DESIGN, SIMULATION AND APPLICATIONS OF INDUCTORS AND TRANSFORMERS FOR SI RF ICS

ALI M. NIKNEJAD
Silicon Laboratories, Austin TX

ROBERT G. MEYER
University of California, Berkeley



Kluwer Academic Publishers
Boston/Dordrecht/London

DEL
TK 7872
I63 N55
2000

Distributors for North, Central and South America:
Kluwer Academic Publishers
101 Philip Drive
Assinippi Park
Norwell, Massachusetts 02061 USA
Telephone (781) 871-6600
Fax (781) 871-6528
E-Mail <kluwer@wkap.com>

Distributors for all other countries:
Kluwer Academic Publishers Group
Distribution Centre
Post Office Box 322
3300 AH Dordrecht, THE NETHERLANDS
Telephone 31 78 6392 392
Fax 31 78 6546 474
E-Mail <orderdept@wkap.nl>
Electronic Services <http://www.wkap.nl>

Library of Congress Cataloging-in-Publication Data

Niknejad, Ali M., 1972-
Design, simulation and applications of inductors and transformers for Si RF ICs / Ali M. Niknejad, Robert G. Meyer.
p. cm. -- (Kluwer international series in engineering and computer science. Analog circuits and signal processing)
Includes bibliographical references and index.
ISBN 0-7923-7986-1 (alk. paper)
1. Electric inductors--Computer-aided design. 2. Electronic transformers--Computer-aided design. 3. Silicon-on-insulator technology. 4. Microwave integrated circuits--Computer-aided design.
I. Meyer, Robert G., 1942- II. Title. III. Series

TK7872.I63 N55 2000
621.381--dc21

00-062480
CIP

Copyright © 2000 by Kluwer Academic Publishers..

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, mechanical, photo-copying, recording, or otherwise, without the prior written permission of the publisher, Kluwer Academic Publishers, 101 Philip Drive, Assinippi Park, Norwell, Massachusetts 02061

Printed on acid-free paper.
Printed in the United States of America

This material may be protected by Copyright law (Title 17 U.S. Code)

Contents

List of Figures ix
List of Tables xv
Preface xvii
Acknowledgments xix

Part I Analysis and Simulation of Passive Devices

1. INTRODUCTION 3
   1. Introduction 3
   2. Passive Devices in Early Integrated Circuits 3
   3. Applications of Passive Devices 4
   4. Wireless Communication 6
   5. Si Integrated Circuit Technology 8
   6. Contributions of this Research 9

2. PROBLEM DESCRIPTION 11
   1. Definition of Passive Devices 11
   2. Loss Mechanisms 15
   3. Device Layout 21
   4. Substrate Coupling 35

3. PREVIOUS WORK 39
   1. Early Work 39
   2. Passive Devices on the GaAs substrate 40
   3. Passive Devices on the Si substrate 40
   4. Passive Devices on Highly Conductive Si Substrate 43

4. ELECTROMAGNETIC FORMULATION 45
   1. Introduction 45
   2. Maxwell's Equations 45
   3. Calculating Substrate Induced Losses 48

4. Inversion of Maxwell's Differential Equations 50
5. Numerical Solutions of Electromagnetic Fields 52
6. Discretization of Maxwell's Equations 53

5. INDUCTANCE CALCULATIONS 59
1. Introduction 59
2. Definition of Inductance 60
3. Parallel and Series Inductors 65
4. Filamental Inductance Formulae for Common Configurations 66
5. Calculation of Self and Mutual Inductance for Conductors 67
6. High Frequency Inductance Calculation 68

6. CALCULATION OF EDDY CURRENT LOSSES 75
1. Introduction 75
2. Electromagnetic Formulation 76
3. Eddy Current Losses at Low Frequency 84
4. Eddy Currents at High Frequency 88
5. Examples 93

7. ASITIC 97
1. Introduction 97
2. ASITIC Organization 99
3. Numerical Calculations 100
4. Circuit Analysis 101

8. EXPERIMENTAL STUDY 109
1. Measurement Results 109
2. Device Calibration 110
3. Single Layer Inductor 113
4. Multi-Layer Inductor 119

Part II Applications of Passive Devices

9. VOLTAGE CONTROLLED OSCILLATORS 125
1. Introduction 125
2. Motivation 127
3. Passive Device Design and Optimization 128
4. VCO Circuit Design 134
5. VCO Implementation 141
6. Measurements 144
7. Conclusion 146

10. DISTRIBUTED AMPLIFIERS 149
1. Introduction 149
2. Image Parameter Method 150

3. Distributed Amplifier Gain 156

11.CONCLUSION 163

1. Future Research 164

Appendix

A– Distributed Capacitance 167

Chapter 1

INTRODUCTION

1. INTRODUCTION

The wireless communication revolution has spawned a revival of interest in the design and optimization of radio transceivers. Passive elements such as inductors, capacitors, and transformers play a critical part in today's transceivers. In this book we will focus on the analysis and applications of such devices. In particular, we will demonstrate techniques for calculating the loss when such devices are fabricated in the vicinity of conductive substrates such as Silicon.

2. PASSIVE DEVICES IN EARLY INTEGRATED CIRCUITS

Until recently, passive devices, especially in integrated form, played a relatively minor role in Si integrated circuits in comparison with active devices such as transistors. The most important reason for this can be attributed to the size difference. Active devices were continually shrinking and occupying less and less chip area whereas passive devices remained large. At low frequencies, circuit designers employed simulated passive devices as much as possible to make their products more compact and reliable.

While it was possible to fabricate small values of capacitance on-chip, inductors were virtually impossible due to the large physical area required to obtain sufficient inductance at a given frequency. This was compounded by the losses in the substrate which made it virtually impossible to fabricate high quality devices. Small Si die were desirable to keep costs low and to improve reliability since larger die resulted in lower yields [Gray and Meyer, 1993].

When passive devices were needed, usually they were connected externally on-board rather than on-chip. This is possible as long as few external components are needed and the package parasitics are negligible in comparison with

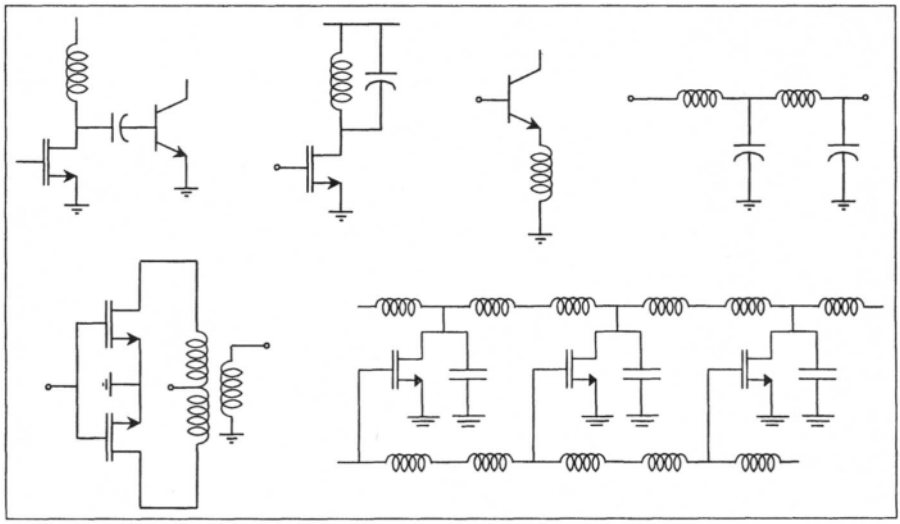


Figure 1.1. Applications of passive devices in Si IC building blocks. (a) Impedance matching. (b) Tuned load. (c) Emitter degeneration. (d) Filtering. (e) Balun. (f) Distributed amplifier.

the external electrical characteristics of the device. Take, for instance, a VCO at 100 MHz versus a VCO at 10 GHz. At 100 MHz a typical tank inductance value will be on the order of 100 nH whereas at 10 GHz the tank inductance is around 1 nH. To access a 1 nH inductor externally is impossible in standard low cost packaging since the package pin and bond wire inductance can exceed 1 nH. Also, as more and more functionality is integrated on-chip, more and more passive devices are needed requiring larger packages which increase the cost. These issues have led to a surge of recent interest in integrated passive devices.

3. APPLICATIONS OF PASSIVE DEVICES

Passive devices, such as inductors and transformers, play an integral part in the performance of circuit building blocks, especially at high frequency. Inductors can be avoided at lower frequencies by using simulated inductances employing active devices. Simulated inductors are more difficult to realize at higher frequencies as active device gain drops. In addition, simulated inductors have finite dynamic range, require voltage headroom to operate, and inject additional noise into the circuits. These limitations place a severe restriction on their application, especially in highly sensitive analog building blocks.

In Fig. 1.1 we see several common applications of inductors, capacitors, and transformers in wireless building block circuits. In (a) we see a narrow-band impedance matching example. Here the input impedance of the second transis-

tor is matched to an optimal impedance value desired by the driving transistor. For instance, in a power amplifier the input impedance of a large output stage device is low due to the capacitance and to obtain sufficient power gain this low impedance is transformed into a larger value. Impedance matching allows circuit designers to obtain minimal noise, maximum gain, minimal reflections, and optimal efficiency when designing circuit building blocks such as low-noise amplifiers (LNAs), frequency-translation circuits (mixers), and amplifiers.

In (b) we see an LC tuned load. A tuned load can take the place of a resistive load to obtain gain at high frequency. The advantages are clear as an LC passive is less noisy than a resistor, consumes less voltage headroom, and obtains a larger impedance at high frequency. A resistive load is always limited by the RC time constant which limits the frequency response. Tuned loads are also a critical component of oscillators. The LC tank tunes the center frequency of the oscillator and the intrinsic Q allows the tank to oscillate with minimal power injection (and hence noise) from the driving transistor.

In (c) an inductor is used as a series-feedback element. Series feedback can be used to increase the input impedance, stabilize the gain, and lower the non-linearity of the amplifier. By using an inductor in place of a resistor, less voltage headroom is consumed, and less additional noise is injected into the circuit. The inductance can also be used to obtain a real input impedance at a particular frequency, thus providing an impedance match at the input of the amplifier.

In (d) inductors and capacitors are used to realize a low-pass filter. Filters of this type are superior to active filter realizations such as gm-C or MOSFET-C filters as they operate at higher frequencies, have higher dynamic range due to the intrinsic linearity of the passive devices, and inject less noise while requiring no DC power to operate.

In (e) we see a center-tapped transformer serving as a balun, a device which converts a differential signal into a single-ended signal to drive external components. Differential operation is advantageous in the on-chip environment due to the intrinsic noise rejection and isolation. Off-chip components, such as SAW filters, though, are single-ended and a balun is needed to convert external single-ended signals to on-chip differential signals.

Finally, in (f) we see inductors and capacitors forming an artificial transmission line in a distributed (traveling-wave) amplifier. Since the LC network acts like a transmission line, it has a broadband response. A wave propagating on the gate-line is amplified and transferred onto the drain line. If the wave speed on the drain line matches the gate line, the signals on the drain line add in phase and the drain line delivers power into a matched load.

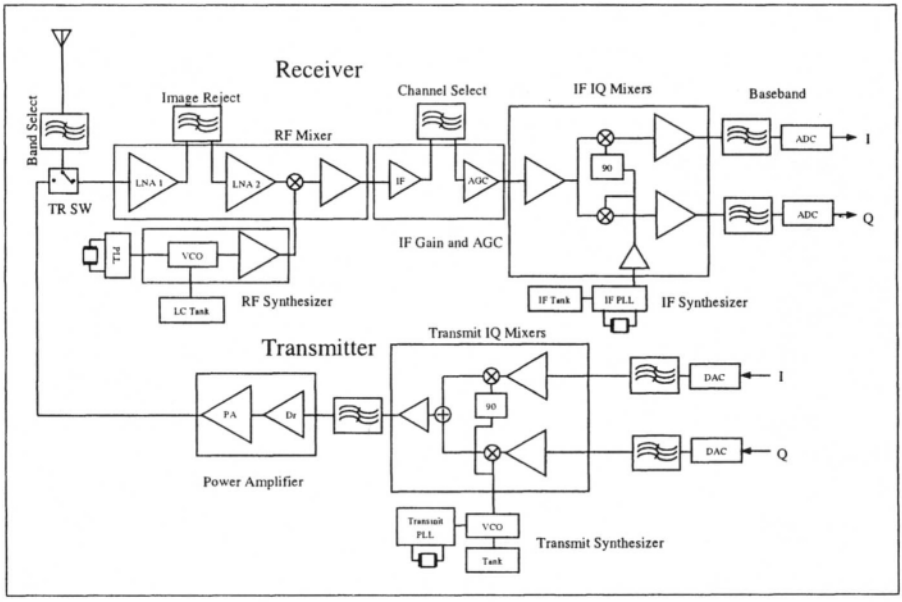


Figure 1.2. Traditional superheterodyne transceiver architecture.

4. WIRELESS COMMUNICATION

The wireless transceiver serves as an excellent example of a system which employs passive devices. In recent times, several factors have contributed to the possibility of portable wireless communications. Continuing technology improvements have enabled low-cost Si circuits to operate in the 1–l0 GHz frequency range. In this frequency range efficient portable antennas can be realized since the free-space wavelength is on the order of centimeters. Higher frequencies also allow higher bandwidths to be realized for increased throughput or an increase in the number of users sharing the spectrum. Furthermore, by limiting the transmit power, transceivers which are physically remote can reuse the same spectrum with minimal interference leading to the cellular concept of communication. Finally, at these higher frequencies the critical passive elements, such as inductors and capacitors, are small enough to be realized on-chip.

To see the importance of passive devices, consider the simplified block diagram of a traditional superheterodyne transceiver shown in Fig. 1.2. Note that this transceiver is realized as several different chips or modules and many components of this transceiver are off-chip discrete components. This transceiver is thus bulky and expensive. The goal of many research projects has been to realize this transceiver in a more integrated form [Rudell et al., 1997, Rofougaran

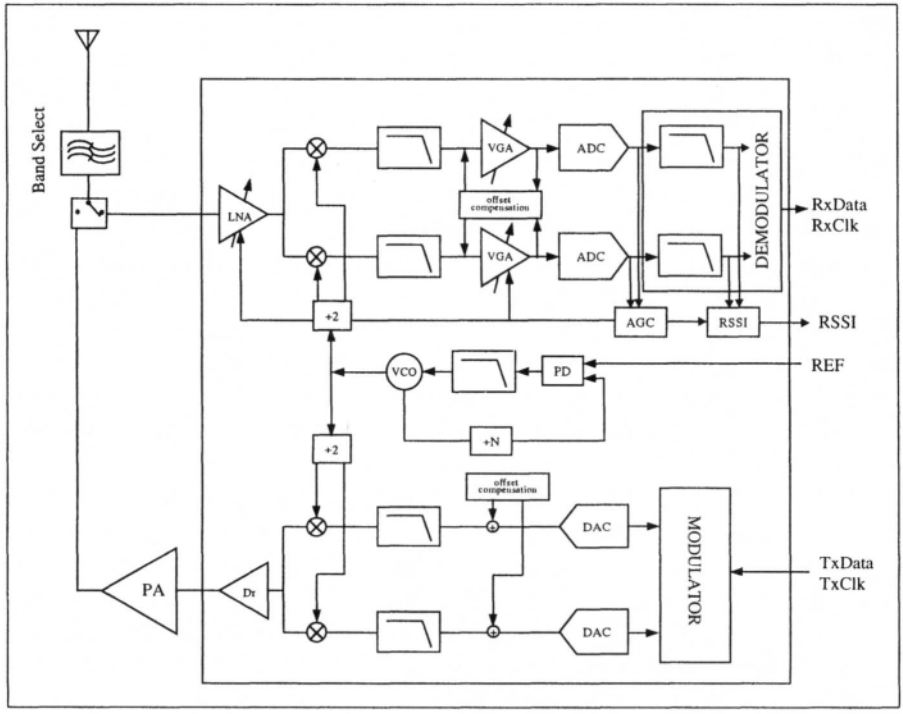


Figure 1.3. Zero-IF direct conversion receiver architecture transceiver.

et al., 1998a, Rofougaran et al., 1998b, Steyaert et al., 1998]. To realize this goal, many off-chip components must be integrated and the inductor is one such key element.

Now contrast Fig. 1.2 with Fig. 1.3 [Eynde et al., 2000], a modern integrated transceiver with minimal off-chip components. Here system architecture innovations such as a zero-IF direct conversion receiver topology eliminates or reduces the need for external filtering. The LNA and VCO tank and the external PLL have also been integrated onto a single chip. The IF is sampled at baseband and digital processing of the signal is performed on the same die. The integrated inductor plays an important role in impedance matching, gain, tuning, and filtering. For instance, in the low noise amplifier (LNA) inductors are used for a real input impedance match and as a narrow-band high impedance load. Here, a resistor is precluded because of the extra noise and voltage headroom required. In the voltage controlled oscillator (VCO) inductors and capacitors form the tuned load of the tank while varactors are used to tune the center frequency to provide a fixed IF frequency. High quality (Q) factor passives limit the phase noise of the VCO which can degrade the receiver sensitivity through

reciprocal mixing. The high Q passives also minimize power leaking into adjacent channels which desensitizes other receiver units in nearby channels. On the transmit side, passives are used in the driver and power amplifier (PA) for impedance matching and gain. Impedance matching allows the power hungry PAs to operate at maximum efficiency and thus extends battery life.

One of the key difficulties in realizing an integrated transceiver lies in the difficult specifications that a transceiver must meet. The harsh wireless environment puts stringent requirements on the dynamic range requirements of the receiver. Since signal strength can vary by many orders of magnitude as mobile users move close to and far from the base station antenna, the transceiver must be able to operate with both extremely small signals close to the intrinsic noise floor of the transceiver and with extremely large signals which excite the desensitizing non-linearities of active devices. Furthermore, these constraints must be met while consuming very low power levels to extend battery life.

Thus, the fully integrated transceiver introduces several new problems. Low power and low noise operation requires the realization of high Q passive devices. Integration introduces parasitic coupling between the various circuit blocks. Within the analog portion parasitic coupling from the powerful transmit circuitry to the sensitive receive path is particularly worrisome. The coupling from the noisy digital blocks to the sensitive analog blocks is also a big concern. Note that this coupling occurs primarily through the package and through the substrate.

