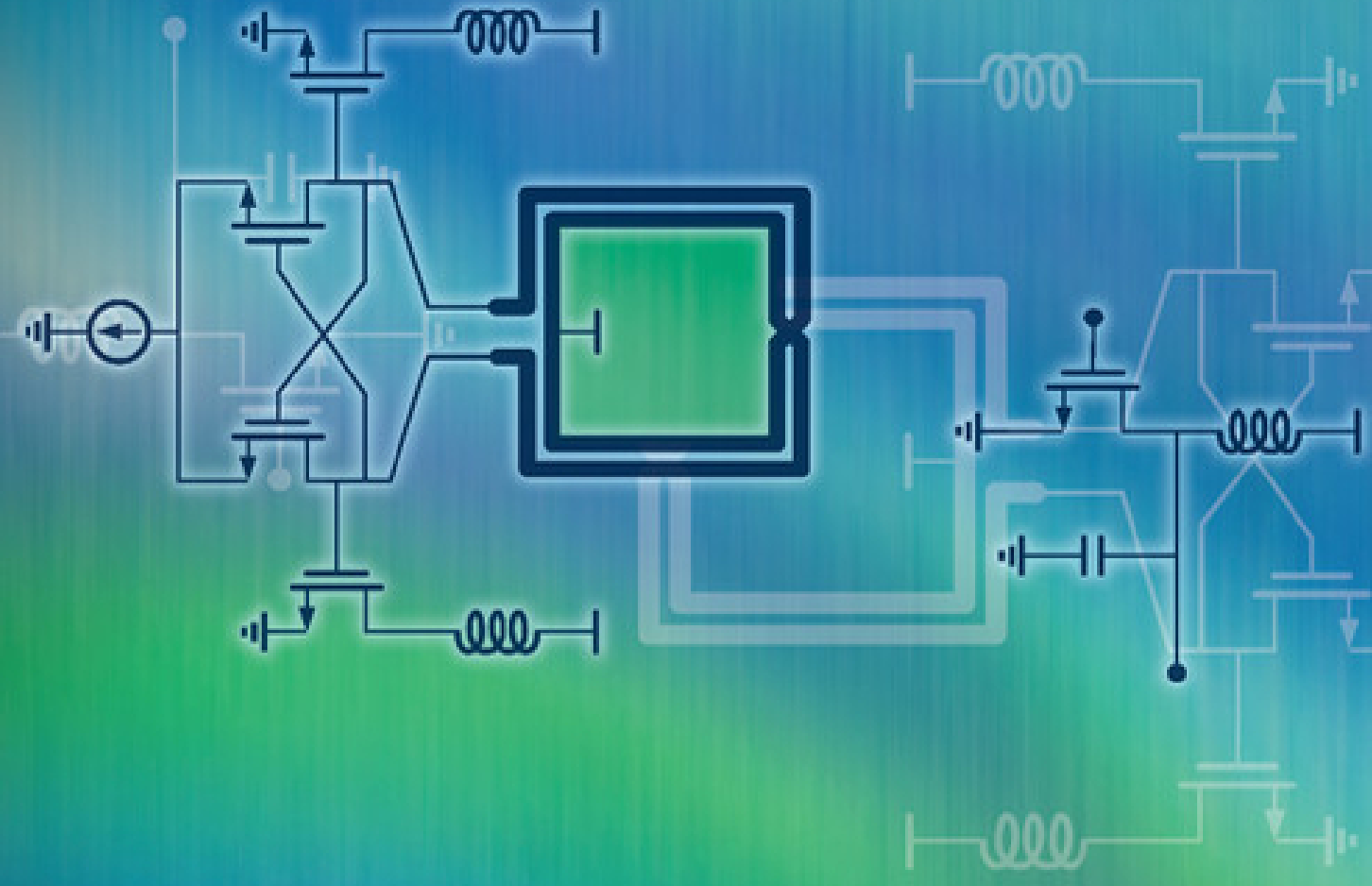


RADIO FREQUENCY INTEGRATED CIRCUITS AND SYSTEMS

Hooman Darabi

Radio Frequency Integrated Circuits and Systems

Focusing on the core topics of RF IC and system design, this textbook provides the in-depth coverage and detailed mathematical analyses needed to gain a thorough understanding of the subject. Throughout, theory is linked to practice with real-world application examples; practical design guidance is also offered, covering the pros and cons of various topologies, and preparing students for future work in industry. Written for graduate courses on RFICs, this uniquely intuitive and practical book will also be of value to practicing RF IC and system designers.

Key topics covered include:

- RF components, signals and systems
- Two-ports
- Noise
- Distortion
- Low-noise amplifiers
- Mixers
- Oscillators
- Power amplifiers
- Transceiver architectures

Lecture slides and a solutions manual for instructors are provided online to complete the course package.

Hooman Darabi is a Senior Technical Director and Fellow of Broadcom Corporation, California, and an Adjunct Professor at the University of California, Los Angeles. He is an IEEE Solid State Circuits Society distinguished lecturer, and an IEEE Fellow. His research interests include analog and RF IC design for wireless communications.

"Excellent textbook for students keenly interested to learn how to design IC transceivers. Lots of focused and relevant fundamental background from the viewpoint of the well-known industrial researcher. The book is also good as a convenient refresher for seasoned IC designers."

