.73)$$

where the integral is over the entire volume space of the problem where the fields are non-zero. In a good “inductor,” the magnetic energy dominates over the electrical energy, or $\mu^{-1} |\mathbf{B}|^2 \gg \epsilon |\mathbf{E}|^2$. Thus, the above parenthetical expression reduces to the total stored magnetic energy

$$E_L = \frac{1}{2} \int_V \frac{1}{\mu} |\mathbf{B}|^2 dV \quad (5.74)$$

Using equation E_L we derived an alternate expression for inductance

$$L = \frac{1}{I^2} \int_V \frac{1}{\mu} |\mathbf{B}|^2 dV \quad (5.75)$$

This viewpoint emphasizes the fact that the voltages induced across an inductor terminal or around a circuit loop are due to the magnetic stored energy. Thus an induced voltage can only sustain itself as long as $\mathbf{B}$ is non-zero. In a time periodic circuit the magnetic energy is continuously “pumped” as the induced voltage changes sign. When $dI/dt > 0$ energy is supplied to the inductor. Alternatively, when $dI/dt < 0$ energy returns to the rest of the circuit. Naturally, an inductor constructed from materials with zero resistance must conserve energy.

Copyright © 2007, Cambridge University Press. All rights reserved.

The inductance computed from the definition of flux must equal the value computed from the energy of the field. This implies that

$$\frac{1}{I} \int_V \frac{1}{\mu} |\mathbf{B}|^2 dV = \int_S \mathbf{B} \cdot d\mathbf{S} \quad (5.76)$$

Example 10

In this example we compute the partial inductance of a two wire transmission line. Assume each wire is round and of radius a and the origins of the wires are separated by a distance d .

From previous examples we already have the necessary tools to compute the external inductance. Since the magnetic field outside of a round wire is equivalent to the field of a filament placed at the origin of the wire, we can reuse the results from the previous example for the magnetic field of two filaments. For the external inductance, we integrate this expression between the wires (excluding the conductors)

$$L = \int_a^{d-a} B_y(y=0) dx \quad (5.77)$$

Recall that $B_y(0)$ is given by

$$B_y = \frac{-\mu I(x+d)}{2\pi((x+d)^2 + y^2)} - \frac{\mu I(x-d)}{2\pi((x-d)^2 + y^2)} \Big|_{y=0} \quad (5.78)$$

Substituting for B_y we have the following easy integral

$$L = \frac{\mu I}{2\pi} \int_a^{d-a} \left(\frac{1}{x} - \frac{1}{x-d} \right) dx \quad (5.79)$$

so that the external inductance per unit length is given by

$$L = \frac{\mu I}{\pi} \ln \left(\frac{d-a}{a} \right) \quad (5.80)$$

The above expression grows logarithmically as a function of separation d . To compute the internal inductance, let us first consider an isolated conductor, or equivalently take $d \gg a$. Then the volume integral of $|B|^2$ inside the wire is given by

$$\int_V B^2 dV = \int_0^{2\pi} \int_0^a B(r)^2 r dr d\theta \quad (5.81)$$

where

$$B(r)^2 = \frac{\mu^2 I^2 r^2}{4\pi^2 a^4} \quad (5.82)$$

Performing the trivial integral (there aren't too many in this business!), we have

$$\int_V B^2 dV = \frac{\mu^2 I^2}{8\pi} \quad (5.83)$$

and from the energy definition of the inductance, the internal inductance per unit length for the conductor is given by

$$L = \frac{\mu}{8\pi} \quad (5.84)$$

This expression is interestingly independent of the radius of the wire. So for two widely separated wires we double the above result for the internal inductance. The exact solution is more complicated but follows in a similar fashion.

Magnetic vector potential

As we have seen, the concept of the magnetic vector potential comes about directly as a consequence of Maxwell's equations. The magnetic vector $\mathbf{B}$ is divergenceless

$$\nabla \cdot \mathbf{B} \equiv 0 \quad (5.85)$$

due to the absence of magnetic monopoles. Since $\nabla \cdot \nabla \times \mathbf{F} \equiv 0$ is an identity for any vector function $\mathbf{F}$, it follows that

$$\mathbf{B} = \nabla \times \mathbf{A} \quad (5.86)$$

We call the vector $\mathbf{A}$ the vector potential in analogy with the electric scalar potential ψ . At first, though, it seems like the vector $\mathbf{A}$ is superfluous since it too is a vector and not a scalar as in the case of ψ . Furthermore there is some physical ambiguity in $\mathbf{A}$ since its divergence is arbitrary.

The utility in defining $\mathbf{A}$, though, is clear when we compare the defining equation for $\mathbf{A}$ compared with $\mathbf{B}$. Under magnetostatic conditions the vector function $\mathbf{A}(\mathbf{r})$ is given by

$$\mathbf{A}(\mathbf{r}) = \mu \int_{V'} \frac{\mathbf{J}(\mathbf{r}')}{4\pi R} dV' \quad (5.87)$$

As before R is the distance between the source and field point and the volume of integration is over all space where $\mathbf{J} \neq 0$.

Compare this equation with the Biot-Savart Law shown in Eq. (5.65). The vector $\mathbf{A}$ is in the direction of the current and no cross products are involved in the computation. In fact, if we partition the vector $\mathbf{J} = J_x \hat{\mathbf{x}} + J_y \hat{\mathbf{y}} + J_z \hat{\mathbf{z}}$ into Cartesian coordinates, then Eq. (5.87) reduces to three scalar integrals identical in form to the calculation of the electric scalar potential in terms of charge density.

The vector potential $\mathbf{A}$ is related to inductance as follows. Substitute $\mathbf{B} = \nabla \times \mathbf{A}$ into Maxwell's equation

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} = -\nabla \times \frac{\partial \mathbf{A}}{\partial t} \quad (5.88)$$

Since the curl operator is associative, we have

$$\nabla \times \left(\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} \right) = 0 \quad (5.89)$$

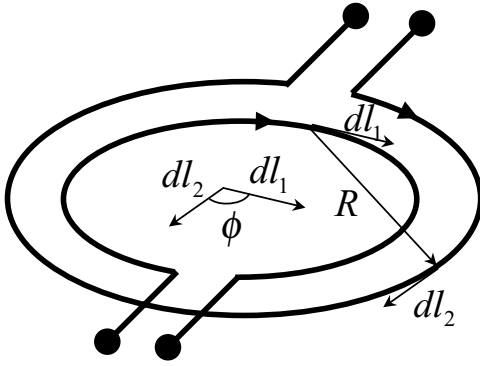


Figure 5.18 Two magnetically coupled filamentary loops.

and this implies that the curl-free expression in the parentheses, similar to the electrostatic field, can be written as the gradient of a scalar function

$$\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} = -\nabla \phi \quad (5.90)$$

The sign of the scalar potential is arbitrarily chosen to correspond to the electrostatic case. In other words, the electric field at any point in space has two components

$$\mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t} - \nabla \phi \quad (5.91)$$

If we integrate the above expression around a closed loop, the second term is zero and we are only left with the integral of the first term

$$V_L = \oint_C \mathbf{E} \cdot d\mathbf{l} = -\frac{d}{dt} \oint_C \mathbf{A} \cdot d\mathbf{l} \quad (5.92)$$

leading to an expression for the inductance in terms of the magnetic vector potential

$$L = \frac{1}{I} \oint_C \mathbf{A} \cdot d\mathbf{l} \quad (5.93)$$

We could have derived the above result directly from Stoke's Theorem

$$\int_S \mathbf{B} \cdot d\mathbf{S} = \oint_C (\nabla \times \mathbf{B}) \cdot d\mathbf{l} \quad (5.94)$$

but the above derivation gives us more physical insight into the magnetic vector potential. In fact, Eq. (5.91) justifies the name *potential*. Under general circumstances, the electric field is the gradient of the scalar potential and the time derivative of the vector potential. Both operations involve differentiation. The time derivative with the vector potential reminds us that magnetic voltages are only induced under dynamic conditions. From Eq. (5.91) the SI units of $\mathbf{A}$ must be volt seconds per meter ($\text{V} \cdot \text{s}/\text{m}$, or since the units of $\mathbf{B}$ can be specified in Tesla (T) or weber per square meter (W/m^2), we also have Tesla meter ($\text{T} \cdot \text{m}$) or weber per meter (W/m).

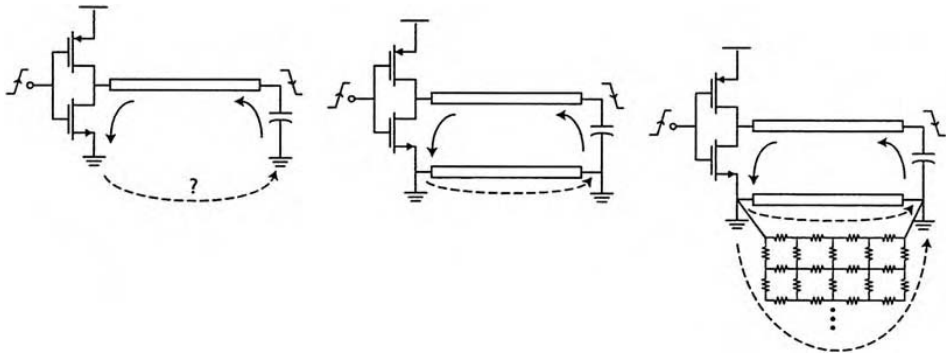


Figure 5.19 (a) An integrated circuit wire carries AC current to a capacitor load. (b) The return current can flow through an explicit path, or (c) through a distributed path formed by the substrate and ground network.

5.5 Partial inductance and return currents

Often people speak of the inductance of an “open” path, such as a piece of wire or a metal trace in an integrated circuit. Quite often, you hear about the inductance of a bond wire or a lead on a package. What are they talking about?

In fact, people are speaking of *partial* inductance. In other words, consider the inductance of the integrated circuit trace shown in Fig. 5.19a. Imagine that we drive this wire with a voltage source at one end through a capacitor load at the other end. Where does the “return current” come from? This depends on how the circuit is laid out. For instance, a separate trace in parallel to the first can carry the return current back to the source, and thus the inductance is defined by this loop (Fig. 5.19b). But often an explicit ground return is not included, and the return current flows through the ground network. Since the substrate is often grounded, the return current can also flow partially through the substrate and thus the inductance is now different since the “loop” is different (Fig. 5.19c). In fact, with the ground and substrate we have a dilemma since we don’t know the exact current distribution. At high frequency, we imagine that most of the current will flow close to the surface in the vicinity of the trace. But at low frequencies, the current flow is more widespread.

Nevertheless, people still say that this trace has “so and so” inductance. They are talking about the inductance of the loop assuming the return current is at infinity. In other words, imagine calculating the inductance through the vector potential as follows. As suggested by Fig. 5.20, if we integrate Eq. (5.93) *partially* through several open paths, then the voltage induced at the terminals of the i th curve is given by

$$V_i = \int_{C_i} \mathbf{E} \cdot d\mathbf{l} \tag{5.95}$$

If we ignore the contribution from ϕ in Eq. (5.91) and only consider the contribution from the vector potential, we then have

$$\hat{V}_i = -\frac{d}{dt} \int_{C_i} \mathbf{A} \cdot d\mathbf{l} \tag{5.96}$$

Copyright © 2007. Cambridge University Press. All rights reserved.

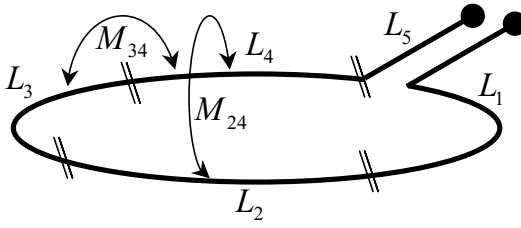


Figure 5.20 Breaking the inductance into partial self and mutual inductance terms.

Due to linearity, we can partition $\mathbf{A}$ into several terms

$$\mathbf{A} = \mathbf{A}_1 + \mathbf{A}_2 + \cdots \quad (5.97)$$

where each term represents the contribution of segment i to the magnetic vector potential

$$\mathbf{A}_i = \frac{\mu}{4\pi} \int_{V_i} \frac{\mathbf{J}_i}{R} dV_i \quad (5.98)$$

The total current I_i flowing into segment i can be taken out of the integral as long as the current distribution is retained in the integral, $\mathbf{J}(\mathbf{r})_i = \mathbf{k}(\mathbf{r})I_i$. Let us define the *partial* mutual inductance between segment i and j as follows

$$M_{ij} = \frac{1}{I_j} \int_{C_i} \mathbf{A}_j \cdot d\mathbf{l} \quad (5.99)$$

so that the voltage around the entire structure is simply given by

$$V = \sum_i \sum_j M_{ij} \frac{dI_j}{dt} \quad (5.100)$$

From the above definition we define the *partial inductance matrix* as the $N \times N$ matrix M with elements M_{ij} . Given vectors v and i representing the segment voltages and currents, we extend Eq. (5.1)

$$v = M \frac{di}{dt} \quad (5.101)$$

5.6 Impedance and quality factor

The circuit concept of impedance can be stated in terms of power

$$Z_{in} = \frac{V}{I} = \frac{V \cdot I^*}{|I|^2} = \frac{P}{\frac{1}{2}|I|^2} = R + jX \quad (5.102)$$

Applying Poynting's Theorem to a "black box," we can write this as

$$Z_{in} = \frac{P_0 + P_\ell + 2j\omega(W_m - W_e)}{\frac{1}{2}|I|^2} \quad (5.103)$$

Note that the resistive component has a radiation term, an ohmic loss term, and possibly a dielectric or permeability loss term

$$R = \frac{P_0 + P_\ell}{\frac{1}{2}|I|^2} \quad (5.104)$$

The reactance is positive if $W_m > W_e$, and negative otherwise

$$X = \frac{2\omega(W_m - W_e)}{\frac{1}{2}|I|^2} \quad (5.105)$$

The quality factor for a “black box” (usually resonator) is defined as follows

$$Q = 2\pi \frac{\text{Peak energy stored}}{\text{Energy loss per cycle}} \quad (5.106)$$

The denominator can be reformulated in terms of the average power loss to give

$$Q = \omega \frac{\text{Peak energy stored}}{P_\ell} \quad (5.107)$$

From Poynting’s Theorem, the net stored energy is given by $W_m - W_e$, so we may be tempted to write

$$Q \stackrel{?}{=} \omega \frac{W_m - W_e}{P_\ell} = \frac{1}{2} \frac{\Im(Z_{in})}{\Re(Z_{in})} \quad (5.108)$$

But the peak energy is different from the *net* energy. In a LC resonator, the energy is shared between the inductor and capacitor in a time-varying fashion. The peak energy can be computed from the capacitor when $i = 0$ or from the inductor when $v_C = 0$.

$$Q = \omega \frac{|W_m| + |W_e|}{P_\ell} = \omega \frac{W_m^{peak}}{P_\ell} = \omega \frac{W_e^{peak}}{P_\ell} = \frac{\omega L}{R} = \frac{1}{\omega C R} = \frac{X}{R} \quad (5.109)$$

For a single one-port element *not* in resonance, we often define the Q factor as

$$Q = \frac{\Im(Z_{in})}{\Re(Z_{in})} = \frac{X}{R} \quad (5.110)$$

We see that this is correct under the assumption that the one-port forms a resonant circuit!

5.7 Frequency response of inductors

Skin effect

At low frequencies when the DC resistance of an inductor dominates over the reactance, for example when $R \gg \omega L$, then current flows evenly through the cross section of a uniform material of a given conductivity. Thus when computing the low-frequency inductance (strictly the DC inductance¹) we have non-zero flux in the volume of conductors and to

¹ The concept of DC may seem meaningless since at DC there is no induction and thus inductance plays no part. However, we shall see that inductance is a weak function of frequency so that even at moderate frequencies, the DC inductance remains valid.

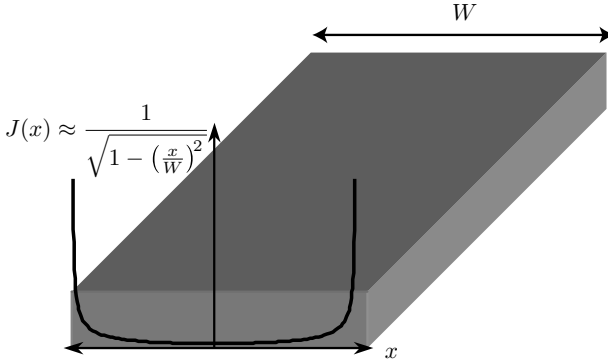


Figure 5.21 Non-uniform current flow through conductors at high frequency. The shown functional dependence is an approximation to the edge discontinuity for rectangular conductors.

calculate the inductance correctly we have to take this into account. At increasing frequency, the internal current distribution redistributes, changing the *internal impedance* of conductors. Ignoring displacement current, the internal impedance is given by

$$Z_{int} = \frac{V_{res}}{I} = \frac{1}{I} \int_C \frac{\mathbf{J}}{\sigma} \cdot d\mathbf{l} \quad (5.111)$$

The above equation is simply the line integral of the electric field $\mathbf{E}$ along the path of a resistive material of conductivity σ . At low frequency when $J = I/A$ is uniform we have the familiar Ohm's Law expression

$$Z_{int} = \frac{1}{I} \int_C \frac{Idl}{\sigma A} = R \quad (5.112)$$

At high frequency $\mathbf{J} = \sigma \mathbf{E}_{tot}$ where $\mathbf{E}_{tot}$ is the total vectorial sum field, including the applied and induced fields. As we alluded to earlier when discussing the internal and external inductance, the induced field is stronger along the outer edge of a loop (which encloses more flux) than the inner loop. Since the induced field opposes the applied field, as shown in Fig. 5.21, we expect the current density to peak near the inner edge of the conductor. In fact, this is exactly what is observed experimentally or by solving Maxwell's equations directly.

We shall show that at increasingly high frequencies the non-uniformity of the current distribution dominates completely and the internal flux is driven to zero. Thus, the infinite frequency asymptote for inductance should equal the value obtained by the external flux

$$L_{\infty} = \frac{\psi_{ext}}{I} \quad (5.113)$$

Note that the DC inductance

$$L_{DC} = \frac{\psi_{ext} + \psi_{int}}{I} \quad (5.114)$$

is larger only by ψ_{int} . Since for usual geometries $\psi_{ext} \gg \psi_{int}$ and $\psi_{int}(f)$ is a decreasing function of frequency, the inductance decreases only slightly with frequency. For instance,

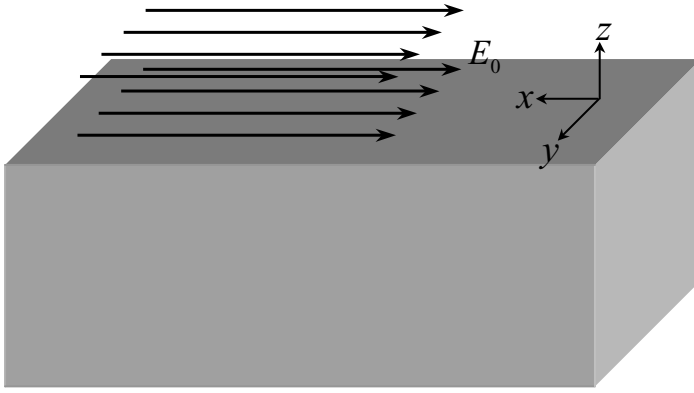


Figure 5.22 Current flow along the surface of a semi-infinite plane.

if the internal flux only accounts for 5% of the total flux, then the high-frequency inductance will only differ by 5% from the DC value.

Surface impedance

In this section we will justify some of the above arguments by solving Maxwell's equations directly. For a simple analytic solution we have to limit ourselves to an unrealistic one-dimensional geometry shown in Fig. 5.22 but the insight afforded by this simple example applies in a surprising number of more general situations.

For a semi-infinite plane supporting a conductive current, we have the following time-harmonic Maxwell equations

$$\nabla \times \mathbf{H} = (\sigma + j\omega\epsilon)\mathbf{E} \quad (5.115)$$

The displacement current $j\omega\epsilon\mathbf{E}$ is negligible in comparison to the conductive current $\sigma\mathbf{E}$. This is true for good conductors and in fact a useful definition for a good conductor. Taking the curl of Maxwell's curl equation for an electric field we have

$$\nabla \times \nabla \times \mathbf{E} = -j\omega\mu\nabla \times \mathbf{H} \quad (5.116)$$

and substituting for the curl of $\mathbf{H}$ from above while dropping the displacement current

$$\nabla \times \nabla \times \mathbf{E} = -j\omega\mu\sigma\mathbf{E} \quad (5.117)$$

The above differential equation describes the spacial variation of the components of the vector $\mathbf{E}$. For simplicity we assume that $\mathbf{E}$ is x -directed so that $\mathbf{E} = \hat{x}E_x$ and furthermore we assume that $\mathbf{E}$ is uniformly incident on the plane. Simplifying the above equation we are led to the following scalar differential equation

$$-\left(\frac{\partial^2 E_x}{\partial y^2} + \frac{\partial^2 E_x}{\partial z^2}\right) = -j\omega\mu\sigma E_x \quad (5.118)$$

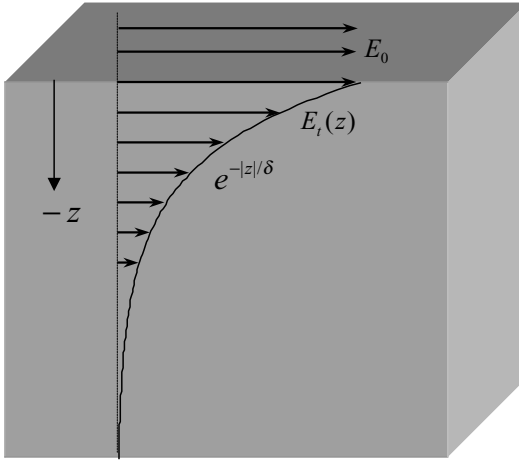


Figure 5.23 The decay of electric field into the depth of a semi-infinite plane conductor.

We assume a uniform wave incident on the plane and thus $\partial E_x / \partial y = 0$ for $z > 0$. By the symmetry of the plane, we can extend this condition to the interior of the plane and we are thus faced with the following simple first-order equation

$$\frac{d^2 E_x}{dz^2} = j\omega\mu\sigma E_x \quad (5.119)$$

The above equation is readily solved by substituting the trial solution $E_0 \exp(\alpha z)$

$$\alpha^2 E_x = j\omega\mu\sigma E_x \quad (5.120)$$

And thus

$$\alpha = \pm \sqrt{j\omega\mu\sigma} = \pm \sqrt{j} \sqrt{\omega\mu\sigma} \quad (5.121)$$

Since the fields must be finite as $z \rightarrow -\infty$, we keep the positive root. Working with polar coordinates, verify that the first term simplifies to $\sqrt{j} = (1 + j)/\sqrt{2}$. Let us define the constant

$$\delta \equiv \sqrt{\frac{2}{\omega\mu\sigma}} \quad (5.122)$$

With this definition we have the following form for the electric field inside the material at depth $z < 0$

$$E_x(z) = E_0 e^{z/\delta} e^{jz/\delta} \quad (5.123)$$

We call δ the *penetration depth* parameter or more commonly the *skin depth*. The motivation for this is clear since δ has units of length (m) and represents the decay rate of the field into the conductor. This is illustrated in Fig. 5.23. Since $\mathbf{J} = \sigma \mathbf{E}$ it is easy to see that the current density decays in a similar manner to the electric field. It can be shown that the magnetic field also diminishes exponentially.

Table 5.1 *The conductivity and penetration depth of various conductors at room temperature*

Material	Conductivity (S/m)	Skin Depth (μ)		
		100 MHz	1 GHz	10 GHz
Copper	5.80×10^7	6.6	2.1	0.66
Aluminum	3.72×10^7	8.2	2.6	0.82
Gold	4.44×10^7	7.6	2.4	0.76
Silver	6.17×10^7	6.4	2.0	0.64
Brass	1.57×10^7	12.7	4.0	1.27

For static fields it is easy to show that $E \equiv 0$ inside perfect conductors. Now we see more generally that AC fields decay to zero in *good* conductors. The rate of decay is quite rapid since σ is large for good conductors. In Table 5.1 we quantify δ for typical materials and some typical frequencies. For instance, for Cu at 1 GHz, the skin depth is only 2.1μ . For this reason, δ is also called the *skin depth* since current flows along the outer skin of good conductors.

It is interesting to note that the phase term in Eq. (5.123) causes the electric field to flip direction inside the material at depth $z = \pi\delta$, thus causing the current to travel “backwards”! What causes this funny behavior? It is the magnetically induced currents. To observe this, note that time-varying conduction current creates a time-varying magnetic field which in turn induces a time-varying current opposing the original current. In the steady state, therefore, the current distribution takes on the form we have described.

It is reasonable to assume that the above solution can be applied to the *interior* of a *finite* rectangular conductor of width w and thickness t as long as the thickness is much greater than the skin depth. We do not expect the solution to apply to the edges of the rectangular conductor and in fact we shall later see that the field behavior at the edges is complicated. A round conductor of large radius (compared with the skin depth) should also have approximately the same field distribution, and, due to the absence of the edge discontinuity, the results should be even more applicable. In the next section we shall have occasion to make this comparison quantitative.