5. SI INTEGRATED CIRCUIT TECHNOLOGY

Si technology is a prime candidate for realizing future integrated circuits. While GaAs offers superior gain, higher frequencies of operation, and an insulating substrate, the difference in cost alone favors Si. In addition, emerging advances in Si technology (such as SiGe) are closing the gap between Si and GaAs in performance for the mass commercial market target in the 1–10 GHz frequency range.

When we also consider the potential to integrate digital functionality in CMOS or BiCMOS technology, Si is again the clear winner. Furthermore, due to the mass digital microprocessor and DSP markets, CMOS technology continues to improve. Thus, there are great benefits to integrating passive devices in CMOS technology. Consider the unity gain frequency of an NMOS transistor [Gray and Meyer, 1993]

$$f_T = \frac{1}{2\pi}\frac{g_m}{C_{gs}} = 1.5\frac{\mu_n}{2\pi L^2}(V_{GS} - V_t) \tag{1.1}$$

and the similar expression for a bipolar transistor

$$f_T = 2\frac{\mu_n}{2\pi W_B^2}V_T \tag{1.2}$$

Since the vertical base width of a bipolar transistor W_B is determined by a diffusion process whereas the lateral channel length L of an MOS transistor is determined by lithographic processes, bipolar transistors have enjoyed a superiority in speed. New advances in lithographic technology, though, have narrowed the channel length tremendously as MOS technology is the core of the worldwide digital market. Even taking into account that narrow channel device f_T have 1/L dependence as opposed to the $1/L^2$ dependence, CMOS technology is today a viable and cost-effective alternative to both bipolar and GaAs and will probably continue to be so for the next decade or longer.

From the perspective of cost, CMOS is the clear winner. But from the perspective of integrated passive devices, GaAs is clearly superior to standard Si. The reason for this stems from the insulating nature of the GaAs substrate which allows very high Q factor passives to be realized on-chip. The dominant limitation in performance is the metal conductivity whereas in Si the dominant loss mechanism at high frequency is the conductive substrate. Electromagnetic energy couples to the substrate and the lossy nature of the Si substrate limits the Q severely.

This is especially the case when the substrate is heavily conductive, as with epi CMOS substrates. Here, magnetically induced eddy currents in the substrate can be a dominant loss mechanism. For moderately conductive substrates, though, electrically induced substrate currents are the dominant loss mechanisms at frequencies below the self-resonant frequency of the device.

6. CONTRIBUTIONS OF THIS RESEARCH

This research has focused on the analysis, design, and applications of passive devices. In Part I emphasis is placed on analyzing inductors and transformers. The solution techniques developed can be easily applied to integrated capacitors and resistors. A general approach is developed from Maxwell's equations to determine the partial inductance and capacitance matrix of an arbitrary arrangement of conductors situated on top of a stratified conductive substrate. The lossy nature of the underlying substrate and losses in the conductors leads to complex capacitance and inductance matrices. These techniques have culminated in a design and analysis software tool called *ASITIC,* "Analysis and Simulation of Inductors and Transformers for ICs." These general techniques can also be applied to extracting the substrate coupling occurring between metal structures residing in the substrate or in the oxide layers on top of the Silicon. This technique is a direct extension of [Gharpurey and Meyer, 1996].

In Part II of the book, we focus on some key applications of passive devices, such as voltage-controlled oscillators and power amplifiers. We show that key performance parameters such as phase noise and power amplifier efficiency depend on the quality of the passive devices. Thus, it is critical to be able to predict the performance of such structures.

Chapter 2

PROBLEM DESCRIPTION

1. DEFINITION OF PASSIVE DEVICES

Consider an arbitrary black box shown in Fig. 2.1 with an arbitrary number of externally accessible terminals. Sample "black" boxes are also shown in the figure. We would like to categorize each black box as passive or active.

Intuitively, the distinction between passive and active devices is clear. While passive devices can only consume or store energy, active devices can also supply energy. Thus, with active devices one can obtain power gain though this is impossible for a passive device. In other words, if we inject power into any terminal of a passive device and observe the power flowing into an arbitrary impedance at another terminal of the device, the real power is necessarily less than or equal to the power injected. If we lift this restriction, the device is active.

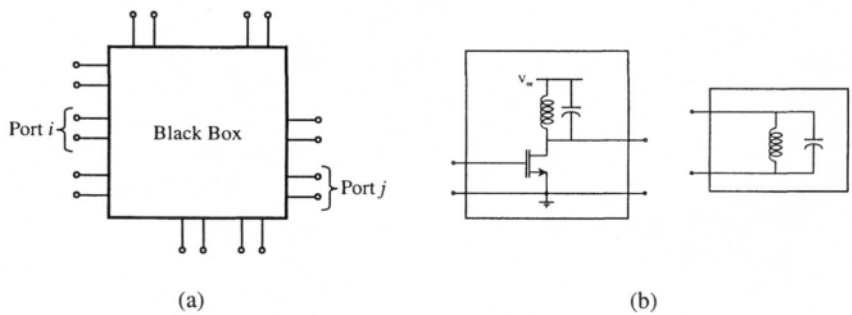


Figure 2.1. (a) An arbitrary black box with externally accessible ports, (b) The contents of two example "black boxes".

Typical examples of passive devices include linear time-invariant resistors of finite resistance $R > 0$ which dissipate energy, ideal time-invariant linear capacitors and inductors, which only store energy. While the above elements are linear elements, passive elements such as diodes can be non-linear[1]. Even though passive devices can have voltage or current gain, they cannot have power gain.

The archetypical example of an active element is of course the transistor and its predecessor, the vacuum tube. Unlike passive devices, such a device can be used to obtain power gain. These devices can be linear or non-linear, but most real world examples are non-linear. The distinction between active and passive is more difficult to make when time-varying elements are employed. For instance, a simple parametric amplifier constructed from a sinusoidally time-varying capacitor in shunt with an RLC tank forms an active device.

A precise definition of passive elements is given in [Desoer and Kuh, 1969]. Given a one-port with port voltage $v(t)$ and port current $i(t)$, the one-port is said to be passive if

$$\int_{t_0}^{t} v(t')i(t')dt' + \mathcal{E}(t_0) \geq 0 \tag{2.1}$$

where $\mathcal{E}(t_0)$ is the energy stored by the one-port at time t_0. With application of this definition, one can clearly delineate an element as passive or active.

At high frequencies it is usually more convenient to discuss the scattering matrix S [Carlin and Giordano, 1964]. The power dissipated by an arbitrary port of an n-port is given by

$$P_k = P_{ik} - P_{rk} \tag{2.2}$$

where the subscript i denotes incident power and r denotes reflected power. In terms of the incident and reflected waves we have

$$P_k = a_k\bar{a}_k - b_k\bar{b}_k \tag{2.3}$$

summing over all ports we obtain the total power

$$P = \sum_{k=1}^{n} P_k = \bar{a}^T a - \bar{b}^T b \tag{2.4}$$

but by the definition of the S matrix $b = Sa$ so that

$$P = \bar{a}^T a - \bar{a}^T\bar{S}^T Sa = \bar{a}^T(I - \bar{S}^T S)a = \bar{a}^T Qa \tag{2.5}$$

where Q is the dissipation matrix. Note that

$$\bar{Q}^T = \bar{I}^T - \overline{(\bar{S}^T S)}^T = I - \bar{S}^T S = Q \tag{2.6}$$

and so Q is a Hermitian matrix. By definition of passivity, we have

$$P = \bar{a}^T Q a \geq 0 \tag{2.7}$$

Clearly, then, if the matrix Q is positive definite or positive semi-definite, the matrix S corresponds to a passive network. One can also show that this condition implies [Carlin and Giordano, 1964]

$$0 \leq |s_{ij}| \leq 1 \tag{2.8}$$

This can be deduced intuitively in the following manner. Observe that each diagonal entry of the S matrix is the reflection coefficient when all other ports are matched and each off-diagonal component is the transmission coefficient under matched conditions. For a passive network, the conservation of energy implies that the power reflected from any port must be less than the power injected, thus $|\rho|^2 \leq 1$ and the power transmitted similarly must satisfy the same condition, $|\tau|^2 \leq 1$. This implies all matrix elements have magnitude less than or equal to unity and thus reside in the unit circle in the complex plane [2]. An active device consists of a black box with sources, and thus the matrix elements may have magnitude greater than unity. It follows that the real part of the input port impedances may be negative, something that may never occur for a passive network.

1.1 STABILITY AND PASSIVITY

Passivity is closely related to stability [Desoer and Kuh, 1969]. Again, this is clear intuitively as any passive device must be stable by the conservation of energy. In other words, one can never construct an unstable device with purely passive devices. Conditional stability, as with an oscillating LC tank, is possible as long as the initial conditions supply some energy to the tank. It can be shown that a passive circuit is indeed stable and this further implies that any natural frequency of a passive network must lie in the closed left-hand plane, and any $j\omega$-axis natural frequency must be simple [Desoer and Kuh, 1969].

1.2 RECIPROCITY

Consider the n-port parameters of the black box. If the volume of the black box of Fig. 2.1 is isotropic and encloses no sources, then by the Lorenz reciprocity theorem of electromagnetics it can be shown [Pozar, 1997] that the impedance matrix Z is symmetric. Such a network is called a reciprocal network. A reciprocal network is also described by a symmetric S matrix since in such a case one can show that

$$S = (\frac{Z}{Z_0} + I)^{-1}(\frac{Z}{Z_0} - I) = (\frac{Z}{Z_0} - I)(\frac{Z}{Z_0} + I)^{-1} \tag{2.9}$$

where Z_0 is the system impedance, usually **50 Ω**, and I is the identity matrix. If Z is symmetric, then the symmetry of S follows.

It should be noted that reciprocity has nothing to do with passivity. While many passive networks are indeed reciprocal, the connection is not obvious. Any linear time-invariant RLCM network is reciprocal. In fact, some of the elements may be active and reciprocity is still satisfied. A gyrator, a passive device, in non-reciprocal. The reciprocity theorem in circuits follows from excluding any network with gyrators, dependent and independent sources.

1.3 THE QUALITY OF PASSIVE DEVICES

In general, the complex power delivered to a one-port black box network at some frequency ω is given by [Pozar, 1997]

$$P = \frac{1}{2}\oint_S \mathbf{E} \times \mathbf{H}^* \cdot d\mathbf{s} = P_l + 2j\omega(W_m - W_e) \tag{2.10}$$

where P_l represents the average power dissipated by the network and W_m and W_e represent the time average of the stored magnetic and electric energy, respectively. One can define the input impedance as follows [Pozar, 1997]

$$Z_{in} = R + jX = \frac{V}{I} = \frac{VI^*}{|I|^2} = \frac{P}{\frac{1}{2}|I|^2} = \frac{P_l + 2j\omega(W_m - W_e)}{\frac{1}{2}|I|^2} \tag{2.11}$$

If $W_m > W_e$ the device acts inductively whereas if the opposite is true the device acts capacitively.

An important parameter to consider when discussing passive devices is the quality factor. The quality factor has the following general definition

$$Q = 2\pi \frac{E_{\text{store}}}{E_{\text{diss}}} \tag{2.12}$$

where E_{store} is the energy stored per cycle whereas E_{diss} is the energy dissipated per cycle. Implicit in the above definition is that the device is excited sinusoidally. From (2.11) with T equal to the cycle time

$$Q = 2\pi \frac{(W_m + W_e)}{P_l \times T} = \frac{\omega(W_m + W_e)}{P_l} \tag{2.13}$$

The higher the Q factor, the lower the loss of a passive device. This definition is most pertinent when discussing inductors or capacitors as such devices are meant to store energy while dissipating little to no energy in the process. Thus ideal inductors and capacitors have infinite Q whereas practical devices have finite Q. Applying the above definition to an ideal inductor L where $W_e \equiv 0$ in series with a resistor R, one obtains $Q = \omega L/R$ and similarly to an ideal capacitor C where $W_m \equiv 0$ in series with a resistor, $Q = (\omega CR)^{-1}$.

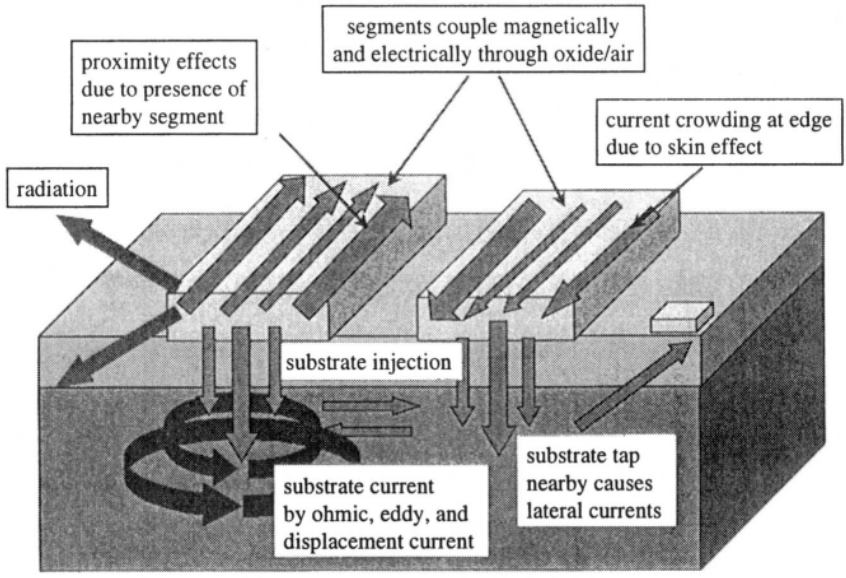


Figure 2.2. Various loss mechanisms present in an IC process.

Physically, the lossy nature of passive devices is rooted in physical phenomena which convert electrical energy into other, unrecoverable forms of energy. Processes which increase entropy are not reversible. For instance, a resistor converts electrical energy into heat. A light bulb converts electrical energy into light and heat. An antenna also converts electrical energy into radiating electromagnetic energy. One can therefore distinguish between passive devices which increase entropy while conserving energy, like a resistor or an incoherent light source, and other devices which conserve energy but do not increase entropy, such as an ideal laser. An ideal laser converts electrical energy into a coherent emission of monochromatic photons. In reality, any physical laser will emit photons with a Lorentzian distribution of energies and thus the entropy of the system increases.

2. LOSS MECHANISMS

The Q factor of integrated passive devices is largely a function of the material properties used to construct the ICs. Specifically, the semiconductor substrate and metal layers used to build the device play the most important roles. The various loss mechanism are summarized in Fig. 2.2 and discussed further below.

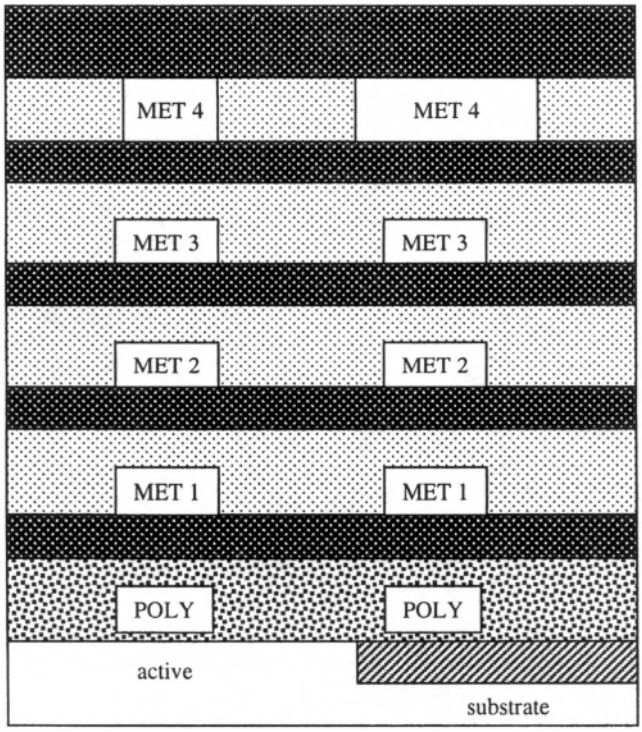


Figure 2.3. Cross-section of metal and polysilicon layers in a typical IC process.

2.1 METAL LOSSES

Passive devices such as inductors and capacitors are constructed from layers of metal, typically aluminum, and polysilicon layers. Hence, the conductivity of such layers plays an integral part in determining the Q factor of such devices, especially at lower frequencies. For instance, a capacitor is constructed by placing two metal conductors in close proximity. Reactive energy is stored in the electric field formed by the charges on such conductors. Since the metal layers are not infinitely conductive, energy is lost to heat in the volume of the conductors. This loss can be represented by a resistor placed in series with the capacitor. Similarly, an inductor is wound using metal conductors of finite conductivity. Most of the reactive energy is stored in the magnetic field of the device, but energy is also lost to heat in the volume of the conductors.

Fig. 2.3 shows a cross-section of the metal layers of a typical modern IC process. Most processes come with three or more interconnection metal layers. This may include one or two layers of polysilicon as well. Some modern CMOS

processes include up to eight metal layers, with the top layer separated from the substrate by $\sim 10\ \mu m$ of oxide.

Most IC metal layers are constructed from aluminum which has a room temperature conductivity of $\sigma = 3.65 \times 10^7\ S/m$. Typical metal layers have a thickness ranging from $.5\ \mu m$ to $4\ \mu m$, resulting in sheet resistance values from $55\ m\Omega/\square$ to $7\ m\Omega/\square$. Even though silver, copper, and gold are more conductive, at $\sigma_{Ag} = 6.21 \times 10^7\ S/m$, $\sigma_{Cu} = 5.88 \times 10^7\ S/m$, and $\sigma_{Au} = 4.55 \times 10^7\ S/m$, aluminum is the more compatible metal in the IC process. Even though Aluminum is prone to spiking and junction penetration [Jaeger, 1993], it is usually mixed with other metals such platinum, palladium, titanium, and tungsten to overcome these limitations. Electromigration in Al is another problem, setting an upper bound on the maximum safe current density. Although electromigration with AC currents is less problematic, it remains one of the important limitations preventing integration of "high-power" passives on Si, such as the matching networks at the output of a power amplifier. The necessary metal width would require excessively large areas resulting in low self-resonant frequencies.

Many IC processes geared for wireless communication applications are now providing a thick top-metal layer option for constructing inductors. Such a metal layer is also useful for high-speed digital building blocks and clock lines and thus this option is widely available in digital processes as well. This top metal layer may also reside on top of an extra thick insulator for minimum capacitance [Kamogawa et al., 1999]. On the other hand, the wealth of interconnection opens up the possibility of designing structures with many different metal layers. This has the added benefit of requiring no extra processing steps as such "3D" interconnection is available with most modern CMOS processes.