Bogdan Staszewski, *Delft University of Technology*

"Although this is not the first book on RF circuits, it is the most up-to-date one that I know. For instance, it includes recent circuit insights to make CMOS radio receivers more interference robust. I especially like the intuition that Hooman Darabi develops and the depth of coverage without becoming overly mathematical."

Eric Klumperink, *University of Twente*

Radio Frequency Integrated Circuits and Systems

HOOMAN DARABI

Broadcom Corporation, California

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

Cambridge University Press is part of the University of Cambridge.

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

www.cambridge.org
Information on this title: www.cambridge.org/9780521190794

First published 2015

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

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

ISBN 978-0-521-19079-4 Hardback

Additional resources for this publication at www.cambridge.org/darabi

To my family

CONTENTS

PREFACE

In the past twenty years, radio frequency (RF) integrated circuits in CMOS have evolved dramatically, and matured. Started as a pure research topic in mid-1990 at several universities, they have made their way into complex systems-on-a-chip (SoCs) for wireless connectivity and mobile applications. The reason for this dramatic evolution comes primarily from two main factors: the rapid improvement of CMOS technology, and innovative circuits and architectures. In contrary to the common belief that RF and analog circuits do not improve much with technology, a faster CMOS process has enabled a number of topologies that have led to substantially lower cost and power consumption. In fact, many of the recent inventions may not have been possible if it were not for a better and faster technology. This rapid change has caused the modern RF design to be somewhat industry-based and, consequently, it is timely, and perhaps necessary, to provide an industry perspective. To that extent, the main goal of this book has been to cover possibly fewer topics, but in a much deeper fashion. Even for RF engineers working on routine products in industry, a deep understanding of fundamental concepts of RF IC design is very critical, and it is the intention of this work to break this gap. During the course of writing the book, I have tried to address the topics that I would have wanted as a *wish list* for my fellow future colleagues. Our main focus then has been to elaborate the basic definitions and fundamental concepts, while an interested designer with strong background can explore other variations on his or her own.

The contents of this book stem largely from the RF courses taught at the University of California Los Angeles and Irvine, as well as many years of product experience at Broadcom. Accordingly, the book is intended to be useful both as a text for students and as a reference book for practicing RF circuit and system engineers. Each chapter includes several worked examples that illustrate the practical applications of the material discussed, with examples of real life products, and a problem set at the end (with solutions manual) to complement that.

RF circuit design is a *multi-disciplinary* field where a deep knowledge of analog integrated circuits, as well as communication theory, signal processing, electromagnetics, and microwave engineering is crucial. Consequently, the first three chapters as well as parts of Chapter 4 cover selected topics from the aforementioned fields, but customized and shaped properly to fit the principles of RF design. It is, however, necessary for the interested students or RF engineers to have already taken appropriate senior level undergraduate courses.

An outline of each chapter is given below along with suggestions for the material to be covered if the book is to be used for a 20-lecture quarter based course. Furthermore, at the beginning of each chapter a list of specific items to be covered, as well as more detailed suggestion of which sections to include for the class use, are outlined. For beginner and intermediate practicing

engineers we recommend following the selected topics suggested for class use, while more advanced readers may focus on the other topics assigned for reading.

Chapter 1 contains a review of basic electromagnetic concepts and particularly inductors and capacitors. Among many students and RF engineers, often the basic definition of capacitors and inductors is neglected, despite using them very regularly. A short reminder is deemed necessary. Furthermore, some basic understanding of Maxwell's equations is needed to understand transmission lines, electromagnetic waves, the antenna concept, and scattering parameters. These are discussed in Chapter 3. The chapter also gives an overview of integrated inductors and capacitors in modern CMOS. A total of two lectures is expected to be needed to cover the key ideas.

Chapter 2 deals with basic communication and signal processing concepts, which are a crucial part of RF design. The majority of the material is gathered to provide a reminder, and may be left to be studied outside the class, depending on the students' background. However, we cannot emphasize enough the importance of them. Spending a total of two or three lectures on the stochastic processes, modulation section, as well as a brief general reminder of passive filters and the Hilbert transform may be helpful.

Chapter 3 is concerned with several key concepts in RF design such as available power, matching topologies, transmission lines, as well as scattering parameters, and complements Chapter 1. Two lectures may be dedicated to cover the first three sections, while the more advanced material on transmission lines, the Smith chart, and scattering parameters may be very briefly touched, or omitted altogether depending on the students' background.

In Chapter 4 we discuss noise, noise figure, sensitivity, and an introduction to phase noise. The introductory part on types of noise may be assigned as reading, but the noise figure definition, minimum noise figure, and sensitivity sections must be covered in full. A total of two to three lectures suffices.

Chapter 5 covers distortion and blockers. A large portion of this chapter (as well as Chapter 10) may be left for a more advanced course, and one lecture should suffice to cover only the basic concepts. However, the material may be very appealing to RF circuit and system engineers who work in industry. A thorough knowledge of this chapter is crucial to understand Chapter 10.

Chapters 6 to 9 deal with RF circuit design. Chapter 6 is mostly built upon the concepts covered in Chapters 3 and 4, and deals with low-noise amplifiers. Three lectures may be dedicated to cover most of the topics presented in this chapter.