Since the current distribution is non-uniform, we employ Eq. (5.111) to obtain the internal impedance of the rectangular conductor. For the path of integration, we take the path along the surface at $z = 0$, where $\mathbf{E} = E_x \hat{x}$

$$Z_{int} = \frac{1}{I} \int_0^L E_x dx = \frac{E_x L}{I} \quad (5.124)$$

so the internal impedance per unit length, or the *surface impedance* is given by

$$Z_s = \frac{W}{L} Z_{int} = W \frac{E_x}{I} \quad (5.125)$$

The total current is given by integrating $\mathbf{J}$ along the depth of the conductor

$$I = W \int_{-\infty}^0 J(z) dz = W \sigma E_x \int_0^{\infty} e^{-(1+j)z/\delta} dz = W \delta \sigma E_x / (1+j) \quad (5.126)$$

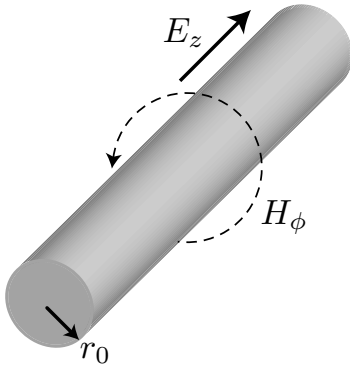


Figure 5.24 A round wire of infinite length has cylindrical symmetry.

Substituting in Eq. (5.125) we have

$$Z_s = \frac{(1+j)}{\sigma \delta} = (1+j)R_s \quad (5.127)$$

Does the above equation look familiar? The real part looks like the *sheet resistance* ($\Omega/\square$) of a material of conductivity σ and thickness δ . In other words, at frequency ω our rectangular slab appears the same as a slab at zero frequency with uniform current flowing through a thickness δ . Thus the resistance of the slab increases $\sqrt{\omega}$ as a function of frequency since $\delta \propto \sqrt{\omega}$. We can also observe from the imaginary part of Z_s that the *internal inductance* decreases at high frequency since the magnetic strength is reduced in the volume of the conductor.

Impedance of round wires

For a long round wire, the current J_z is invariant with z and the angle θ as shown in Fig. 5.24. Therefore, the Helmholtz equation simplifies

$$\nabla^2 \mathbf{J} = j\omega\mu\sigma \mathbf{J} = \tau^2 \mathbf{J} \quad (5.128)$$

$$\frac{\partial^2 J_z}{\partial r^2} + \frac{1}{r} \frac{\partial J_z}{\partial r} + \tau^2 J_z = 0 \quad (5.129)$$

Two linearity independent solutions are the Bessel function and the Hankel function of the first kind

$$J_z = AJ_0(\tau r) + BH_0^{(1)}(\tau r) \quad (5.130)$$

Since the Hankel function has a singularity at $r = 0$, it cannot be a solution. Normalizing to the current at the surface of the wire

$$J_z = \frac{\sigma E_0}{J_0(\tau r_0)} J_0(\tau r) \quad (5.131)$$

A plot of the current density in the wire is shown in Fig. 5.25. In the plot, a Cu wire with 1 mm diameter is used. Note that at low frequencies the current is essentially

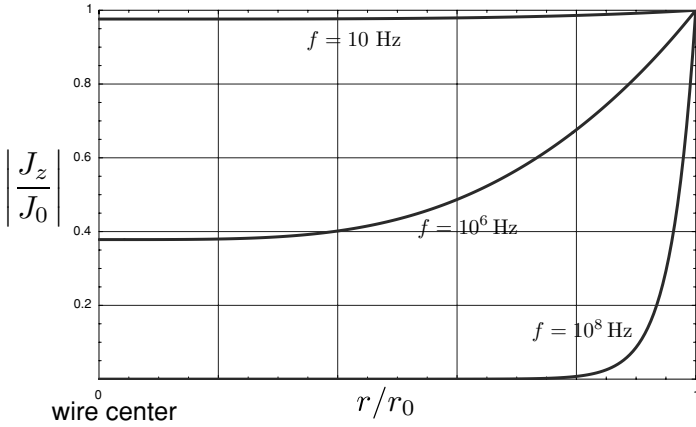


Figure 5.25 The current density inside of a round wire as a function of the radius. The origin represents the center of the wire. The current is uniform at low frequency but concentrated at the skin at high frequency.

uniform. At high frequency, though, the current decays exponentially as we penetrate the conductor.

In the limit that the radius is large, or equivalently $r_0/\delta \gg 1$, then the wire should behave like our plane conductor. In fact

$$\left| \frac{J_z}{\sigma E_0} \right| \approx e^{-(r_0-r)/\delta} \quad (5.132)$$

The impedance of a round wire can be computed by noting that only E_z and H_ϕ are present. Furthermore, we have

$$\oint \mathbf{H} \cdot d\mathbf{l} = I = 2\pi r_0 H_\phi \quad (5.133)$$

By $\nabla \times \mathbf{E} = -j\omega\mu\mathbf{H}$, it is easy to show that

$$H_\phi = \frac{1}{j\omega\mu} \frac{\partial E_z}{\partial r} \quad (5.134)$$

Using $E_z = J_z/\sigma = E_0 J_0(\tau r)/J_0(\tau r_0)$, the magnetic field is given by

$$H_\phi = \frac{E_0 \tau}{j\omega\mu} \frac{J'_0(\tau r)}{J_0(\tau r_0)} \quad (5.135)$$

Recall that $J'_0(x) = -J_1(x)$. Solving for the current

$$I = \frac{2\pi r_0 \sigma E_0}{\tau} \frac{J_1(\tau r_0)}{J_0(\tau r_0)} \quad (5.136)$$

Finally, we can write the internal impedance of the wire

$$Z_i = \frac{E_z(r_0)}{I} = \frac{\tau J_0(\tau r_0)}{2\pi r_0 \sigma J_1(\tau r_0)} \quad (5.137)$$

The above expression involves the Bessel function and its derivative of complex argument. While these functions have been amply tabulated, any modern computation environment such as *matlab* or *mathematica* can be used to evaluate the above expression.

In the low-frequency limit, the internal impedance of the wire reduces to

$$Z_i \approx \frac{1}{\pi r_0^2 \sigma} \left[1 + \frac{1}{48} \left(\frac{r_0}{\delta} \right)^2 \right] + j\omega \frac{\mu}{8\pi} \quad (5.138)$$

The real part corresponds to a correction to the DC resistance of the wire (per unit length). The imaginary term corresponds exactly to the static internal inductance of the wire. As expected, the high-frequency limit matches the analysis for the semi-infinite plane

$$Z_i = \frac{(1+j)R_s}{2\pi r_0} \quad (5.139)$$

This is because at high frequency the perimeter of the conductor is much larger than the penetration depth δ and thus the plane conductor analysis becomes valid.

Approximate impedance of rectangular wires

The resistance of a rectangular wire as a function of frequency can be obtained numerically and fitted to the following equation [25]

$$R = \frac{l}{\sigma w t} \left[\frac{0.43093x_w}{1 + 0.041 \left(\frac{w}{t} \right)^{1.19}} + \frac{1.1147 + 1.2868x_w}{1.2296 + 1.287x_w^3} + 0.0035 \left(\frac{w}{t} - 1 \right)^{1.8} \right] \quad (5.140)$$

where $x_w = \sqrt{2f\sigma\mu wt} \geq 2.5$ and for $x_w < 2.5$

$$R = \frac{l}{\sigma w t} \left[1 + 0.0122x_w^{3+0.01x_w^2} \right] \quad (5.141)$$

The above expression has an accuracy between 3% and 5% in the range of $w/t < 12$ and $x_w < 20$. It is interesting to compare the above result with the surface impedance calculation to see how much the edge effects alter the resistance.

Lossless inductors

The above discussion only applies to materials with finite conductivity. For lossless materials, the fields diminish in the volume of conductors and an infinitely dense current flows on the surface of conductors. The tangential electric field is zero inside and just outside of the conductors. The tangential magnetic field H_t is also zero inside of the volume of conductors but has a jump discontinuity equal to the surface current density J_s just outside the space of conductors

$$H_t = J_s \quad (5.142)$$

Thus the flux is independent of frequency and we do not expect the same variation in inductance due to the skin effect. On the other hand, the effective inductance will vary due to capacitive effects.

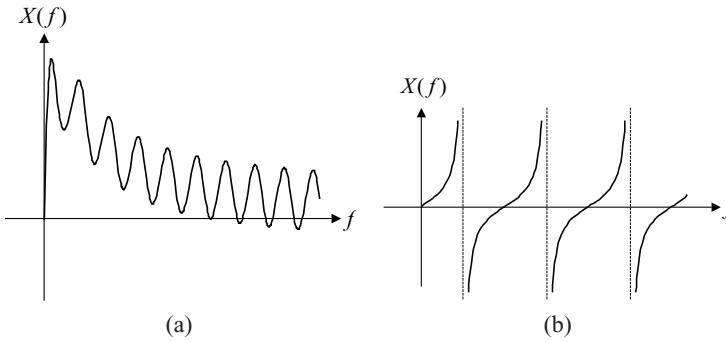


Figure 5.26 (a) The reactance of a lossy passive device as a function of frequency. (b) The reactance of a lossless passive device as a function of frequency.

Consider the complex power flowing into a black box shown in Fig. 5.9. If we integrate the Poynting vector $\mathbf{S} = \mathbf{E} \times \mathbf{H}^*$ across a surface S enclosing the black box

$$P = \frac{1}{2} \oint_S \mathbf{E} \times \mathbf{H}^* \cdot d\mathbf{S} = P_l + 2j\omega(W_m - W_e) \quad (5.143)$$

From electromagnetic theory [52] we have written the imaginary part of the above expression as the difference between the magnetic and electric stored energy. The reactance of the black box is thus related to the energy stored as follows

$$X = \frac{4\omega(W_m - W_e)}{II^*} \quad (5.144)$$

For an inductor $W_m > W_e$ and ideally $W_m \gg W_e$ so that the reactance of an inductor is positive. Furthermore it can be shown [52] that

$$\frac{dX}{d\omega} = \frac{4(W_m + W_e)}{II^*} \quad (5.145)$$

Please note that the above result is not a simple derivative of Eq. (5.144) (notice the change in sign). This shows that the derivative of the reactance is positive. This is a surprising result since it says that a lossless inductor cannot have a frequency-dependent profile such as shown in Fig. 5.26a, but it must resemble Fig. 5.26b. In other words, the reactance of an inductor is positive and *always* increasing as a function of frequency. The inductance as a function of frequency

$$L(\omega) = \frac{X}{\omega} \quad (5.146)$$

has slope

$$\frac{dL}{d\omega} = \frac{1}{\omega} \frac{dX}{d\omega} - \frac{X}{\omega^2} \quad (5.147)$$

Notice that every term in the above equation is positive for $\omega > 0$. From circuit theory, at low frequencies we expect

$$\frac{dX}{d\omega} = \frac{X}{\omega} \quad (5.148)$$

so that $dL/d\omega = 0$, implying a constant inductance. If the slope of the reactance is ever smaller than the reactance divided by frequency, the inductance will have a negative slope. But comparing Eq. (5.144) to Eq. (5.145), it is clear that $\frac{dX}{d\omega} \geq \frac{X}{\omega}$ so that

$$\frac{dL}{d\omega} \geq 0 \quad (5.149)$$

It is important to qualify the above result. First, we have assumed lossless materials in the construction of the inductor. Second, we have defined inductance by (5.146). Earlier we defined the inductance in terms of the magnetic energy alone. We should really call this the *effective* inductance since at high frequencies it will contain both magnetic and electrical energy. It is nevertheless interesting to notice that the effective inductance has a positive slope as a function of frequency somewhat counterintuitively. This result, though, is consistent with circuit theory. If we calculate the effective inductance of an LC tank circuit, we have

$$L_{eff} = \frac{L}{1 - LC\omega^2} \quad (5.150)$$

and for $\omega < \omega_0 \equiv 1/\sqrt{LC}$ the inductance is positive and increasing. As $\omega \rightarrow \omega_0$, the effective inductance takes on arbitrarily high values.

5.8 Quality factor of inductors

Using the fundamental definition of the Q factor as the ratio of energy storage to energy loss in a practical inductor over a cycle of excitation

$$Q \equiv \frac{\text{energy stored}}{\text{energy loss}} \quad (5.151)$$

For a physical inductor that has resistance R at frequency ω , in one cycle the inductor dissipates

$$E_{loss} = \frac{1}{2}|I|^2 RT = \frac{1}{2}|I|^2 R \frac{2\pi}{\omega} \quad (5.152)$$

Similarly, the energy stored in an inductor is

$$E_{stored} = \Im \left[\frac{1}{2} I V^* T \right] = \frac{1}{2} |I|^2 \omega L \frac{2\pi}{\omega} = \frac{1}{2} |I|^2 L 2\pi \quad (5.153)$$

and thus

$$Q = \frac{\omega L}{R} \quad (5.154)$$

The quality factor of inductors is important in many applications. For instance, a resonant tank exchanges energy between the magnetic and electric fields every cycle. If a fraction $1/Q$ of the stored energy is lost every cycle, then the system can sustain itself for about Q cycles. This determines, for instance, the power that must be replenished to the tank to sustain oscillations. While active devices such as transistors can be used to do this, these devices always add undesirable noise to the system.

To improve an inductor, therefore, we must minimize the various loss mechanisms associated with the inductor.

Inductor loss mechanisms

Conductive losses

At low frequencies, the resistance of an inductor is simply related to the DC resistance of the windings that make up the inductor. Since the resistance is an increasing function of temperature, the on-chip temperature should be estimated in determining the losses.

Internally induced losses

At high frequencies the effective resistance of an inductor is complicated by non-uniform current distribution in the interior of conductors. We have already seen that the *skin effect* tends to decrease the cross-sectional area of the device and thus we expect a resistance increase proportional to $\sqrt{f}$ due to this effect alone.

The physical asymmetry in the geometry of the structure will only exacerbate this. For instance, we observed that in a loop inductor the current tends to crowd to the inner radius of the conductor.

Externally induced losses

The proximity of nearby conductors also influences loss in an inductor. Due to the time-varying flux, *eddy currents* are induced to flow in conductors. As shown in Fig. 5.27, the effective resistance of these eddy current loops appears in series with the inductor. If we assume that the magnetic field of the induced currents interacts weakly with the magnetic field of the inductor, then we can neglect the change in inductance due to these currents and calculate the eddy currents directly as $j\omega\mathbf{A}$, where $\mathbf{A}$ is determined by the inductor currents alone. Otherwise, we must compute $\mathbf{A}$ by including the inductor current as well as the induced currents and this is a complicated situation.

Magnetization losses

Hitherto we have implicitly assumed that the material properties are linear, isotropic, and real. For most non-magnetic or weakly magnetic materials these are good assumptions. For magnetic materials, though, the magnetic permeability μ is non-linear and frequency dependent and dissipative. The exact physical cause of this is beyond the scope of this chapter

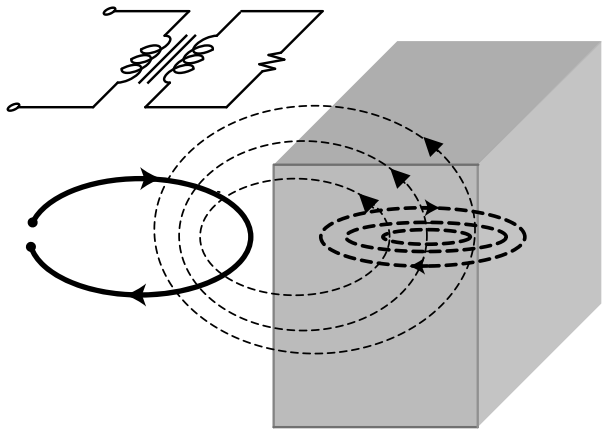


Figure 5.27 Model for eddy current losses as effective series losses.

and requires analysis of the quantum mechanics and thermodynamics of the constituent atoms. We can sweep such details under the rug and assume that $\mu = \mu' + j\mu''$ is a complex frequency-dependent quantity due to losses. The Q factor is thus bounded by μ'/μ'' due to the lossy part of the volume integral of $\mathbf{B} \cdot \mathbf{H}$ over all space.

Displacement current losses

Inevitably every real inductor stores some electric energy. We have up to now neglected the small displacement current $\mathbf{J}_d = j\omega\epsilon\mathbf{E}$ that flows in the volume of the problem space. If the material in question is lossy, then the permittivity has an imaginary component $\epsilon = \epsilon' - j\epsilon''$ as well. Physically, the electrically induced polarization in the material drains energy from the field and thus acts like a loss mechanism. The ratio of the imaginary part of the permittivity to the real part is known as the *loss tangent*

$$\tan \delta = \frac{\epsilon''}{\epsilon'} \tag{5.155}$$

and this is specified for various materials at various frequencies. We will defer a detailed discussion of the loss tangent. For now we will merely point out that the loss tangent acts as an additional loss mechanism when an inductor stores electric energy in addition to magnetic energy.

Radiative losses

For inductors that are large relative to the wavelength $\lambda = c/f$, additional losses can occur. For instance, consider breaking the inductor into many small sections and employing the partial inductance concept, the induced field at segment i due to current in segment j is simply $j\omega\mathbf{A}_j$. When segments i and j are in close proximity, the vector $\mathbf{A}$ is purely real and so the induced field only stores energy. For points far removed, though, there is a

Copyright © 2007, Cambridge University Press. All rights reserved.

propagation delay between these two points which adds an additional phase to $\mathbf{A}_j$ when evaluated at segment i . This additional phase delay causes the induced fields to have a *real* component, which appears as additional resistance. Physically, energy is lost to *radiation*, an electromagnetic wave which travels into space.

Overall quality factor due to multiple loss mechanisms

Consider the basic definition of the Q factor as the ratio of energy storage to energy loss. If we expand the various loss mechanisms we have

$$Q = \frac{W_{\text{stored}}}{W_{\text{loss}}} = \frac{W_{\text{stored}}}{W_{\text{loss},1} + W_{\text{loss},2} + \dots} \quad (5.156)$$

Taking the reciprocal of the above we have

$$\frac{1}{Q} = \frac{W_{\text{loss},1} + W_{\text{loss},2} + \dots}{W_{\text{stored}}} = \frac{1}{Q_1} + \frac{1}{Q_2} + \dots \quad (5.157)$$

And thus the overall Q can be computed as the “parallel” combination of the individual Q s. The overall Q factor is therefore limited by the smallest Q factor and is always less than this quantity. As expected, when one loss mechanism dominates, the Q factor is approximately determined by this loss term. For instance, for an inductor constructed over an insulating substrate (such as GaAs) the displacement current losses may be orders of magnitude smaller than the ohmic losses and can be safely neglected. This is in stark contrast to Si, where the conductive substrate contributes significantly to the loss at high frequency.

5.9 Inductors and switching circuits

Ideal inductors are DC shorts. With the absence of resistive loss, the voltage drop along the path of an inductor is equal to the time rate of change of flux, or

$$V_L(t) = \frac{d\psi}{dt} = L \frac{dI}{dt}$$

which is zero for static fields. More interesting, though, is that this implies that the average voltage across an inductor is almost always zero. To see this, consider the Laplace Transform of the voltage. The average value can be obtained by taking the limit as $s \rightarrow 0$. Assuming the Laplace Transform of the current does not have any zeros we have

$$V_L(s \rightarrow 0) = \lim_{s \rightarrow 0} I_L s L = 0 \quad (5.158)$$

A similar conclusion applies for period excitation if we replace $s = j\omega$. Thus the average voltage across an inductor is zero. Therefore, for most periodic waveforms across an inductor, the average value of the steady-state voltage inductor must be zero.

As an application of this concept, consider Fig. 5.28a where an inductor terminal is periodically shorted to ground by an ideal switch with a duty cycle of say 50%. What is the steady-state voltage at node v_1 as a function of time? For 50% of the period, the voltage v_1 is zero due to the ideal switch applying the full rail voltage across the inductor terminals.

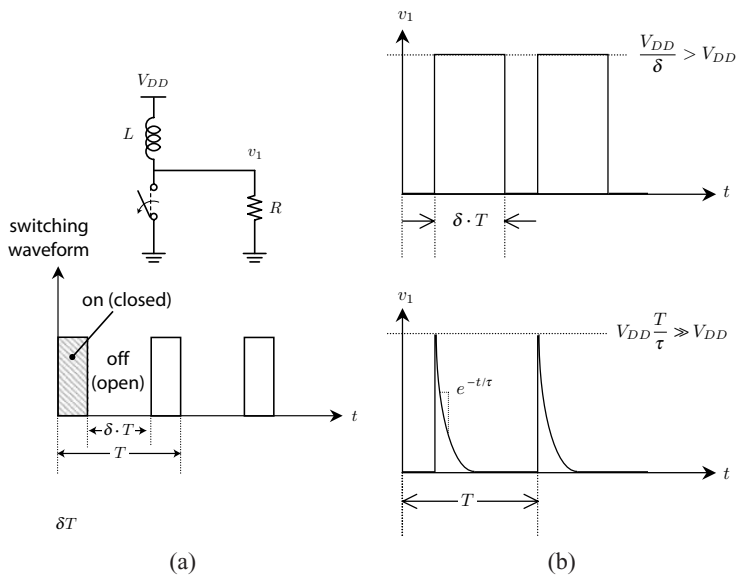


Figure 5.28 (a) Simple LR circuit switching circuit. (b) Approximate voltage waveforms at node v_1 for the case of $\tau \gg 1$ (top) and $\tau \ll 1$ (bottom) where $\tau = L/R$ is the characteristic time constant of the circuit.

In order to maintain an average value of zero, then the voltage v_1 must increase to at least a voltage of $2 \times V_{dd}$ for the remaining off cycle, a voltage above the supply! In fact, if we assume that the inductor is sufficiently large such that $\tau = L/R \gg 1$, then we will observe a steady-state voltage waveform similar to the top plot of Fig. 5.28b. In this case the energy stored in the inductor is sufficiently large that the energy drain during the off-period does not significantly reduce the magnetic energy of the inductor. Stated differently, the inductor is able to maintain a constant current.

On the other hand, if the inductor is small such that $\tau = L/R \ll 1$, then the magnetic energy in the inductor will exponentially drain and the steady-state voltage waveform will resemble the bottom graph of Fig. 5.28b.

Since the output voltage is larger than the supply voltage, this simple circuit principle can be used to build a voltage multiplier. In fact, as the duty cycle of the on period is increased, the voltage during the off-period has to take on correspondingly larger values to sustain a zero average voltage. By using feedback, we can generate a constant voltage by regulating the duty cycle of the waveform to maintain a constant DC voltage over the changing loads.

While at DC, inductors are shorts, at AC they in fact look like “bad” current sources. Just as capacitors at high frequency look like bad voltage sources, inductors have the dual behavior. A constant current drawn from an inductor produces no voltage drop. Inductors present increasingly more reactance to any attempts at changing the current. In many high-frequency circuits this behavior is exploited and large inductors, or RF chokes, are employed to deliver bias currents without disturbing the high-frequency behavior of circuits.

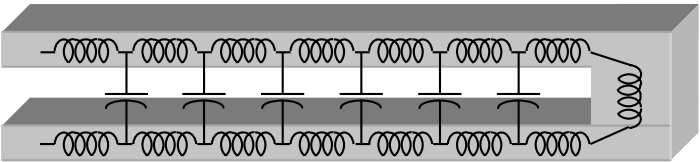


Figure 5.29 The capacitive coupling in a simple loop inductor.

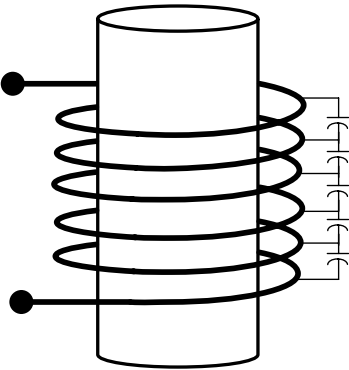


Figure 5.30 The capacitive coupling in a coil inductor.

5.10 Preview: how inductors mutate into capacitors

In Section 5.7 we observed that a structure appears inductive as long as the stored magnetic energy dominates over the electrically stored energy, $W_m > W_e$. Consider the loop inductor shown in Fig. 5.29. Schematically we have shown two paths between the terminals, a *conductive* path and a *displacement* path. The displacement path occurs by the time-varying electric field between the conductors. This electric field only exists if there is a net charge buildup on the two “plates” of the capacitors. It is important to observe that the actual displacement and conductive current paths are distributed and occur side by side. Thus our structure is really both an inductor and a capacitor. In Fig. 5.30 we have shown a common coil inductor and some representative capacitive elements that couple energy electrically from turn to turn of the inductor. While the coupling between adjacent turns is shown, every turn couples to every other turn in the coil.

Define the inductance per unit length as L' and the capacitance per unit length as C' . At low frequencies the loop presents a low impedance since $\omega L'$ is small, whereas $1/(\omega C')$ is large and thus the bulk of the current flows through the loop. Overall the behavior appears as inductive. At high frequencies, though, the inductive reactance is increasing, whereas the capacitive reactance is decreasing, so that there will be a frequency point when $1/(\omega C') = \omega L'$. At this point we may guess that the structure will no longer behave as an inductor and will in fact mutate into a capacitor. While the overall conclusion is correct, the frequency when this occurs is predicted incorrectly due to distributed effects.

Copyright © 2007. Cambridge University Press. All rights reserved.

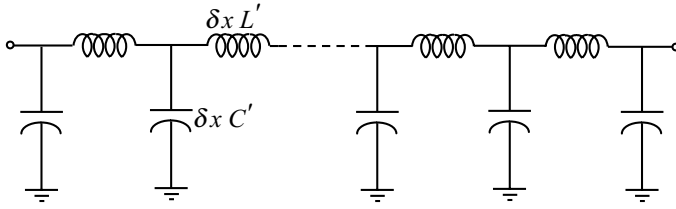


Figure 5.31 A distributed LC line.