At increasingly higher frequencies, even in the absence of the substrate, the current distribution in the metal layers changes due to eddy currents in the metallization, also known as skin and proximity effects, current constriction, and current crowding. At any given frequency, alternating currents take the path of least impedance. Currents tend to accumulate at the outer layer or skin of conductors since magnetic fields of the device penetrate the conductors and produce opposing electric fields within the volume of conductors. When the effective cross-sectional area of the conductors decreases at increasing frequencies, the current density increases, converting more energy into heat. For an isolated conductor, the magnetic fields originate from the conductor itself (the self-inductance). This increase in AC resistance is know as skin effect and typically follows a $\sqrt{f}$ functional dependence. This rate of increase can be traced to the effective depth of penetration δ of the current since the effective area is

a function of the skin depth[3]

$$\delta = \sqrt{\frac{2}{\omega\mu\sigma}} \tag{2.14}$$

In a multi-conductor system, the magnetic field in the vicinity of a particular conductor can be written as the sum of two terms, the self-magnetic field and the neighbor-magnetic field[4]. Thus, the increase in resistance of any particular conductor can be attributed not only to skin effect but also to proximity effects, the effect of nearby conductors. If nearby conductors enhance the magnetic field near a given conductor, the AC resistance will increase even further and this is the case for a spiral inductor. On the other hand, if the nearby fields oppose the field of a given conductor, as is the case in a transformer, the AC resistance will decrease as a result.

2.2 SUBSTRATE INDUCED LOSSES

Integrated passive devices must reside near a conductive Si substrate. The substrate is a major source of loss and frequency limitation and this is a direct consequence of the conductive nature of Si as opposed to the insulating nature of GaAs. The Si substrate resistivity varies from **10 $k\Omega$-cm** for lightly doped Si **(10^{13}atoms/cm^3)** to **.001 Ω-cm** for heavily doped **Si (10^{20}atoms/cm^3)**. In fact, to combat these substrate induced losses, some researchers propose removing the substrate from under the device by selective etching [Chang et al., 1993] [Lopez-Villegas et al., 2000].

The conducting nature of the Si substrate leads to various forms of loss, namely conversion of electromagnetic energy into heat in the volume of the substrate. To gain physical insight into the problem, we can delineate between three separate loss mechanisms. First, electric energy is coupled to the substrate through displacement current. This displacement current flows through the substrate to nearby grounds, either at the surface of the substrate or at the back-plane of the substrate. Second, induced currents flow in the substrate due to the time-varying magnetic fields penetrating the substrate. These magnetic fields produce time-varying solenoidal electric fields which induce substrate currents. These currents are show in Fig. 2.4 for the case of a spiral inductor. Note that electrically induced currents flow vertically or laterally, but perpendicular to the spiral segments. Eddy currents, though, flow parallel to the device segments[5].

Finally, all other loss mechanisms can be lumped into radiation. Electromagnetically induced losses occur at much higher frequencies where the physical dimensions of the device approach the wavelength at the frequency of propagation in the medium of interest. This frequency is actually difficult to quantify due to the various propagation mechanisms of the substrate. For instance, if we consider propagation into air, the the free-space wavelength is the appropriate

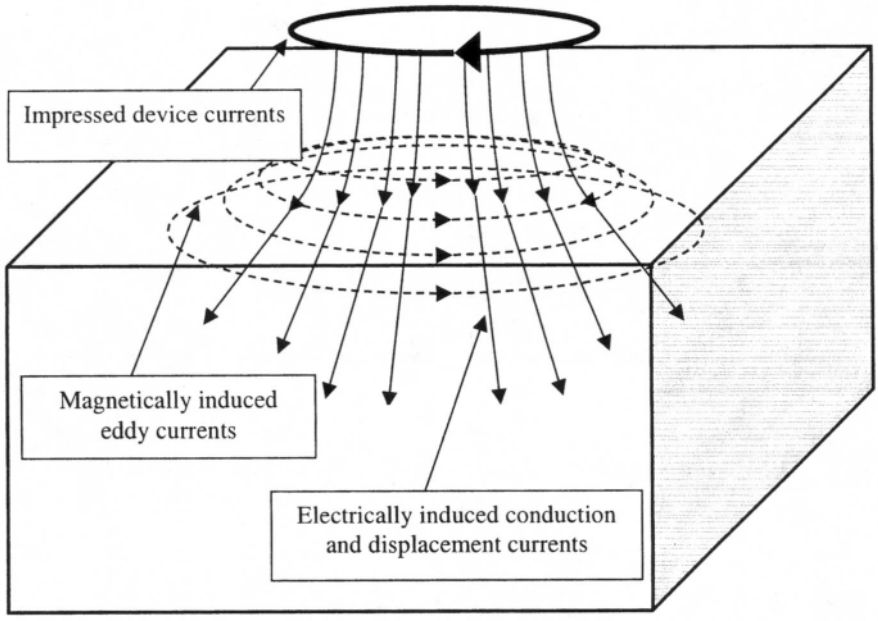


Figure 2.4. Schematic representation of substrate currents. Eddy currents are represented by the dashed lines and electrically induced currents by the solid lines.

factor. Even at 10 GHz, the wavelength in air is 3 cm, much larger than any RF inductor or capacitor. Even at 100 GHz, the wavelength is now 3 mm, still much larger than any device at this frequency. Thus, we can safely ignore the electromagnetic propagation into the air.

Efficient electromagnetic propagation into the substrate, though, occurs at lower frequencies due to the lower propagation speed, roughly at a factor of $\sqrt{\epsilon_{Si}}$ lower due to the diamagnetic nature of Si. Since $\epsilon \approx 11.9$ in Si, this is slightly slower from propagation in air. Furthermore, due to the lossy nature of the substrate, waves traveling vertically into the surface of the substrate are heavily attenuated. Waves traveling along the surface of the substrate, though, can propagate partially in the lossless oxide and partially in the substrate. For a lightly doped substrate, the wave propagation behaves like a "quasi-TEM" mode. As the substrate is made heavily conductive, the wave is constrained to the oxide and the substrate acts like a lossy ground plane. This is the so-called "skin effect" mode of propagation [Hasegawa et al., 1971]. There is a third kind of possible excitation, the "slow-wave" mode of propagation, where the effective speed of propagation is orders of magnitude slower than propagation in free space [Hasegawa et al., 1971].

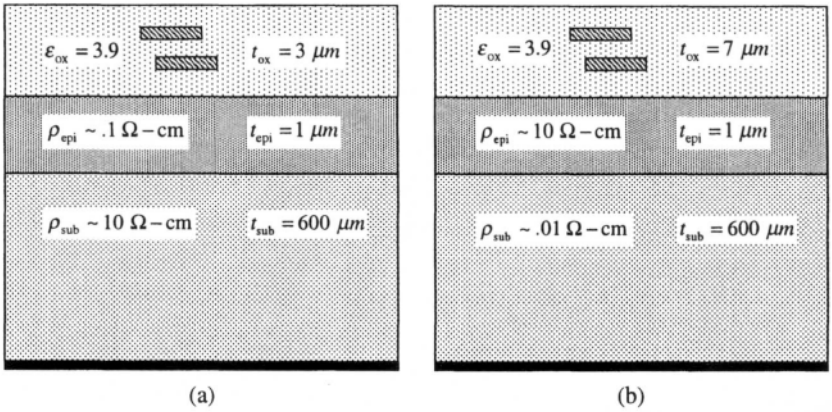


Figure 2.5. Cross-section of typical (a) bipolar and (b) CMOS substrate layers.

2.2.1 SI IC PROCESS SUBSTRATE PROFILE

In Fig. 2.5 typical bipolar and CMOS process substrate profiles are shown. Each substrate consists of one or more layers of Si or a compatible material[6]. Layers of varying conductivity are added to the bulk substrate by various g processes, such as diffusion, chemical vapor deposition and growth, epitaxy, and ion implantation. Various layers of oxide (SiO_2 for instance) and polyimide are also grown to provide insulation from the substrate and between metal layers.

In general, the more conductive the substrate layers, the more detrimental the resulting losses. It is therefore no surprise that intrinsic Si substrates[7] result in the lowest losses [Park et al., 1997b]. Due to the close proximity of the Si substrate to the inductors and transformers residing in the metal layers, the case of an infinitely conductive substrate is also problematic. For a heavily conductive substrate, the magnetic and electric fields do not penetrate the substrate appreciably and even though no substrate induced losses occur, the surface currents flowing in the substrate, acting like "ground-plane" currents, produce opposing magnetic fields which tend to drive the inductance value of coils to low non-usable values.

Therefore, given the choice, designers of ICs and process engineers should ensure that as few as possible conductive substrate layers appear under or near an inductor. This is unfortunately not always possible due to planarization constraints. Furthermore, the thickest possible oxide should be realized under the device to minimize the substrate capacitance. This not only minimizes the losses, but also maximizes the self-resonant frequency of the device. In the limit, self-resonance will occur due to interwinding capacitance as opposed to substrate capacitance. Since interwinding capacitance can be controlled by

increasing the metal spacing, this gives the IC designer more control over the passive device behavior.

Most bipolar and BiCMOS substrates come with a standard 10–20 Ω-cm substrate. With this value of resistivity, electrically induced losses dominate the substrate losses in the 1–10 GHz frequency range [Ruehli, 1972]. This is also the case for bulk CMOS substrates with the same range of resistivity. In such a case, one must ensure that no conductive *n*- or *p*-wells appear below the device. This may require a special mask to block the dopants in the well creation process, especially for a twin-well CMOS process. To minimize the chance of latch-up, many modern CMOS processes begin with a heavily conductive thick substrate about 700 μm thick and grow a thin epitaxial layer of resistive Si on the surface to house the wells. This is unfortunate for RF/microwave circuits as the bulk substrate can be as conductive as 10^4 *S/m* and this can be a major source of substrate induced losses due to eddy currents.

The back-plane of the substrate may or may not be grounded. Even if it is physically grounded for DC signals, AC signals are constrained to flow within several skin depths δ and this factor is a strong function of the conductivity. For heavily conductive substrates, currents are constrained to flow at the surface of the substrate at high frequencies whereas for moderately conductive substrates currents flows deep into the substrate and into the back-plane ground.

A physical ground may be realized if the die (chip) resides in a package with a conductive ground plane, or if the die is bonded directly onto a board, and a conductive epoxy cement glue is used. Although the epoxy is not conductive, metals can be mixed in to produce a conductive solution. In the modeling of the substrate this can be an important factor in enforcing the boundary conditions surrounding the chip.

Unless the substrate thickness is reduced substantially, the conductive back-plane ground is sufficiently distant not to appreciably influence the electromagnetic behavior of the inductor. On the other hand, some packages use "down-bonds," bond wires from the Si die to the package grounded "paddle". To minimize the bond wire length, the substrate thickness is reduced in post-processing steps. This has further benefits for a packaged power amplifier since a thinner substrate also has better thermal conductivity. For such a thin, moderately conductive grounded substrate, one must take into account ground "image" currents which can reduce the inductance value and serve as a further loss mechanism [Krafcsik and Dawson, 1986]. If the substrate is sufficiently conductive such that the skin depth δ is much less than the substrate thickness, then image currents will be confined to flow at the substrate surface.

3. DEVICE LAYOUT

In this section we will discuss various ways to lay out inductors using the planar metallization layers of a typical IC process. Off-chip inductors are usu-

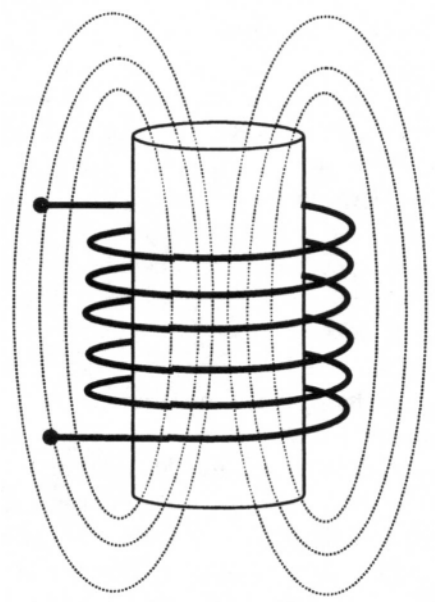


Figure 2.6. The typical coil inductor.

ally realized as a solenoidal coil or toroid, as shown in Fig. 2.6. Each additional turn adds to the magnetic field in phase with the previous turn. The magnetic energy is stored mostly in the inner core of each winding. The inductance is largely a function of the area of the loop and the number of turns in the winding typically resulting in an N^2 dependence.

3.1 PLANAR INDUCTOR STRUCTURES

Since on-chip inductors are constrained to be planar, the typical solution is to form a spiral, as shown in Fig. 2.7. Since some IC processes constrain all angles to be 90°, a square version of the spiral, shown in Fig. 2.8, is a popular alternative. A polygon spiral, as shown in Fig. 2.9, is a compromise between a purely circular spiral and a square spiral.

In designing integrated circuits it is sometimes convenient to tap an inductor at some arbitrary point. While this is certainly possible, as more than one metal layer is present, it is sometimes necessary to tap a spiral in the center, especially for differential circuits. In a spiral it is difficult to find such a symmetric center point since the electric fields on the outer turns tend to fringe and thus the "inductive" center does not correspond to the "capacitive" center. Also, the "inductive" center does not correspond to the "resistive" center due to the non-uniform mutual magnetic coupling. To solve this problem, some researchers

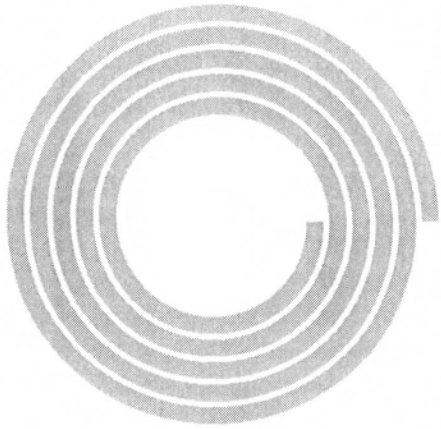


Figure 2.7. A circular spiral inductor.

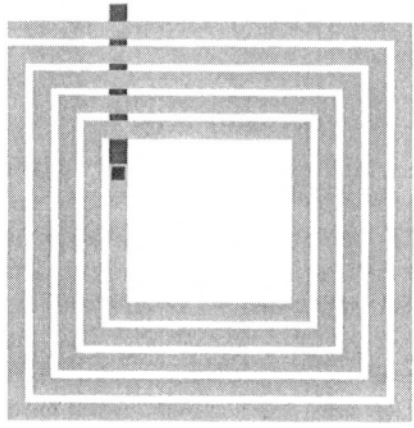


Figure 2.8. A square spiral inductor.

have proposed symmetric structures, such as that shown in Fig. 2.10. Note that each turn involves a metal-level interchange, a process that requires vias. A different center-tapped structure proposed by [Kuhn et al., 1995] requires only one metal interchange. This structure is very similar to a inter-digited planar transformer structure shown in Fig. 2.13. These structures have a natural geometric center which coincides with the electrical center point. This is needed in differential circuits as such points can be grounded or connected to supply

ADVANCED MICRO DEVICES, INC. IPR2023-00789 Ex. 1009, p. 25 of 69

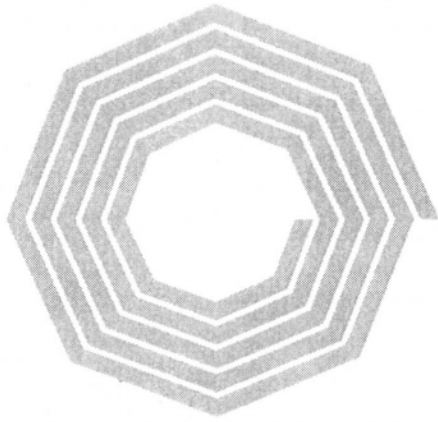


Figure 2.9. A polygon spiral inductor.

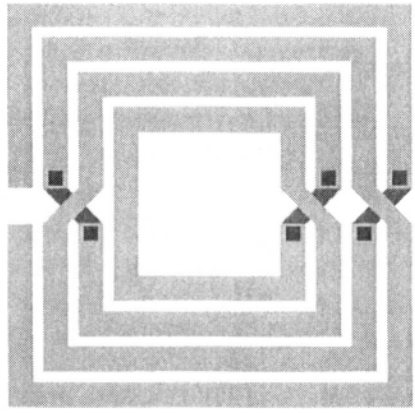


Figure 2.10. A symmetric spiral inductor.

without disturbing the differential signal. Circular or polygon versions are also possible, as shown in Fig. 2.11.

3.2 NON-PLANAR INDUCTOR STRUCTURES

Up to now we have only considered planar structures even though modern IC processes offer many metal layers. Two simple approaches in utilizing the metal layers are to connect multiple spiral inductors in series or in shunt. While N spirals in series increase the series resistance by a factor of approximately N



Figure 2.11. A symmetric polygon spiral inductor.

(neglecting via resistance), the inductance value increases faster due to the mutual magnetic coupling. At low frequencies where the current flowing through each series connected spiral I_j is equal, the effective inductance of N series connected coupled inductors is

$$L_{se} = \sum_{i=1}^{N} L_i + 2\sum_{i=1}^{N}\sum_{j\neq i} M_{ij} \tag{2.15}$$

where M_{ij} is the mutual magnetic coupling between each series connected spiral i and j. Thus, the series connection approaches an N^2 increase in inductance and the Q factor can potentially increase by a factor of N for the case of perfectly coupled spirals $(k = 1)$.

Alternatively, in the shunt connection, the series resistance drops by a factor of N (assuming equal resistivity in each metal layer and uniform current distribution among the coils) whereas the inductance of mutually coupled inductors drops to[8]

$$L_{sh} = \frac{1}{\sum_{i=1,j=1}^{N} K_{ij}} \tag{2.16}$$

where the matrix K is the inverse of the partial inductance matrix M. This result certainly agrees with the case of a diagonal matrix M corresponding to zero coupling since in such a case we have the familiar result for parallel inductors

$$L_{sh,k=0} = \frac{1}{\sum_{i=1}^{N} \frac{1}{M_{ii}}} \tag{2.17}$$

For the case of two coupled inductors we have a simpler relation

$$L_{sh} = \frac{1}{2}\frac{L_1 L_2 - M^2}{L_1 + L_2 - 2M} \tag{2.18}$$

For the case of perfectly coupled equal value inductors, with $k = +1$, the above result yields $L_{sh} = L$. For the general case, the matrix M is singular but by symmetry the current flowing through all inductors is equal so the voltage across the jth inductor gives

$$V_j = \sum_{k=1}^{N} sM_{jk}I_k = \frac{I}{N}\sum_{k=1}^{N} sM_{jk} = \frac{I}{N}sL\sum_{k=1}^{N} 1 = sLI \tag{2.19}$$

and so we have

$$L_{sh} = L_1 = L_2 = \cdots = L_N \tag{2.20}$$

In this limit, the Q factor also improves by a factor of N due to the drop in series resistance. In practice, both the series and shunt connection offer a Q improvement close to the theoretical limit due to the tight coupling achievable in the on-chip environment. These benefits, though, only occur at low frequencies where the above assumptions hold.