Chapter 7 provides a detailed discussion on receiver and transmitter mixers. Roughly two lectures may be dedicated to this chapter to cover basic active and passive topologies with some limited discussion on noise. The majority of the material on *M*-phase and upconversion mixers may be assigned as reading.

Chapter 8 discusses oscillators, including LC, ring, and crystal oscillators, and an introduction to phase-locked loops. The chapter is long, and the latter three topics may be assigned as reading, while two lectures could be dedicated to LC oscillators, and a brief introduction to phase noise. A detailed discussion of phase noise is very math intensive, and may be beyond the scope of an introductory RF course. Thus, it may be sufficient to focus mostly on the premises of an abstract linear oscillator, and summarize Bank's general results to provide a more practical perspective.

Power amplifiers are discussed in Chapter 9. Basic PA classes are presented in the first few sections, followed by efficiency improvement and linearization techniques. Most of the material on the latter subject may be skipped, and one or two lectures may be assigned to cover a few examples of classes (perhaps only classes A, B, and F), as well as the introductory material on general concerns and tradeoffs.

Finally, in Chapter 10 transceiver architectures are presented. This is one of the longest chapters of the book, and much of the material can be assigned as reading. The last section covers some practical aspects of the design, such as packaging and production issues. It presents a few case studies as well. The topics may be appealing for practicing RF engineers, but the entire section may be skipped for class use. A maximum of two lectures is sufficient to cover selected key transceiver architectures.

I have been very fortunate to have been working with many talented RF designers and instructors throughout my career at UCLA, and subsequently at Broadcom. They have had an impact on this book in one way or another. However, I wish to acknowledge the following individuals who have directly contributed to the book: Dr. David Murphy from Broadcom who co-wrote most of Chapter 8, and provided very helpful insight on Chapter 6, particularly the LNA topologies; Dr. Ahmad Mirzaei from Broadcom as well, who helped on the writing of some sections of Chapters 9 and 10, and proofread the entire book painstakingly. They both have been major contributors to this book beyond the chapters mentioned. I am very grateful to my advisor Prof. Asad Abidi from UCLA who has been a great inspiration on writing this book in general, and particularly for his insights and unique analysis which has been used in sections 1.7, 1.9.3, and 4.2/4.4 (FET equivalent noise and NF). I would also like to thank Dr. Hwan Yoon from Broadcom with whom I had numerous helpful discussions on the Chapter 1 material, and particularly the integrated inductors. My sincere thanks go to Professor Eric Klumperink of University of Twente, who proofread most of the book diligently, and provided valuable insight on various topics. I would also like to acknowledge my sister Hannah, who helped in the design of the book cover. Lastly, I wish to thank my wife, Shahrzad Tadjpour, who not only provided technical feedback on the book, but for her general support throughout all these years.

1 RF components

In this chapter basic components used in RF design are discussed. A detailed modeling and analysis of MOS transistors at high frequency may be found in [1], [2]. Although mainly developed for analog and high-speed circuits, the model is good enough for most RF applications operating at several GHz, especially for the nanometer CMOS processes used today. Thus, we will offer a more detailed study of inductors, capacitors, and LC resonators instead in this chapter. We will also present a brief discussion on the fundamental operation of distributed circuits and transmission lines, and follow up on that in Chapter 3. In Chapters 4 and 6, we will discuss some of the RF related aspects of transistors, such as more detailed noise analysis as well as substrate and gate resistance impact.

LC circuits are widely used in RF design, with applications ranging from tuned amplifiers, matching circuits, and LC oscillators. Inspired by superior noise and linearity compared to transistors, historically, radios have relied heavily on inductors and capacitors, with large portions of the RF blocks occupied by them. Although this dependence has been reduced in modern radios, mostly for cost concerns, still RF designers deal with integrated inductors and capacitors quite often.

We start the chapter with a brief introduction to electromagnetic fields, and take a closer look at capacitors and inductors from the electromagnetic field perspective. We will then discuss capacitors, inductors, and LC resonators from a circuit point of view. We conclude the chapter by presenting the principles and design tradeoffs of integrated inductors and capacitors.

The specific topics covered in this chapter are:

- capacitance and inductance electromagnetic and circuit definitions;
- Maxwell's equations;
- distributed elements and introduction to transmission lines;
- energy, power, and quality factor;
- lossless and low-loss resonance circuits;
- integrated capacitors and inductors.

For class teaching, we recommend focusing on Sections 1.7, 1.8, and 1.9, while Sections 1.1–1.6 may be assigned as reading, or only a brief summary presented if deemed necessary.

1.1 ELECTRIC FIELDS AND CAPACITANCE

Let us start with a brief overview of electric fields and electric potential. We shall define the concept of capacitance accordingly.