But at what frequency does the structure *resonate*? At resonance the energy stored in the magnetic field is equal to the energy stored in the electric field. The answer to the question comes from *distributed* analysis of the structure. For instance, the loop inductor of Fig. 5.31 can be analyzed with transmission line theory (Chapter 9). The structure shown can be modeled as a shorted transmission line. Defining the inductance per unit length L' and the capacitance per unit length C' , the impedance of a shorted transmission line is

$$Z_s = j Z_0 \tan(\omega l/c) \tag{5.159}$$

where l is the length of the line, $c = \frac{1}{\sqrt{L'C'}}$ is the propagation velocity, and $Z_0 = \sqrt{\frac{L'}{C'}}$ is the characteristic impedance of the line. Thus the impedance of the line is inductive until $\omega l/c < \pi/2$. But the tangent function is periodic and thus there are an infinite number of frequency ranges where the above device appears inductive. In terms of the inductance and capacitance per unit length, the circuit resonates at frequency ω if

$$n \frac{\pi}{2} = \beta l = \omega \sqrt{L'C'} l = \omega \sqrt{LC} \tag{5.160}$$

for n an integer. Solving for ω

$$\omega = n \frac{\pi}{2} \sqrt{\frac{1}{LC}} \tag{5.161}$$

The above result contradicts circuit theory. Circuit theory predicts a single resonant frequency at a frequency $2/\pi$ lower than distributed analysis. This is not surprising since circuit theory neglects any spacial variation in the fields. We shall return to this subject in Chapter 11. Since any physical inductor stores some parasitic capacitive energy, no matter how small, there is a frequency limit to the inductive behavior.

5.11 References

The ideas for this chapter come from working in the field for many years and so it is hard to trace the material back to the source. I know many of the ideas were inspired by Feynman’s Lectures [18], and a careful reading of *Fields and Waves in Communication Electronics* [52]. In fact, the idea of writing this book is partly a result of reading Chapter 4 of this book and wanting to turn the chapter into an entire book.

Copyright © 2007, Cambridge University Press. All rights reserved.

6 Passive device design and layout

In this chapter we focus on the design and layout of passive devices. In particular we will emphasize the realization of inductors and capacitors in an integrated circuit process. We begin the discussion with a simple loop inductor. Since the loop is a core building block for more complicated structures, we will spend more time on this structure. By using loops in series and parallel, we can build spirals, solenoids, and other more complex structures. The design and layout of transformers for ICs builds on this background and is covered in detail in Chapter 10. Finally, we cover the design of MIM capacitors in an IC process.

6.1 Ring inductor

A loop or ring inductor is a natural inductor structure because we can cut the loop at any point to introduce the leads into the inductor, as shown in Fig. 6.1. Since the structure is planar, it can be easily incorporated onto a PCB or integrated circuit. A circular loop is approximated in an IC process with a polygon.¹ In an IC process allowing only “Manhattan” geometries, or orthogonal interconnect angles, then a rectangular loop inductor is a good choice. Intuitively, the inductance is a function of the loop area A since by the mean value theorem

$$\psi = \int_A \mathbf{B} \cdot d\mathbf{A} = B_{av} \cdot A$$

The resistance, though, depends on the perimeter of the loop. Thus, to maximize the Q of the structure, we should utilize a shape with maximum area to perimeter. It is well known that such a shape is indeed a circle. But a rectangular loop has only slightly smaller Q factors and is more convenient for layout.

The external inductance of a loop is computed by assuming a filamental current flows at the center of the loop, while computing the flux crossing the inner area enclosed by the loop.

The self-resonant frequency of the coil inductor can be very high, since there is little capacitive coupling in the structure. The lead capacitance will determine the self-resonant frequency for an isolated ring in free space, as shown in Fig. 6.2a. Unfortunately, if the

¹ A polygon with 64 sides is practically a circle. Some IC fabrications only allow 45° angles limiting us to an octagon. See Fig. 6.12.

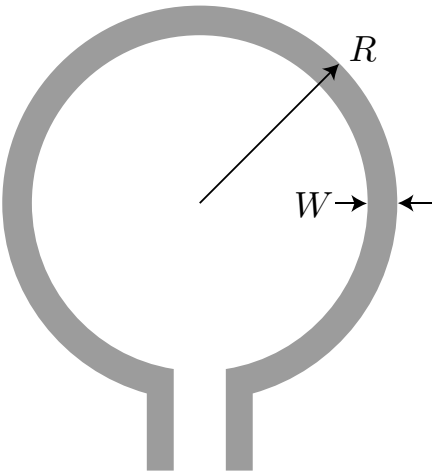


Figure 6.1 A single loop ring inductor structure.

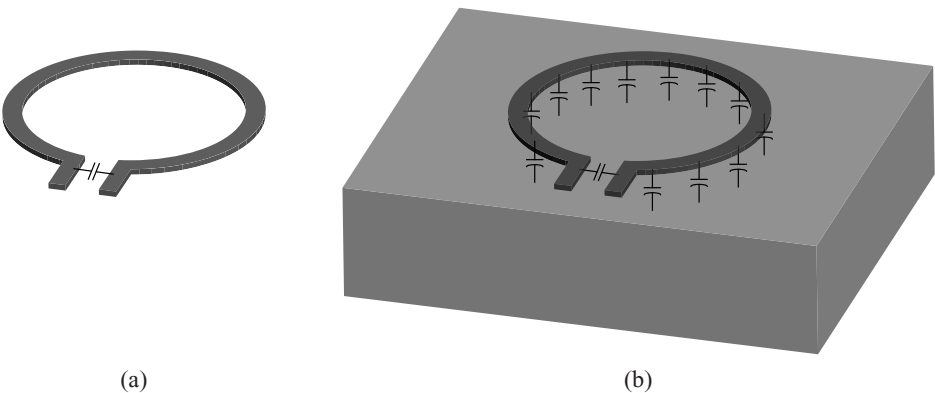


Figure 6.2 (a) In an isolated ring the self-capitance, especially around the leads, sets the self-resonant frequency. (b) When the ring is fabricated on top of a conductive substrate, the distributed capacitance to the ground dominates the self-resonant frequency.

structure is implemented on-chip in Si, there is significant capacitance to the substrate due to the close proximity of the conductive substrate, as shown in Fig. 6.2b. This capacitance can be estimated by the parallel plate portion (proportional to area) and the fringing component (proportional to perimeter).

Skin effect

The resistance of a loop can be lowered by employing wider conductor traces in the loop. Naturally, though, at high frequencies the AC currents take the path of least impedance, as

Copyright © 2007. Cambridge University Press. All rights reserved.



Figure 6.3 Current constriction in a square ring inductor.

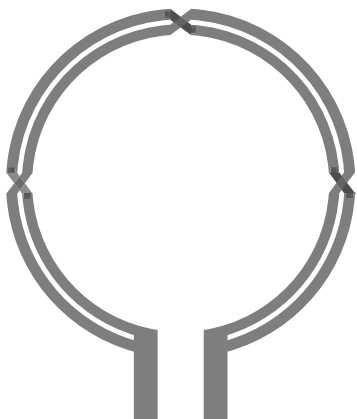


Figure 6.4 A loop inductor employing two symmetric traces in parallel to minimize current constriction.

illustrated in the current density plot of Fig. 6.3. The current flow is non-uniform and tends to flow along the inner side, or lower inductance path. This is easily modeled as two loops in parallel. Since the loop inductance of the inner loop is smaller, more current will take this path. As we split the loop into more and more coupled loops, we can in fact predict this current distribution more accurately.

One way to mitigate this non-uniform current flow is illustrated in Fig. 6.4, where the trace is split into two parallel paths. Each path is balanced as it spends an equal amount of time on the inside and outside of the loop. Thus the current should split uniformly between the two paths, leading to less current crowding in the structure.

Copyright © 2007, Cambridge University Press. All rights reserved.

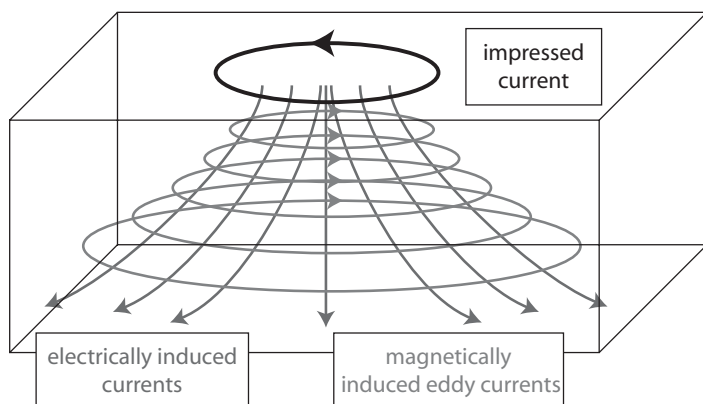


Figure 6.5 The electrically induced current flow in a typical inductor fabricated over a conductive substrate. The magnetically induced eddy currents flow in loops.

Substrate losses

Since the substrate is conductive, the substrate coupling in an inductor introduces loss. The substrate resistance depends strongly on the placement of substrate contacts, the ground connection at the substrate back plane, and the frequency of operation. At low frequencies the current spreads and can reach the back side. At high frequencies, the nearby substrate contacts dominate as current sinks. The flow of capacitive substrate current is shown in Fig. 6.5. It is important to distinguish between the capacitive substrate currents, arising from displacement current, from the substrate-induced eddy currents, flowing due to changing magnetic fields. Since the time-varying magnetic field penetrates the substrate, as shown in Fig. 6.5, the induced emf in the substrate leads to current flow. This current flow causes additional power loss to occur and appears as an effective increase in the series resistance of the ring inductor. For heavily conductive substrates, with $\rho < 0.01\Omega\text{-cm}$, this substrate loss can be a significant fraction of the overall losses. For moderately conductive substrate, though, $\rho > 1\Omega\text{-cm}$, the induced eddy currents are small and can be neglected at moderately high frequencies ($L < 20\text{ nH}$, $f < 10\text{ GHz}$).

Observe that there is a worst-case value of substrate resistance that results in the optimal power transfer from the ring to the substrate. Such a value is simply equal to the capacitor reactance (by the maximum power transfer theorem). The only way to eliminate capacitive substrate loss is to employ a substrate with resistivity of zero or infinity. An infinitely resistive substrate, or an insulating substrate, is ideal, but not widely available in Si IC processes. An infinitely conductive substrate, on the other hand, is undesirable since the induced eddy currents would render very low (practically zero) inductance from the ring. This is because the eddy currents generate a magnetic flux that opposes the inductor flux. So in practice, with finite non-zero substrate resistivity, we should strive to place our substrate contacts in such a way to move away from the “optimal” substrate resistance value.

One such approach is to attempt to move the substrate resistance to zero by employing a shield. The shield cannot be solid since the eddy currents would again reduce the inductance value, but a patterned shield can minimize the effect of the eddy currents without disturbing



Figure 6.6 (a) A patterned shield is employed to prevent electrical fields from penetrating the substrate while preventing eddy current (closed loops) flow over a large area. (b) A co-planar ground shield improves the quality factor of the ring inductor by intercepting a fraction of the fringing fields from reaching the lossy ground.

the displacement current flow, as shown in Fig. 6.6a. The shield can be constructed with a polysilicon or metal layer and patterned in such a way as to minimize any “loops” in the shield.

The main drawback of the shield is that it increases the substrate capacitance of the structure, leading to a lower self-resonant frequency. Furthermore, it cannot shield the devices from magnetically induced substrate losses. On the other hand, the shield can potentially reduce the amount of energy that is injected into or captured from the substrate, leading to better isolation. This will be discussed in more depth in Chapter 13.

A compromise to the shield structure is the “co-planar” ground shown in Fig. 6.6b. Here a ground plane surrounds the shield and captures most of the fringing fields from reaching the substrate. Since the ground does not form a closed loop around the ring, though, the ground plane does not form a transmission line with the inductor. In fact, if such a transmission line were formed, our ring would turn into a short section of a co-planar transmission and the magnetic flux would shift from the inside of the loop to the outside of the loop, between the current and return current. This would lower the inductance but lead to improved magnetic isolation, since the “monopole” currents of the loop are turned into dipole currents. This technique is very effective at high frequencies, where the desired inductance value allows the application of transmission lines.

6.2 The classic coil

The main limitation of a ring inductor is the achievable value of inductance. To realize higher inductance for a fixed area, we need to add more loops to the structure. By adding more loops in parallel, we can boost the value of inductance. This is accomplished by a spiral staircase style structure of a solenoid inductor, shown in Fig. 6.7. If the turns are wound in very close proximity, then each loop of the structure has a magnetic flux N times

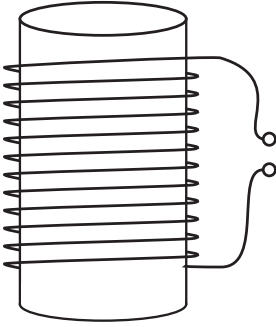


Figure 6.7 A solenoid “coil” inductor structure.

larger than a single loop. Since we traverse N loops, we expect to see an N^2 increase in the inductance. The winding resistance, on the other hand, will only increase like N . So overall the Q is boosted by N .

We can make this simple argument more quantitative as follows. A tightly wound very long solenoid has $B = 0$ outside and $B_x = 0$ inside, so that by Ampère’s Law

$$B_y \ell = NI\mu_0 \quad (6.1)$$

where N is the number of current loops crossing the surface of the path. The vertical magnetic field is therefore constant

$$B_y = \frac{NI\mu_0}{\ell} = \mu_0 I n \quad (6.2)$$

The flux per turn is therefore simply given by

$$\Psi_{\text{turn}} = \pi a^2 B_y \quad (6.3)$$

The total flux through N turns is thus

$$\Psi = N\Psi_{\text{turn}} = N\pi a^2 B_y \quad (6.4)$$

$$\Psi = \mu_0 \frac{N^2 \pi a^2}{\ell} I \quad (6.5)$$

The solenoid inductance is thus

$$L = \frac{\Psi}{I} = \frac{\mu_0 N^2 \pi a^2}{\ell} \quad (6.6)$$

The above analysis is very simple and there are many more detailed and sophisticated calculations for the inductance of a coil. The interested reader is encouraged to consult some of the classic references [60] [22].

Rectangular coils

To implement a coil in an IC process, we can create a rectangular coil as shown in Fig. 6.8. Such a structure suffers from a very small cross-sectional area due to the small thickness of

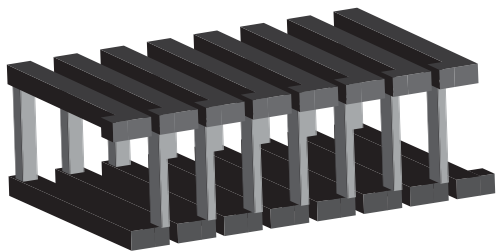


Figure 6.8 A rectangular coil inductor fabricated in a standard IC process. The dimensions of the vias have been exaggerated for illustration purposes.

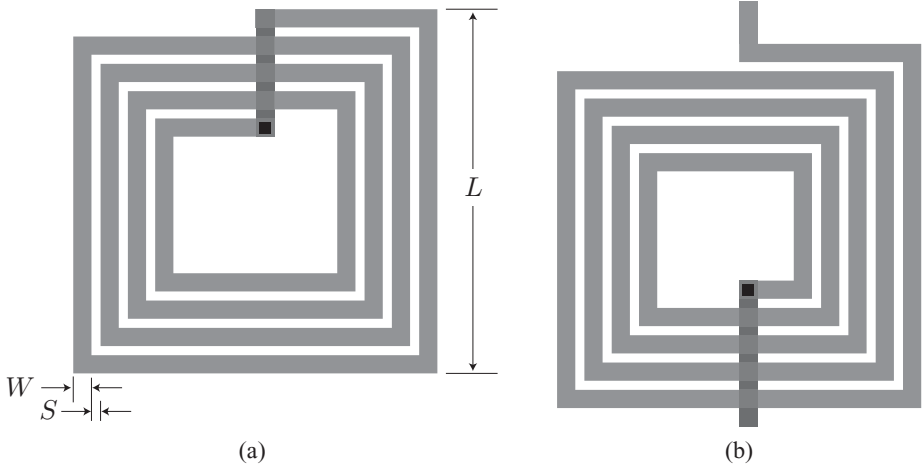


Figure 6.9 Square spiral inductor layout with (a) joined terminals, and (b) separated terminals.

the oxide layers. One advantage of such a structure, though, is that the core magnetic field lines are parallel to the substrate. Only the fringing fields penetrate the substrate and thus the eddy currents generated in the substrate are much smaller. More elaborate structures can be constructed with post processing. MEMS post processing has been demonstrated by several researchers [33] [67], allowing us to integrate high-quality inductors into an integrated circuit environment.

6.3 Spirals

Square spiral inductors are the most popular planar inductor geometries. Spiral inductors are wound on the top metal layer and a second layer is used to connect the innermost winding to the outside. As shown in Fig. 6.9, the inner turn can lead to several different points, depending on the application of the inductor. The layout of Fig. 6.9a forms a closed loop and is practical when both terminals of the inductor are accessed (say to a parallel capacitor

Copyright © 2007. Cambridge University Press. All rights reserved.

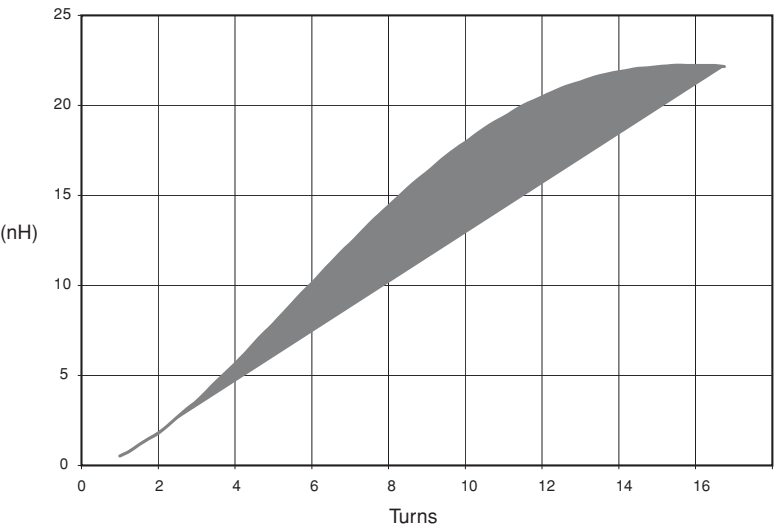


Figure 6.10 Variation of spiral inductance as a function of the number of turns N .

to form a resonator). On the other hand, if the spiral is in series with another element, the layout of Fig. 6.9b is preferred. It is important to note, however, that this layout is a “partial” inductor and the total inductance will depend on the loop in which the spiral is placed.

Due to the planar constraint, instead of adding a loop in the third dimension to spiral up (as in a solenoid), we spiral inward by adding loops inside each successive loop to form a spiral. Since the radius of each spiral is decreasing, the effective inductance does not increase like N^2 , but like N^p with $p < 2$. A typical variation of inductance with turns is shown in Fig. 6.10. As long as $p > 1$, then the quality factor increases with increasing turns. In practice, though, the inner turns of a spiral suffer from increased losses due to enhanced skin/proximity effects. For this reason, to realize the highest Q inductor, it is better to use a single turn inductor with larger area. The advantage of a spiral over a ring inductor is that much higher inductance values can be generated in the same area.

Very similar comments with regards to the substrate losses of a loop inductor also apply to the spiral inductor, since it is essentially a series connection of loops. There are additional capacitive and loss currents, owing to the close proximity of the turns. For most IC structures, though, the substrate capacitance tends to dominate over the distributed winding capacitance.

The design of an optimal spiral is highly frequency and process dependent. This is because of the multiple loss mechanisms that arise from the distributed effects in the structure. In general, for a fixed area, we can employ many different values of trace width W , spacing S , and turns N , to achieve the same value of inductance. As we increase W for instance, the resistance drops. This drop is proportional to $1/W$ at low frequencies but much more gradual at higher frequencies, due to the skin effect, especially with $W \gg \delta$, where δ is

Copyright © 2007. Cambridge University Press. All rights reserved.

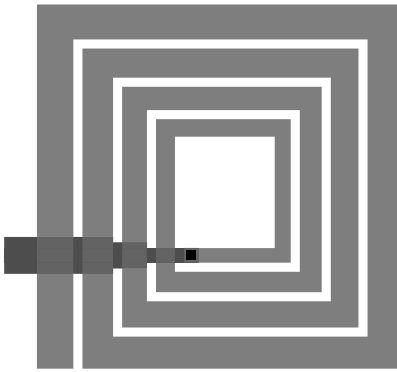


Figure 6.11 A square spiral inductor with tapered trace width. Inner turns do not need to be as wide due to enhanced current constriction.

the effective skin “width” (not depth). The substrate losses, though, tend to increase with W , since this increases the capacitance of the structure. At the same time, we observe that structures with more turns, N , tend to have higher resistive losses in the inner turns. For this reason “hollow” spirals are preferred [28]. The origin of these enhanced losses is due to the fact that the magnetic flux increases as we move towards the center of the spiral, due to the additive nature of the flux from each successive loop of the spiral. If the width of the structure is tapered, as shown in Fig. 6.11, then the performance of the spiral can be improved. Since the wide inner turns do not lower the resistance (due to current constriction), it is better to transfer the width to the outer turns, while keeping the total area of the spiral constant.

Finally, for low values of S , which can be extremely small in a modern IC process, we would expect excessive winding capacitance. For large values of S we find that the reduced proximity effects lead to more uniform current distributions, similar to an isolated conductor. But higher values of S lead to lower inductance, due to the reduction in coupling from turn-to-turn, and due to the smaller area of inner loops. There is thus an optimal value of S that allows a fair balance to occur.

Other variations on the layout include a rectangular layout, as opposed to a square outer area. A rectangular layout has lower Q than a square layout, but other constraints may impose a rectangular “opening” over which the spiral must be wound. Circular and polygon structures, shown in Fig. 6.12, have marginally higher Q factors and are preferred for this reason. The reason for the higher-quality factor follows the same line of argument for the ring inductor given previously.

6.4 Symmetric inductors

The main limitation of a spiral inductor is the asymmetry between the outer and inner lead of the structure when $N > 1$. This is evident in the Y parameters of the device which

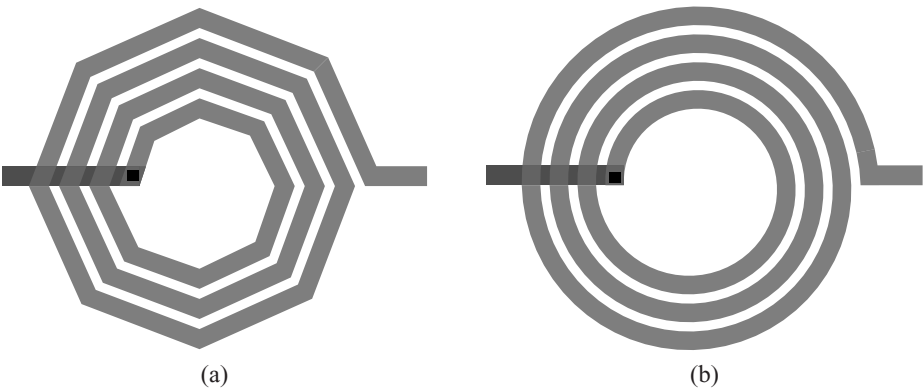


Figure 6.12 (a) Polygon spiral inductor layout with eight sides. This polygon has performance very similar to a (b) 64 sided polygon, which very accurately approaches the performance of a circular spiral.

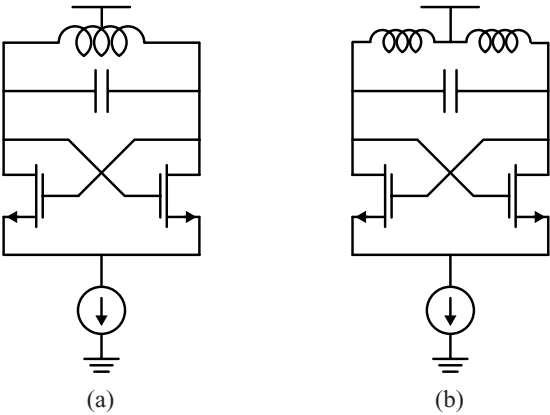


Figure 6.13 (a) A differential oscillator requires a perfectly symmetric inductor. The center “tap” of this inductor is used to deliver the bias voltage and current. (b) With a conventional spiral, two spirals must be connected in series for a fully symmetric layout.

shows that $Y_{11} \neq Y_{22}$. Since the outer-most turn interfaces to an open area, it tends to have more fringing capacitance. The inner turns, though, fringe much less, since the spiral turns are at approximately the same potential, and thus the fields have more “even” mode than “odd” mode. This asymmetry in the inductor is undesirable for fully balanced differential circuits. Furthermore, the electrical “center” of the spiral is hard to define. This center point is important to identify because if the structure is excited with a fully balanced signal, this point will not move electrically thus forming a virtual ground. We can thus use this point to connect a bias voltage without incurring any change into the operation of the inductor. In a differential oscillator, shown in Fig. 6.13a, for example, we can use this point to deliver the V_{DD} or ground connection. One symmetric layout is shown in Fig. 6.14a. Here the geometric center of the structure is also the electrical center, and thus this point can be used

Copyright © 2007. Cambridge University Press. All rights reserved.