At high frequencies, both approaches are liable to reduce the Q factor over the case of a single layer coil due to the capacitive and substrate effects. The series connection suffers from high interwinding capacitance which lowers the self-resonance frequency lowering the maximum frequency of operation of the device. The shunt connection moves the devices closer to the substrate where capacitive current injection into the substrate may dominate the loss of the device.

Another approach is to attempt to realize a lateral coil on-chip by using the top and bottom metal layers and vias as the side. This approach has been successfully demonstrated in [Young et al., 1997] using special post-processing steps to realize sufficient cross-sectional area in the coil. This approach has the added advantage of positioning the magnetic fields laterally to the substrate where eddy currents are reduced. In a vertical coil or spiral, the magnetic field is strongest at the center of the coil. Due to the finite conductivity of the substrate, these changing magnetic fields leak into the substrate and produce eddy currents. Eddy currents can be a significant source of loss and this technique might be an effective method to combat this loss mechanism. Standard monolithic integration, though, has failed to produce a coil with sufficient cross-sectional area to produce significant Q. Some innovative strategies around this are to employ a combination of metal and bond wires to realize the coil [Lee et al., 1998b]. The work of [Lee et al., 1998b] claims tight tolerance on the inductance value which is a primary concern of using bond wires alone[Craninckx and Steyaert, 1995].

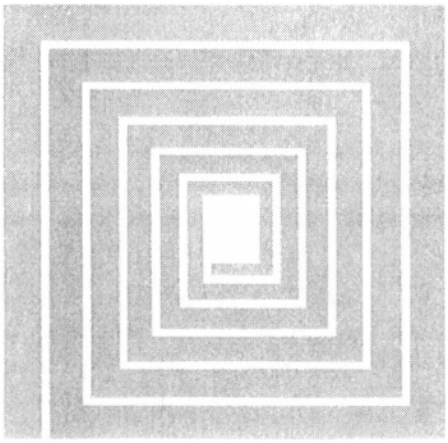


Figure 2.12. A tapered spiral inductor.

3.3 TAPERED SPIRALS

A tapered spiral is shown in Fig. 2.12. The metal pitch and spacing are varied to minimize the current constriction at high frequency [Niknejad, 1997, Lopez-Villegas et al., 2000]. Current constriction, or skin effect, is non-uniform as a function of the location in the spiral due to the non-uniformity of the magnetic field. At low frequencies the current is nearly uniform whereas at high frequency non-uniform current flows due to proximity effects.

The magnetic field is strongest in the center of the spiral [Craninckx and Steyaert, 1997] and thus the time-varying magnetic field produces eddy currents of greatest strength in the volume of conductors near the center of the device. Since at high frequency current constriction limits the current to the outer edges of the conductors, conductor width does not have as strong an influence on minimizing metal losses as at low frequency. For this reason [Craninckx and Steyaert, 1997] advocate removing the inner turns to produce a "hollow" spiral. Another approach is to decrease the width of the inner turns and to effectively move these turns closer to the outer edge. This approach contrasts with the approach suggested by [Lopez-Villegas et al., 2000] where the sum of the metal pitch and spacing, $W + S$, is kept constant.

Tapering is most effective when substrate losses are negligible, as in the case of an insulating substrate. This stems from the fact that current constriction dominates the metallization losses at high frequencies when the substrate losses are also significant. On the other hand, tapering can be an effective means of increasing the self-resonant frequency of a device by decreasing the cross-sectional area of the device.

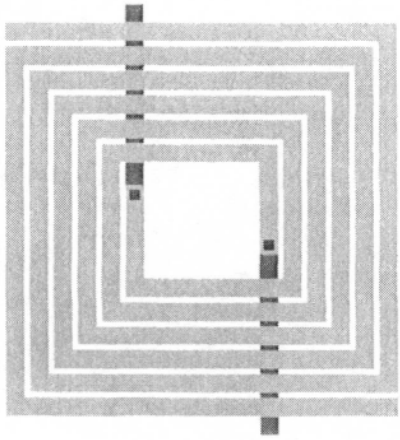


Figure 2.13. A planar square spiral transformer.

3.4 TRANSFORMERS

On-chip transformers are realized very similarly to inductors. To maximize the coupling factor k, two inductors can be interwound as shown in Fig. 2.13. Polygon spirals can be similarly interwound to form transformers. These transformers have equal turns ratio at the primary and secondary. Since the turns ratio n is given by

$$n = k\sqrt{\frac{L_2}{L_1}} \tag{2.21}$$

one way to realize $n \neq 1$ is to alter the number of turns and metal pitch in the secondary. In addition, to lower the losses in the secondary, turns may be connected in shunt [Long and Copeland, 1997]. Typical practical coupling values in the range of $.6 < k < .8$ can be achieved with planar spirals. Metal-metal transformers utilizing two or more metal layers can save area, but cause asymmetry and increased capacitive coupling between the primary and secondary. The asymmetry can be reduced by utilizing a structure such as the one in Fig. 2.14.

If center-tapped transformers are desired, such as in a balun, a structure such as Fig. 2.15 serves well. This structure is formed by interwinding two center-tapped coils of Fig. 2.10.

3.5 SHIELDED STRUCTURES

In an attempt to shield a device from substrate losses, researchers have proposed building a shield with lower metal layers or polysilicon layers to block

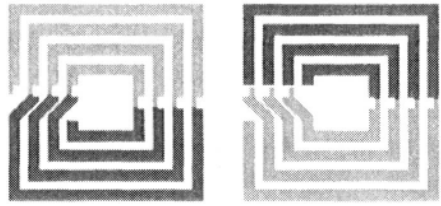


Figure 2.14. An expanded view of a non-planar symmetric spiral transformer. The primary is shown on the left and the secondary on the right. These inductors actually reside on top of one another.

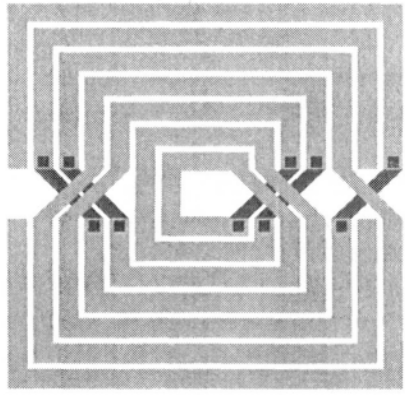


Figure 2.15. A planar symmetric balun transformer.

electromagnetic energy from coupling to the substrate [Yue and Wong, 1997]. While electrostatic shielding works well for capacitors and RF pads, especially at "low-frequencies", the shield must be patterned in the case of inductors so as to avoid or reduce the effects of eddy currents. Due to the close proximity of the device and the shield, using solid metallization would allow "image" eddy currents to flow which would produce an opposing magnetic field. This would reduce the device magnetic energy storage and hence the Q factor. A patterned shield, similar to Fig. 2.16, only allows shield currents to flow perpendicular to the conductive paths of a spiral inductor, thereby preventing the majority of eddy currents which flow parallel to the device[9].

The effects of a shield, or any other "grounding" structure, can be analyzed with the techniques presented in this research. At first glance a patterned shield seems like a very effective means of improving the device Q. This, however,

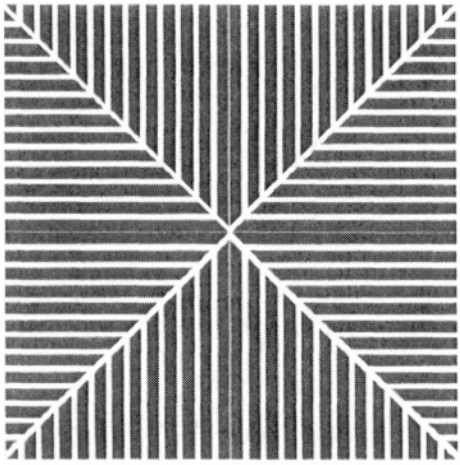


Figure 2.16. A patterned ground shield.

must be carefully examined for each process and product. There are many obvious problems with a shield, such as drastically reduced self-resonant frequency. There are also more subtle problems with a shield.

To see this, consider the inductor layout as a lossy inductor and a parasitic lossy capacitor, C_s in series with a loss resistor R_s. This frequency-independent representation is certainly valid as the network two-port parameters can be transformed into this equivalent circuit uniquely at a given frequency. In practice, the addition of a few elements, such as a lossless interwinding capacitor, broadbands the model, especially below self-resonance [Niknejad, 1997]. Then the action of the shield is to effectively increase the Q_C factor of the capacitor portion of the equivalent circuit. Consider a series to parallel transformation of this lossy capacitor into C_p in shunt with R_p

$$R_p = (1 + Q_C^2)R_s \tag{2.22}$$

$$C_p = \frac{Q_C^2}{1 + Q_C^2}C_s \tag{2.23}$$

Suppose that without a shield the Q_C factor of this capacitor is very low, $Q_C \ll 1$, such that

$$R_{low} \approx R_s \tag{2.24}$$

$$C_{low} \approx Q_C^2 C_s \tag{2.25}$$

If R_{low} is then already larger than the parallel equivalent inductor loss resistor, then this capacitor plays a minor role in determining the overall Q and shielding will actually deteriorate the performance of the device. Note that since

$C_{low} \ll C_s$, the effective shunt capacitance is small and does not lower the self-resonant frequency of the device. Clearly, in this case shielding does not help. Now consider the other extreme where the resistor R_{low} actually loads the tank significantly. Then clearly increasing the capacitor Q_C factor helps since for $Q_C \gg 1$

$$R_{high} \approx Q_C^2 R_s \tag{2.26}$$

$$C_{high} \approx C_s \tag{2.27}$$

This occurs, on the other hand, at the expense of lowered self-resonance since now the parasitic capacitor loads the tank. So we see the shield may potentially improve the overall tank Q but at the cost of reducing the usable frequency range of the device. Another approach, discussed in [Burghartz et al., 1997], surrounds the device with an open halo of substrate contacts. This has the added benefit of increasing the capacitor Q_C without loading the tank capacitance significantly. Measurements by other researchers have also corroborated these findings [Yoshitomi et al., 9998].

Another potential problem with the shield arises due to finite non-zero ground inductance. Since most packaged ICs suffer both package and bond wire inductance, the actual zero potential resides off-chip and there is considerable "ground bounce" on-chip. A typical IC has several inductors and transformers and if they are shielded, they are all effectively tied to a common non-zero impedance point. Thus a parasitic coupling path exists from device to device.

For instance, in an amplifier, this intra-block leakage can either lower the gain (as negative feedback) or cause instability (as positive feedback). Inter-block coupling, on the other hand, can produce spectral leakage and spurs, jamming and reduced SNR (gain compression), or mode-locking.

Any IC, of course, suffers from parasitic substrate coupling but shielding increases the substrate capacitance value and removes the resistive isolation between the devices.

3.6 VARACTORS (REVERSE-BIASED DIODES)

Varactors, or variable capacitors, are usually constructed as back-biased diodes. Diodes are realized as junctions between *p*-type and *n*-type doped regions of Si. Briefly, the presence of an *n*-type region abutting a *p*-type region creates a large concentration gradient which leads to diffusion. This diffusion forms the so-called "space-charge" or "depletion" region, a charged volume of space surrounding the junction. The charge of this region is due to the charge of immobile dopant sites. Donor or acceptor atoms are easily ionized at room temperature resulting in free mobile carriers. The diffusion process is balanced by the conduction current resulting from the built-in electric field since the charge buildup from the ionized immobile dopants produces an electric field

opposing the diffusion. Thus a natural barrier is formed against charge crossing the junction boundary.

If the diode is forward biased beyond a threshold, carriers obtain sufficient energy to cross the barrier. On the other hand, if the diode is reverse-biased, no DC conduction current flows. In reality, a small current will flow even under reverse-biased conditions. One contribution to these "leakage currents" is due to minority carriers which are created on average within a diffusion length of the junction. These carriers are actually propelled by the built-in electric field and cross the junction. Note that this current is relatively bias-independent[10].

While little to no DC current flows, AC displacement currents can flow since the reverse-biased diode acts very much like a capacitor. The thin space-charge region serves as the "insulator" and the *n*-type and *p*-type Si regions act as capacitor plates. Electric energy is stored in the built-in field of the space-charge region. Under an abrupt junction assumption, the depletion region thickness is given by [Neudeck, 1989]

$$d_j = \left[\frac{2K_S\epsilon_0(V_{bi} - V_A)}{q}\frac{(N_A + N_D)}{N_A N_D}\right]^{1/2} \tag{2.28}$$

N_D and N_A are the donor and acceptor volume densities and V_{bi} is the built-in potential given by

$$V_{bi} = \frac{kT}{q} ln\left[\frac{N_D N_A}{n_i^2}\right] \tag{2.29}$$

where complete ionization of the donors and acceptor is assumed. Note that under reverse-biased conditions, the spacing d_j is a function of the applied voltage $V_A < 0$. This effect can be exploited to produce a variable capacitor. The physical origin of the dependence of d_j on V_{rev} is simply due to the fact that increasing the voltage difference increases the electric field which requires more charge which in turn comes from extending the space charge region. If $N_D \gg N_A$ then the fractional increase in d_j from the donor side is small compared to the acceptor sites. The opposite, of course, also applies. Thus, in such a case the series resistance of the diode in reverse bias will be dominated by the acceptor side, or the *p*-type Si, due to its higher resistivity. Note that if both the donor and acceptor densities are increased in an attempt to lower the series losses, the capacitance gets large. While this is desirable, the breakdown voltage of the reverse-biased junction decreases. For avalanche breakdown we have [Neudeck, 1989]

$$V_{BR} = \left(\frac{\mathcal{E}_C T^2 \epsilon_r}{2q}\right)\left[\frac{N_A + N_D}{N_A N_D}\right] \tag{2.30}$$

The same applies for Zener breakdown since decreasing d_j increases the probability that a carrier will tunnel through the potential barrier. Since the operating

voltage of ICs is reducing, a proportional decrease in the breakdown voltage is tolerable allowing lower series resistance and thus higher Q values to be realized.

The equivalent capacitance of a junction with cross-sectional area A is thus given by

$$C_j = \frac{\epsilon_r \epsilon_0 A}{d_j} = \frac{\epsilon_r \epsilon_0}{\left[\frac{2K_S \epsilon_0 (V_{bi} - V_A)}{q} \frac{(N_A + N_D)}{N_A N_D}\right]^{1/2}} \tag{2.31}$$

this can be rewritten as

$$C_j = \frac{C_{j0}}{\left[1 - \frac{V_A}{V_{bi}}\right]^m} \tag{2.32}$$

where C_{j0} is the junction capacitance for the zero applied bias case, $V_A = 0$. The factor m is equal to 1/2 for an abrupt junction but takes on the value of 1/3 for a linearly graded junction. For real junctions, one finds that $1/3 < m < 1/2$.

A pertinent factor when designing VCOs and filters with variable cutoff is the derivative of (2.32) with respect to the applied voltage V_A

$$K_{CV} = \frac{C_{j0} m}{V_{bi}(1 - \frac{V_A}{V_{bi}})^{m+1}} \tag{2.33}$$

Increasing this factor is important for maximizing the change in capacitance for a change in the reverse-biased voltage. This maximizes the tuning range of an LC tank, for instance, since

$$\frac{d\omega}{dV} = \frac{-K_{CV}}{2C_j} \omega_0 \tag{2.34}$$

where $\omega_0 = \sqrt{1/LC_j}$.

3.7 MOS CAPACITORS

Another technique to realize a variable capacitor is through an MOS capacitor. The MOS capacitor has four distinct regions of operation: accumulation, depletion, weak inversion, and strong inversion. In accumulation mode, the capacitance is equal to the oxide capacitance

$$C_{acc} = C_{ox} = \frac{\epsilon_r \epsilon_0 A}{t_{ox}} \tag{2.35}$$

where t_{ox} is the gate oxide thickness and A is the capacitor plate area. This assumes that majority carrier distribution equilibrates much faster than the period of the applied signal. We also assume that all the charge resides at the surface of the substrate (delta-function approximation for charge density) and we neglect the fringing fields. In reality, the charge distribution will not peak

at the surface but slightly away from the surface due to quantum mechanical considerations.

In depletion mode, majority carriers at the surface of Si are repelled by the applied electric field and the region is thus depleted of mobile carriers. A depletion region forms in a similar manner to the reverse-biased junction diode considered previously. The capacitance from the gate to the substrate, therefore, has two components, an oxide capacitance in series with a depletion capacitance [Pierret, 1990]

$$C_{dep} = \frac{C_{acc}C_{dep}}{C_{acc} + C_{dep}} = \frac{C_{ox}}{1 + \frac{\epsilon_{r,SiO_2} d_j}{\epsilon_{r,Si} t_{ox}}} \tag{2.36}$$

As the channel region begins to invert, minority carrier charge accumulates at the surface and the capacitance once again approaches the C_{ox} value. This assumes that the rate of minority carrier generation is sufficiently fast to follow the time-varying AC signal. On the other hand, if a MOSFET-C structure is used instead, a large source of minority carrier charge is available from the surrounding source and drain regions. Between inversion and depletion, or in weak inversion, the situation is more complicated as an incremental increase in the applied voltage can result in new field lines terminating on both immobile charges (by an increase in the depletion junction depth) and terminating on newly generated minority carriers.

In the weak inversion region, the capacitance changes rapidly from C_{dep} to C_{ox} due to the exponential build-up of charge at the surface. This results in a very large K_{CV} factor. The quality factor of such a capacitor is limited mainly by the gate series resistance and can be shown to be approximately equal to $R_G/3$[11].