Published first in 1875 by Charles Coulomb, the French army officer, Coulomb's law states that the force between two point charges, separated in vacuum or free space by a distance, is proportional to each charge, and inversely proportional to the square of the distance between them (Figure 1.1). It bears a great similarity to Newton's gravitational law, discovered about a hundred years earlier. Writing the force ($\boldsymbol{F}_\mathrm{t}$) as a force per unit charge gives the electric field intensity, $\boldsymbol{E}$ measured in V/m (or volt per meter) as follows:

$$\boldsymbol{E} = \frac{\boldsymbol{F}_\mathrm{t}}{Q_\mathrm{t}} = \frac{Q}{4\pi\epsilon_0 r^2}\boldsymbol{a}_r,$$

where the ***bold*** notation indicates a vector in 3D space, $\epsilon_0 = \frac{1}{36\pi} \times 10^{-9}$ F/m (or farads per meter) is the permittivity in free space, and Q is the charge in C (or coulomb).[1] $\boldsymbol{a}_r$ is the unit vector pointing in the direction of the field, which is in the same direction as the vector $\boldsymbol{r}$ connecting the charge Q to the point of interest P in space (see Figure 1.1). Q_t is a test charge to which the force (or field) created by Q is applied.

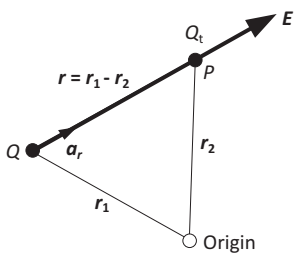


Figure 1.1 Coulomb's law

In many cases the electric field can be calculated more easily by applying Gauss's law instead. It states that the electric flux density,[2] $\boldsymbol{D} = \epsilon_0\boldsymbol{E}$ (measured in C/m^2) passing through any closed surface is equal to the total charge enclosed by that surface,[3] and mathematically expressed as:

$$\oint_S \boldsymbol{D} \cdot \mathrm{d}\boldsymbol{S} = Q,$$

where $\oint_S \boldsymbol{D} \cdot \mathrm{d}\boldsymbol{S}$ indicates the integral over a closed surface. The *dot* product $(\cdot)$ indicates the product of the magnitude and the cosine of the smaller angle. The charge Q could be the sum of several charge points, that is: $Q = \sum Q_i$, a volume charge distribution: $Q = \int_V \rho_V \mathrm{d}V$, or a surface distribution, . . . The nature of the surface integral implies that only the normal component of $\boldsymbol{D}$ at the surface contributes to the charge, whereas the tangential component leads to $\boldsymbol{D} \cdot \mathrm{d}\boldsymbol{S}$ equal to zero.

[1] Not to be confused with Q used as quality factor later in this chapter.
[2] Only in free space.
[3] The expression itself is a result of Michael Faraday's experiment. Gauss's contribution was providing the mathematical tools to formulate it.

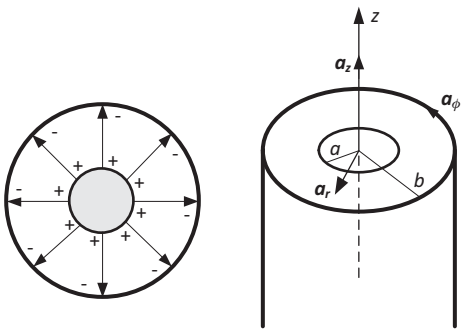


Figure 1.2 Electric flux in a coaxial cable

For example, consider a long coaxial cable with inner radius a and outer radius b, carrying a uniform charge distribution ρ_S on the outer surface of the inner conductor (and $-\frac{b}{a}\rho_S$ on the inner surface of the outer conductor) as shown in Figure 1.2. For convenience, let us use cylindrical coordinates [3].

The flux will have components in the $\boldsymbol{a}_r$ direction, normal to the surface. For an arbitrary length L in the z-axis direction, we can write

$$\int_{z=0}^{L}\int_{\phi=0}^{2\pi} D_r\,(r\mathrm{d}\phi\mathrm{d}z) = Q = \rho_S(2\pi aL).$$

Thus, inside the cable, that is for $a < r < b$:

$$\boldsymbol{D} = \frac{\rho_S a}{r}\boldsymbol{a}_r.$$

The electric field and flux density are both zero outside the cable as the net charge is equal to zero.

Based on the electric energy definition,[4] the potential difference between points A and B (V_{AB}) is defined as

$$V_{\mathrm{AB}} = \frac{W}{Q} = -\int_{\mathrm{B}}^{\mathrm{A}} \boldsymbol{E}\cdot\mathrm{d}\boldsymbol{L},$$

where W is the energy in J (or joule), and the right side is the *line integral* of the electric field. The physical interpretation of potential is such that moving a charge Q along with the electric field from point A to B results in energy reduction (or the charge releases energy), and accordingly we expect point A to be at a *higher potential*. By definition of the line integral, the sum of *static* potentials in a closed path must be equal to zero, that is to say $\oint \boldsymbol{E}\cdot\mathrm{d}\boldsymbol{L} = 0$, which is a general representation of *Kirchhoff's voltage law* or KVL. This is physically understood by noting that when the charge is moved around a closed path, the total energy received and released balance each other, thus no *net* work is done.

[4] We shall discuss the electric energy shortly.

We close this section by defining the capacitance. Suppose we have two oppositely charged (each with a charge of Q) conductors M1 and M2 within a given dielectric with permittivity[5] of $\epsilon = \epsilon_r \epsilon_0$ (Figure 1.3). Assuming a potential difference of V_0 between the conductors, we define capacitance C measured in Farad as

$$C = \frac{Q}{V_0}.$$

Alternatively, one can re-write C as

$$C = \epsilon \frac{\oint_S \boldsymbol{E} \cdot \mathrm{d}\boldsymbol{S}}{-\int \boldsymbol{E} \cdot \mathrm{d}\boldsymbol{L}},$$

which indicates that capacitance (if linear), is independent of the charge or potential, as $\boldsymbol{E}$ (or $\boldsymbol{D}$) depends linearly on Q according to Gauss's law.