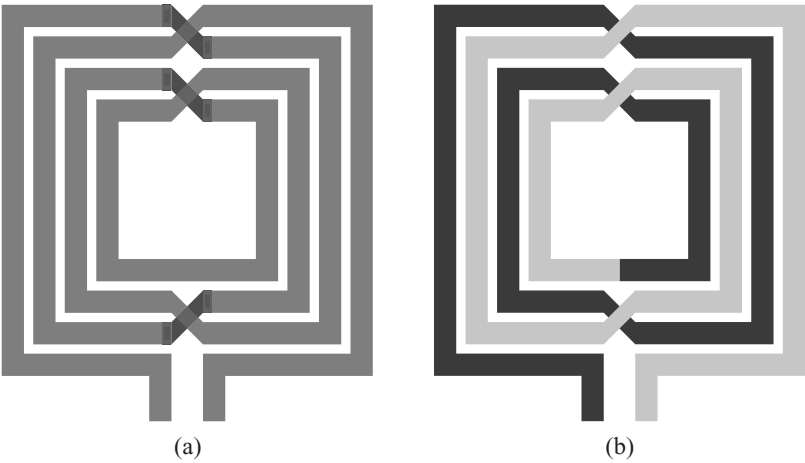


Figure 6.14 (a) A more symmetric inductor layout can be (b) decomposed into two coupled inductors.

as a “center tap” for the inductor. As shown in Fig. 6.14b, we can decompose this structure into two closely wound coupled inductors connected in series. In practice the coupling factor $k \sim 0.7$.

A center tapped inductor has a big advantage over using two series connected inductors shown in Fig. 6.13b. Since two separate inductors are only weakly coupled, the total inductance is simply $L + L + 2M \approx 2L$. If we wind the inductors together, on the other hand, a coupling factor of $k \sim 0.7$ can be utilized and the inductance is boosted to $2L(1 + k)$. Thus for equal inductance values, the area of the inductors can be substantially reduced leading to a higher-quality factor. In fact, at higher frequencies when the quality factor is dominated by the substrate, the two interwound inductors can have substantially higher Q values.

6.5 Multilayer inductors

Most modern IC processes offer many metal layers that can be employed to design more complicated three-dimensional structures. For instance, two or more spirals on different layers can be placed in series to increase the inductance, as shown in Fig. 6.15 and in three dimensions in Fig. 6.16. Note that the spirals must be wound correctly in order for the magnetic flux to be additive. Thus as the spiral winds inward on top, it should wind outward on the second layer, and then wind inward again on the third layer. For three equal inductance spirals, the inductance is increased to

$$L_t = L + M_2 + M_3 + L + M_2 + M_2 + L + M_2 + M_3 = 3L + 4M_2 + 2M_3 \quad (6.7)$$

where M_2 is the mutual inductance between adjacent layers and M_3 is the mutual inductance between two spirals separated by another spiral in between. Since the coupling factors $k_2 \sim 0.8$ and $k_3 \sim 0.6$ are significant (due to the small separation of the layers), we can approach the upper bound of $L_t \leq 9L$, a N^2 boost in inductance value.

Copyright © 2007. Cambridge University Press. All rights reserved.

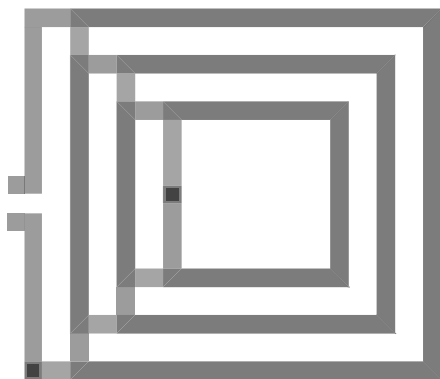


Figure 6.15 A multi-layer square spiral inductor is formed by connecting a spiral from the top layer to a spiral on the next level. Each spiral is wound so as to reinforce the magnetic field of the other. This is easily generalized to three or more layers.

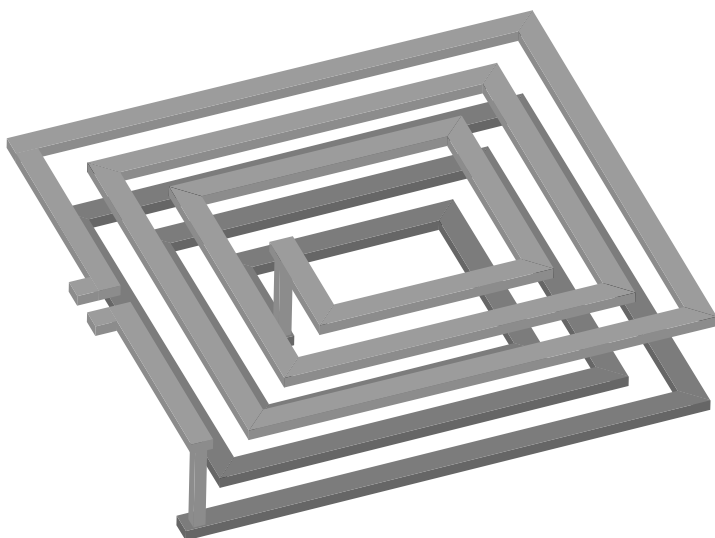


Figure 6.16 A three-dimensional view of the multi-layer series connected spiral shown in Fig. 6.15. The via thickness is exaggerated for illustration.

Even though very large and compact inductors can be realized with this approach, the self-resonant frequency suffers from the large spiral-to-spiral winding capacitance. The successive layers can be staggered somewhat to reduce the capacitance, but, overall, while these inductors can save area over a single layer structure, due to the closer proximity to the substrate, the quality factor of such structures suffers.

Another more obvious application of multiple layers is to simply connect the spirals in parallel, in order to realize an effectively thicker conductor. Many vias are employed to strap the layers together. This will indeed lower the resistance with a small impact on the

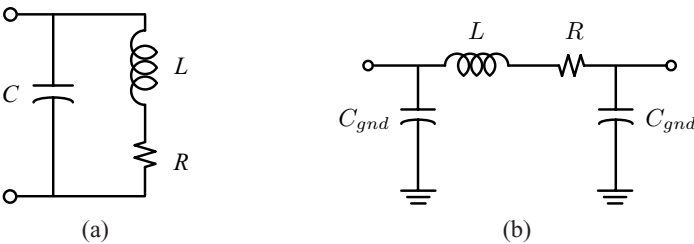


Figure 6.17 (a) A simple air-core inductor is modeled by a simple LCR network. (b) An inductor near a ground plane has a lower self-resonant frequency due to the extra capacitance.

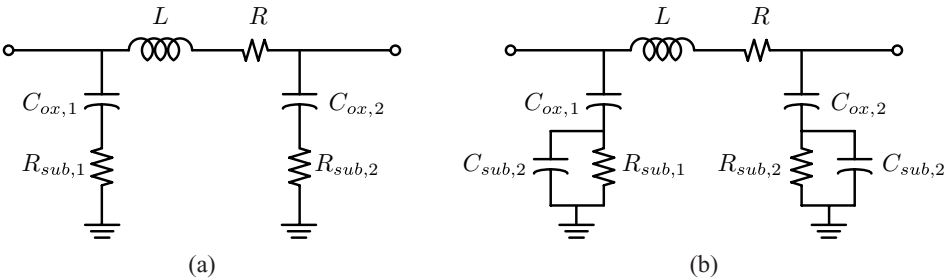


Figure 6.18 (a) An inductor fabricated over a lossy substrate can be modeled in a narrow band of frequencies with a Π model. (b) Additional capacitors can model displacement current in the substrate.

inductance value (since $k \sim 0.7$), but the lower layers are closer to the substrate and thus the improvement in the lower series resistance is eventually outweighed by the increased shunt losses through the substrate.

6.6 Inductor equivalent circuit models

The behavior of inductors can be modeled in practice with equivalent circuits consisting of magnetic and electric storage elements as well as loss elements. For instance, a simple air wound inductor is modeled as shown in Fig. 6.17a, where the series winding resistance is modeled by R and the winding capacitance is modeled by C . Even though the actual capacitance is a distributed across the windings, in practice the lumped value of C is adjusted to capture the first resonant frequency of the structure. The frequency dependence in the resistance R is simple to model in AC simulations, but difficult to model for time-domain simulations. For structures fabricated over an ideal ground plane, the capacitance to substrate can be modeled as shown in Fig. 6.17b. This is a simple model that can capture the behavior up to the first resonant frequency.

If the substrate is lossy, resistors in series with the capacitors can model the electrical substrate loss, as shown in Fig. 6.18a. In fact, it is easy to show that such a Π model is a one-to-one equivalent circuit for a two-port passive network

$$R_x + j\omega L = -\frac{1}{Y_{12}} \tag{6.8}$$

Copyright © 2007. Cambridge University Press. All rights reserved.

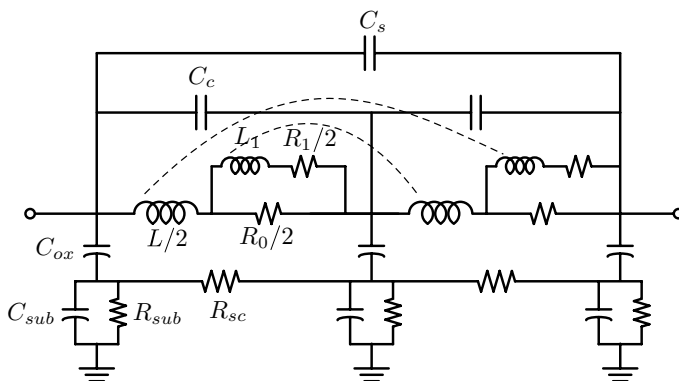


Figure 6.19 A multi-section model can be used to fit the inductor behavior over a wide frequency range.

and

$$(R_x + j\omega L) \parallel \left(R_{s1} + \frac{1}{j\omega C_{s1}} \right) = \frac{1}{Y_{11}} \quad (6.9)$$

$$(R_x + j\omega L) \parallel \left(R_{s2} + \frac{1}{j\omega C_{s2}} \right) = \frac{1}{Y_{22}} \quad (6.10)$$

It should be noted that in practice the values of the equivalent circuit elements are derived from optimization rather than from physical calculations. The physical values of the elements can be used as a starting point in the optimization. But slight adjustments in these component values can be used to obtain a better fit over a wide frequency range. For a broader band model, multiple Π sections can be cascaded. In the limit, the cascaded Π sections form an artificial transmission line that closely mimics the behavior of the inductor.

For moderately conductive substrates, the displacement current $\propto \omega \epsilon$ may be of the order of the conductive currents $\propto \sigma$. In this case, the model of Fig. 6.18b can be used for more accuracy. To model the frequency dependence of the series resistance due to skin and proximity effects, the model of Fig. 6.19 can be used [3]. Since at low frequencies the inductor windings are short circuits, the DC resistance of the structure is simply $R_0 \parallel R_1$. At high frequency, though, the resistance asymptotically approaches R_0 . Thus the series resistance of the model increases as a function of frequency. An extra degree of freedom is added by adding mutual inductance between the winding inductor and the skin effect modeling inductor.

6.7 Integrated capacitors

High-quality capacitors are critical in RF circuits. As shown in the schematic of Fig. 6.20, these capacitors are employed in matching networks, as AC coupling capacitors (to isolate the bias of an amplifier from the driver), and as bypass capacitors to provide low inductance

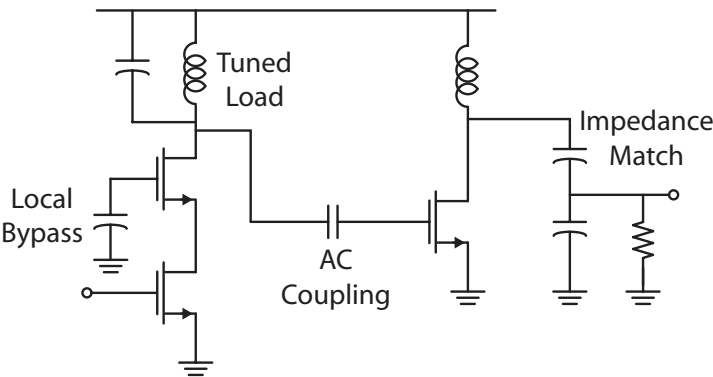


Figure 6.20 Integrated capacitors in a typical RF amplifier.

return to ground. While some specialized IC processes offer “MIM” (or metal-insulator-metal) capacitors as a special high-density capacitor device, many processes do not have this option. Otherwise, this option adds extra expense due to the required masking layers. But due to the high density of metalization in a modern IC, we can construct capacitors employing the standard metalization layers.

Metal-insulator-metal (MIM) capacitors

If a MIM capacitor is offered in a process, it is usually constructed from a parallel plate capacitor employing a thin oxide with a high dielectric constant. While the standard metals can be used to form a MIM capacitor, the capacitor density of such structures is quite low. For instance, if the metals are spaced at $1\ \mu\text{m}$ apart, then capacitor density is only

$$C_{ox} = \frac{\epsilon_{ox}}{t} \approx \frac{3.9 \times 8.854 \times 10^{-12} \text{ F/m}}{1 \times 10^{-6} \text{ m}} = 35 \text{ aF}/\mu^2 \quad (6.11)$$

To lay out a $1\ \text{pF}$ capacitor requires an area of approximately $170\ \mu \times 170\ \mu$. Alternatively, if we exploit the small lateral spacing of the process, we can realize much higher capacitor densities. For instance, a “finger” capacitor structure, shown in Fig. 6.21, utilizes the lateral fringing capacitance between two fingers. The spacing between two conductors is now of the order of $100\ \text{nm}$ in a modern CMOS process. Since multiple layers are available, the capacitance increases with the number of layers without increasing the area. Moreover, a dense array of vias can create a virtual “wall,” which boosts the capacitance greatly. For simplicity, assume that the process has six metal layers that can be routed with minimum spacing of $100\ \text{nm}$. Assume that vias require at least $200\ \text{nm}$ of metal width. Thus each finger requires $300\ \text{nm}$ of length, with the average spacing between fingers about $D_{eff} = 300\ \text{nm}$. The capacitance per unit length between each finger is given by

$$C_f/L = T_{eff} \times \frac{\epsilon_{ox}}{D_{eff}} \approx 6.5\ \mu \cdot \frac{3.9 \times 8.854 \times 10^{-12} \text{ F/m}}{0.3 \times 10^{-6} \text{ m}} = 0.75 \text{ fF}/\mu \quad (6.12)$$

Copyright © 2007. Cambridge University Press. All rights reserved.

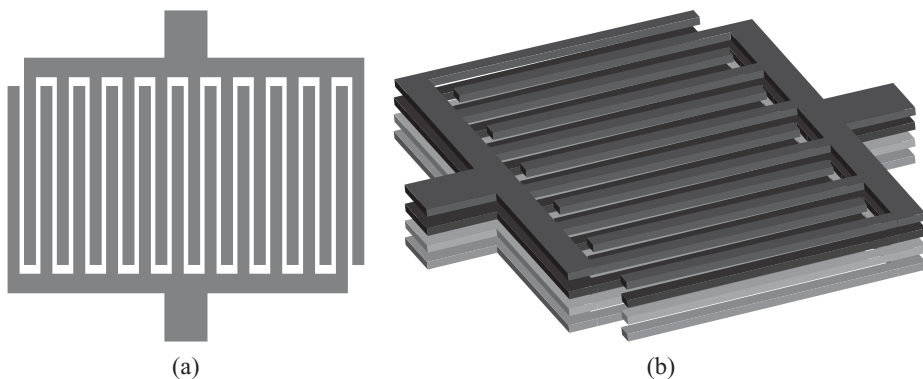


Figure 6.21 (a) Integrated “finger” capacitors exploit the high metal density to realize lateral capacitor structures. (b) Three-dimensional view of multi-layer finger capacitor structure.

where we assume that all six layers are employed, and each layer is $0.25\ \mu$ thick and spaced at $1\ \mu$ apart, giving $T_{eff} = 6.5\ \mu$. Thus the capacitance density is approximately

$$C_f/(W \cdot L) = 2.2\ \text{fF}/\mu/(0.2\ \mu + 0.1\ \mu) = 7.5\ \text{fF}/\mu^2 \quad (6.13)$$

With such a lateral capacitor structure, the 1 pF capacitor can now be realized in an area of only $67\ \mu \times 67\ \mu$.

Capacitor Q

This high density allows relatively compact capacitors to be integrated into the RF chip. The finger length L must be chosen carefully to avoid de- Q 'ing the capacitor with series resistance. Since the conductors have a resistivity $\rho \sim 100\ \text{m}\Omega/\square$, one finger has a resistance per length of $0.5\ \Omega/\mu$. But due to the distributed nature of the current flow (see Chapter 11), the actual resistance is only 1/3 of the DC resistance. Intuitively this occurs since the full current does not flow through the entire finger. In fact, the current drops to zero by the time we reach the end of the finger.

Many short fingers in parallel must be employed to retain a high-quality factor. But in general we find it much easier to design a high-quality factor capacitor than a high Q inductor. To see this, consider a capacitor at 1 GHz. If we use $100\ \mu$ long fingers, the resistance per finger is $0.5\ \Omega \times 100\ \mu \times 2 \times \frac{1}{3}$ and the capacitance per finger is 75 fF, corresponding to a quality factor of nearly 64. As more fingers are added in parallel to realize a higher capacitance value, the resistance drops and the Q remains the same. So the Q factor is to first order only dependent on the length of the fingers.

The quality factor also depends on the quality of the dielectric layers. Native SiO_2 is an excellent dielectric with a very small loss tangent. Values as low as 1×10^{-6} have been reported. With such a small loss tangent, the series resistance would continue to dominate the quality factor of integrated capacitors. To see this note that the losses due to the parallel

plate capacitance are given by

$$P_{Ld} = \frac{1}{4} \int_V |E|^2 \omega \epsilon'' dV \quad (6.14)$$

where ϵ'' is the imaginary component of the dielectric constant $\epsilon = \epsilon' + j\epsilon''$ and $\epsilon'' < 0$. For a uniform field the above expression simplifies

$$P_{Ld} = \frac{\omega \epsilon''}{4} E_0^2 W \times L \times t \quad (6.15)$$

where the applied voltage $V_0 = E_0 \times t$. Substituting we have

$$P_{Ld} = \frac{\omega \epsilon''}{4} \frac{V_0^2}{t} W \times L = \frac{\omega C}{4} \frac{\epsilon''}{\epsilon'} |V_0|^2 = \frac{G |V_0|^2}{4} \quad (6.16)$$

where $G = \omega C \frac{\epsilon''}{\epsilon'} = \omega C \tan \delta$.

The energy storage in the capacitor, though, is given by

$$E_C = \frac{1}{4} C |V_0|^2 \quad (6.17)$$

where the AC current $I = j\omega C V_0$ flows through the device. The losses through the capacitor plates are given by

$$P_{Lc} = 2 \times \frac{1}{4} |I|^2 \frac{R_p}{3} = \frac{1}{6} C^2 |V_0|^2 \omega^2 R_p \quad (6.18)$$

where the factor of $1/3$ accounts for the distributed nature of the current flow and the leading factor of 2 accounts for the presence of two plates of equal resistance.

Given the two loss mechanisms, we can compute the Q as follows

$$Q_C = \frac{\omega \frac{1}{4} C |V_0|^2}{\frac{1}{6} R_p C^2 \omega^2 R_p + \frac{1}{4} \omega C \tan \delta |V_0|^2} = \frac{1}{\frac{2}{3} \omega C R_p + \tan \delta} \quad (6.19)$$

For $\tan \delta \approx 10^{-6}$, the Q is bounded by 10^6 , and so the plate resistance dominates the losses.

6.8 Calculation by means of the vector potential

Partial inductance

In the previous chapter, we introduced the concept of partial inductance, or the inductance of an “open” path. Physically, we know that inductance can only be associated with a closed path, so, when we speak of a partial inductance, we must exercise caution. A classic example of a confusing statement is the “inductance of a bond wire,” which is a meaningless concept unless the return current is carefully considered. This is shown schematically in Fig. 6.22, where the inductance of the first path is defined by the current flow through the ground plane, whereas the inductance of the second path is defined by a loop that does not flow through the ground plane. The ground plane will lower the inductance of the second path due the image current flows (see Section 6.8).

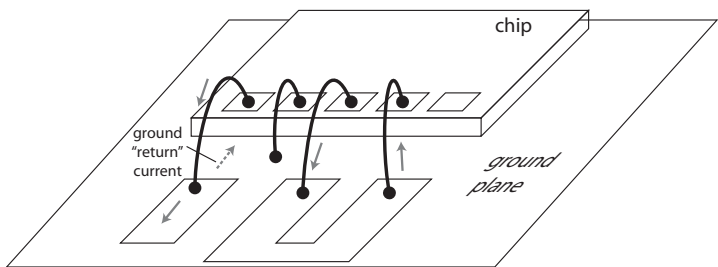


Figure 6.22 The “inductance” of a bondwire depends on the complete closed loop of the current flow.

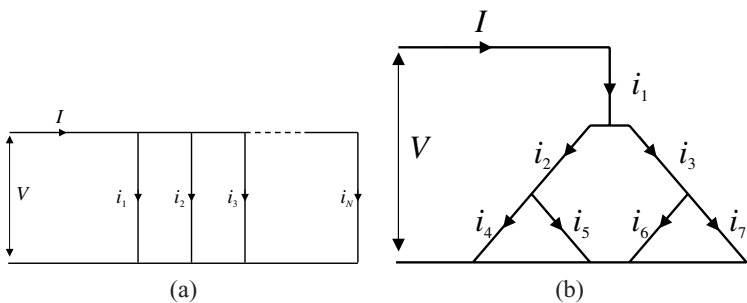


Figure 6.23 (a) A geometric structure with multiple loops but only two nodes. (b) A tree-like geometric structure.

But there is definite utility in the partial inductance concept when we look at efficient ways to calculate the inductance of practical circuits. For instance, consider the inductance of an arbitrary structure consisting of several loops and one pair of terminals, as shown in Fig. 6.23a. The structure has many mesh currents but only one node pair. Thus the voltage across every segment is equal and by KCL the sum of the currents must equal the terminal currents. The following expression, then, is the inverse of inductance of the entire structure

$$L^{-1} = \frac{I}{V} = \frac{1}{V} \sum_i I_i = \sum_{ij} [M^{-1}]_{ij} \tag{6.20}$$

where an inversion of the partial inductance matrix M is required.

As a counter example, Fig. 6.23b grows in a tree-like fashion and the number of meshes is less than the number of nodes. Setting up this system of equations is a bit more complicated using the partial inductance concept. To calculate the inductance of the above structure, we only have to note that if a current I is applied to the terminals of the structure, the voltage at the terminals is given by Eq. (5.100) and the currents are related by KCL

$$I = i_1 = i_2 + i_3 \tag{6.21}$$

$$i_2 = i_4 + i_5 \tag{6.22}$$

$$i_3 = i_6 + i_7 \tag{6.23}$$

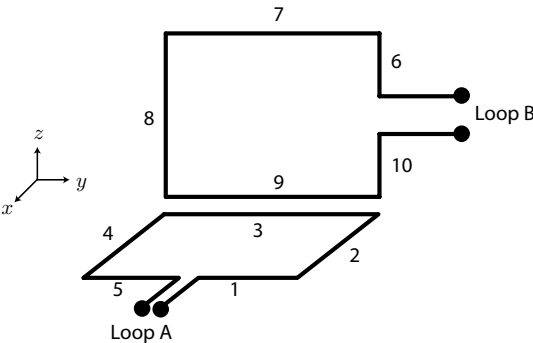


Figure 6.24 Two loops with several orthogonal sections.

and the voltages are related by KVL

$$V = v_1 + v_2 + v_4 \quad (6.24)$$

$$V = v_1 + v_2 + v_5 \quad (6.25)$$

$$V = v_1 + v_3 + v_6 \quad (6.26)$$

$$V = v_1 + v_3 + v_7 \quad (6.27)$$

$$(6.28)$$

The above linear system of equations, along with Eq. (5.101), can be solved and expressed compactly in matrix notation

$$\begin{pmatrix} 1 & -1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 & -1 \end{pmatrix} i = Bi = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad (6.29)$$

and

$$\begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix} v = Av = V \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \quad (6.30)$$

It is not too hard to show that

$$L = \hat{e}_1^T AZB^{-1}\hat{e}_1 \quad (6.31)$$

where $\hat{e}_1^T = (10 \dots 0)$ and the impedance matrix $Z = j\omega M$. This problem can also be solved using a mesh equation formulation.