3.8 RESISTORS AND CAPACITORS

Resistors can be implemented in several ways but all essentially involve moving current through a length of material with relatively high resistivity. In a standard bipolar process, diffusion resistors are common. The resistance value can be varied by the application of a bias such as in base-pinch and epitaxial pinch resistors [Gray and Meyer, 1993]. In a standard MOS process we also find diffusion resistors, polysilicon resistors, well-resistors, and the MOS device itself as variable resistors. In a modern bipolar, BiCMOS, or MOS process, most of the above options are available.

Capacitors are realized as either MOS capacitors, as poly-poly "metal" capacitors, metal-insulator-metal (MIM) capacitors, and as poly-diffusion capacitors. For RF applications MIM capacitors are preferred since they result in the lowest losses. Even though standard metal interconnection can be used to construct MIM capacitors, this has two negative consequences. First, it will result in a large uncertainty in the capacitance value due to wide oxide thickness variation

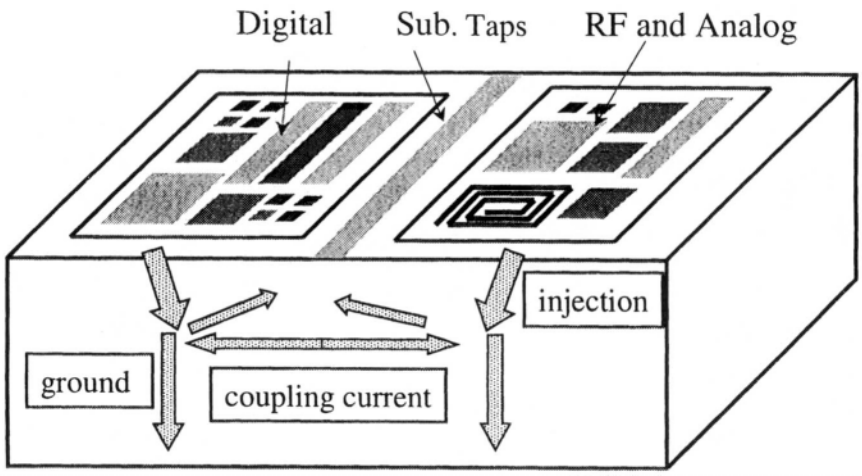


Figure 2.17. Substrate coupling in a mixed-signal system on a chip.

across the wafer. Second, the capacitance value per unit area will be very small due to the thick oxide separating the metal layers. A special MIM capacitor constructed with a well-controlled thin oxide is therefore preferred.

Resistors and capacitors have always been a part of the standard IC process. Thus, great effort has already gone into the analysis and modeling of such structures at lower frequencies. At higher frequencies, though, these structures have many non-ideal effects which must be taken into consideration, such as substrate coupling, self-resonance, and inductance. In this work we will focus on a general interconnection of metal structures above a lossy Si substrate and this generality will allow us to apply our analysis technique to inductors as easily as to capacitors, transformers, and resistors.

4. SUBSTRATE COUPLING

An issue that permeates this research is substrate coupling. Figure 2.17 illustrates the various substrate coupling mechanisms present in an IC environment. Current is injected into the substrate through various mechanisms. Physically large passive devices such as inductors, capacitors, transformers, interconnect and bonding pads inject displacement current in the substrate. This current flows vertically and horizontally to points of low potential in the substrate, such as substrate taps and the back-plane ground. This current couples to other large passive structures in a similar manner.

Active devices also inject current into the substrate, directly and capacitively. Since most active devices are isolated from the substrate by either reverse-biased

pn junctions or oxide, capacitive substrate current injection and reception occur at high frequencies. Direct current can also be injected into the substrate due to hot electron effects. In a short-channel MOS transistor, for instance, electron-hole pair creation takes place in the high-field pinch-off region near the drain due to collisions. The electric field lines lead one set of the carriers into the substrate. Under high field conditions, a multiplication process (avalanche) can also lead to additional hole-pair creation and result in substantial currents.

Coupling between sensitive analog nodes can lead to instability or gain reduction. Furthermore, in a mixed signal IC the coupling between analog and digital portions of the chip can be very problematic. Digital gates switch in a pseudo-random manner and each gate transition results in some energy transfer to the substrate. This energy can couple to sensitive analog nodes and result in reduced signal-to-noise (SNR) ratios.

Though careful layout techniques and differential operation minimize them, these effects can never be totally eliminated or ignored in the design of mixed-signal ICs. Substrate coupling is treated extensively in [Gharpurey and Meyer, 1996]. In this book we draw from an extension of this work, presented in [Niknejad et al., 1998], to analyze passive device substrate coupling.

Notes

1 This depends on the I-V curve of a diode. A tunnel diode can certainly be biased as to supply energy, becoming an active device.

2 This is a necessary condition for a passive device. However, this condition alone is not sufficient since power gain may occur for a non-matched load or source impedance.

3 This is not a rigorous argument since the skin-depth concept of surface impedance applies strictly to a semi-infinite conductor.

4 This is due to the linearity of Maxwell's Equations.

5 In general the eddy currents are solenoidal whereas the displacement and conductive currents are curl free

6 By this we mean the crystal structures of the adjacent layers are compatible.

7 By intrinsic we mean no intentional dopants are introduced into the substrate and the only conduction occurs through thermionic emission into the conduction band.

8 This result will be established in Chapter 5.

9 Current distributions with zero curl lead to zero magnetic fields.

10 Consider the current associated with a stream of suicidal lemmings. Given that the the cliff is sufficiently high, the suicide rate of lemmings is independent of the height of the cliff.

11 See Appendix A.

Chapter 5

INDUCTANCE CALCULATIONS

1. INTRODUCTION

In this chapter we will define inductance, calculate the low frequency inductance of several common configurations of conductors, and extend the results of our calculations to high frequency.

Up until recently, the calculation of inductance has received relatively little attention from the integrated circuit community. Only very specialized fields, such as power engineering, have dealt directly with inductors. The reason for this is simply related to the relatively minor role that inductors have played at lower frequencies, especially in the IC environment. In such cases, parasitic capacitance plays a dominant role in determining circuit behavior whereas the effects of parasitic inductance are minor. The physical origin of this stems from the relatively small size of integrated circuits. The typical inductance associated with long integrated circuit metal traces is on the order of 1 nH, a small enough value of inductance as to produce negligible reactance at typical IC frequencies less than 100 MHz. Today, RF, microwave, and digital ICs are operating at 1 - 10 GHz and these inductive effects are no longer negligible.

Comparatively, magnetic forces tend to have longer range effects than electric forces. This can be attributed to the absence of magnetic monopoles. Whereas electric field lines terminate on charges in nearby conductors and the substrate, magnetic field lines are divergenceless and only wane in the presence of induced eddy currents.

At very high frequencies, frequencies often encountered by the microwave community, the skin depth δ is small and magnetic fields are confined to the external volume of conductors. In such a case, the inductances of a system of coupled multi-conductor transmission lines can be derived from the capacitance matrix. Therefore, more emphasis has been placed on deriving the capacitance

matrix as the source of capacitance is charge derived from a scalar Poisson's equation whereas the source of inductance is from currents satisfying the vector Poisson's equation.

The situation has changed as on-chip frequencies have increased to the point where magnetic effects are appreciable. In his pioneering work, Ruehli [Ruehli, 1972] focused attention on efficient techniques for predicting inductive effects in an integrated circuit environment at moderate frequencies where internal inductance cannot be ignored. These techniques are widely applicable, even to board and package environments.

2. DEFINITION OF INDUCTANCE

Inductance is the electric dual of capacitance. While capacitors store electric energy, inductors store magnetic energy. The self-inductance of a circuit can be thought of as the "mass", or resistance to change in "motion", of the circuit. The origin of this "mass" comes from the magnetic field of a circuit, which originates from the currents in the circuit. Any change in the current of a circuit induces a change in the magnetic field. From Faraday's law we know that a changing magnetic field induces an electric field. From Lenz's law we can deduce further that this induced electric field always opposes further change in the current.

2.1 ENERGY DEFINITION

From circuit theory the total magnetic energy stored by an inductor is given by

$$W_m = \frac{1}{2} L I^2 \tag{5.1}$$

If we thus calculate the total magnetic energy in a volume V by a physical inductor

$$W_m = \frac{1}{2} \int_V \mathbf{B} \cdot \mathbf{H} dV \tag{5.2}$$

we can equate the above quantities to obtain the inductance.

We have already alluded to the energy definition of an inductor in Chapter 2 when we examined the complex power flowing into a black box

$$P = \frac{1}{2} \oint_S \mathbf{E} \times \mathbf{H}^* \cdot d\mathbf{s} = P_l + 2j\omega(W_m - W_e) \tag{5.3}$$

implying that the device acts inductively if $W_m > W_e$. This also has some interesting implications for the impedance of a lossless passive device as a function of frequency. It can be shown that [Ramo et al., 1994]

$$\frac{dX}{d\omega} = 4\frac{W_m + W_e}{II^*} \tag{5.4}$$

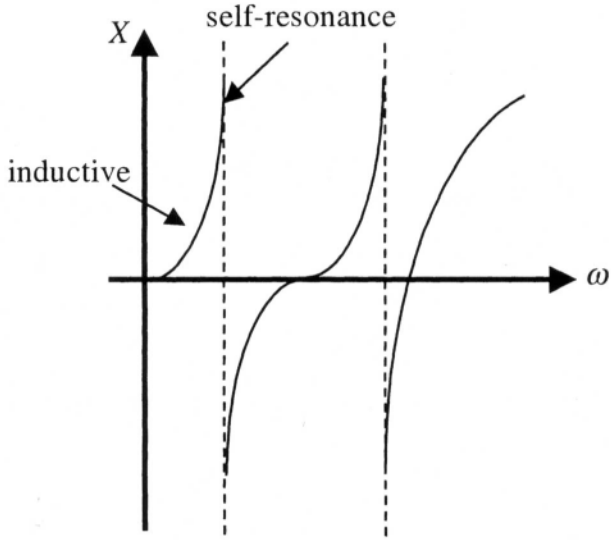


Figure 5.1. The reactance of a lossless passive device as a function of frequency.

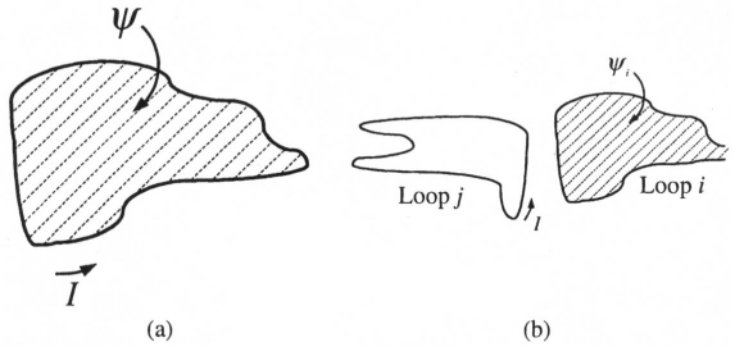


Figure 5.2. (a) An isolated current loop. (b) A magnetically coupled pair of loops. Current only flows in loop *j*.

which implies that the slope of the impedance versus frequency is never negative. Such a typical profile as a function of frequency is shown in Fig. 5.1.

2.2 MAGNETIC FLUX OF A CIRCUIT

Consider an arbitrary closed-circuit formed by conductors, as shown in Fig. 5.2a. The magnetic flux of this circuit is defined as the magnetic field

crossing the cross-sectional area of the circuit

$$\psi = \oint_S \mathbf{B} \cdot d\mathbf{S} \tag{5.5}$$

The origin of the magnetic field **B** is from the circuit itself since there are no other currents in the system. Now, the self-inductance of the circuit is simply defined as

$$L = \frac{\psi}{I} \tag{5.6}$$

where I is the total current flowing in the circuit. If we now consider an arrangement of loops as shown in Fig. 5.2b, then if we let current flow in loop j and measure the impinging flux on loop i, we have the following definition of mutual inductance

$$M_{ij} = \frac{\psi_i}{I_j} \tag{5.7}$$

where

$$\psi_i = \oint_{S_i} \mathbf{B} \cdot d\mathbf{S}\Big|_{I_k=0 \,\forall\, k\neq j} \tag{5.8}$$

By Faraday's well-known law, the voltage induced on a loop is related to the flux as follows

$$V = \frac{d\psi}{dt} \tag{5.9}$$

and by the definition of inductance this is simply

$$V = L\frac{dI}{dt} \tag{5.10}$$

2.3 MAGNETIC VECTOR POTENTIAL

From Maxwell's equation, we know that all physically observable magnetic fields are solenoidal, i.e. $\nabla \cdot \mathbf{B} = 0$ and thus we can write $\mathbf{B} = \nabla \times \mathbf{A}$ for some vector potential function A. Note that we are free to choose the divergence of **A** as we wish. From the definition of flux, we can invoke Stoke's theorem to obtain

$$\psi = \oint_S (\nabla \times \mathbf{A}) \cdot d\mathbf{S} = \oint_C \mathbf{A} \cdot d\mathbf{l} \tag{5.11}$$

In many practical cases, it is more convenient to work with the magnetic vector potential **A** than the magnetic field **B**. Although the magnetic field has a more physical interpretation, the magnetic vector potential is often simpler to derive.[1] The governing equation for the vector potential is the vector Poisson's equation derived in Chapter 4

$$\nabla^2 \mathbf{A} = \mu \mathbf{J} \tag{5.12}$$

The solution of this equation can be written with the aid of a dyadic Green function $\bar{\mathbf{G}}$

$$\mathbf{A}(\mathbf{r}) = \mu \int_{V'} \bar{\mathbf{G}}(\mathbf{r}, \mathbf{r}') \cdot \mathbf{J}(\mathbf{r}') dV' \tag{5.13}$$

where for instance the component of $\bar{\mathbf{G}}$ for a test source in the direction of $\hat{x}_i$ is a solution of

$$\nabla^2 \mathbf{G}_i = \mu \delta(\mathbf{r} - \mathbf{r}') \hat{x}_i \tag{5.14}$$

In an isotropic medium, such as free-space or a spherically symmetric arrangement of conductors, $G_x = G_y = G_z$ and $G_{km} = 0 \; \forall \; k \neq m$ and thus the equation can be simplified

$$\mathbf{A}(\mathbf{r}) = \mu \int_{V'} G(\mathbf{r}, \mathbf{r}') \mathbf{J}(\mathbf{r}') dV' \tag{5.15}$$

By definition, then, the mutual inductance between two loops can be derived from the voltage induced in one loop (unprimed coordinates) for a changing current flowing in some "source" loop (primed coordinates). This voltage is given by

$$V_{\text{fld}} = - \oint_C \mathbf{E} \cdot d\mathbf{l} \tag{5.16}$$

From Maxwell's equation, the electric field is given by

$$\mathbf{E} = -\nabla \phi - \frac{dA}{dt} \tag{5.17}$$

Since the first term is "conservative" in nature, its line integral along a closed path yields zero. Thus, we have

$$V_{\text{fld}} = \frac{d}{dt} \oint_C \mathbf{A} \cdot d\mathbf{l} = M \frac{dI}{dt} \tag{5.18}$$

Hence

$$M = \frac{1}{I} \oint_C (\mu \int_{V'} G(\mathbf{r}, \mathbf{r}') \mathbf{J}(\mathbf{r}') dV') \cdot d\mathbf{l} \tag{5.19}$$

At low frequencies, the current distribution $\mathbf{J}$ is uniform and the second integration therefore only involves the function G. We can view this volume integral as the line integral of the cross-sectional area of the conductor along the path of the conductor. Therefore

$$\int_{V'} G(\mathbf{r}, \mathbf{r}') J(\mathbf{r}') dV' = I \oint_{C'} (\int_{A'} G(\mathbf{r}, \mathbf{r}') dA') dl' = I \oint_{C'} \tilde{G}(\mathbf{r}, \mathbf{r}') dl' \tag{5.20}$$

where $\tilde{G}$ is the value of G averaged over the cross-sectional area of the "source" circuit at any location along the loop. Substituting the above in the equation for inductance we have

$$M = \mu \oint_C \oint_{C'} \tilde{G}(\mathbf{r}, \mathbf{r}') d\mathbf{l} \cdot d\mathbf{l}' \tag{5.21}$$

The current factors out in the above equation and thus we see that the inductance at low frequency is simply related to the geometry of conductors in question. For the case of free-space

$$G(\mathbf{r},\mathbf{r}') = \frac{1}{4\pi|\mathbf{r}-\mathbf{r}'|} \tag{5.22}$$

If the source loop is filamental, then we have Neumann's famous result [Ramo et al., 1994]

$$M = \frac{\mu}{4\pi}\oint\oint \frac{d\mathbf{l}\cdot d\mathbf{l}'}{R} \tag{5.23}$$

which incidentally proves the reciprocity relation $M_{ij} = M_{ji}$.

Evidently, the above result is dependent on the path of integration in the "source" loop. For instance, if we take the inner most turn in a loop of finite width, the value of inductance is different from the outer most loop, due to path length difference and the non-uniformity of the magnetic field. This path dependence is in fact a paramount property of a solenoidal field. Thus, to obtain a path independent mutual inductance, we should average M over all paths. Consider the set of all paths parameterized by $s \in [0,1]$ tangent to the current flow $K = \{\mathbf{r}_1(s) : \mathbf{n}(\mathbf{s})\cdot\mathbf{J} = 0\}$. The vector $\mathbf{r}_1(\mathbf{s})$ traces one particular path and the vector $\mathbf{n}(\mathbf{s})$ is always normal to the path. Let the average path be described by the following convergent infinite sum

$$\mathbf{r}(s) = \sum_{j\in K} \mathbf{r}_j(s) \tag{5.24}$$

and let this path be described by C_1. Now let us integrate the field quantity $\mathbf{E}'$ along this path to obtain a scalar value with units of voltage

$$V' = -\oint_{C_1} \mathbf{E}'\cdot d\mathbf{l}_1 \tag{5.25}$$

where the field quantity $\mathbf{E}'$ is the average value of the electric field along the cross-section of the path C_1

$$\mathbf{E}'(r) = \int_{A_1} \mathbf{E}(\mathbf{r})dA_1 \tag{5.26}$$

We can thus interpret (5.25) as the average voltage induced in our "source" loop. Thus, the average mutual inductance becomes

$$M = \mu\oint_{C_1}\oint_{C'} H(\mathbf{r},\mathbf{r}')d\mathbf{l}\cdot d\mathbf{l}' \tag{5.27}$$

where H is given by

$$H(\mathbf{r},\mathbf{r}') = \int_{A_{\text{src}}}\int_{A_{\text{fld}}} G(\mathbf{r},\mathbf{r}')dA_{\text{src}}dA_{\text{fld}} \tag{5.28}$$

We can interpret the function $H(\mathbf{r}, \mathbf{r}')$ as the average value of G evaluated for all points in the cross-sectional areas of the source and field loops at a given position $\mathbf{r}$ in the field and $\mathbf{r}'$ in the source loops.