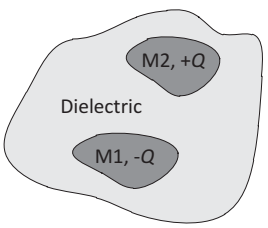


Figure 1.3 Capacitance definition

Physically, the capacitance indicates the capability of energy or equivalently electric flux storage in electrical systems, analogous to inductors that store magnetic flux.

Returning to our previous example of the coaxial cable, the potential between the inner and outer conductors is calculated by taking the line integral of $\boldsymbol{E} = \boldsymbol{D}/\epsilon$, where $\boldsymbol{D}$ was obtained previously. This yields

$$V_0 = -\frac{1}{\epsilon} \int_b^a \frac{a\rho_S}{r} \mathrm{d}r = \frac{a\rho_S \ln\frac{b}{a}}{\epsilon},$$

and thus the capacitance per unit length is equal to

$$C = \frac{2\pi\epsilon}{\ln\frac{b}{a}}.$$

Clearly, the capacitance is only a function of the coaxial cable radii and the dielectric.

1.2 MAGNETIC FIELDS AND INDUCTANCE

A steady magnetic field can be created in one of three ways: through a permanent magnet, a linear time-varying electric field, or simply due to a direct current. The permanent magnet has several applications in RF and microwave devices, such as passive gyrators used in a lossless circulator, which is a *passive*, but *non-reciprocal* circuit, [4], [5]. However, we will mostly focus

[5] $\epsilon_r = 1$ for free space.

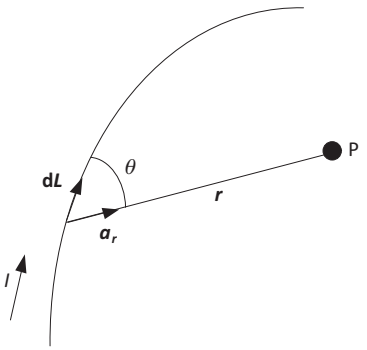


Figure 1.4 Biot–Savart law expression

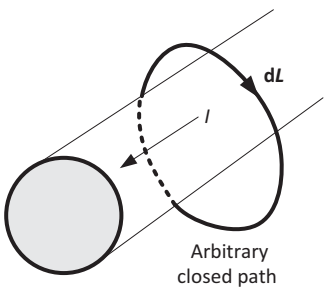


Figure 1.5 Ampère's law

on the latter two methods of creating the magnetic fields, and defer the gyrator and circulator discussion to [4].

In 1820, the law of Biot-Savart was proposed as follows, which associates magnetic field intensity $\boldsymbol{H}$ (expressed in A/m) at a given point P to a current of I flowing in a differential vector length $\mathrm{d}\boldsymbol{L}$ of an ideal filament (Figure 1.4):

$$\mathrm{d}\boldsymbol{H} = \frac{I\,\mathrm{d}\boldsymbol{L} \times \boldsymbol{a}_r}{4\pi r^2}.$$

The *cross* product ($\times$) indicates the product of the magnitude and the sine of the smaller angle. The magnetic field will then be perpendicular to the plane containing the current filament and the vector $\boldsymbol{r}$, and whose direction is determined based on the right hand rule. The law states that the magnetic field intensity is directly proportional to the current (I), but inversely proportional to the square of the distance (r) between P and the differential length, and also proportional to the magnitude of the differential element times the sine of the angle θ shown in Figure 1.4.

A more familiar law describing the magnetic field was proposed by Ampère shortly after in 1823, widely known as Ampère's circuital law,[6] and is mathematically expressed as:

$$\oint \boldsymbol{H} \cdot \mathrm{d}\boldsymbol{L} = I,$$

indicating that the line integral of the magnetic field ($\boldsymbol{H}$) about any *closed path* is exactly equal to the current enclosed by that path (Figure 1.5). This law proves to be more useful as it allows us

[6] Ampère's law may be derived from Biot-Savart's law.

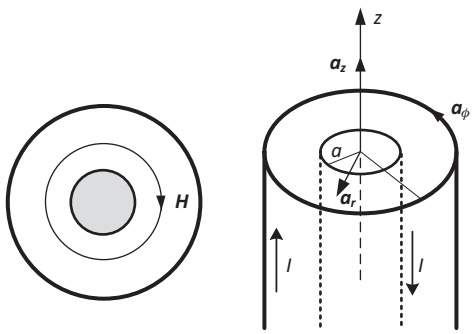


Figure 1.6 Magnetic field in a coaxial cable, top and side views

to calculate the magnetic field more easily as long as which components of the field are present is properly determined, and the symmetry is invoked appropriately. By comparison, Ampère's circuital law is more analogous to Gauss's law, whereas the law of Biot-Savart could be considered similar to Coulomb's law.