In addition to serving as a computation aid, the partial inductance concept also adds a new vantage point in designing inductors. For instance, consider the mutual inductance between two orthogonal loops shown in Fig 6.24. Using the partial mutual inductance concepts we can immediately ignore several coupling terms since the vectors are orthogonal and $\mathbf{dl}_1 \cdot \mathbf{dl}_2 = 0$. Also, since several mutual coupling terms are identical we can quickly estimate M . This is not so obvious from the flux or energy concepts. As an example, consider

the partial mutual coupling between segment 3 of loop A and loop B (ignoring the leads)

$$M_{3,B} = M_{3,6} + M_{3,7} + M_{3,8} + M_{3,9} + M_{3,10}$$

Dropping all the perpendicular segments ($M_{3,6} = M_{3,8} = M_{3,10} = 0$), we are left with

$$M_{3,B} = M_{3,7} + M_{3,9}$$

Magnetic vector integral and differential equations

In theory we can compute the self or mutual inductance of any structure as long as we calculate vector potential given by Eq. (5.87) and then follow with a line integral about the path of interest. In practice, though, the integrand of Eq. (5.87) is unknown since the high-frequency current distribution is unknown. How is the high-frequency current distribution obtained? From the magnetic vector potential

$$\mathbf{J} = \sigma \mathbf{E} = j\omega\sigma \mathbf{A} + \sigma \nabla \phi \quad (6.32)$$

In other words, at any given frequency to solve for $\mathbf{A}$ we must solve the following integral equation²

$$\mathbf{A}(\mathbf{r}) = \mu \int_V \frac{j\omega\sigma \mathbf{A}(\mathbf{r}')}{4\pi R} dV' \quad (6.33)$$

For simplicity we have dropped the electric potential ϕ from the above equation for the magnetostatic case but in general we must solve a pair of integral equations: one for the electrostatic potential and one for the magnetostatic potential. Current continuity can be used to relate charge and current. By the definition of magnetic vector potential and from Maxwell's equations we have

$$\nabla \times \nabla \times \mathbf{A} = \nabla \times \mathbf{B} = \mu \mathbf{J} + j\omega\mu\epsilon \mathbf{E} \quad (6.34)$$

Employing the vector identity

$$\nabla \times \nabla \times \mathbf{A} \equiv \nabla \nabla \cdot \mathbf{A} - \nabla^2 \mathbf{A} \quad (6.35)$$

and the simple Coulomb gauge $\nabla \cdot \mathbf{A} \equiv 0$ ³ we have

$$-\nabla^2 \mathbf{A} = j\omega\mu\epsilon(-j\omega\mathbf{A} - \nabla\phi) + \mu \mathbf{J} \quad (6.36)$$

For magnetostatic problems we neglect the first two terms since we take the limit as $\omega \rightarrow 0$ and we are left with the spatial differential equation

$$\nabla^2 \mathbf{A} = -\mu \mathbf{J} \quad (6.37)$$

In Cartesian coordinates, the above equation is identical to three scalar Poisson's equations. In principle we can solve these differential equations to solve any magnetostatic

² Note that $\mathbf{A}$ appears on both sides of the *integral equation*. Solving integral equations is analogous to (and as difficult as) solving differential equations.

³ Recall that we are free to choose the divergence of $\mathbf{A}$ as we see fit.

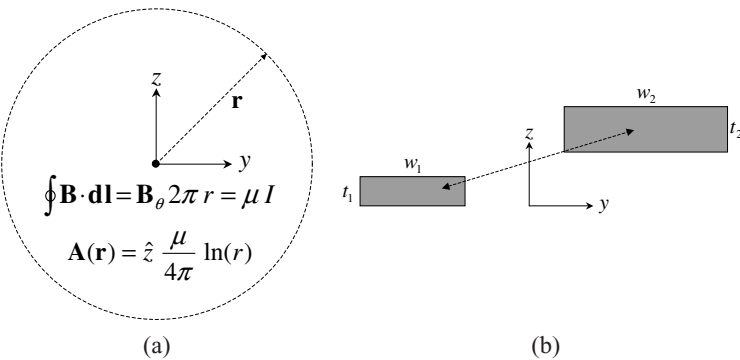


Figure 6.25 (a) The magnetic potential of a filament in two dimensions. (b) The mutual inductance of two rectangular conductors in two dimensions.

problem. In reality, the solution process is complex and requires numerical computations. In the following sections we focus on problems that admit analytic solutions.

Filamental calculations

For filamental circuits, the current distribution is known (trivially) and the volume integral of Eq. (5.87) simplifies to a line integral over the path of the current

$$\mathbf{A}(\mathbf{r}) = I \mu \int_C \frac{d\mathbf{l}}{R} \quad (6.38)$$

There are no unknowns in the above integral, but there is a problem with the $1/R$ discontinuity. If we try to integrate the above over the path of the current, $R = 0$. However, as noted in Section 5.4 in deriving Neumann's formula, the above integral can be evaluated without any difficulty when computing the mutual inductance.

The most important case that we shall consider is the mutual inductance between two filaments in free space. To begin, consider an infinitely long filament. The magnetic vector potential in the vicinity of such a filament is readily computed

$$A(r) = \frac{\mu_0 I}{4\pi} \ln(r) \quad (6.39)$$

The magnetic vector potential falls off slowly as the natural log of distance from the filament. Therefore, the mutual inductance between two filaments separated by a radial distance r is thus

$$M(r) = \frac{\mu_0}{4\pi} \ln(r) \quad (6.40)$$

Using the above result, let us derive the mutual inductance between two conductors of the rectangular cross section shown in Fig. 6.25b. Each conductor has width w_i and thickness t_i and is located within the specified Cartesian coordinates. Since we have a solution for any two filaments in the cross sections, we can average $M(r)$ from above over every pair of

constituent filaments in the cross sections to obtain the mutual inductance

$$M = \frac{\mu_0}{4\pi} \frac{1}{w_1 w_2 t_1 t_2} \int \int_{A_1} \int \int_{A_2} \ln \sqrt{(y - y')^2 + (z - z')^2} dy' dz' dy dz \quad (6.41)$$

where $A_1 = w_1 t_1$ and $A_2 = w_2 t_2$ are the cross-sectional areas of the rectangular conductors. The above quadruple integral is difficult to evaluate, but fortunately has a closed-form solution given by [66]

$$-\frac{1}{2} F(y - y', z - z') \Big|_{y_1}^{y_2} \Big|_{y_3}^{y_4} \Big|_{z_1}^{z_2} \Big|_{z_3}^{z_4} - \frac{25}{12} w_1 w_2 t_1 t_2 \quad (6.42)$$

where the function F is given by

$$F(y, z) = \frac{y^4 - 6y^2 z^2 + z^4}{24} \ln(y^2 + z^2) - \frac{yz}{3} \left(y^2 \tan^{-1} \frac{z}{y} + z^2 \tan^{-1} \frac{y}{z} \right) \quad (6.43)$$

As is usually the case, it is not too difficult to verify that the above is the solution to the integral since

$$\frac{\partial^4 F(y - y', z - z')}{\partial y \partial y' \partial z \partial z'} = \ln((y - y')^2 + (z - z')^2) - \frac{25}{6} \quad (6.44)$$

From the definition of F , the mutual inductance between two filaments can be readily calculated (preferably by a computer). What about the self inductance of a rectangular conductor? If we carefully take limits to let $y_3 \rightarrow y_1$ and $y_4 \rightarrow y_2$ we arrive at

$$L(w, t) = \frac{\mu_0}{4\pi} \left(\frac{1}{12w^2 t^2} 8t w^3 \tan^{-1} \left(\frac{t}{w} \right) + 8t^3 w \tan^{-1} \left(\frac{w}{t} \right) + t^4 \ln(t^2) + w^4 \ln(w^2) + (t^4 - 6t^2 w^2 + w^4) \ln(t^2 + w^2) \right) - \frac{25}{12} \quad (6.45)$$

Self inductance of conductors

Using the above techniques we can derive the self inductance of a flat conductor (zero-thickness) of length l and width w [46]

$$L = \frac{\mu l}{2\pi} \left(\sinh^{-1} \left(\frac{l}{w} \right) + \frac{l}{w} \sinh^{-1} \left(\frac{w}{l} \right) + \frac{w}{3l} - \frac{1}{3} \left(\frac{l}{w} \right)^2 \left[\left(1 + \left(\frac{w}{l} \right)^2 \right)^{\frac{3}{2}} - 1 \right] \right) \quad (6.46)$$

If the conductor has thickness t much less than w , then the above equation can be employed where w is replaced by $w + t$.

Geometric mean distance and the Grover/Greenhouse methods

Even for the simple two-dimensional geometry shown above, the computations are exorbitantly complicated. While three-dimensional expressions are even more complicated, we can often employ numerical integration techniques. One practical difficulty, though, is due

to the singularity in the self-inductance computation. In this section we will discuss the *geometric mean distance* approximation. We shall see this technique is essentially a refinement of the two-dimensional approximation.

First, consider the mutual inductance between two parallel filaments of length l . It is easy to show that the following integral expression must be evaluated for the vector potential at (x, y, z) due to a filament at the origin

$$\mathbf{A}(x, y, z) = \frac{I\mu_0}{4\pi} \int_{-l/2}^{+l/2} \frac{1}{\sqrt{x^2 + y^2 + (z - z')^2}} dz' \quad (6.47)$$

Evaluating the integral

$$A_z(x, y, z) = \hat{\mathbf{z}} \frac{\mu_0 I}{4\pi} \ln \left[\frac{(z + l/2) + \sqrt{x^2 + y^2 + (z + l/2)^2}}{(z - l/2) + \sqrt{x^2 + y^2 + (z - l/2)^2}} \right] \quad (6.48)$$

Integrating the above expression along a parallel path at $x = d$

$$M(l, d) = \frac{1}{l} \int_{-l/2}^{l/2} A_z(d, 0, z') dz' \quad (6.49)$$

This gives the following geometric expression for the mutual inductance

$$M(l, d) = \frac{\mu_0}{4\pi} l \left[\ln \left(\sqrt{1 + \left(\frac{l}{d}\right)^2} + \frac{l}{d} \right) - \sqrt{1 + \left(\frac{d}{l}\right)^2} + \frac{d}{l} \right] \quad (6.50)$$

Using an approach similar to the previous section, we can average the above expression over a cross-sectional area of two rectangular conductors to obtain the *partial* mutual inductance between two rectangular slabs

$$M_{12} = \frac{1}{A_1 A_2} \int_{A_1} \int_{A_2} M(L, r) dA_1 dA_2 \quad (6.51)$$

where $r = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}$. For the case of self inductance, we can set $A_1 = A_2$ and perform the integration. This will result in integrable singularity in the calculation.

If the filaments under consideration are of length such that $L \gg d$, then the results for $M(L, d)$ can be simplified if we neglect the term $\left(\frac{d}{L}\right)^2$

$$M(L, d) \approx \frac{\mu_0}{4\pi} L \left(\frac{d}{L} - \ln d + \ln 2L - 1 \right) \quad (6.52)$$

Integration of the above expression across the cross section yields a simpler calculation. The only difficult term to integrate is the $\ln(d)$. We have already met this integral in the two-dimensional case. This term, the *geometric mean distance* (or GMD) is the average value of the logarithm of the distance between all points between the conductors. The integral also contains a simpler term, the *arithmetic mean distance* (AMD), simply the average distance between the conductors.

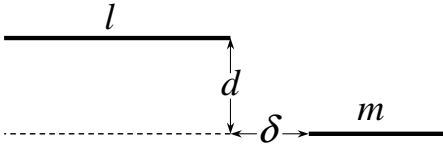


Figure 6.26 Two parallel filaments of unequal length.

In general, for any two parallel conductors of an arbitrary cross section, the GMD approximation gives the following result for the mutual inductance

$$M(L, d) \approx \frac{\mu L}{4\pi} (\text{AMD}/L - \text{GMD} + \ln(2L) - 1) \quad (6.53)$$

as long as $(\frac{d}{L})^2 \ll 1$.

In many practical situations the conductors are isolated so that $d \gg l$. The GMD approximation does not hold in these cases but the simple filamental approximation is valid even up to high frequencies. This is plausible since at considerable distance from the source, the exact current distribution is immaterial and an equivalent filament serves as an excellent approximation. Since we have seen that at high frequencies only the current distribution changes, the approximation is good even at AC frequencies.

General orthogonal geometries

The partial mutual inductance between unequal parallel filaments shown in Fig. 6.26 is given by the following expression [22]

$$M = \frac{\mu_0}{4\pi} \left[\alpha \sinh^{-1} \frac{\alpha}{d} - \beta \sinh^{-1} \frac{\beta}{d} - \gamma \sinh^{-1} \frac{\gamma}{d} + \delta \sinh^{-1} \frac{\delta}{d} - \sqrt{\alpha^2 + d^2} + \sqrt{\beta^2 + d^2} + \sqrt{\gamma^2 + d^2} - \sqrt{\delta^2 + d^2} \right] \quad (6.54)$$

where

$$\alpha = l + m + \delta, \quad \beta = l + \delta, \quad \gamma = m + \delta \quad (6.55)$$

The distance δ can go negative if the filaments overlap. The above result is generated by integrating the magnetic vector potential given by Eq. (6.48) over the shown path.

Using some clever symmetry arguments, we can also derive this result without doing any integrals. Consider the case shown in Fig. 6.27a, where we begin with two equal length parallel segments of length ℓ and break one segment to a length m . With reference to Fig. 6.27b, the mutual inductance for the equal segments can be written

$$M(l) = M(m) + M(l - m) + 2M_x \quad (6.56)$$

The mutual inductance M_x is shown in Fig. 6.27c. The notation $M(p)$ is the filamental mutual inductance between two parallel equal length filaments of length p . But we can also express M_x in terms of our desired mutual inductance M

$$M_x = M - M(m) \quad (6.57)$$

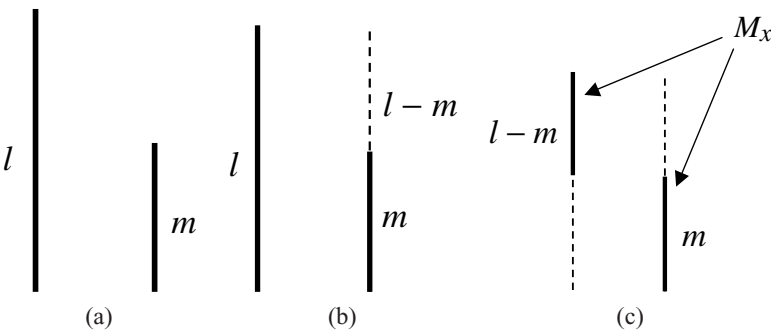


Figure 6.27 (a) Two parallel unequal length filaments with ends at equal positions. (b) An equivalent system of real and fictitious segments used in computing the mutual inductance. (c) The mutual inductance M_x can be expressed in terms of the desired result (see text).

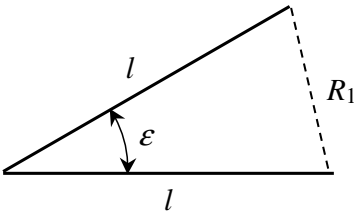


Figure 6.28 Two equal length filaments meeting at a point.

Substituting we have an expression for the desired mutual inductance

$$M = \frac{1}{2} (M(l) + M(m) - M(l - m)) \tag{6.58}$$

where each term in the above expression is given by Eq. (6.52). In the problems (online) you are asked to use such symmetry arguments to compute the mutual inductance of various filamental configurations. Using the GMD argument, you can substitute the GMD for l and generalize the above calculations to appropriate cases.

Arbitrary geometry

The filamental partial mutual inductance between any arbitrary arrangements of filaments can be computed using the above recipe. Compute the magnetic vector potential for the first filament (Eq. (6.48)). Next, integrate this expression over the path of the second filament. In theory these computations are easy but in practice they are tedious. The results are from Grover and repeated here since this source is out of print [22].

Consider the simple case of equal filaments meeting at a point (Fig. 6.28).

$$M = \frac{\mu_0}{\pi} l \cos \epsilon \tanh^{-1} \frac{l}{l + R_1} \tag{6.59}$$

Copyright © 2007, Cambridge University Press. All rights reserved.

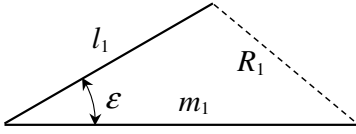


Figure 6.29 Two unequal length filaments meeting at a point.

where

$$R_1^2 = 2l^2(l - \cos \epsilon) \quad (6.60)$$

For unequal filaments meeting at a point use

$$M = \frac{\mu_0}{2\pi} \cos \epsilon \left[l_1 \tanh^{-1} \frac{m_1}{l_1 + R} + m_1 \tanh^{-1} \frac{l_1}{l_1 + R} \right] \quad (6.61)$$

As shown in Fig. 6.29, the filament lengths are l_1 and m_1 and subtend an angle ϵ . The distance between the filament ends is R . The following relations can be used to substitute for either R or ϵ

$$\cos \epsilon = \frac{l_1^2 + m_1^2 - R^2}{2l_1 m_1} \quad (6.62)$$

and

$$\frac{R^2}{l_1^2} = 1 + \frac{m_1^2}{l_1^2} - 2 \frac{m_1}{l_1} \cos \epsilon \quad (6.63)$$

Generalizing further we present the formula for unequal filaments in the same plane, not meeting. As shown in Fig. 6.30 the filaments lengths are l and m and the following geometric relations apply

$$2 \cos \epsilon = \frac{\alpha^2}{lm} \quad (6.64)$$

$$\alpha^2 = R_4^2 - R_3^2 + R_2^2 - R_1^2 \quad (6.65)$$

$$\mu = \frac{l [2m^2(R_2^2 - R_3^2 - l^2) + \alpha^2(R_4^2 - R_3^2 - m^2)]}{4l^2 m^2 - \alpha^4} \quad (6.66)$$

$$v = \frac{m [2l^2(R_4^2 - R_3^2 - m^2) + \alpha^2(R_2^2 - R_3^2 - l^2)]}{4l^2 m^2 - \alpha^4} \quad (6.67)$$

$$R_1^2 = (\mu + l)^2 + (v + m)^2 - 2(\mu + l)(v + m) \cos \epsilon \quad (6.68)$$

$$R_2^2 = (\mu + l)^2 + v^2 - 2v(\mu + l) \cos \epsilon \quad (6.69)$$

$$R_3^2 = \mu^2 + v^2 - 2\mu v \cos \epsilon \quad (6.70)$$

$$R_4^2 = \mu^2 + (v + m)^2 - 2\mu(v + m) \cos \epsilon \quad (6.71)$$

And at last, with the preliminaries out of the way, the partial mutual inductance is given by

$$\begin{aligned} \frac{M}{2 \cos \epsilon} = \frac{\mu_0}{4\pi} \left[(\mu + l) \tanh^{-1} \frac{m}{R_1 + R_2} + (v + m) \tanh^{-1} \frac{l}{R_1 + R_4} \right. \\ \left. - \mu \tanh^{-1} \frac{m}{R_3 + R_4} - v \tanh^{-1} \frac{l}{R_2 + R_3} \right] \quad (6.72) \end{aligned}$$

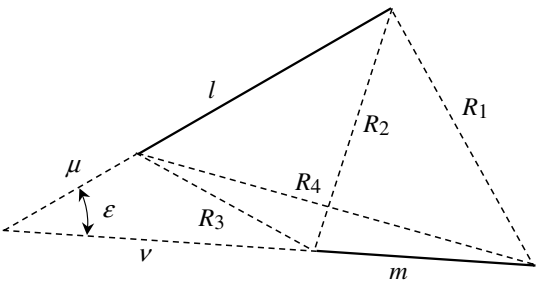


Figure 6.30 Two arbitrary filaments lying in a common plane.

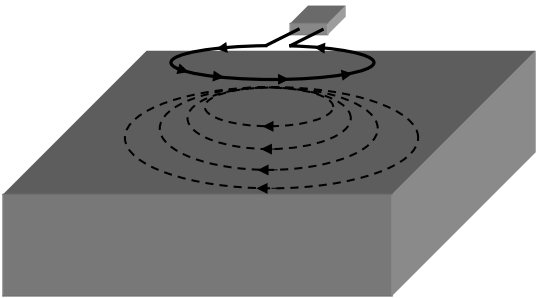


Figure 6.31 A loop inductor near an ideal ground plane.

Using symmetry arguments and Eq. (6.59) we have

$$M_{lm} = (M_{\mu+l, \nu+m} + M_{\nu\mu}) - (M_{\mu+l, \nu} + M_{\nu+m, \mu}) \quad (6.73)$$

The most general result, the partial mutual inductance of two straight filaments placed in any desired position, is presented in Appendix 6.10. You may wonder why we present the above results when we could have simply quoted this most general case from the outset. The reason is obvious when we write the equation down – it is a *long* formula. Thus it is wise to use the simplest equation possible, not only to save computation time but also to enhance numerical stability. Using the above formulas, the inductance of any piecewise linear structure can be approximated. The self inductance for each section can be computed directly and the mutual inductance can be approximated from the above filamental equations. Given enough segments and a modern workstation, any smooth curve can also be approximated.

Effect of an ideal ground plane

An ideal ground plane is defined as an infinitely conductive plane. If an inductive circuit resides *near* the ground plane, then from our experience thus far we can see that induced currents will flow in the ground plane in such a way as to oppose the impinging magnetic field. This is shown schematically in Fig. 6.31. Since the induced currents *reduce* the magnetic flux, the effective inductance seen at the terminals of the inductor is lowered.

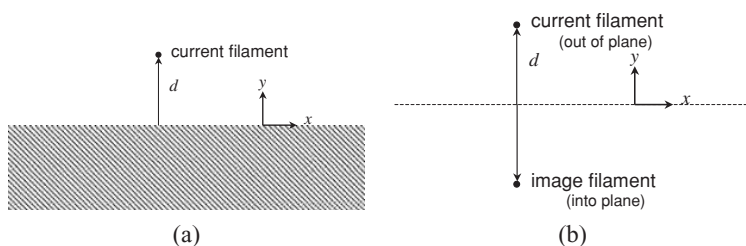


Figure 6.32 (a) An infinitely long filament residing parallel to a ground plane. (b) An equivalent system of an image filament without the ground plane.

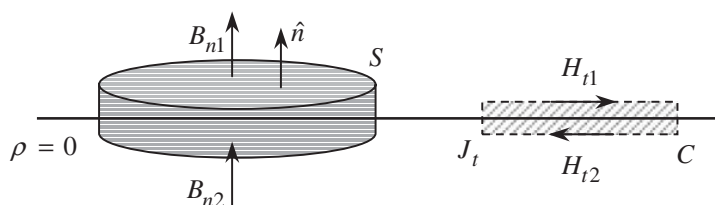


Figure 6.33 A surface and curve enclosing regions inside and outside of the conductive region.

Often, magnetically induced currents are referred to as *eddy currents*. Since magnetic fields arise due to currents, we must include eddy currents in Eq. (5.87) when computing the magnetic vector potential or the magnetic field. Thus we must either solve the resulting integral equation (5.87) or the differential Eq. (6.37).

For now, we will present a powerful alternative method of solution that does not require solving any integral or differential equations, the *method of images*. For many symmetric situations, we can invoke this approximation to include the induced ground currents. For instance, consider an infinitely long filamental current at a distance d over an infinitely conductive ground plane shown in Fig. 6.32a. Now consider an equivalent system of two conductors separated by a distance $2d$ shown in Fig. 6.32b, where the introduced current filament at $z = -d$ has a current flowing in a direction opposite to the original filament. We would like to show that the field configurations of the two systems are identical for $z > 0$.

Since the ground plane is infinitely conductive, the fields decay to zero infinitely fast and thus $\mathbf{E} = 0$ and $\mathbf{H} = 0$ inside the conductor. The only induced currents, then, are on the surface of the conductor. The vector fields in Fig. 6.32a are uniquely determined by the sources (the filamental current) and the boundary conditions. Let us take the boundary to be the surface of the ground plane. Imposing boundary conditions on the normal component of the magnetic field, consider the surface integral of $\mathbf{B}$ over the pillbox surface at the boundary shown in Fig. 6.33a. Let the height of the cylinder go to zero so the only contribution to the integral is the normal component of the field at the top and bottom surface

$$\int_S \mathbf{B} \cdot d\mathbf{S} = B_{n1} - B_{n2} = 0 \quad (6.74)$$

The right-hand side is zero by Maxwell's equations. Since B_{n2} is zero inside the conductive region, $B_{n1} = 0$ as well. Note that the current filaments of Fig. 6.32b produce a zero

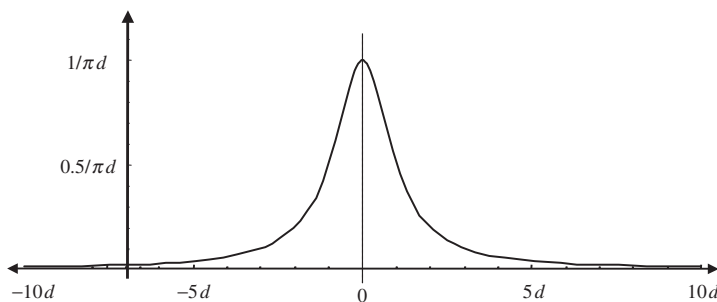


Figure 6.34 The current density at the surface of the ground plane.

normal field at $z = 0$. By using the Biot-Savart Law or the magnetic vector potential of a filament we see that the magnetic field is tangential at the surface. The tangential component is given by

$$B_t = B_x(x, 0) = \frac{\mu I}{2\pi} \left[\frac{y-d}{x^2 + (y-d)^2} - \frac{y+d}{x^2 + (y+d)^2} \right]_{y=0} \quad (6.75)$$

If the systems are equivalent, then we can calculate the functional form of the current density in the ground plane from the above expression

$$J_t = H_t = \frac{1}{\mu} B_t = \frac{I}{\pi} \frac{d}{x^2 + d^2} \quad (6.76)$$

We have plotted this surface current in Fig. 6.34. As expected, the current peaks near the filamental source. The peak value is $I/(\pi d)$ and increases inversely as the filament is moved closer to the ground plane.

Thus, the effect of the ground plane is to lower the self inductance by the mutual inductance between the source and its image. For a conductor with inductance L_0 in the absence of the ground plane, then, the effective inductance near the ground plane is reduced to $L_{eff} = L_0 - M$, where M is the mutual inductance between the conductor and its image. Or, alternatively, $L_{eff} = L(1 - k)$, where k is the coupling coefficient.

6.9 References

This chapter is an extension of a previous book I wrote as a graduate student at Berkeley [43]. Many of the equations come from the trusty reference by Grover [22] and Ternan [60]. The vector potential approach and partial inductance techniques were popularized by researchers such as Greenhouse [21] and Ruelli [54].