3. PARALLEL AND SERIES INDUCTORS

When N inductors are connected in series, the effective mutual inductance can be computed easily since the voltage across the series connection gives

$$V = \sum_{i=1}^{N} V_i = \sum_{i=1}^{N}\sum_{j=1}^{N} sM_{ij}I_j = Is\sum_{i=1}^{N}\sum_{j=1}^{N} M_{ij} \tag{5.29}$$

where we have assumed that the branch current is equal to I for all branches since we are neglecting displacement current at low frequencies. Clearly, then

$$L_{se} = \sum_{i=1}^{N} L_i + 2\sum_{i=1}^{N}\sum_{j\neq i} M_{ij} \tag{5.30}$$

The case of parallel connected inductors is more complicated. The result has already been stated in Chapter 2

$$L_{sh} = \frac{1}{\sum_{i=1,j=1}^{N} K_{ij}} \tag{5.31}$$

where the matrix K is the inverse of the partial inductance matrix M. To see that this is true in general, consider N-shunt connected coupled inductors described by the partial inductance matrix M,

$$\mathbf{v} = j\omega M\mathbf{i} \tag{5.32}$$

where element v_j of vector $\mathbf{v}$ is the voltage across jth inductor and the element i_j of vector $\mathbf{i}$ is the current through the jth inductor. Since the voltages are connected in shunt, all voltages are equal

$$v_x = v_1 = v_2 = \cdots = v_N \tag{5.33}$$

and thus

$$\mathbf{v} = v_x\mathbf{s} \tag{5.34}$$

where $\mathbf{s} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ and the total current is given by

$$i_x = i_1 + i_2 + \cdots + i_N = \mathbf{s}^T\mathbf{i} \tag{5.35}$$

Since the effective input inductance is given by

$$L_{sh} = \frac{v_x}{j\omega i_x} \tag{5.36}$$

we have from (5.32)

$$(j\omega M)^{-1}\mathbf{v} = (j\omega M)^{-1}\mathbf{s}v_x \tag{5.37}$$

and further

$$\mathbf{s}^T(j\omega M)^{-1}\mathbf{s}v_x = \mathbf{s}^T\mathbf{i} = i_x \tag{5.38}$$

and

$$L_{sh} = \frac{1}{\mathbf{s}^T M^{-1}\mathbf{s}} \tag{5.39}$$

which establishes (5.31). In [Ruehli, 1972] the above result is stated in the following form

$$L_{sh} = \left[2\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\operatorname{cof}M_{ij} + \sum_{i=1}^{N}\operatorname{cof}M_{ii}\right]^{-1}\det M \tag{5.40}$$

4. FILAMENTAL INDUCTANCE FORMULAE FOR COMMON CONFIGURATIONS

Consider a $\hat{\mathbf{z}}$ directed filament of finite length l carrying a current I in free-space located at the origin of a rectangular coordinate system. The magnetic vector potential at any point in space is given by (5.15)

$$\mathbf{A}(x,y,z) = \hat{\mathbf{z}}\frac{\mu_0 I}{4\pi}\int_{-l/2}^{l/2}\frac{1}{\sqrt{x^2+y^2+(z-z')^2}}dz' \tag{5.41}$$

Performing the above integration we have

$$A_z(x,y,z) = \hat{\mathbf{z}}\frac{\mu_0 I}{4\pi}\log\left[\frac{(z+l/2)+\sqrt{x^2+y^2+(z+l/2)^2}}{(z-l/2)+\sqrt{x^2+y^2+(z-l/2)^2}}\right] \tag{5.42}$$

Now consider the mutual magnetic inductance of two such parallel filaments separated by a distance d. If we integrate the expression for A_z along the path of the second filament we obtain

$$M(l,d) = \int_{-l/2}^{l/2} A_z(d,0,z')dz' \tag{5.43}$$

The above integral can be simplified into the following form

$$M(l,d) = \frac{\mu_0}{4\pi}l\left[\log\left(\sqrt{1+\left(\frac{l}{d}\right)^2}+\frac{l}{d}\right) - \sqrt{1+\left(\frac{d}{l}\right)^2}+\frac{d}{l}\right] \tag{5.44}$$

The above expression is the starting point for calculating the mutual inductance between two arbitrary parallel filaments. The technique is described in Grover [Grover, 1946] and Ruehli [Ruehli, 1972].

5. CALCULATION OF SELF AND MUTUAL INDUCTANCE FOR CONDUCTORS

Following the averaging procedure described in (5.21), we can find the mutual inductance between two parallel conductors of arbitrary cross-section by integrating the filamental formula (5.44) over the area of the cross-section of the conductors

$$M_{12} = \frac{1}{A_1 A_2} \int_{A_1} \int_{A_2} M(L,r) dA_1 dA_2 \tag{5.45}$$

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 an integrable singularity in the calculation.

5.1 THE GEOMETRIC MEAN DISTANCE (GMD) APPROXIMATION

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

$$M(L,d) \approx \frac{\mu_0}{4\pi} L \left(\frac{d}{L} - \log d + \log 2L - 1 \right) \tag{5.46}$$

Integration of the above expression across the cross-section yields a simpler calculation. The only difficult term to integrate is the $\log(d)$. This term, the *Geometric Mean Distance*, or GMD, is the average value of the logarithm of the distance between all points between the conductors. This integral can be calculated for several simple configurations, especially for arbitrary rectangular cross-sections. Some of these results can be found in [Niknejad, 1997].

A simple interpretation of the above result can be obtained if we make the $L \gg d$ approximation from the outset. If we let $L \to \infty$ we have the following 2D magnetic vector potential

$$\mathbf{A}(x,y) = \hat{z} \frac{\mu_0 I}{4\pi} \log \sqrt{x^2 + y^2} \tag{5.47}$$

And thus the mutual inductance per unit length between two infinite parallel conductors is given by

$$M_{12} = \frac{\mu_0}{4\pi} \int_{A_1} \int_{A_2} \log(r) dA_1 dA_2 \tag{5.48}$$

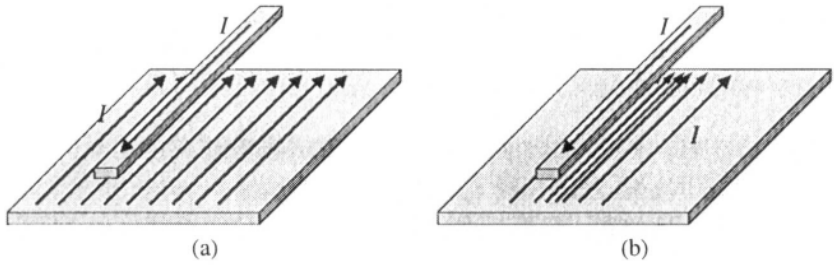


Figure 5.3. (a) Ground current at low frequency. (b) Ground current at high frequency.

where r takes on the value of the distance between every pair of points in the cross-section of each conductor. With the exception of the Arithmetic Mean Distance (AMD) correction factor, this is exactly the same expression found with the GMD approximation.

6. HIGH FREQUENCY INDUCTANCE CALCULATION

6.1 BACKGROUND

All the equations derived thus far for inductance have made the assumption of uniform current distribution within the volume of conductors. While this applies to low frequencies, we know that at high frequencies currents redistribute to minimize the energy of the system. Consider, for instance, a conductor carrying a low frequency current above a ground plane as shown Fig. 5.3a. At low frequencies (strictly DC) the return current flowing in the ground plane flows uniformly and so the resistance per unit length for the entire system can be computed by

$$R_{\text{tot}} = R_{\square}\left(\frac{1}{W} + \frac{1}{W_{gnd}}\right) \tag{5.49}$$

At higher frequencies, though, currents redistribute to minimize the energy of the system as shown in part (b) of Fig. 5.3. The magnetic energy of the system can be minimized by minimizing the inductance value per unit length. Thus, the ground currents redistribute and concentrate under the signal carrying conductor.[2] The ground current will flow as close as possible to the signal carrying current in order to minimize the magnetic field (since ground currents flow in the opposite direction). The high frequency resistance is thus closer to

$$R_{\text{tot}} \approx 2\frac{R_{\square}}{W} \tag{5.50}$$

Another perspective is to note that AC current takes the path of least impedance. Now consider the inductance value of the various paths through the ground plane. As shown in Fig. 5.4 a path far removed from the substrate has a large

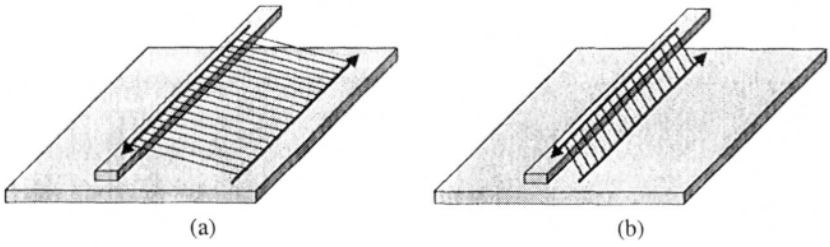


Figure 5.4. The cross-sectional area of a return path loop (a) far removed from the conductor and (b) near the conductor.

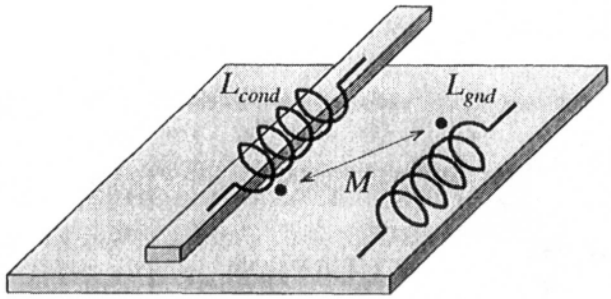


Figure 5.5. The self and mutual inductance of the conductor and return path.

cross-sectional area and thus a large value of inductance. A path directly under the inductor involves the smallest area and thus the smaller value of inductance. Therefore, AC current will flow in paths of low inductance.

A final perspective to this problem can be drawn from our study of "partial" inductance. As shown in Fig. 5.5, the total inductance for a typical path in the ground plane is given by

$$L_{\text{tot}} = L_{cond} + L_{gnd} - 2M \tag{5.51}$$

Since M varies from path to path whereas the L for a typical path is constant, nature chooses a path to maximize M so as to minimize the total inductance.

It is interesting to note that the high frequency current distribution is a result of the electric field term of

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \tag{5.52}$$

rather than the magnetic field term. This is a result of the electromagnetic interaction or Faraday's law. To see this is true consider the static magnetic force on two current carrying wires. It is well known that if the currents are in opposite directions, the wires repel one another. So based on the magnetic

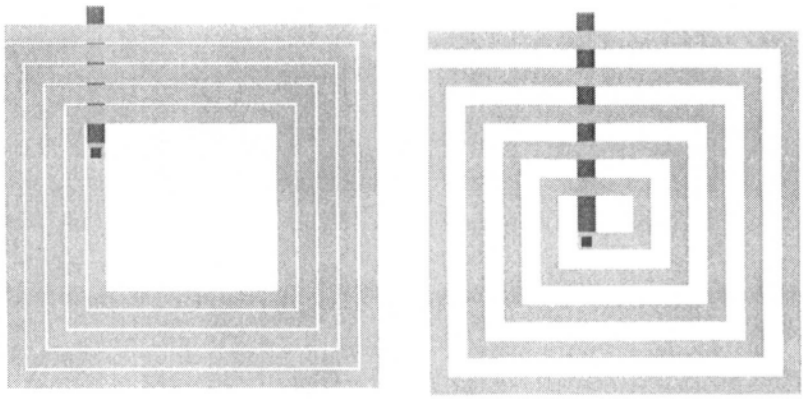


Figure 5.6. Two inductors realized with different metal spacing values of 1 μm and 10 μm.

forces alone the current distribution in our example should be the exact opposite of Fig. 5.3b.

6.2 EXAMPLE CALCULATION

We will now illustrate the importance of calculating the partial inductance matrix at high frequency with an example. Consider the two spiral inductor layouts shown in Fig. 5.6. The inductor geometries are identical with outer length $L = 200\ \mu$m, and five turns of metal pitch $W = 10\ \mu$m. In (a), though, the metal spacing $S = 1\ \mu$m whereas in (b) $S = 10\ \mu m$.

We analyze both devices by computing the partial inductance matrix $\tilde{Z}$. Each segment of the spirals is sub-sectioned into twenty separate conductors producing a 400 x 400 matrix $\tilde{Z}$. The matrix is reduced in order by computing $Z^M = (S\tilde{Z}^{-1}S^T)^{-1}$ and then by summing over the elements to obtain

$$R(f) + j\omega L(f) = \sum_{i,j} Z^M_{i,j}(f) \tag{5.53}$$

From an inductance point of view, using the smaller value of spacing results in a higher DC inductance value of 5.2 nH as opposed to 2.6 nH. From Fig. 5.7 we see that as a function of frequency, the inductance values of both devices decrease slightly approaching the external inductance limit[3], falling about 4% from the DC value at 5 GHz.

The series resistance of the device, though, increases much more rapidly for the smaller value of spacing, as shown in Fig. 5.7b, where the ratio of AC to DC resistance is plotted. While the device with large spacing shows a 40% increase in series resistance at 5 GHz compared to DC, the device with small spacing suffers nearly twice as much loss at 5 GHz compared to low frequency. This

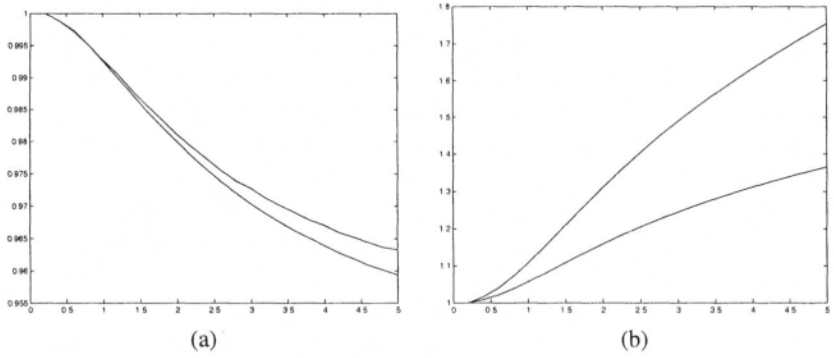


Figure 5.7. The ratio of AC to DC inductance and resistance as a function of frequency plotted for two inductors of Fig. 5.6.

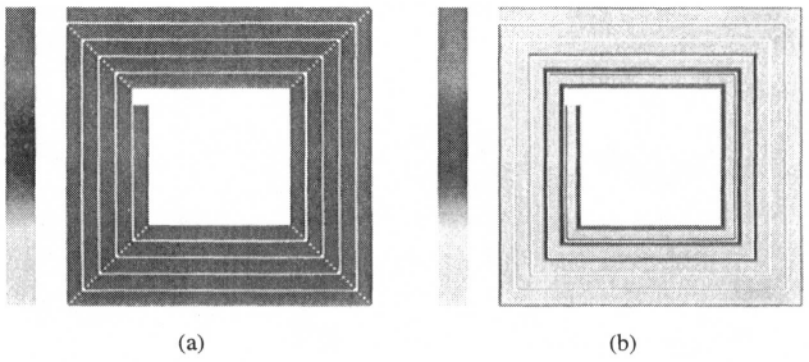


Figure 5.8. The current density at (a) 1 GHz and (b) 5 GHz.

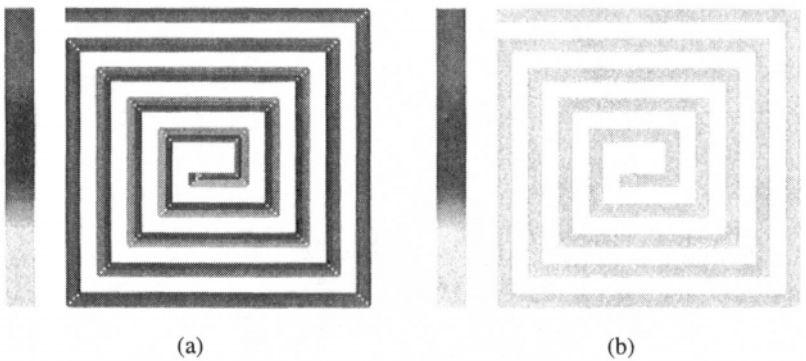


Figure 5.9. The current density at (a) 1 GHz and (b) 5 GHz.

increase is faster than the $\sqrt{f}$ increase predicted by simple skin effect theory. To see why, note that the magnetic field is non-uniform across the cross-section of the inductor and increases rapidly towards the center of the device. Thus the magnetic field penetrating the conductors results in eddy currents with a more profound influence over the loss at the inner core of the spiral. A plot of the current density[4], shown in Fig. 5.8, shows that the 1 μm design suffers from current constriction at 5 GHz in the innermost turn. On the other contrary, as shown in Fig. 5.9, the 10 μm design has uniform current distribution even at 5 GHz, more uniform than at 1 GHz.

Now consider an optimization routine searching for a value of spacing S to optimize the Q of the device. If the optimizer makes the assumption that resistance increases like $\sqrt{f}$ independent of spacing, then clearly the optimizer will select the smallest possible value of S to maximize the inductance. In other words, an optimizer will find an incorrect optimal solution, an artifact created by the weakness of the modeling process.

Notes

1 Magnetic vector potential also has physical meaning in the study of quantum mechanics [Feynman et al., 1963].

2 For simplicity we have ignored the current distribution in the signal conductor.

3 Please note that this is the physical inductance of the device and not the effective inductance. In other words, the capacitive effects have been completely ignored in computing (5.53).

4 See Chapter 7 for an explanation of how these plots were generated. Note that these plots show the current distribution after solving the electric and magnetic problem.

Chapter 8

EXPERIMENTAL STUDY

In this chapter we will compare measured device characteristics to simulated predictions based on the techniques described in the previous chapters. In a previous report [Niknejad, 1997], experimental results confirmed the validity of the our approach. Several structures were fabricated and measured in a BiCMOS process including square and polygon spiral inductors, planar and non-planar transformers, and coupled inductors.