As an example, consider a long coaxial cable carrying a uniform current of I in the center conductor and $-I$ in the outer one, as shown in Figure 1.6. Clearly the field cannot have any component in the z direction, as it must be normal to the current direction. Moreover, the symmetry shows that $\boldsymbol{H}$ cannot be a function of ϕ or z, and thus could be expressed as a general form of $\boldsymbol{H} = H_r \boldsymbol{a}_\phi$. Inside the coaxial cable, that is $a < r < b$, applying the line integral then leads to

$$\boldsymbol{H} = \frac{I}{2\pi r}\boldsymbol{a}_\phi.$$

Moreover, similarly to the electric field, the magnetic field is zero outside the cable as the net current flow is zero, showing the concept of *shielding* provided by the coaxial cable. Note that, inside the cable, the magnetic field consists of closed lines circling around the current, as opposed to the electric field lines that start on a positive charge and end on a negative one.

In free space, magnetic flux density $\boldsymbol{B}$ (measured in Webber/m^2 or tesla), is defined as

$$\boldsymbol{B} = \mu_0 \boldsymbol{H},$$

where $\mu_0 = 4\pi \times 10^{-7}$ H/m (or Henry per meter) in free space and is the *permeability*. The magnetic flux, ϕ, is then the flux passing through a designated area S, measured in Webber, and is defined as

$$\phi = \int_S \boldsymbol{B} \cdot \mathrm{d}\boldsymbol{S}.$$

Generally the magnetic flux is a linear function of the current (I), that is, $\phi = LI$, where the proportionality constant, L, is known as the *inductance*, and is measured in Henry. We can thus say:

$$L = \mu_0 \frac{\int_S \boldsymbol{H} \cdot \mathrm{d}\boldsymbol{S}}{\oint \boldsymbol{H} \cdot \mathrm{d}\boldsymbol{L}},$$

and since $\boldsymbol{H}$ is a linear function of I, as established by Ampère's (or Biot-Savart's) law, the inductance is a function of the conductor geometry and the distribution of the current, but not

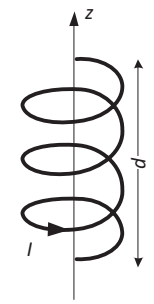


Figure 1.7 An N-turn solenoid

the current itself. As an example, calculating the total flux inside the coaxial cable of the previous example, one can show that the cable inductance per unit length is

$$L = \frac{\mu_0}{2\pi} \ln \frac{b}{a},$$

whereas the capacitance per unit length of the same coaxial cable was calculated before by applying Gauss's law, and is equal to

$$C = \frac{2\pi\epsilon}{\ln \frac{b}{a}}.$$

Clearly,

$$LC = \mu_0 \epsilon.$$

We conclude this section by defining the mutual inductance M_{12} between circuits 1 and 2 in terms of their *flux linkage*:

$$M_{12} = \frac{N_2 \phi_{12}}{I_1},$$

where ϕ_{12} signifies the flux produced by I_1 which links the path of the filamentary current I_2, and N_2 is the number of turns in circuit 2. The mutual inductance therefore depends on the magnetic interaction between the two currents.

As an example, consider an N-turn solenoid with finite length d, consisting of N closely wound filaments that carry a current I, as shown in Figure 1.7. We assume the solenoid is long relative to its diameter.

The magnetic field is in the $\boldsymbol{a}_z$ direction, as the current is in the $\boldsymbol{a}_\phi$ direction, and Ampère's law readily shows that within the solenoid,

$$\boldsymbol{H} = \frac{NI}{d} \boldsymbol{a}_z.$$

If the radius is r, corresponding to an area of $A = \pi r^2$, the self-inductance is

$$L = \frac{N\phi}{I} = \mu_0 N^2 \frac{A}{d}.$$

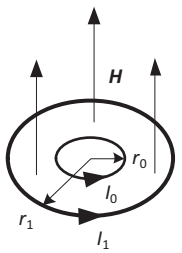


Figure 1.8 Two coaxial solenoids, top view

Now consider two coaxial solenoids, with radii r_1 and $r_0 < r_1$, carrying currents of I_1 and I_0, and with different numbers of turns N_1 and N_0, respectively. The top view is shown in Figure 1.8.

To find the mutual inductance M_{01}, we can write

$$\phi_{01} = \mu_0 A_0 H_0,$$

where $H_0 = \frac{N_0 I_0}{d}$ is the magnetic field intensity created by the smaller solenoid. Since H_0 is zero outside the radius of the smaller solenoid, we have

$$M_{01} = \frac{N_1}{I_0}\mu_0 A_0 H_0 = \mu_0 N_0 N_1 \frac{A_0}{d}.$$

A similar procedure leads to M_{10}, which comes out to be equal to M_{01}. This is in agreement with reciprocity, as expected.

1.3 TIME-VARYING FIELDS AND MAXWELL'S EQUATIONS

As described earlier, time-varying fields could also be sources of electric or magnetic field creation. In 1831, Faraday published his findings which resulted from performing the following experiment where he proved that a time-varying magnetic field does indeed result in a current. He wound two separate coils on an iron toroid, placed a galvanometer in one and a battery and a switch in the other (Figure 1.9). Upon closing the switch, he realized that the galvanometer

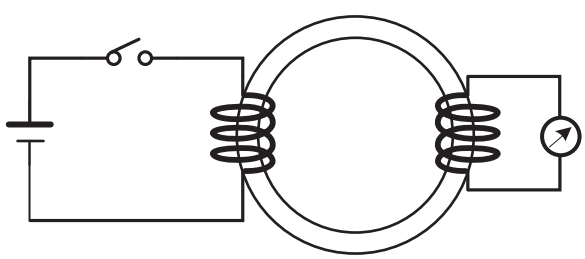


Figure 1.9 Faraday's experiment

was momentarily deflected. He observed the same deflection but in an opposite direction when the battery was disconnected. In terms of fields, we can say that a time-varying magnetic field (or flux) produces an *electromotive force* (emf, measured in volts) that may establish a current in a closed circuit. A time-varying magnetic field may be a result of a time-varying current, or the relative motion of a steady flux and a closed path, or a combination of the two.