6.10 Appendix: Filamental partial mutual inductance

Here we present the partial mutual inductance of two straight filaments placed in any desired position. The geometry is shown in Fig. 6.35. Let the distances between the segment ends

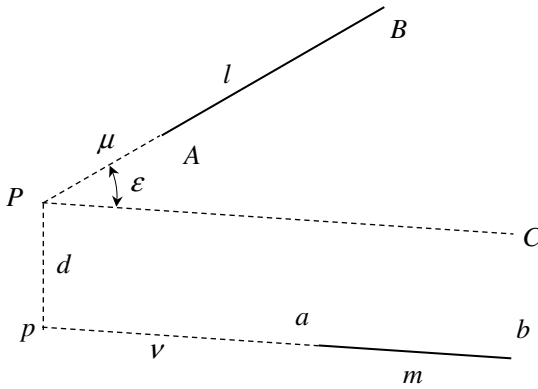


Figure 6.35 Two arbitrarily placed filaments.

be denoted by R_i

$$Bb = R_1 \quad (6.77)$$

$$Ba = R_2 \quad (6.78)$$

$$Aa = R_3 \quad (6.79)$$

$$Ab = R_4 \quad (6.80)$$

$$\begin{aligned} \frac{M}{\cos \epsilon} = \frac{\mu}{2\pi} & \left((l + \mu) \tanh^{-1} \left(\frac{m}{R_1 + R_2} \right) - v \tanh^{-1} \left(\frac{l}{R_2 + R_3} \right) \right. \\ & \left. + (m + v) \tanh^{-1} \left(\frac{l}{R_1 + R_4} \right) - \mu \tanh^{-1} \left(\frac{m}{R_3 + R_4} \right) \right) - d \Omega \csc(\epsilon) \end{aligned} \quad (6.81)$$

where

$$\begin{aligned} \Omega = \tan^{-1} & \left(\frac{d^2 \cos \epsilon + (\mu + l)(v + m) \sin^2 \epsilon}{d R_1 \sin \epsilon} \right) - \tan^{-1} \left(\frac{d^2 \cos \epsilon + (\mu + l)v \sin^2 \epsilon}{d R_2 \sin \epsilon} \right) \\ & + \tan^{-1} \left(\frac{d^2 \cos \epsilon + \mu v \sin^2 \epsilon}{d R_3 \sin \epsilon} \right) - \tan^{-1} \left(\frac{d^2 \cos \epsilon + \mu(v + m) \sin^2 \epsilon}{d R_4 \sin \epsilon} \right) \end{aligned} \quad (6.82)$$

The various parameters can be written in terms of the geometric parameters

$$\cos \epsilon = \frac{\alpha^2}{2lm} \quad (6.83)$$

$$\alpha^2 = R_4^2 - R_3^2 + R_2^2 - R_1^2 \quad (6.84)$$

$$d^2 = R_3^2 - \mu^2 - v^2 + 2\mu v \cos \epsilon \quad (6.85)$$

$$\frac{\mu}{l} = \frac{2m^2(R_2^2 - R_3^2 - l^2) + \alpha^2(R_4^2 - R_3^2 - m^2)}{4l^2m^2 - \alpha^4} \quad (6.86)$$

$$\frac{v}{m} = \frac{2l^2(R_4^2 - R_3^2 - m^2) + \alpha^2(R_2^2 - R_3^2 - l^2)}{4l^2m^2 - \alpha^4} \quad (6.87)$$

Alternatively, the above equations can be inverted to obtain

$$R_1^2 = d^2 + (\mu + l)^2 + (v + m)^2 - 2(\mu + l)(v + m) \cos \epsilon \quad (6.88)$$

$$R_2^2 = d^2 + (\mu + l)^2 + v^2 - 2v(\mu + l) \cos \epsilon \quad (6.89)$$

$$R_3^2 = d^2 + \mu^2 + v^2 - 2\mu v \cos \epsilon \quad (6.90)$$

$$R_4^2 = d^2 + \mu^2 + (v + m)^2 - 2\mu(v + m) \cos \epsilon \quad (6.91)$$

10 Transformers

Transformers find wide and important applications in RF circuits. Historically, transformers were used in power systems for voltage step-up and step-down. Since power transmission is more efficient at high voltages, transformers are used to boost signals to tens of thousands of volts for long-range transmission. For safety, though, we prefer to work with much lower voltage levels, and thus a transformer is used to step down the voltage to hundreds of volts before delivery into homes and factories. Electronic components, though, even operate at lower voltages. Before the widespread use of switching power supplies, transformers were ubiquitous in performing this task. For ultra low-noise applications, such as sensitive measurements, transformer-based designs still reign supreme as the method of choice.

The name “transformer,” in fact, derives from this function. As such, it’s an ideal element for transforming voltages. Consequently, due to conservation of energy, transformers are equally good at stepping down/up currents. Thus, transformers are doubly good at impedance transformation. A desirable quality of transformers is their broadband operation. Whereas passive LC circuits can easily double as impedance transformation circuits, they do so only over a narrow bandwidth.

An ideal transformer is broadband, faithfully duplicating, inverting, and scaling voltages and currents independent of frequency with one important exception. Since its behavior derives from magnetic induction, it cannot perform its function for static or DC signals. This simple fact is probably responsible for the worldwide deployment of an AC electrical network. Early proponents of AC power successfully argued that high voltage AC power delivery was more efficient and flexible due to the existence of the transformer. If an efficient DC transformer existed at the dawn of power, we would all no doubt be much less accustomed to AC voltages.

10.1 Ideal transformers

An ideal two-winding transformer is shown in Fig. 10.1a. Each winding defines a port, or a terminal pair with voltage $V_{1,2}$ and current $I_{1,2}$. An ideal transformer is characterized by the turns ratio N , which defines the voltage step up/down from one port to another

$$V_2 = NV_1 \quad (10.1)$$

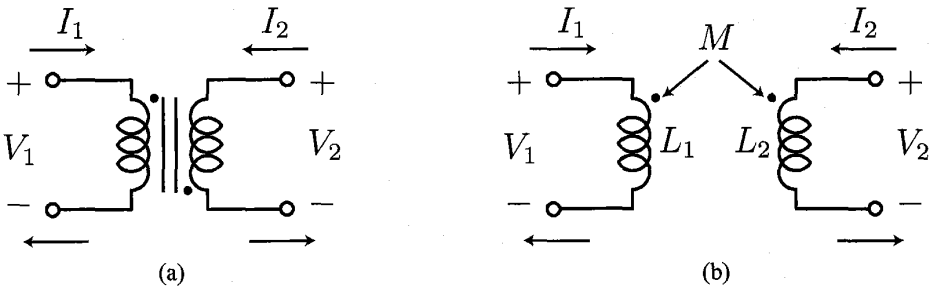


Figure 10.1 (a) Schematic symbol for an ideal transformer. (b) Two coupled inductors form a non-ideal transformer.

As already noted, this relationship is valid for any AC frequency $\omega \neq 0$. Since the transformer is a passive element, it conserves power and thus

$$V_1 I_1 + V_2 I_2 = 0 \quad (10.2)$$

which implies

$$I_2 = -\frac{1}{N} I_1 \quad (10.3)$$

The impedance transformation property is therefore a function of N^2

$$Z_1 = \frac{V_1}{I_1} = \frac{\frac{1}{N} V_2}{-N I_2} = \left(\frac{1}{N^2} \right) \frac{V_2}{-I_2} = \frac{Z_2}{N^2} \quad (10.4)$$

Since the ports of an ideal transformer are isolated, we can easily redefine a port of opposite polarity $V'_2 = -V_2$ and thus perform voltage inversion with no additional effort.

10.2 Dot convention

A common point of confusion with transformers is the dot convention. As shown in Fig. 10.1, it's standard practice to place dots on one side of the terminals of transformers or coupled inductors. The dot convention can be stated in the following way [12]: If current is sent *into* the dotted terminals of the coupled inductors, the magnetic fluxes linking the coils will reinforce each other. With this convention, the mutual inductance M is always positive.

When confusion over notation arises, it's often fruitful to go back to the fundamental physics. In general, when a time-varying magnetic field impinges on a conductor, the induced currents will always flow in a direction so as to oppose the magnetic field. This is generally valid, as the current generates its own magnetic field to oppose the incoming field so as to lower the overall energy of the system. This is known as Lenz's Law. The same is true for electric fields. When an electric field impinges on a conductor, the charges redistribute so as to cancel the incoming field. This is why the magnetic and electric fields in a perfect conductor are driven to zero.

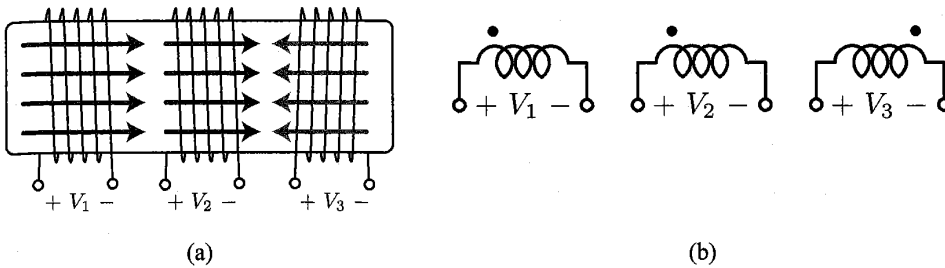


Figure 10.2 (a) Three coils wound on to a common core. (b) Equivalent circuit representations of the coils using the dot convention.

An example will illustrate the dot convention more clearly. Three coils are wound on to a common core as shown in Fig. 10.2a. The terminals have been defined as positive, as indicated in the figure. Since coils 1 and 2 have a similar clockwise rotation, their magnetic fields will always add (when current flows into the positive terminals). So they should have dots placed on the same terminal. We place the dots at the positive terminal as shown in Fig. 10.2b. Equivalently we could have chosen both of the negative terminals. The third coil, though, is wound counterclockwise. Thus, the magnetic field of coil 3 with positive current will oppose the fields of the first two coils. As such, we must place the dot at the negative terminal to indicate that a negative current in the third coil is necessary in order to reinforce the fields of coils 1 and 2.

10.3 Coupled inductors as transformers

Ideal transformers are an abstraction as all real transformers deviate from ideal behavior in one form or another. As shown in Fig. 10.1b, real transformers are implemented as coupled inductors. As we have learned already, if two inductors are coupled magnetically, then

$$V_1 = j\omega L_1 I_1 + j\omega M I_2 \quad (10.5)$$

$$V_2 = j\omega M I_1 + j\omega L_2 I_2 \quad (10.6)$$

Solving for I_2 from the second equation

$$I_2 = \frac{1}{j\omega L_2} (V_2 - j\omega M I_1) \quad (10.7)$$

and substituting for I_2 from the second equation into the first

$$V_1 = j\omega L_1 I_1 + \frac{j\omega M}{j\omega L_2} (V_2 - j\omega M I_1) \quad (10.8)$$

and collecting terms we have

$$V_1 = j\omega \left(L_1 - \frac{M^2}{L_2} \right) I_1 + \frac{M}{L_2} V_2 \quad (10.9)$$

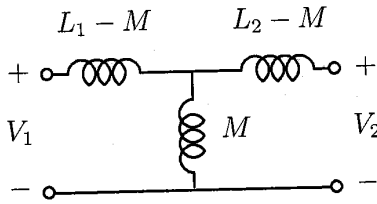


Figure 10.3 T model equivalent circuit for two coupled inductors.

It is interesting to observe that the transformer behavior is contained in the second term, which is frequency independent. The first term, on the other hand, is undesirable since it represents an inductive component. In fact, Eq. (10.8) explicitly emphasizes that an alternating voltage at port 2 generates a voltage through port 1 by first generating an alternating current in the second inductor, which in turn generates a time-varying magnetic field, in turn inducing a solenoidal electric field, which impresses a voltage at port 1.

The mutual inductance is usually parameterized in terms of the magnetic coupling coefficient $M = k\sqrt{L_1 L_2}$. Substituting this into the above equation we arrive at

$$V_1 = j\omega L_1(1 - k^2)I_1 + k\sqrt{\frac{L_1}{L_2}}V_2 \quad (10.10)$$

The transformer turns ratio is defined by coupling factor k and by the square root of the ratio of the inductances of each winding

$$\frac{1}{N} = k\sqrt{\frac{L_1}{L_2}} \quad (10.11)$$

The above equation is very important as it reveals that an ideal transformer can be constructed by two tightly wound coils such that $k \rightarrow 1$.

10.4 Coupled inductor equivalent circuits

It is sometimes more convenient to view two coupled inductors as a T network, as shown in Fig. 10.3. It is easy to show that this model is identical to two coupled inductors. To see this, consider the Z matrix of two coupled inductors

$$\mathbf{v} = \mathbf{Z}\mathbf{i} = j\omega \begin{pmatrix} L_1 & M \\ M & L_2 \end{pmatrix} \mathbf{i} \quad (10.12)$$

Now consider the coupled inductors as a black box. Since a two-port is completely characterized by its Z matrix, any other circuit that has the same two-port behavior, i.e. Z matrix, must be identical and indistinguishable from the external world. The voltage at port 1 of the T network is

$$V_1 = j\omega(L_1 - M)I_1 + j\omega M(I_1 + I_2) = j\omega L_1 I_1 + j\omega M I_2 \quad (10.13)$$

and so it's clear that the circuits are equivalent.

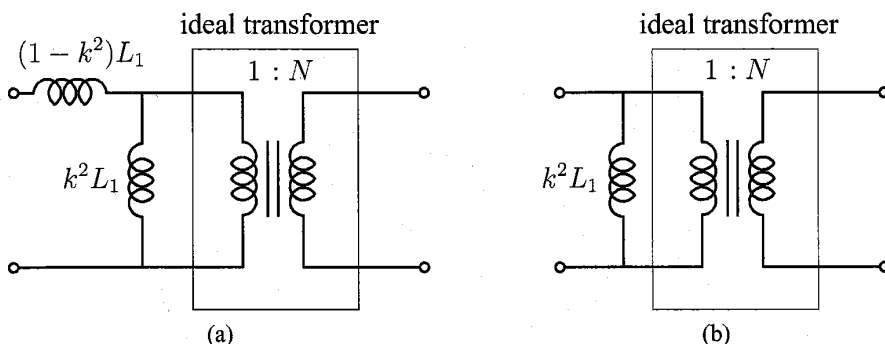


Figure 10.4 (a) Equivalent circuit model for two coupled inductors. (b) Approximation of (a) valid for two tightly coupled inductors ($k \rightarrow 1$).

When two inductors are tightly coupled ($k \approx 1$), then it's sometimes useful to view the coupled inductors as a perfect transformer marred by parasitics. In fact, it's not difficult to show that the equivalent circuit shown in Fig. 10.4a is such an exact representation. At port 1 the current into the shunt inductor k^2L_1 is the sum of the current into port 1 and N times the transformer secondary current

$$V_1 = j\omega(1 - k^2)L_1I_1 + j\omega k^2L_1I_1 + j\omega k^2NL_1I_2 \quad (10.14)$$

and recalling the definition of the magnetic coupling coefficient k , we have $k^2NL_1 = M$. Similarly, the voltage at port two is just N times the voltage V_p at the transformer primary

$$V_2 = NV_p = j\omega k^2NL_1(I_1 + NI_2) \quad (10.15)$$

Simplifying the above and noting that $N^2k^2L_1 = L_2$ we have the desired result.

In the limit that $k \rightarrow 1$, we arrive at the circuit shown in Fig. 10.4b. It's interesting to note that even if $k = 1$, all real transformers have *leakage inductance*.

The leakage inductance reminds us that a real transformer cannot operate at DC and that it always stores magnetic energy. At DC this shunt inductance prevents any current from flowing into the ideal transformer. The leakage inductance $L_{leak} = k^2L_1$ will form a high-pass network with the source resistance driving port 1 with cutoff frequency R_s/L_{leak} . A decade or so past the cutoff frequency, the leakage inductance has little effect on the performance of the transformer. For a sinusoidal excitation, the magnetic energy stored in this inductor, though, is proportional to the current through the inductor

$$E_L = \frac{1}{2}L_1I_1^2 \approx \frac{1}{2}L_1\left(\frac{V_1}{\omega L_1}\right)^2 \quad (10.16)$$

so the reactive power stored in the inductor is inversely proportional to frequency

$$P_L = \omega E_L \approx \frac{1}{2}\frac{V_1^2}{\omega L} \quad (10.17)$$

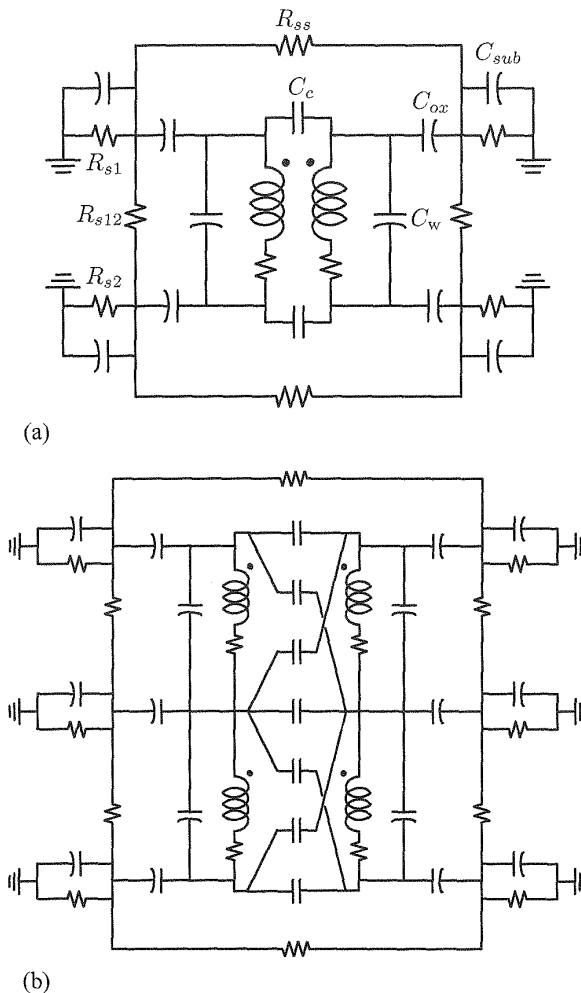


Figure 10.5 (a) Compact model for an integrated transformer including the parasitic electrical coupling. (b) A more distributed model for wideband applications.

In narrow band applications, we may tune out this inductance with a shunt capacitor in order to avoid delivering reactive power to the transformer.

At a given frequency $\omega \neq 0$, we may reduce the effects of the leakage inductance by winding more turns into the coils. Eventually, we are limited by higher-order parasitics, such as winding capacitance, that will limit the frequency response of the transformer in the frequency range of interest.

A more complete model, more useful for computer simulation, is shown in Fig. 10.5a. Here the magnetically coupled inductors are at the core of a more complicated model, which includes the oxide capacitance to the substrate, C_{ox} , substrate resistors, R_{s1} and R_{s2} ,

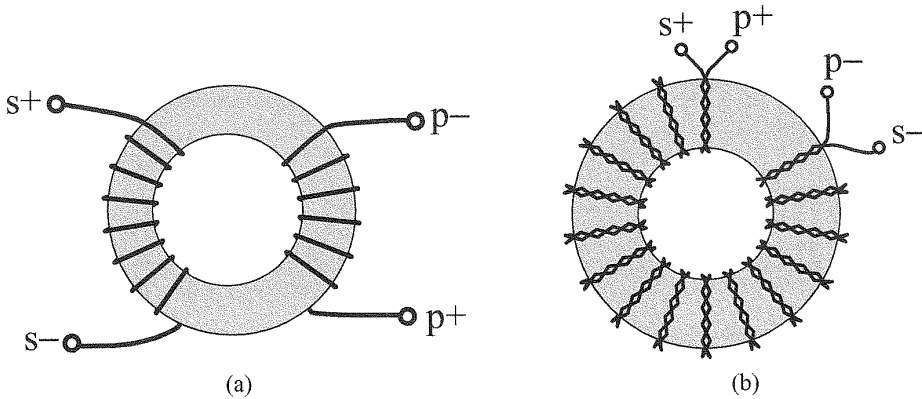


Figure 10.6 (a) A transformer wound on a toroid magnetic core. (b) A 1:1 high-frequency transformer with primary and secondary windings twisted together.

and winding capacitance C_w . All of these parasitic effects were also seen in the inductor compact model. The model includes additional coupling mechanisms, such as the interwinding capacitance C_c , and the coil-to-coil coupling resistance R_{ss} . To capture an even larger frequency range, the model can be broken up into multiple sections in order to model the distributed effects. Such a two-section model is shown in Fig. 10.5b. Such a model can fit measured or 3D EM simulation results over a wide bandwidth. This is important because transformers are often employed for their wideband matching capability, or to add signals at disparate frequencies (say the LO and RF signals), and thus a single compact model should cover a wide frequency range.

10.5 Transformer design and layout

At low frequencies, inductor coils wound on a common magnetic core produce good transformers. The core confines and magnifies the magnetic fields producing large coupling factors k . A typical transformer using a toroid as a core is shown in Fig. 10.6a. Unequal turns ratios are simply produced by winding more turns on side. Another technique prominent in high-frequency applications, shown in Fig. 10.6b, first twists the primary and secondary wires together, which are then wound around the core to produce a 1:1 transformer. To create other turns ratios using the same technique, the primary or secondary can be connected in parallel multiple times. These bifilar structures are generalized to build multifilar transformers having three or more ports.

When confined to planar structures, turns wind inward to produce spirals. Some common transformer layouts are shown in Figs. 10.7–10.10. Two spirals can be wound side by side to form a bifilar transformer as shown in Fig. 10.7a. To create non-unity turns ratio, turns can be avoided in one of the windings. A variation of this bifilar transformer is shown Fig. 10.7b where the primary and secondary windings are wound in a symmetric fashion. This structure preserves symmetry between the primary and secondary windings but each

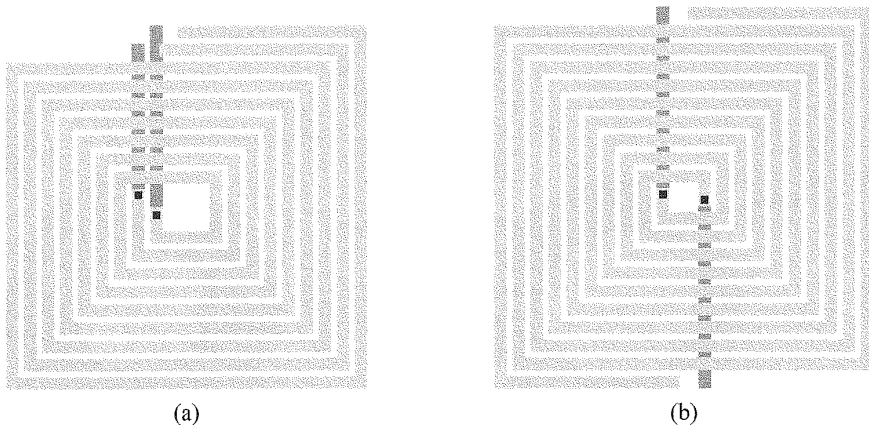


Figure 10.7 (a) Planar bifilar transformer layout. (b) Symmetric bifilar transformer layout.

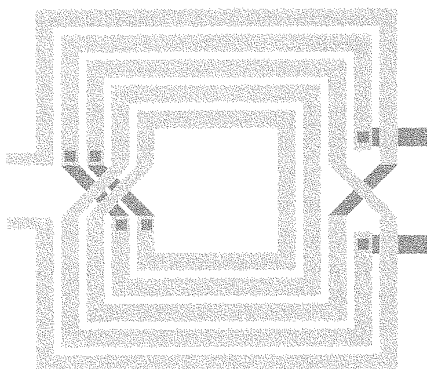


Figure 10.8 Fully symmetric transformer layout. Transformer windings can be tapped at center.

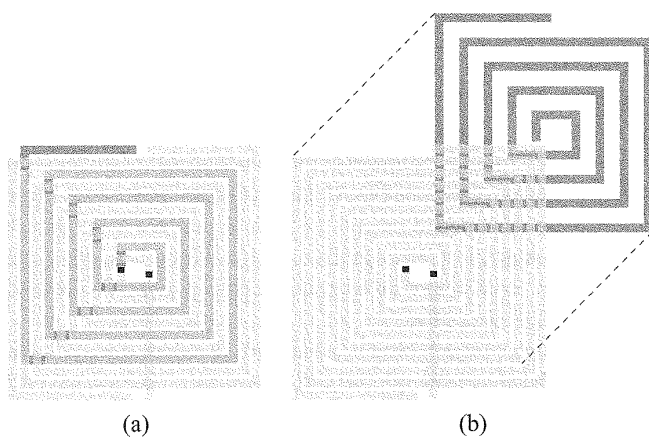


Figure 10.9 (a) Multi-layer transformer layout. The primary is a series-connected structure on two layers. (b) Bottom metal layer is displaced to show structure more clearly.

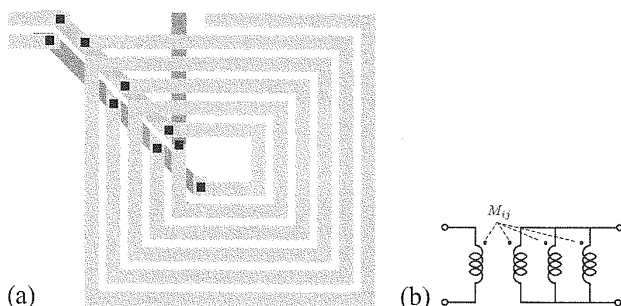


Figure 10.10 Planar spiral transformer with non-unity turns ratio. The secondary turns are connected in shunt with bridge connections.

winding itself is not symmetric. In other words, if a given port is excited by a single-ended signal by grounding the inside port first versus driving it from the inside and grounding the outside port, the input impedance would differ in each case. This is due to the capacitive asymmetries, due the bridge connection and the fringing capacitance (the outer turn sees more fringing capacitance due to the open nature of its surroundings). At higher frequencies, due to current constriction, the inner turns also experience higher resistive losses due to diminishing effective conductor area.