In this chapter, we would like to make the same comparisons but with devices fabricated over a conductive substrate. As previously discussed, a heavily conductive substrate gives rise to new loss mechanisms, in particular magnetically induced eddy current losses. Conductive substrates are typically employed in CMOS processes to minimize latch-up. Latch-up occurs when the parasitic lateral and vertical bipolar transistors of a CMOS inverter turn on in a positive feedback loop [Johns and Martin, 1997]. To minimize voltage drops in the substrate capable of potentially forward biasing these bipolar devices, a heavily conductive highly doped bulk substrate is employed. A more resistive, lightly doped layer is then grown epitaxially over this substrate to house the wells and MOS transistors, hence the name "epi" substrate. A cross-section of such a process is shown in Fig. 2.5. Another variation, similar to the bipolar substrate, is to start with a lightly doped substrate but to grow a heavily conductive surface layer to shunt substrate currents to a low potential.

1. MEASUREMENT RESULTS

As shown in Fig. 8.1, several planar and non-planar spiral inductors have been fabricated in National Semiconductor's $0.25\ \mu m$ CMOS-8 process. This process utilizes a bulk substrate of $10\ \Omega\text{-cm}$, sufficiently resistive that eddy currents play a minor part in the bulk. However, the top layer of Si is fairly conductive at $15 \times 10^{-4}\ \Omega\text{-cm}$. The thickness of this layer is less than $1\ \mu m$ but

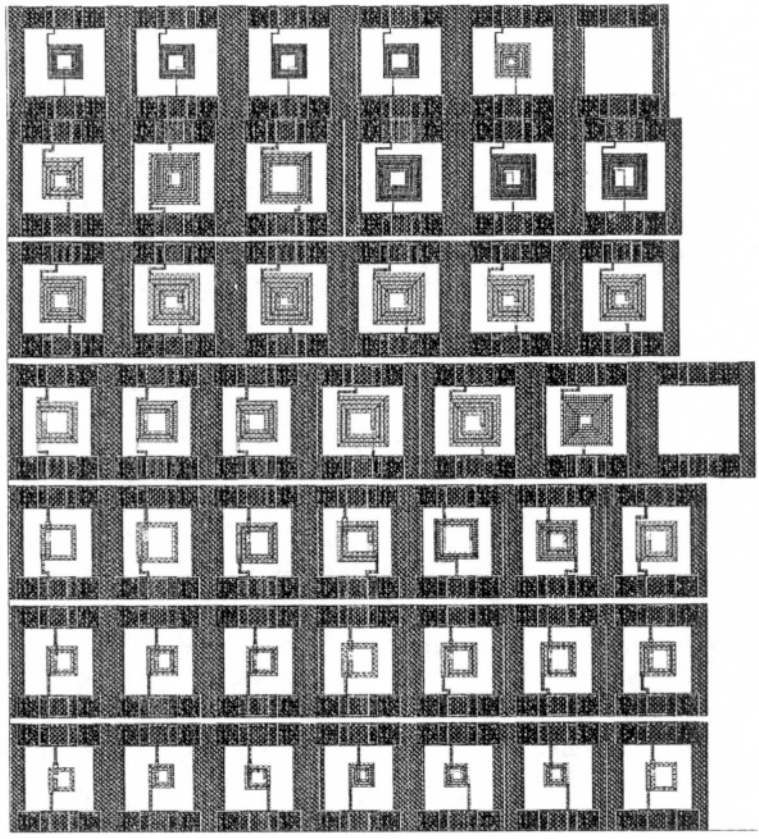


Figure 8.1. Inductor test structures.

this is enough to cause significant eddy current loss. The process parameters are summarized in Table 8.1.

The layout of two representative inductors is summarized in Table 8.2. As shown in Fig. 8.2, spiral inductor L27 is a planar device utilizing the top metal layer. Inductor L19 resides on metal layers M5-M3 connected in series realizing a large inductance value in a relatively small area. The layout of L19 is shown in Fig. 8.3.

2. DEVICE CALIBRATION

Measurements are performed using a ground-signal-ground co-planar waveguide pad configuration. The *s*-parameters are measured using the HP 8719C Network Analyzer. G-S-G co-planar cascade probes are used and the setup is calibrated using the Cascade Microtech 832210 calibration substrate. The

Table 8.1. CMOS Process Parameters Summary

Metal 5	$R_{sh} = 35\,m\Omega/\text{sq}$	$t = 0.91\mu\text{m}$	$C_{sub} = 6.12aF/\mu\text{m}^2$ $C_{M4} = 53.1aF/\mu\text{m}^2$ $C_{M3} = 19.1aF/\mu\text{m}^2$
Metal 4	$R_{sh} = 60\,m\Omega/\text{sq}$	$t = 0.51\mu\text{m}$	$C_{sub} = 7.69aF/\mu\text{m}^2$ $C_{M3} = 53.1aF/\mu\text{m}^2$
Metal 3	$R_{sh} = 60\,m\Omega/\text{sq}$	$t = 0.51\mu\text{m}$	$C_{sub} = 10.4aF/\mu\text{m}^2$
Surface Layer	$\rho = 15 \times 10^{-4}\,\Omega\text{–cm}$	$t = 0.8\mu\text{m}$	p^+ Si
Bulk Substrate	$\rho = 10\,\Omega\text{–cm}$	$t = 725\mu\text{m}$	p^- Si

Table 8.2. Device Physical Dimensions

Spiral Name	L19	L27
Outer Length (μm)	200	250
Metal Width W (μm)	11	10.5
Metal Spacing S (μm)	3	3
No. of Turns N	2.5	7.75
Metal Layer(s)	M3-M5	M5

open-pad y-parameters are also measured and subtracted from the measured y-parameters to remove the pad capacitance and loss.

It is interesting to note that the G-S-G approach has some clear advantages over the S-G approach presented previously in [Niknejad and Meyer, 1998]. To see this, consider the ground pad configurations of each approach, as show in in Fig. 8.4. As noted in [Niknejad, 1997], the S-G approach is prone to calibration errors resulting from the substrate coupling between the pads, since the presence of the device under test alters the substrate coupling. On the other hand, in the G-S-G approach, the substrate pads are widely separated and the presence of the ground de-couples the input from the output port. Thus, the G-S-G calibration step only subtracts out the input and output capacitance and has a minor influence on the coupling capacitance between the input and output ports, in stark contrast to the S-G approach where a large coupling capacitance is involved.

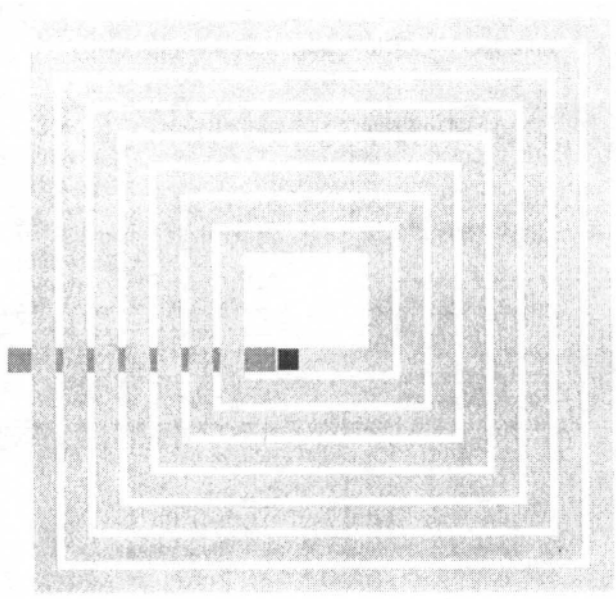


Figure 8.2. Layout of inductor L27.

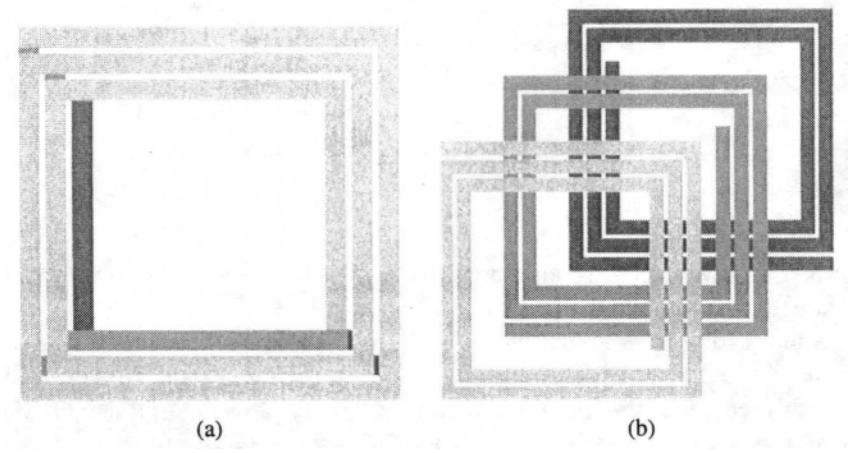


Figure 8.3. Layout of inductor L19. In part (b) the lower metal layers have been staggered to give a clear picture of the device geometry.

For these reasons, the G-S-G calibration procedure was selected. Designing a proper G-S-G calibration test structure in CMOS requires special care. With a device under test connected to the calibration structure, return ground currents flow through the calibration structure. To minimize the losses resulting

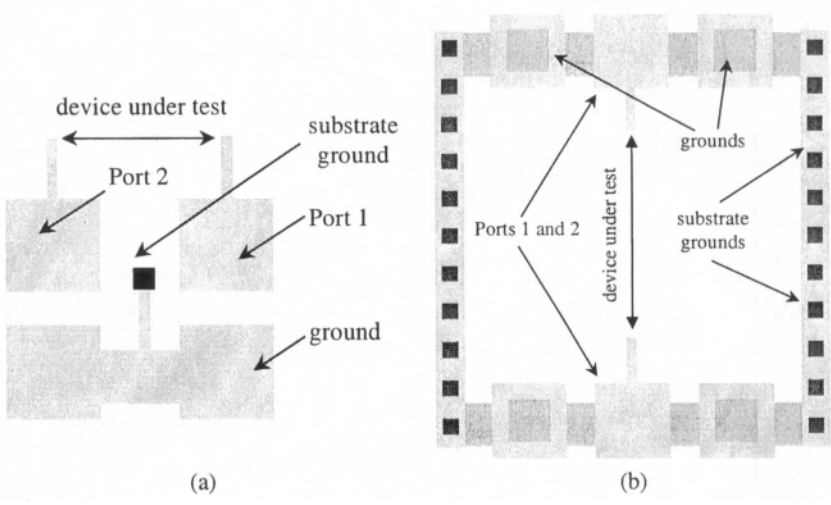


Figure 8.4. Calibration pad structures. (a) S-G structure. (b) G-S-G structure.

from such ground currents, metal layers M1-M5 have been strapped with an ample supply of vias and substrate ground connections to minimize the losses. Furthermore, in the G-S-G approach the signal ports can be isolated from the substrate by using an M1 shield under the pads, minimizing the loss contribution from the pads. In the S-G case, the pads introduce substantial loss which must be calibrated out of the device loss.

3. SINGLE LAYER INDUCTOR

Measured *s*-parameters for inductor L27 are shown in Fig. 8.5. The simulated and measured results match well. The discrepancy above the self-resonant frequency is in the capacitive region where we are less interested in the device. Notice that the inductor self-resonates at a frequency of 4.25 GHz. Simulations using *ASITIC* predicted a self-resonant frequency of 4.15 GHz. The simulations are performed on a Pentium II 400 MHz machine running the Linux operating system. Each frequency point requires less than ten seconds of computation. In Fig. 8.6 we plot the effective value of inductance. This is derived using a one-to-one transformation of the *s*-parameters into π-parameters [3]. Again, a good match is observed between the theory and measurements. The inductance decrease is due mostly to the capacitive effects rather than the inductive effects. Inductance value decreases slightly due to skin effect and eddy currents in the substrate but the main reason for the decrease is that energy is coupled from port to port through the winding capacitance at higher frequencies.

Fig. 8.7 shows the effective value of series resistance as a function of frequency. Two simulations are performed, with and without eddy current losses.

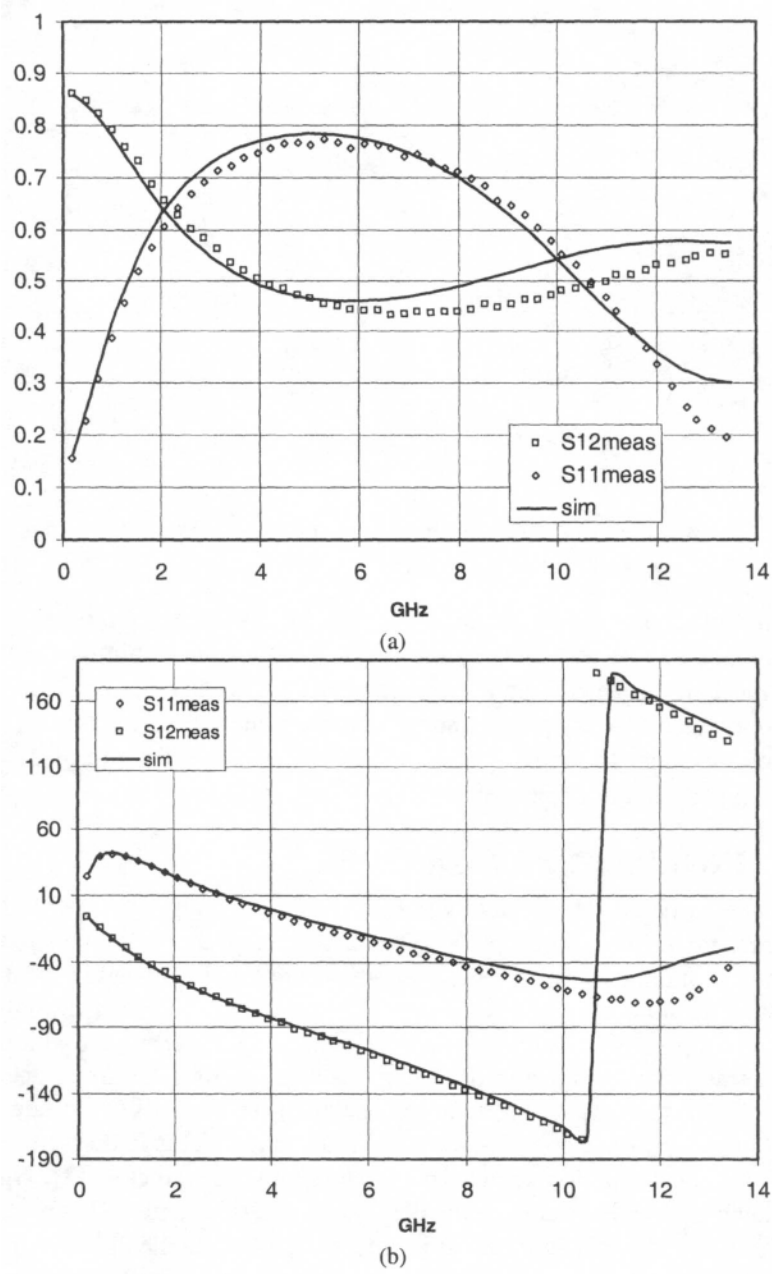


Figure 8.5. Measured and simulated (a) magnitude and (b) phase *s*-parameters of spiral inductor L27.

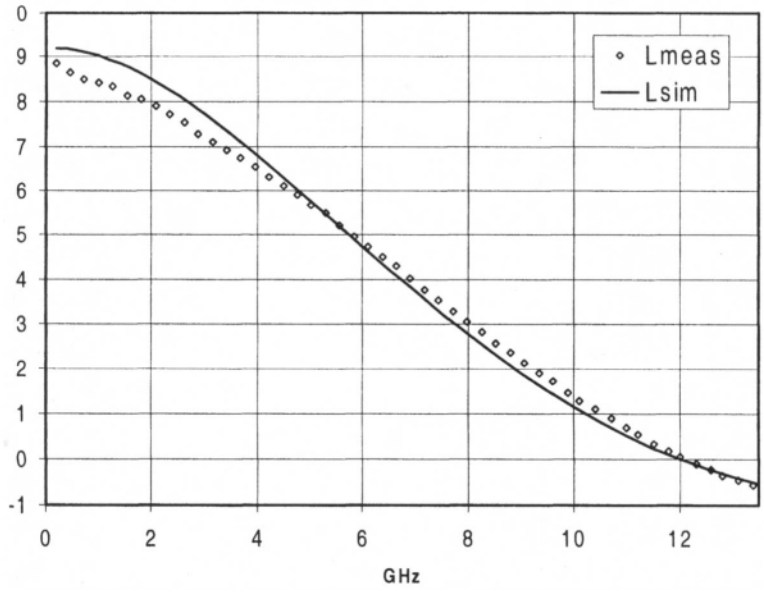


Figure 8.6. Measured and simulated inductance (imaginary component of Y_{21}) of spiral inductor L27.

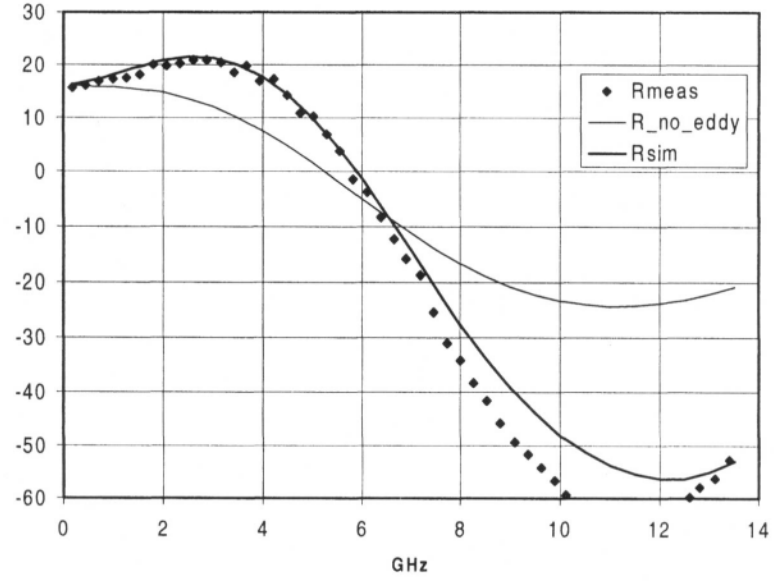


Figure 8.7. Measured and simulated resistance (real component of Y_{21}) of spiral inductor L27.