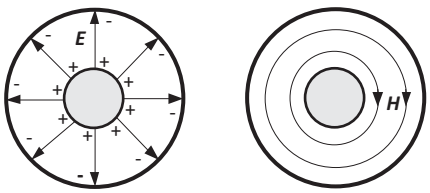


Figure 1.10 Lines of electric and magnetic fields in a coaxial cable

Faraday's law as stated above is customarily formulated as

$$\text{emf} = \oint \boldsymbol{E} \cdot \mathrm{d}\boldsymbol{L} = -\frac{\mathrm{d}\phi}{\mathrm{d}t},$$

where the line integral comes from the basic definition of voltage ($\boldsymbol{E}$ is the electric field intensity). The minus sign indicates that the emf is in such a direction as to produce a current whose flux, if added to the original one, would reduce the magnitude of the emf, and this is generally known as Lenz's law.

Similarly, a time-varying electric flux results in a magnetic field, and is generally formulated by modifying Ampère's circuital law as follows:

$$\oint \boldsymbol{H} \cdot \mathrm{d}\boldsymbol{L} = I + \int_s \frac{\partial \boldsymbol{D}}{\partial t} \cdot \mathrm{d}\boldsymbol{S},$$

where $\boldsymbol{D}$ is the electric flux density, and $\int_S \frac{\partial \boldsymbol{D}}{\partial t} \cdot \mathrm{d}\boldsymbol{S}$ is termed the *displacement current* by Maxwell. To summarize, we can state the four Maxwell's equations in the integral form as follows:

$$\oint \boldsymbol{E} \cdot \mathrm{d}\boldsymbol{L} = -\int_s \frac{\partial \boldsymbol{B}}{\partial t} \cdot \mathrm{d}\boldsymbol{S},$$

$$\oint \boldsymbol{H} \cdot \mathrm{d}\boldsymbol{L} = I + \int_s \frac{\partial \boldsymbol{D}}{\partial t} \cdot \mathrm{d}\boldsymbol{S},$$

$$\oint_S \boldsymbol{D} \cdot \mathrm{d}\boldsymbol{S} = \int_V \rho_V \mathrm{d}V,$$

$$\oint_S \boldsymbol{B} \cdot \mathrm{d}\boldsymbol{S} = 0.$$

The third equation is Gauss's law, as discussed earlier. The fourth equation,[7] states that, unlike the electric fields that begin and terminate on positive and negative charges, the magnetic field forms *concentric circles*. In other words the magnetic flux lines are closed and do not terminate on a magnetic charge[8] (Figure 1.10). Therefore, the closed surface integral of a magnetic field (or magnetic flux density) is zero.

In free space where the medium is source-less, I (or ρ_V) is equal to zero. The first two of Maxwell's equations, when combined, lead to a differential equation relating the second-order

[7] The fourth equation is often known as Gauss's law for magnetism.

[8] Magnetic charges or monopoles are yet to be found in nature, although the magnetic monopole is used in physics as a *hypothetical* elementary particle.

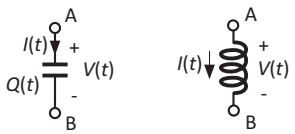


Figure 1.11 Capacitor and inductor circuit representation

derivative of $\boldsymbol{E}$ (or $\boldsymbol{H}$) versus space to its second-order derivative versus time, describing the *wave propagation* in free space. For example, if $\boldsymbol{E} = E_x \boldsymbol{a}_x$, or if electric field is *polarized* only in the x direction, with some straightforward math [6], and using Maxwell's equations in their differential form, it can be shown that[9]

$$\frac{\partial^2 E_x}{\partial z^2} = \mu_0 \epsilon_0 \frac{\partial^2 E_x}{\partial t^2}.$$

The propagation is in the z direction, whose velocity is defined as

$$v = \frac{1}{\sqrt{\mu_0 \epsilon_0}} = c,$$

where $c = 3 \times 10^8$ m/s is the velocity of light in free space.

1.4 CIRCUIT REPRESENTATION OF CAPACITORS AND INDUCTORS

From a circuit point of view, a capacitor is symbolically represented as shown in Figure 1.11, and its voltage and current ($V(t)$ and $I(t)$) as shown satisfy the following relations [7]:

$$I(t) = \frac{\mathrm{d}Q}{\mathrm{d}t},$$

where Q is the charge stored in the capacitor. The above expression is widely known as the *continuity equation*. For the case of a linear and time-invariant capacitor, since $Q = CV$, we can write the well-known expression for the capacitor:

$$I(t) = C\frac{\mathrm{d}V}{\mathrm{d}t}.$$

Note that the continuity equation as expressed in most physics books is $I(t) = -\frac{\mathrm{d}Q}{\mathrm{d}t}$, indicating that the *outward flow* of the positive charge must be balanced by a *decrease* of the charge within the closed surface (that is Q). The minus sign is omitted here, since in Figure 1.11 it is the inward current flow into one terminal of the capacitor with the time rate of increase of charge on that terminal, and not the *outward* current.