A fully symmetric layout as shown in Fig. 10.8 solves these problems by proper inter-connection of concentric rectangles or circles. Due to the bridge connections, the layout is not perfectly symmetric but the geometric center of each winding is nearly the electrical center of the structure and can be tapped to produce more interesting layout structures.

In a modern IC process, multiple metal layers can be utilized to create even more complicated structures. For instance, as shown in Fig. 10.9, one winding can be made of two inductors in series, whereas another winding can be constructed with an ordinary spiral or multiple spirals in parallel. Thus the inductance of the primary goes up (faster than linearly due to magnetic coupling k), whereas the inductance of the secondary windings remains roughly constant (again due to non-zero magnetic coupling k). This can be exploited to create compact transformers of high effective turns ratio.

To create $N : 1$ turns ratios, many different approaches are possible. As already alluded to, the most obvious approach is to simply leave out inner turns in one winding. Another approach that utilizes area more effectively is to include the inner turns but employs shunt bridge connections to place turns in parallel as opposed to in series. This layout technique is shown in Fig. 10.10a and schematically in Fig. 10.10b.

10.6 Baluns

Since transformers can invert signals for free, a common application is to convert signals from differential to single-ended form. Differential signals are immune to common-mode noise and even-order distortion provided that the process of conversion from differential to single-ended form is done with amplitude and phase balance.

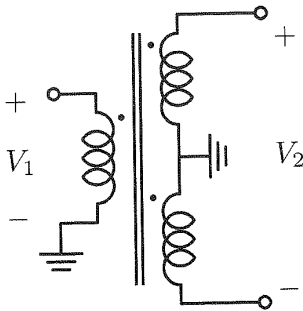


Figure 10.11 A balun converts single-ended signal to a differential or *balanced* signal. The reverse is also easily accomplished.

The balun shown in Fig. 10.11 converts *balanced* signals to *unbalanced* form by employing two coils wound on a common core. The primary coil is grounded and driven by a single-ended signal. The secondary coil, though, is grounded (or tapped) at the center. The voltage pickup on the top half of the secondary coil is therefore in phase with the primary, and the bottom half is 180° out of phase.

The layout of Fig. 10.12a is a common way to layout a balun in an integrated circuit environment, imposing planar geometry. Since it is desirable for ports 2 and 3 to be symmetric, a center tapped spiral is split in half to form two secondary inductors and a third inductor is inter-wound to form the primary. An alternative layout is shown in Fig. 10.12b. Here two metal layers are employed, one layer forming the primary side and a symmetric transformer forming the secondary coil. Laying out a balun with three metal layers, though, would not be symmetric with respect to the primaries due to capacitive and magnetic coupling to the substrate.

10.7 Hybrid transformer

Consider the ideal transformer shown in Fig. 10.13. The turns ratio from primary to secondary is $2N : 1$, but the primary is tapped at the center to produce two transformers of turns ratio $N : 1$. In any practical implementation, the primary windings would also couple to each other, but for now we assume ideal behavior so that $V_1 = V_2 = NV_3$.

Writing loop equations for the transformer, we arrive at [58]

$$V_{s1} - V_{s4} = I_1(Z_1 + Z_4) - I_2Z_4 + NV_3 \quad (10.18)$$

$$-V_{s2} + V_{s4} = -I_1Z_4 + I_2(Z_2 + Z_4) + NV_3 \quad (10.19)$$

$$V_{s3} = I_3Z_3 + V_3 \quad (10.20)$$

Since the transformer is ideal, no power is dissipated in the windings so that

$$V_1I_1 + I_2V_2 + I_3V_3 = 0 \quad (10.21)$$

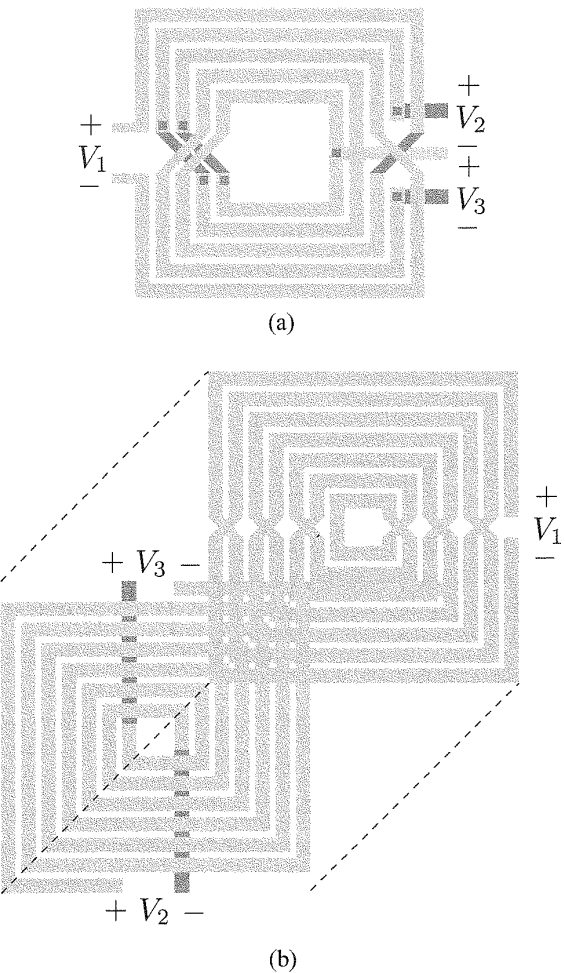


Figure 10.12 (a) A symmetric single-layer transformer balun layout with common center tap. (b) A two-layer balun transformer layout. The primary and secondaries are offset to show geometry more clearly.

substituting for $V_2 = V_1 = NV_3$

$$NV_3(I_1 + I_2) = -I_3V_3 \tag{10.22}$$

or more simply

$$I_3 = -N(I_1 + I_2) \tag{10.23}$$

Equations (10.18), (10.19), and (10.23) can be put into matrix form involving the transformer currents and source voltages. The voltage at the secondary winding can be eliminated by

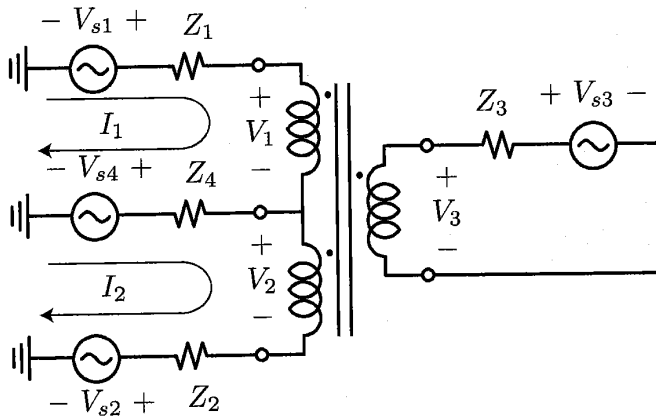


Figure 10.13 A general three-winding transformer.

(10.20).

$$\begin{pmatrix} V_{s1} - V_{s4} - NV_{s3} \\ -V_{s2} + V_{s4} - NV_{s3} \\ 0 \end{pmatrix} = A \begin{pmatrix} I_1 \\ I_2 \\ I_3 \end{pmatrix} \quad (10.24)$$

where

$$A = \begin{pmatrix} Z_1 + Z_4 & -Z_4 & -NZ_3 \\ -Z_4 & Z_2 + Z_4 & -NZ_3 \\ N & N & 1 \end{pmatrix} \quad (10.25)$$

In many applications, it is desirable to isolate certain ports. For instance, in a diode or single transistor mixer, we desire to sum the LO and RF signals to drive a non-linear diode or transistor to perform the mixing function. But if the LO port also couples to the RF port, undesirable LO leakage and radiation will occur. Thus, ideally, a passive device capable of power summation and isolation is desired.

In the three-winding transformer, we can isolate ports 3 and 4 by proper selection of impedances. Without doing the math, it should be fairly obvious that this condition is satisfied if $Z_1 = Z_2$. This occurs since the center tap is a virtual ground and thus isolated. It is not as obvious, though, that ports 1 and 2 can be similarly isolated. Consider the response of I_2 to the source voltage V_{s1} . Using Cramer's rule, we find

$$I_2 = \frac{-V_{s1}(N^2 Z_3 - Z_4)}{|A|} \quad (10.26)$$

The numerator is zero if

$$\frac{Z_4}{Z_3} = N^2 \quad (10.27)$$

By symmetry, of course, port 1 is similarly isolated from port 2.

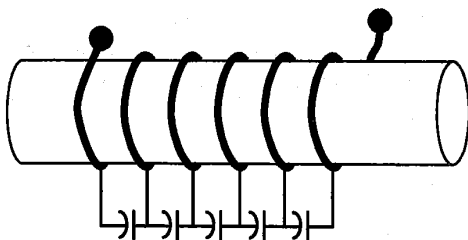


Figure 10.14 The distributed turn-to-turn capacitance in a coil.

Finally, the three-winding transformer can also be designed to provide an impedance match at each port as well as isolation. This occurs when

$$Z_1 = Z_2 = Z \quad (10.28)$$

and

$$Z_3 = \frac{Z_4}{N^2} = \frac{Z^*}{2N^2} \quad (10.29)$$

when these conditions are satisfied, the three-winding structure is commonly known as a hybrid transformer. We discuss hybrid *four-port* devices in more detail in Chapter 15.

10.8 Transformer parasitics

At high frequencies many problems begin to plague transformers. Recall that parasitic inter-turn electric coupling occurs in inductors and ultimately limits the performance or the resonant frequency of inductors. Even air wound inductors have turn-to-turn coupling, as shown in Fig. 10.14, limiting the high-frequency performance. For inductors, though, this problem is mild since these inter-turn capacitors are in effect shorted by the turn-to-turn inductance. As long as the turn-to-turn inductance impedance is small compared with the capacitive impedance, the structure will behave as an inductor.

But for multifilar transformers, the turn-to-turn coupling is not shorted by an inductor since one turn is on the primary and one is on the secondary. Thus the capacitive coupling directly competes with the magnetic coupling. In fact, if we model the system as a distributed circuit, the transformer is essentially a non-uniform transmission line. This idea is exploited by turning the transformer 90° to form a transmission line transformer. See Section 11.2 for details. The Z parameters of the distributed transformer deviate from ideal behavior by creating a real component in Z_{21} . At low frequencies, the coupling is close to the ideal value of $j\omega M$ but at frequencies approaching the resonant frequency of the structure, a large deviation is observed.

10.9 Transformer figures of merit

Transformers are very convenient elements for a host of applications. Their downside is their bulky size and frequency limitations. For on-chip applications, the insertion loss of a

transformer is an important metric. Since the insertion loss of a transformer depends highly on the source and load impedance, it's convenient to find the minimum insertion loss under ideal conditions, a bi-conjugate source and load match. Under such conditions, we know that the gain of the transformer peaks at

$$G_{max} = \frac{y_{21}}{y_{12}} (K_s - \sqrt{K_s^2 - 1}) = K_s - \sqrt{K_s^2 - 1} \quad (10.30)$$

since $y_{21} = y_{12}$ for a passive device. The gain is thus related to the stability factor K_s , which is by definition $K_s \geq 1$ for a passive device. When $K_s = 1$, the insertion loss is 0 dB. The K_s factor is easily computed from the Z parameters of the transformer

$$Z = \begin{pmatrix} R_p + j\omega L_p & j\omega M \\ j\omega M & R_s + j\omega L_s \end{pmatrix} \quad (10.31)$$

where

$$K_s = \frac{2\Re(z_{22})\Re(z_{11}) - \Re(z_{21}z_{12})}{|z_{21}z_{12}|} \quad (10.32)$$

One particularly simple case of interest is a 1 : 1 transformer, where each winding has a resistance R_x

$$K_s = \frac{2R_x^2 + \omega^2 M^2}{\omega^2 M^2} = \frac{2R_x^2 + \omega^2 k^2 L^2}{\omega^2 k^2 L^2} \quad (10.33)$$

The maximum gain can be parameterized by the unitless $Q = \omega L / R_x$ factor and the transformer coupling factor $K = M / L_x$

$$G_{max}(Q, k) = 1 + \frac{2}{Q^2 k^2} - 2\sqrt{\frac{1}{Q^4 k^4} + \frac{1}{Q^2 k^2}} \quad (10.34)$$

A plot of G_{max} is shown in Fig. 10.15, which shows that reasonably low insertion loss can be obtained with relatively modest Q factor and coupling factor k . For instance, a $Q = 10 - 15$ is easily achieved in the 1 – 5 GHz frequency range. On-chip planar transformers have $k \approx 0.75$, which implies an insertion loss of about 0.8, or about 1 dB insertion loss. While the above result is derived for a simple 1 : 1 transformer, the result can be generalized if we simply notice that N windings driven in shunt on the primary and driven in series on the secondary will transform the impedance by a factor of N , as shown in Fig. 10.16. Furthermore, the insertion loss is independent of the number of sections. This is a powerful result because for LC -matching networks we found a strong dependence on the matching ratio (Eq. (7.113)). This was also true for transmission line quarter wave-matching sections (Eq. (9.116)).

Another important figure of merit for the transformer is the maximum frequency that one can safely call a transformer by its name, before it resonates and degenerates into another structure. Consider a pair of coupled inductors forming a transformer with equal capacitive primary and secondary loads, as shown in Fig. 10.17. Let's calculate the Z matrix of this structure directly by solving the three mesh currents $i_1 - i_3$ for a source current i_s at port 1

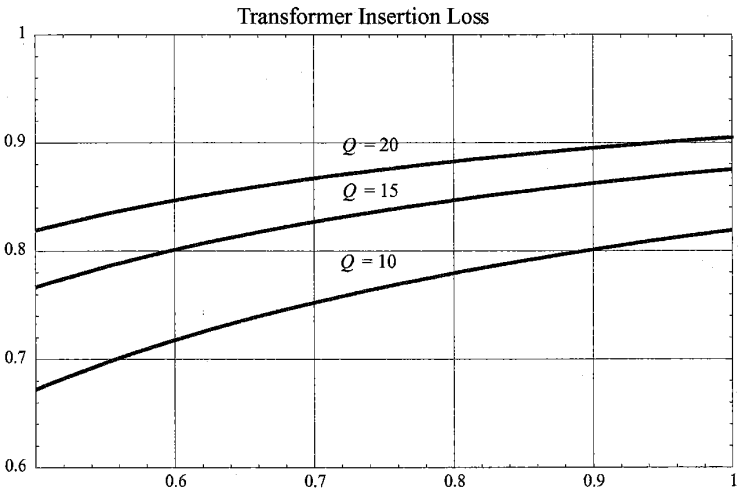


Figure 10.15 The minimum insertion loss of a 1 : 1 transformer as a function of magnetic coupling factor K and winding quality factor Q .

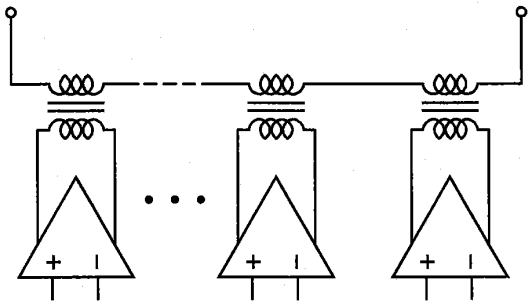


Figure 10.16 Transformer power combiner employing core 1 : 1 sections. For N stages the impedance is transformed down by N for each driver.

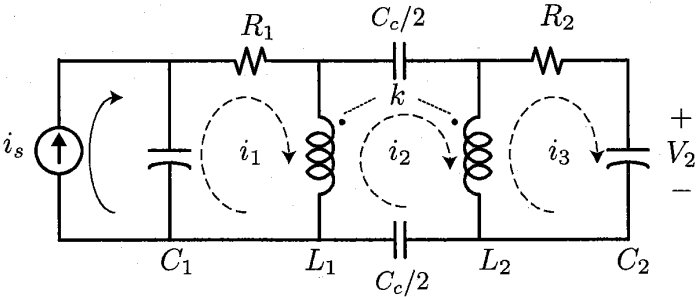


Figure 10.17 Lumped equivalent circuit model for a transformer including the winding capacitance and the coupling capacitance C_c .

and an open-circuited port 2. The mesh equations are given by

$$0 = (i_1 - i_s) \frac{1}{sC_1} + i_1 R_1 + (i_1 - i_2) s L_1 + (i_2 - i_3) s M \quad (10.35)$$

$$0 = (i_2 - i_1) s L_1 - (i_2 - i_3) s M + i_2 \frac{1}{sC_c} + (i_2 - i_3) s L_2 + (i_1 - i_2) s M \quad (10.36)$$

$$0 = (i_3 - i_2) s L_2 - (i_1 - i_2) s M + i_3 R_2 + i_3 \frac{1}{sC_2} \quad (10.37)$$

To find Z_{11} , we calculate the voltage at port 1

$$Z_{11} = \left. \frac{V_1}{I_1} \right|_{I_2=0} = \frac{i_s - i_1}{sC_1 i_s} \quad (10.38)$$

Likewise, we compute Z_{21}

$$Z_{21} = \left. \frac{V_2}{I_1} \right|_{I_2=0} = \frac{i_3}{sC_2 i_s} \quad (10.39)$$

Since the structure is symmetric, Z_{2k} can be obtained by simply interchanging subscripts. The resulting solutions, while straightforward to derive, are too long to derive by hand. Instead, a symbolic analysis program is used.

To build up to the complexity of the full solution, let's start with a simple case with $C_c = 0$, or no capacitive coupling. In particular, the fully symmetric case, with $L_1 = L_2 = L$, $C_1 = C_2 = C$, and $R_1 = R_2 = R$, is easy to solve (only two equations) and particularly insightful. In this case we have

$$Z_{11} = \frac{C(L - M)(L + M)s^3 + 2CLR s^2 + CR^2 s + Ls + R}{(Cs(R + (L - M)s) + 1)(Cs(R + (L + M)s) + 1)} \quad (10.40)$$

and

$$Z_{21} = \frac{sM}{(Cs(R + (L - M)s) + 1)(Cs(R + (L + M)s) + 1)} \quad (10.41)$$

A plot of the magnitude of Z_{11} is shown in Fig. 10.18a. Notice that there two resonant frequencies are evident. A plot of the phase of Z_{21} , which is 90° for an ideal transformer, is shown in Fig. 10.18b, showing a large deviation from ideal behavior near the first resonant frequency.

The poles of the system are given by roots of the denominator of the common term $(Cs(R + (L - M)s) + 1)(Cs(R + (L + M)s) + 1)$. For $R = 0$, we see immediately that these frequencies are given by

$$\omega^+ = \frac{1}{C(L + M)} = \frac{1}{LC(1 + k)} \quad (10.42)$$

and

$$\omega^- = \frac{1}{C(L - M)} = \frac{1}{LC(1 - k)} \quad (10.43)$$

Since this structure has two degrees of freedom, it has two natural modes of oscillation. These fundamental modes are defined by resonance and anti-resonance. These modes are

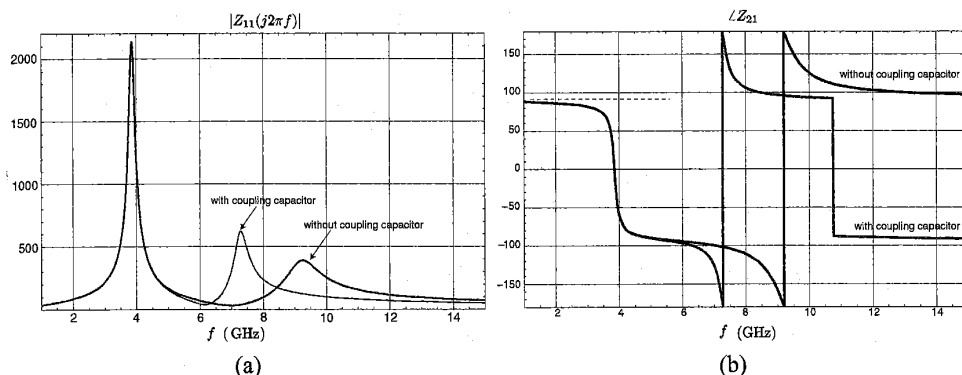


Figure 10.18 (a) The magnitude of Z_{11} for a lumped transformer with winding and coupling capacitance. (b) The phase of Z_{12} for a lumped transformer. The curve with the higher second resonance frequency corresponds to the case with $C_c = 0$, or no coupling capacitance. As expected, the coupling capacitor does not disturb the even mode.

also called the *even* and *odd* mode, due to the symmetry of the excitation. Yet a third name for this is the *differential* or *balanced* mode versus *common* mode.

The lower-frequency resonance occurs in the even mode as the primary and secondary currents flow in phase so as to maximize the magnetic energy in the system. The net inductance, therefore, is $L_1 + L_2 + 2M$. The capacitance in even mode is the winding capacitance to ground, $C_e = C$. It is interesting to note that in even mode we can place a coupling capacitor C_c across the coupled inductors without disturbing the mode voltages and currents, since the net voltage across this capacitor is zero. Thus the lower-frequency resonance mode of the transformer is given by the *self*-winding capacitance and parasitic capacitance to ground, rather than the coupling capacitance. We can thus predict the even mode resonant frequency ω_e as

$$\omega^+ = \sqrt{\frac{1}{C_0(L_1 + L_2 + 2M)}} \quad (10.44)$$

In anti-resonance, or odd mode, the magnetic energy is minimized by circulating current into the windings out of phase. The capacitance of this mode, though, includes the interwinding capacitance $C_o = C + 2C_c$. The factor of two accounts for the Miller effect. The frequency of this mode is therefore

$$\omega^- = \sqrt{\frac{1}{(C + 2C_c)(L_1 + L_2 - 2M)}} \quad (10.45)$$

If $C_c \ll C_0$, we see that the even mode resonance has a lower cutoff frequency. For most integrated transformers, this is indeed the case and so the culprit is the interwinding capacitance.

One of the limitations of the above analysis is the coupling capacitor is treated as a lumped element. In reality, the capacitive coupling is distributed and intermingled with the

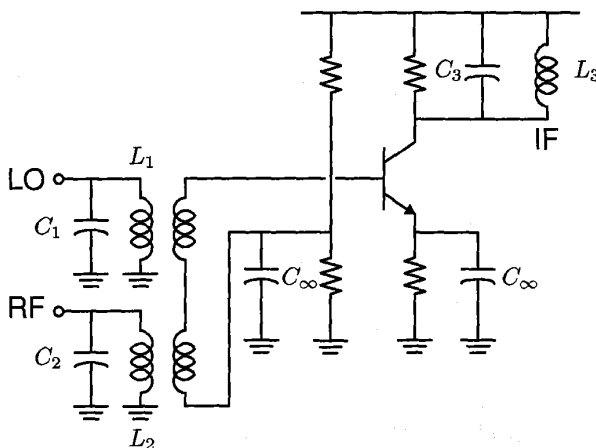


Figure 10.19 A bipolar single transistor mixer employs a transformer to sum the RF and LO signals at the base of the transistor.

magnetic coupling. A full analysis of the problem, therefore, requires distributed analysis, a topic we will cover in Chapter 11.

10.10 Circuits with transformers

Mixers

Integrated transformers are increasingly employed in integrated circuits to solve a variety of problems. The natural differential operation of the transformer is a big advantage in the design of mixers, where balanced operation results in fewer spurious tones and higher power gain. Additionally, the inherent power combining of a transformer can be used to combine the LO and RF signals in a single transistor mixer, as shown in Fig. 10.19. While a simple capacitive transformer is more compact, the transformer can be used as a hybrid to provide LO and RF isolation.

In a Gilbert cell single-balanced mixer, shown in Fig. 10.20a, the transformer can be used to boost the effective transconductance of the input stage M1 by the turns ratio of the device. Furthermore, since the switching pair operates with a DC grounded source, there is plenty of headroom on the drain, which can be employed to boost the gain of the stage by cascoding the stage or by simply increasing the load resistance. In the double-balanced mixer shown in Fig. 10.20b, a balun converts the single-ended RF signal into a balanced signal, eliminating the transconductance stage altogether [36]. Since the transformer is a linear device, the distortion of the mixer is low, especially the $IIP2$ due to the balanced drive. At the output the differential signal is also converted back to a single-ended signal. Single-ended drive is often required to drive an off-chip component, such as a filter.