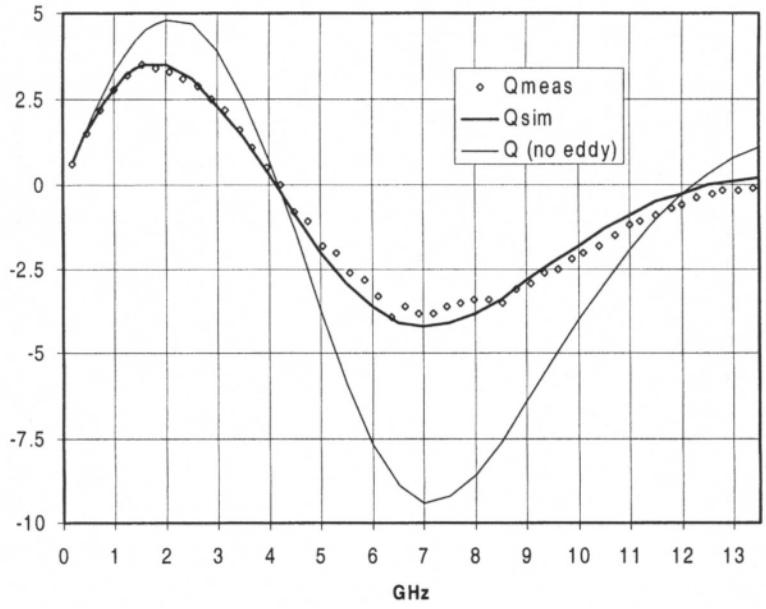


Figure 8.8. Measured and simulated quality factor (imaginary over real component of Y_{21}) of spiral inductor L27.

As evident in the figure, eddy current losses are critical to model. The variation in frequency of the series resistance is due to various competing effects. Skin effect and proximity effects increase the series resistance but beyond 1 GHz this is swamped by the increase from eddy currents. At higher frequencies more energy is transported capacitively, and consequently the resistance decreases and eventually becomes negative. The real part of the total input impedance looking into each port, of course, is positive at all frequencies.

The quality factor Q, the ratio between the imaginary and real part of the input impedance, is plotted in Fig. 8.8. Again, a good match is observed between the theory and measurements. Note that negative quality factor implies that the device is acting as a capacitor rather than an inductor. In reality, this plot is misleading as it implies a Q of zero at self-resonance. A better way to calculate Q is given in [3] but for comparison the given definition is better since it involves a minimal transformation of the measured s-parameters. The substrate resistance and capacitance are also shown in Fig. 8.9 and Fig. 8.10. The overall shape of both curves matches the measurements well. The low frequency substrate resistance measurements are noisy due to measurement error.

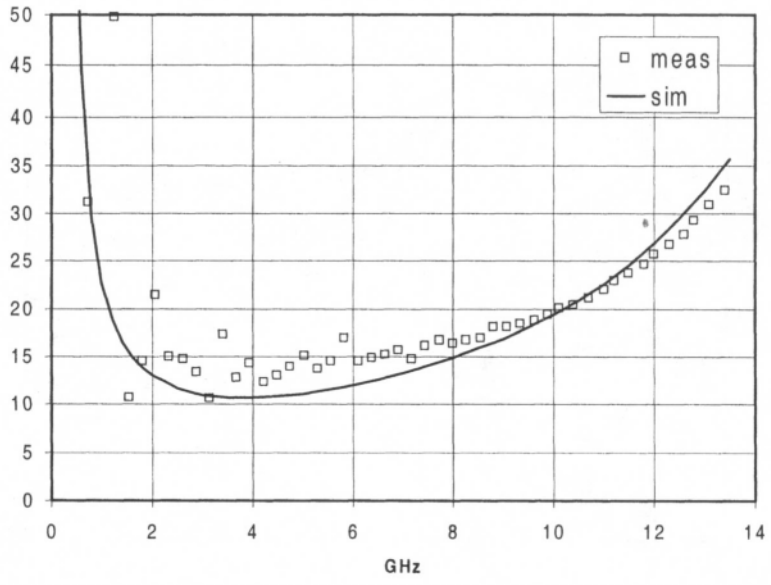


Figure 8.9. Measured and simulated substrate resistance of spiral inductor L27.

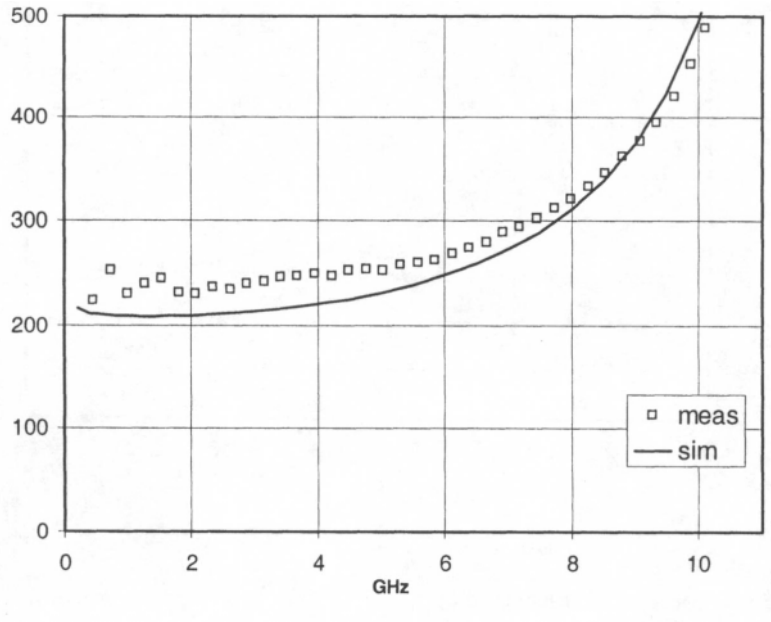


Figure 8.10. Measured and simulated substrate capacitance of spiral inductor L27.

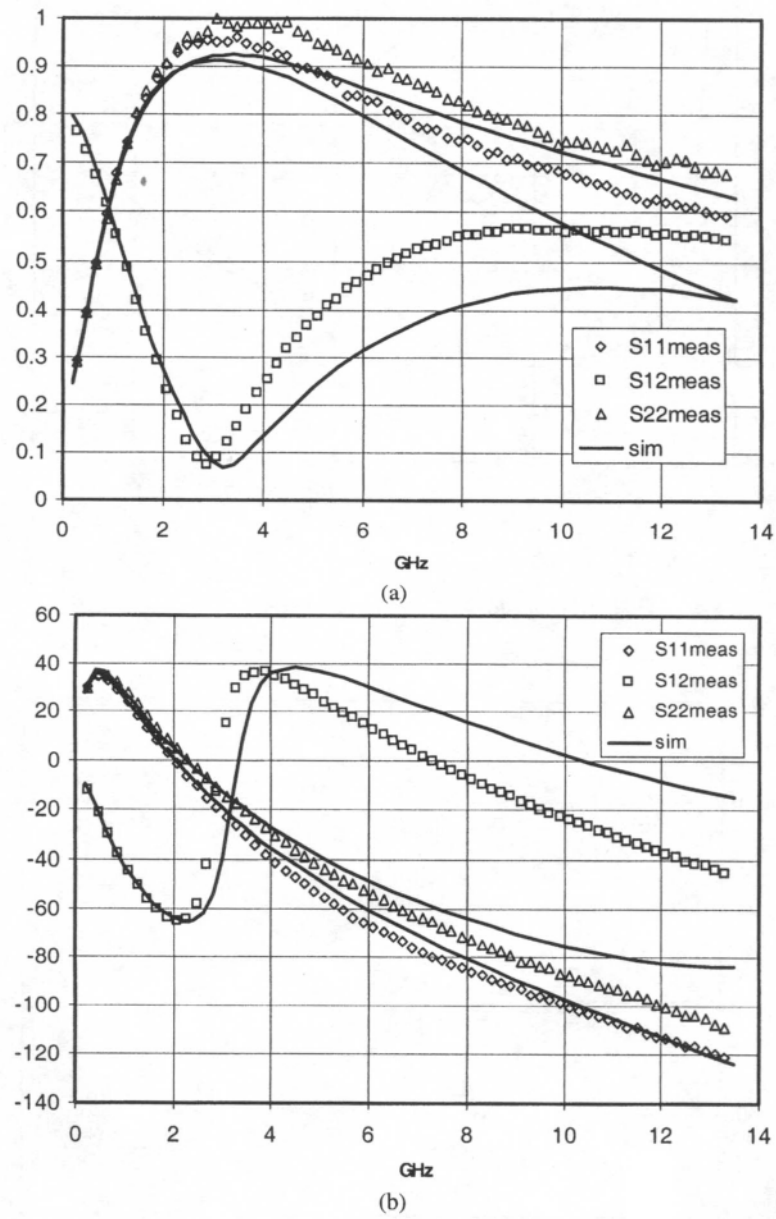


Figure 8.11. Measured and simulated (a) magnitude and (b) phase *s*-parameters of spiral inductor L19.

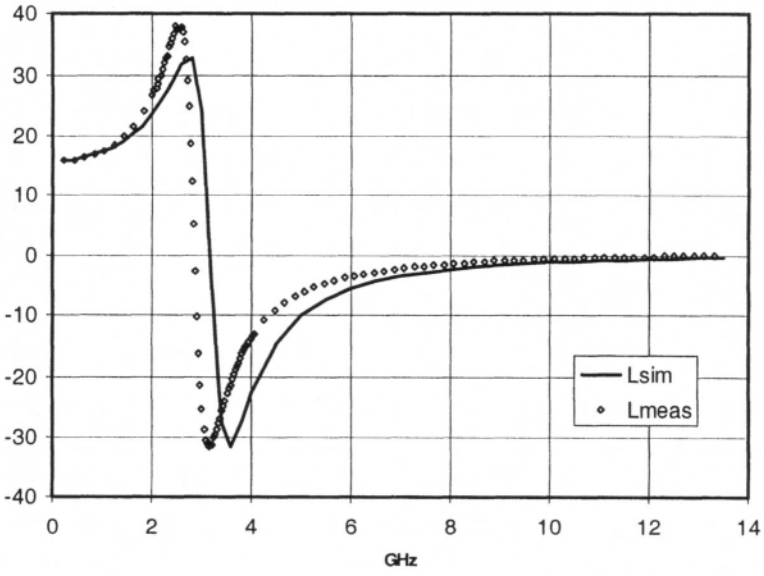


Figure 8.12. Measured and simulated inductance (imaginary component of Y_{21}) of spiral inductor L19.

4. MULTI-LAYER INDUCTOR

Measured *s*-parameters for inductor L19 are shown in Fig. 8.11. A fairly good match between the simulated and measured is observed, especially below the self-resonant frequency. Due to the large interwinding capacitance intrinsic to this series-connected structure, the self-resonant frequency is low. The self-resonant frequency is predicted at 2.57 GHz and measured at 2.47 GHz.

To gain insight into the data we also plot the series inductance and resistance in Figures 8.12 and 8.13. These curves are dramatically different from the previous device. The differences can be accounted for by noting that a planar device self-resonates through the substrate capacitance whereas a multi-metal device self-resonates through the interwinding capacitance. Since this is a relatively high Q capacitor compared to the lossy substrate capacitance, the behavior of L and R can be modeled by the simple equivalent circuit of Fig. 8.14. As we approach self-resonance the inductance and resistance values peak. The maximum value of resistance is approximately $Q^2 R_{series}$. Thus, the lower the losses the higher the value of the peak. Again, simulation is able to predict these effects well. The discrepancy is due mainly to the value of oxide thickness used in the simulation. The process parameters were not adjusted at all in the simulations in order to gauge how well one can predict inductor performance *a priori.* Since oxide thickness is not a well-controlled process parameter, and since the oxide thickness has a significant impact on the device performance, it

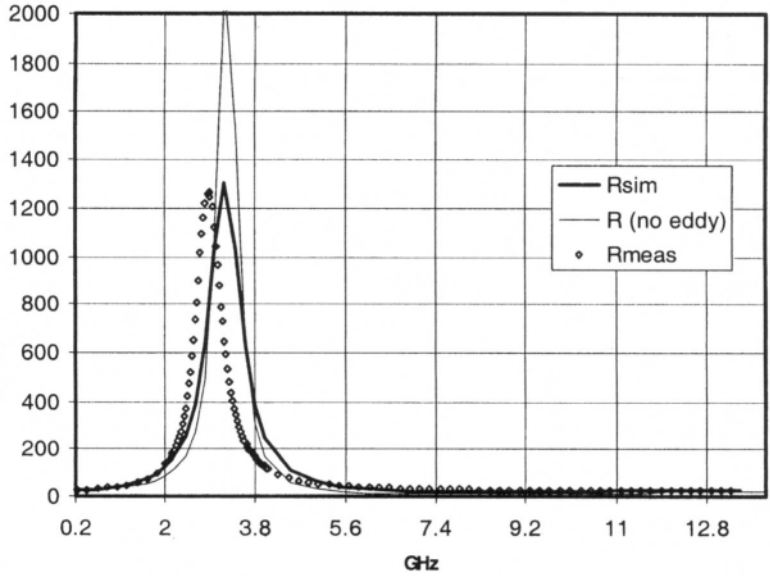


Figure 8.13. Measured and simulated resistance (real component of Y_{21}) of spiral inductor L19.

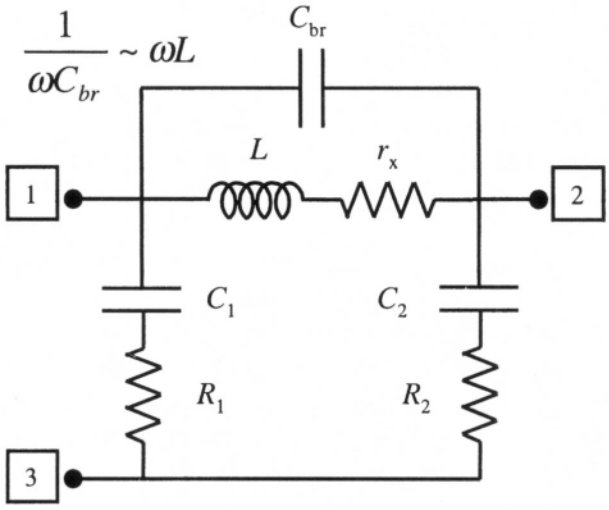


Figure *8.14.* Simple equivalent circuit of inductor L19 close to self-resonance.

seems like this device is not as manufacturable as the planar structure, especially close to self-resonance.

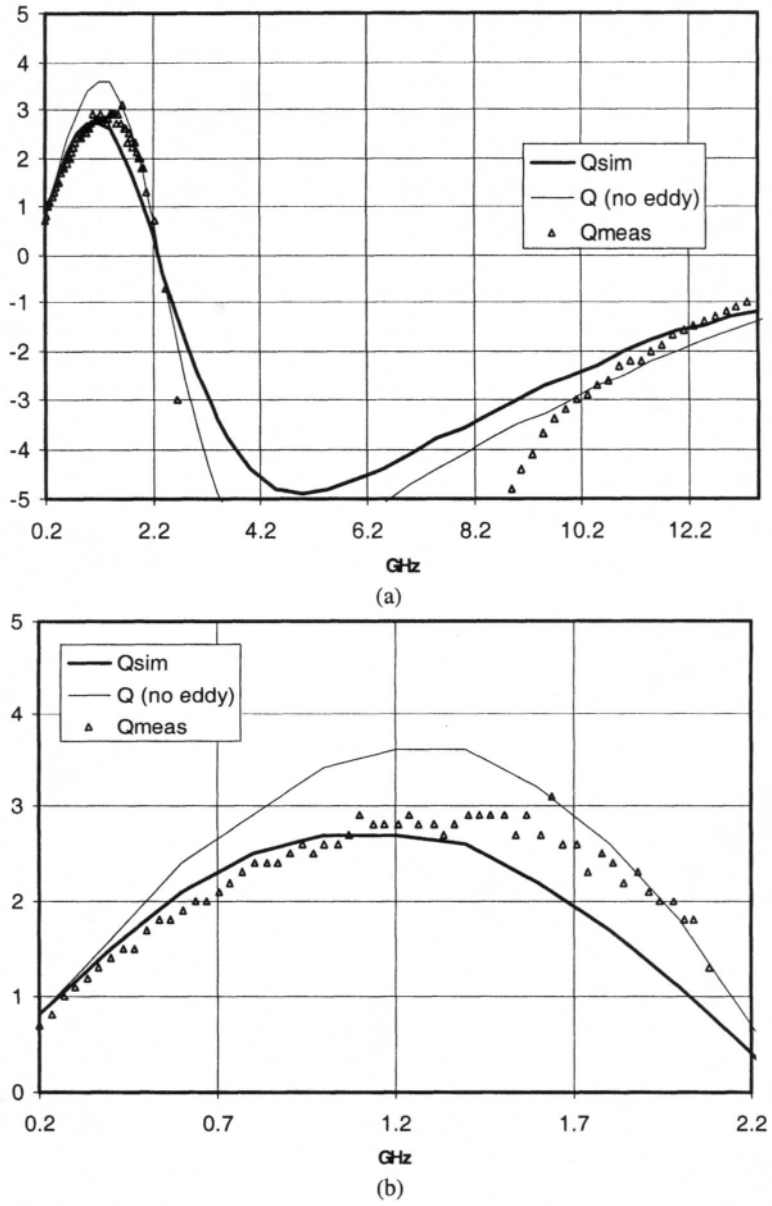


Figure 8.15. (a) Measured and simulated quality factor (imaginary over real component of Y_{21}) of spiral inductor L19. (b) Close-up plot.

Fig. 8.15 shows the quality factor plot for two simulations and the measurements. As before, the simulation results with no eddy currents under-predict the losses, but now the simulations with eddy currents slightly over-predict the losses. This is especially true beyond self-resonance where the device is acting as a capacitor. Fortunately, one is generally not interested as much in frequencies beyond self-resonance. The discrepancy, though, can be easily explained with some physical insight into the problem. Beyond self-resonance the device is acting very much like a metal-insulator-metal capacitor. Since wide metal lines are used to wind the inductor to minimize the low frequency losses, the current distribution in the lines is not constrained to be one-dimensional as we have assumed. In fact, at high frequencies the bottom metal layers act as a solid shield against both the magnetic and electric fields. Thus, eddy currents flow in the shield and prevent the magnetic fields from penetrating the lossy substrate. We could account for this effect by utilizing a two-dimensional mesh but we are less interested in this region of operation.

II
APPLICATIONS OF PASSIVE DEVICES