The classic diode ring mixer shown in Fig. 10.21, is still the preferred choice for extremely linear applications. Furthermore, since diodes work into the THz regime, this mixer can operate at extremely high frequencies. As a passive mixer, the mixer has an insertion loss

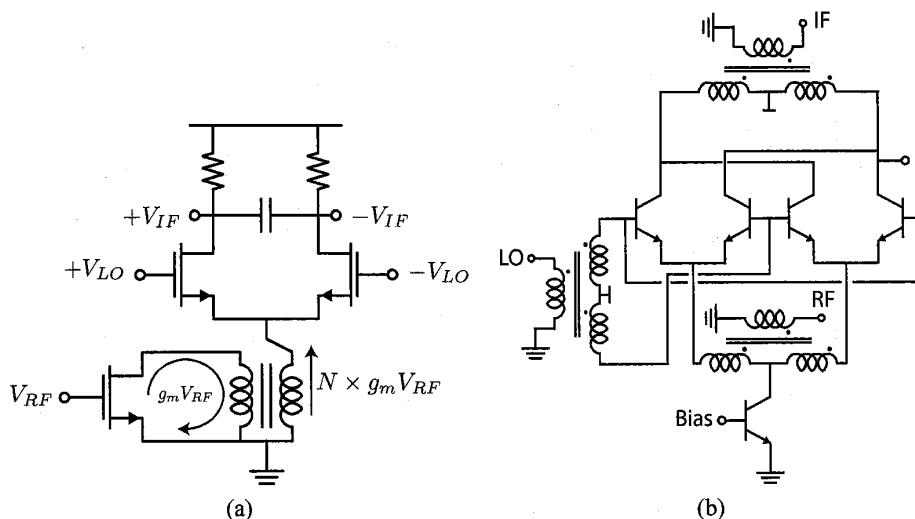


Figure 10.20 (a) A transformer is employed to boost the transconductance of a mixer. At the same time it allows the circuit to operate on a lower supply voltage. (b) A fully balanced version of the Gilbert mixer employs transformers to convert single-ended RF, LO, and IF to a differential circuit. Center taps are used to deliver bias voltages.

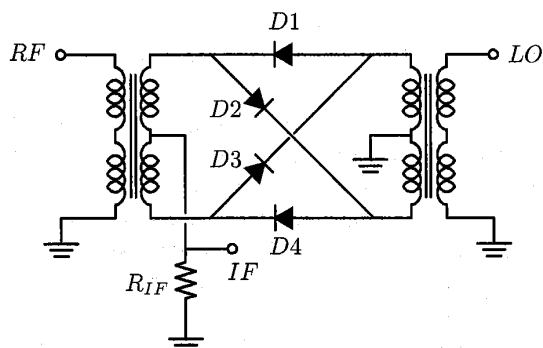


Figure 10.21 A diode ring mixer.

of about 6 dB. The mixer, though, has good isolation, where the $LO \longleftrightarrow RF$ isolation comes from the transformers. The balanced operation also results in RF and LO rejection. The downside is that the LO signal drive must be large since it must turn on/off diodes. Thus the power consumption in the LO port can be quite high.

The ring mixer is a fully balanced mixer but we have the option of driving the LO and RF single ended. As we shall see, the LO signal alternating turns on diodes $D1/D2$ or $D3/D4$, thereby connecting the RF voltage to the IF with alternative polarity. During a positive LO voltage, the secondary of the LO transformer applies a positive voltage across $D1$ in series with the load R , thereby forward biasing the diode. Likewise, the secondary terminal of the

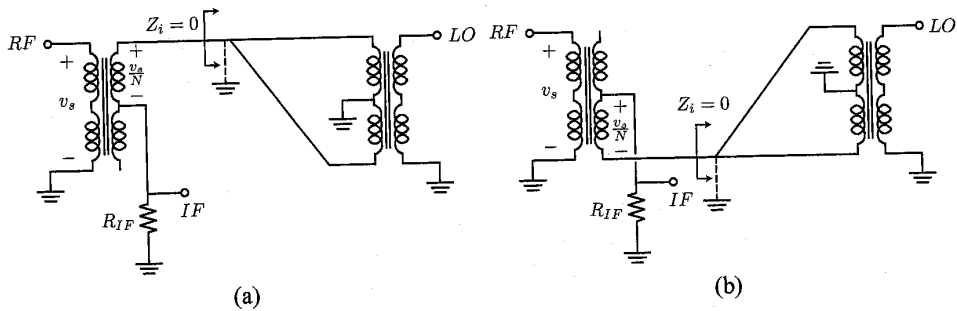


Figure 10.22 Equivalent circuit for the diode ring mixer during the (a) positive and (b) negative LO.

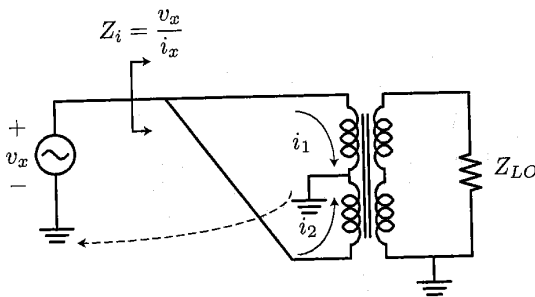


Figure 10.23 Calculation of the input impedance seen by the RF port of the diode ring mixer.

LO transformer applies a negative voltage to the cathode of D2, thereby forward biasing it in series with the load. Diodes D3/D4 are reverse biased and therefore open circuits (ideally).

The equivalent circuit for the positive LO cycle is shown in Fig. 10.22a, where the diode on-resistance is assumed to be zero for simplicity. The RF signal is applied to the load through the center tap of the transformer. We shall show that the impedance looking into the transformer is low (ideally zero), and thus the entire RF signal at the secondary terminal is applied to the load $v_o = -v_s/N$.

Note the input impedance looking into the LO transformer is ideally short circuit, as shown in Fig. 10.23. To see this, apply a test voltage v_x to this node. If we ignore the effect of the LO signal for now, we see that two equal currents $i_1 = i_2 = i_x/2$ flow into the secondary of the LO transformer. The return current comes from the center tap ground.

The induced voltage at the primary is given by

$$v_1 = j\omega(M_{12} - M_{13})i_x/2 \equiv 0 \quad (10.46)$$

This is because $M_{13} = M_{12}$ and the currents on the secondary are out of phase. The voltages induced on the secondary side are likewise zero (assuming perfect coupling)

$$v_2 = j\omega L_2 i_x / 2 + j\omega M_{23}(-i_x/2) \equiv 0 \quad (10.47)$$

$$v_3 = j\omega L_3(-i_x/2) + j\omega M_{32}i_x/2 \equiv 0 \quad (10.48)$$

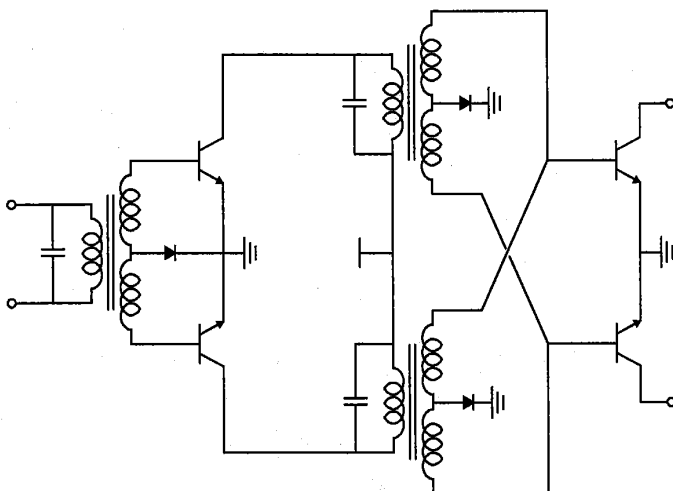


Figure 10.24 A fully balanced bipolar power amplifier employing resonant transformers [65].

By symmetry, during the negative LO cycle, the mixer simplifies to the equivalent circuit shown in Fig. 10.22b. During this cycle, the RF signal is applied to the load through the center tap and the bottom of the transformer, thus producing an output signal in phase with the RF

$$v_o = +v_s/N \quad (10.49)$$

The operation is thus similar to the double-balanced Gilbert cell mixer in voltage mode. The RF voltage is multiplied by ± 1 with a rate of the LO signal. The lack of DC means that the RF is rejected at the IF port. Likewise, the LO/RF signals are isolated by the switches. The only feed-through occurs due to reverse isolation of the diodes. The diode ring mixer is very linear and quite attractive in applications where linearity reigns supreme over power consumption. At microwave frequencies, the transformers can be replaced by couplers (see Chapter 15). Since diodes can operate up to extremely high frequencies (THz), the entire circuit can work up to THz region.

Interstage power matching

In the balanced power amplifier shown in Fig. 10.24, transformers are used for the interstage matching [65] and power combining. The output stage consists of large bipolar devices capable of driving more than 5 W of power into a 50Ω (transformed) load impedance. Large devices have high-input capacitance and require a large base current drive. Two parallel transformer windings at the output provide the current drive from two windings at the input. This allows a high-voltage swing to be maintained at the driver for maximal efficiency. By employing a capacitor in shunt with the transformer, the current gain of the transformer is increased by the Q factor of the network. To see this, consider the equivalent circuit shown

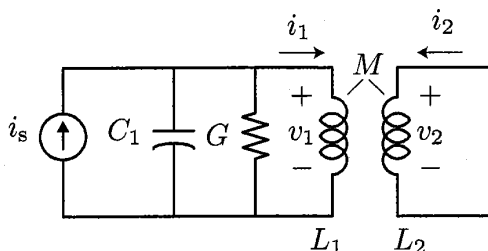


Figure 10.25 A resonant transformer boosts the current gain by the Q factor of the structure.

in Fig. 10.25, where for simplicity the secondary has been shorted. The current gain in the transformer is simply given by solving

$$v_2 = 0 = j\omega M i_1 + j\omega L_2 i_2 \quad (10.50)$$

$$\frac{i_2}{i_1} = -\frac{M}{L_2} = -k \sqrt{\frac{L_1}{L_2}} \quad (10.51)$$

The current i_1 is the source current minus the current in the capacitor

$$i_1 = i_s - G v_1 - j\omega C v_1 = i_s - (G + j\omega C)(j\omega M i_2 + j\omega L_1 i_1) \quad (10.52)$$

The conductor G represents the total loss of the transformer reflected to the primary. In terms of the coupling factor k

$$i_1 = i_s - (G + j\omega C)j\omega L_1(1 - k^2) \quad (10.53)$$

The current i_s is therefore given by

$$i_s = i_1(1 - \omega^2 L_1 C_1(1 - k^2) + j\omega L_1 G) \quad (10.54)$$

At the resonant frequency $\omega_0 = 1/\sqrt{L_1 C_1(1 - k^2)}$, the first two terms cancel and we have

$$i_1 = i_s \frac{R}{j\omega L_1} = -jQ i_s \quad (10.55)$$

If the Q factor of the transformer is large, the effective current boost is given by the turns ratio *times* the Q factor, greatly reducing the drive current.

Power combining

As already noted, 1:1 transformers are also good elements for power combining on-chip. As shown in Fig. 10.26a, the pseudo-differential pairs drive the primary coils in a balanced fashion on a low supply voltage. The output loop is magnetically coupled to differential pair loops and has N times the flux, and hence N times the voltage. This voltage step-up is a very useful property as the breakdown voltage of high-frequency transistors is low, approaching sub-1 V in a modern CMOS process. A potential layout of such a structure is shown in Fig. 10.27a. Two thick overlapping metal layers are used to build the transformer to minimize the losses and to maximize the coupling factor k .

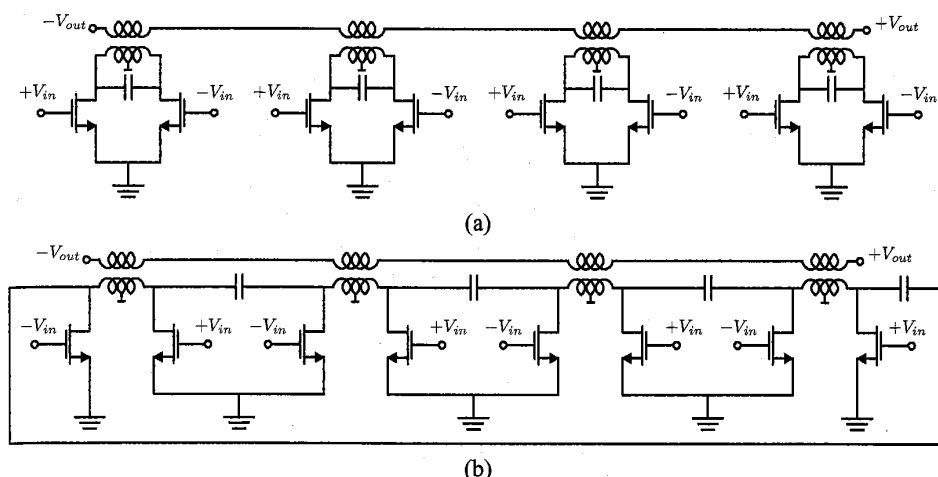


Figure 10.26 (a) The pseudo-differential pairs drive transformer primaries over a low voltage. The secondaries are connected in series to drive a high voltage on to the load. (b) In the DAT technique [1], the adjacent cells are driven differentially to form virtual grounds, allowing one to AC short the transformer connections and thus eliminate the lead inductance.

Notice that the inner turns of Fig. 10.27a do not contribute magnetic flux to the output since the currents are flowing in opposite directions, whereas they contribute to extra loss due to the resistivity of the metal layers. If possible, then, these extra segments should be eliminated. This is precisely the approach shown in Fig. 10.26b, and in the layout of Fig. 10.27b, where the input devices are connected in a circular fashion in order to form a closed loop [1]. By virtue of the virtual ground nodes created from the differential drive, the extra lead inductance in driving each sub-loop is eliminated. This results in enhanced power, combining efficiency and a more symmetric layout.

Magnetic feedback

In Fig. 10.28, the transformer is embedded in a differential LNA [36]. Here the transformer serves two functions. The primary windings are used to degenerate the input device and aid matching. The secondary winding inductance serves as a tuned load for the LNA. Since a cascode device is not used, the linearity of the LNA is very good due to the high voltage swing at the output. Without a cascode, though, the stability is an issue due to the feedback capacitance C_{gd} . Here the magnetic coupling is used to neutralize this feedback, improving the stability of the amplifier. Since the amplifier is inverting, the transformer windings are used in inverting configuration in order to produce a signal in phase with the input.

Another application of magnetic feedback is in the voltage controlled oscillator (VCO) shown in Fig. 10.29 [38] [44]. In a typical cross-coupled VCO, shown Fig. 10.29a, capacitors commonly form the feedback. Since the BJT devices need base current, a separate DC biasing scheme is necessary to bias the devices at the right voltage and to provide DC current. To avoid noise from the biasing circuitry, large resistors are often employed. But

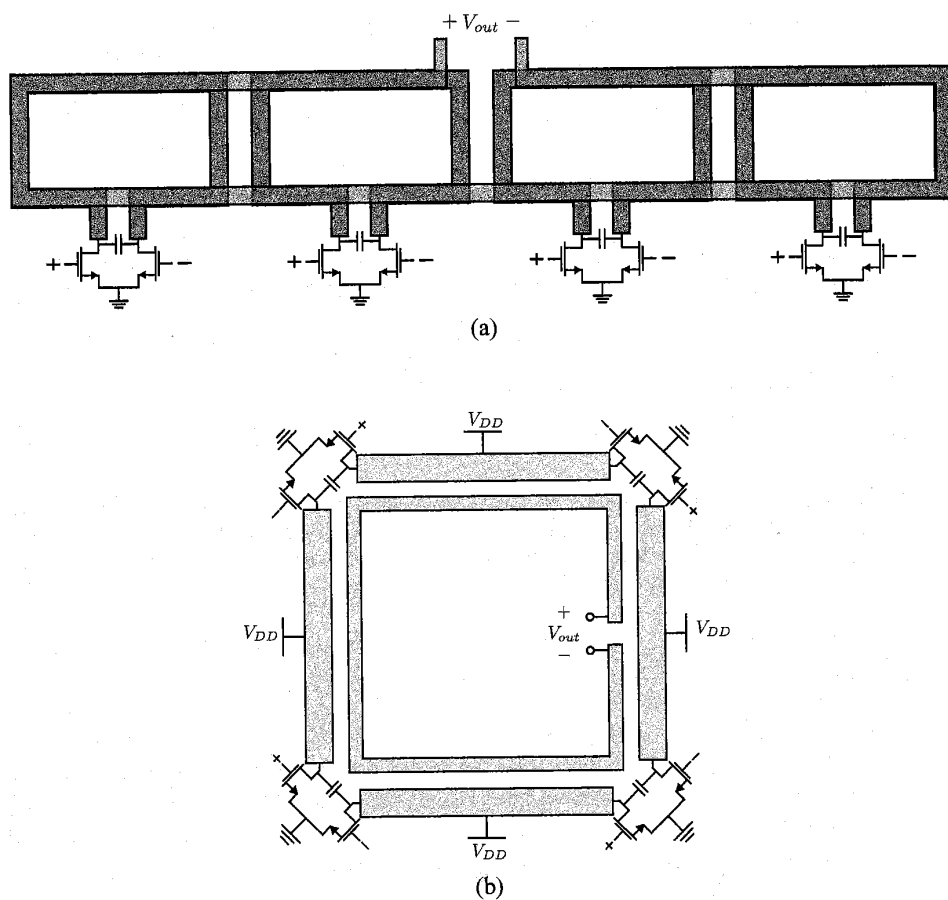


Figure 10.27 (a) A potential layout for the transformer power-combining technique shown in Fig. 10.26a. (b) The ring layout of the DAT structure of Fig. 10.26b.

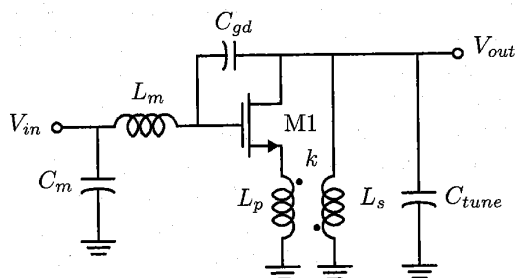


Figure 10.28 AC schematic of LNA neutralized by the transformer magnetic feedback [36].

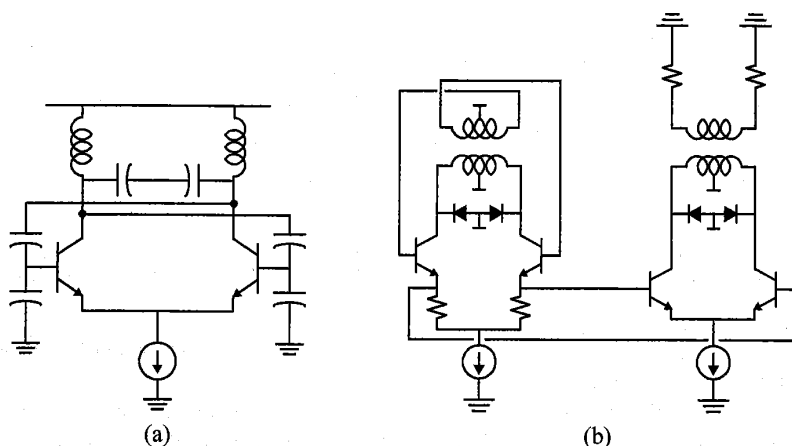


Figure 10.29 (a) A differential oscillator using a capacitive divider for positive feedback. (b) A differential oscillator using a transformer for positive feedback [44].

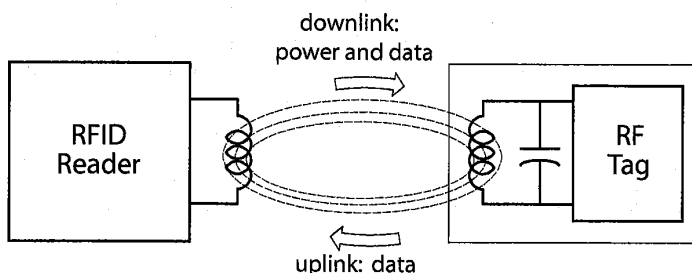


Figure 10.30 A typical RFID system employing magnetic coupling.

since the base current of a BJT varies a lot over process and temperature, to provide a steady voltage at the base is challenging [37]. The transformer feedback circuit of Fig. 10.29b solves this problem since a DC voltage can be easily delivered to the center tap of the primary and secondary. Since the resistance of the windings is small, the variation in base current has a negligible impact on the DC bias voltage. A second differential pair driven by the emitters of the VCO core provides a buffered output. The buffer uses a tuned balun to drive the off-chip load.

RFID

In the RFID system shown in Fig. 10.30, the RF tag is usually a flexible small label that can be attached to an object. The tag “antenna” is usually a printed square spiral inductor which forms the secondary winding of a transformer. The primary coil is a larger coil on a “reader” device. In the near field of the coil, the magnetic field of the “reader” generates sufficient flux to induce a sufficiently large AC voltage at the secondary winding. This voltage is

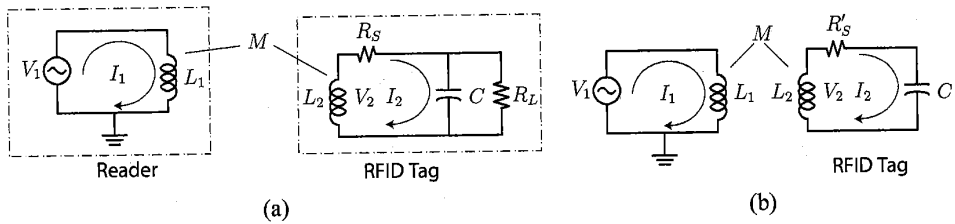


Figure 10.31 (a) Equivalent circuit for RFID tag and reader communication and power link. (b) Simplified circuit where the tag load resistance has been absorbed into the winding resistance.

rectified and used to power the RFID tag, and so the voltage needs to be large enough to turn on an on-chip diode.

To maximize the range, we usually employ a resonant circuit at the secondary. To see the benefit of a resonant secondary, consider the voltage induced on the secondary winding as shown in Fig. 10.31a

$$V_2 = j\omega MI_1 + j\omega L_2 I_2 = I_2(R_S + R_L || j\omega C)$$

It's convenient to transform the load from a shunt RC circuit to a series RC circuit, as shown in Fig. 10.31b. The load resistance is now easily absorbed into the winding resistance R'_S

$$V_2 = j\omega MI_1 + j\omega L_2 \frac{V_2}{R'_S + \frac{1}{j\omega C}}$$

The transfer function from the primary current to the secondary voltage is therefore given by

$$\frac{V_2}{I_1} = \frac{j\omega M(R'_S + \frac{1}{j\omega C})}{R_S + \frac{1}{j\omega C} - j\omega L_2}$$

If the load R_L is relatively large, then the load voltage without resonance ($C = 0$) is given by $j\omega MI_1$. In the resonant circuit, though, the maximum load voltage is at a frequency $\omega_0 = \frac{1}{L_2 C}$

$$\frac{V_2}{I_1} = j\omega_0 M \left(1 + \frac{1}{j\omega_0 C R'_S} \right) = j\omega_0 M (1 + Q')$$

If a large Q' factor can be maintained, there is a substantial improvement in the range of the system. One disadvantage of the resonant system is the potential loss in range due to frequency misalignment. Due to process variations, the resonant frequency of the tags will vary. The reader device can "search" for the tag by sweeping its frequency across a wide range and observing the drive current, shown for a particular case in Fig. 10.32a. The frequency of minimum drive current corresponds to the tag resonant frequency.

Once a power link has been established, the tag can communicate with the reader in various ways. A common approach places a switch across the terminals of the tag to modulate the load resistance of the tag. Since the reader current is a function of the tag resistance R_L ,

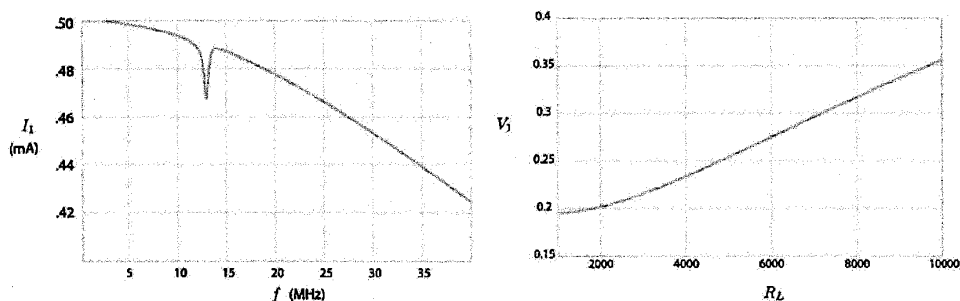


Figure 10.32 (a) The current drive of the primary reader as a function of frequency. Note the dip at resonance. (b) The reader current as a function of the load resistance.

the variation in reader current can be amplified and detected. A plot of the primary current as a function of R_L is shown in Fig. 10.32b.

Example 21

RFID using magnetic coupling

A popular band for RFID is 13.56 MHz. At a distance of 30 cm, the coupling factor k between the primary and secondary is $k = 0.01$. Assume the RFID tag requires 1 mA of current and 1 V DC to operate. The average equivalent resistance of the tag is then 1 k Ω . Let's assume this factor dominates the quality factor. For a tag with $L = 0.5 \mu\text{H}$, the Q factor at resonance is $Q = 1000 / (13.56 \times 0.5 \times 2\pi) = 23.5$. If we assume the reader device has ten times more inductance, then

$$M = k\sqrt{L_1 L_2} = 0.01 \times \sqrt{0.5 \times 5} \mu\text{H} = 0.016 \mu\text{H}$$

For a reader current of $I_1 = 100 \text{ mA}$, the induced voltage at the secondary is given by

$$V_2 \approx I_1 j\omega M(1 + Q') = 100 \text{ mA} \times 13.56 \times 2\pi \times 0.016(1 + 23.5) \text{ V} = 3.3 \text{ V}$$

which is sufficient for the tag to operate. Note that without resonance the induced voltage is only 136 mV, below the diode turn-on voltage.

10.11 References

Some of the material of this chapter comes from *Modern Communication Circuits* [58]. The application of transformers for power combining and impedance matching is drawn from seminal papers by the Cal Tech team [1] and from Simburger [65], as well as our own research efforts at Berkeley [35]. For an in-depth discussion of transformers in an integrated circuit environment, I highly recommend the thesis and research papers by John Long [36].