An inductor is symbolically represented as shown in Figure 1.11, where its voltage and current ($V(t)$ and $I(t)$) as shown satisfy the following relations:

$$V(t) = \frac{\mathrm{d}\phi}{\mathrm{d}t},$$

[9] The more general form of the wave equation is: $\nabla^2 \boldsymbol{E} = \mu_0 \epsilon_0 \frac{\partial^2 \boldsymbol{E}}{\partial t^2}$.

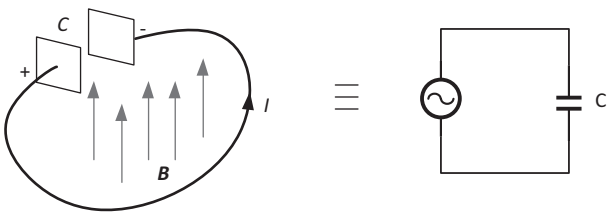


Figure 1.12 Description of Faraday's law in a closed path

where ϕ is the magnetic flux linkage. The above equation is a direct result of Faraday's law, and since $\phi = LI$, we arrive at the well-known expression

$$V(t) = L\frac{\mathrm{d}I}{\mathrm{d}t}.$$

Note that the minus sign is again omitted from the inductor I and V equations, so let us verify if it agrees with Lenz's law. Suppose the current $I(t)$ increases, that is $\mathrm{d}I/\mathrm{d}t > 0$. This indicates that the magnetic field must also increase, hence $\mathrm{d}\phi/\mathrm{d}t > 0$, and it follows that $V(t) > 0$, that is, the potential in node A is greater than that in node B. This is precisely the polarity required to oppose a further increase in current, as required by Lenz's law.

1.5 DISTRIBUTED AND LUMPED CIRCUITS

Kirchhoff's voltage law or KVL states that the sum of electric potentials in a closed path is equal to zero, that is, $\oint \boldsymbol{E} \cdot \mathrm{d}\boldsymbol{L} = 0$, whereas Maxwell's first equation (or Faraday's law as described before) says otherwise. The time-varying term in Maxwell's second equation, that is the displacement current, is similarly in violation of KVL. To clarify further, let us study the simple circuit shown in Figure 1.12, consisting of an ideal (zero inductance and resistance) piece of wire attached to a parallel-plate capacitor, forming a loop around it.

Assume that within the loop an external magnetic field is applied, varying sinusoidally with time. Thus an emf of $V_0 \cos \omega_0 t$ across the capacitor is produced, as predicted by Faraday's law. On the other hand, if the wire is ideal, KVL indicates that the *shorted* capacitor must have a zero voltage across it. Interestingly, the voltage across the capacitor creates a current I in the wire:

$$I = -\omega_0 C V_0 \sin \omega_0 t = -\omega_0 \frac{\epsilon A}{d} V_0 \sin \omega_0 t,$$

where ϵ, A, and d are parallel plate capacitor parameters. In any closed path Ampère's circuital law gives us the magnetic field as a result of this current. Particularly, for a specific closed path which passes between the capacitor plate, we can determine the displacement current. Within the capacitor:

$$D = \epsilon E = \epsilon \left(\frac{V_0}{d} \cos \omega_0 t \right),$$

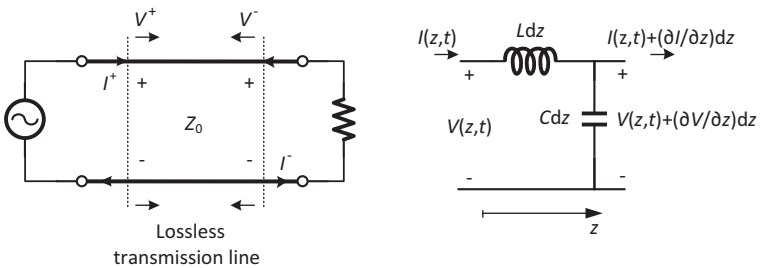


Figure 1.13 Lossless transmission line and its lumped differential equivalent

and according to Maxwell's second equation, the displacement current is

$$I_D = \frac{\partial D}{\partial t} A = -\omega_0 \frac{\epsilon A}{d} V_0 \sin \omega_0 t,$$

which is equal to the earlier result obtained for the current in the loop.

This brings us to a general discussion about *lumped* and *distributed* circuits. We can say that the basic elements in a circuit, and the connections between them, are considered *lumped* (and thus KVL or KCL are applicable) if the time delay in traversing the elements is negligible, and hence they can be treated as static. If the components are large enough, or the frequency is high enough (or equivalently the delays are short enough), one must use *distributed* elements. This means that the resistive, capacitive, or inductive characteristics must be evaluated on a unit distance basis. Common examples of the distributed circuits are transmission lines or waveguides, which are intended to deliver electromagnetic energy from one point to another, and naturally are separated by long (relative to the wavelength) distances. To illustrate an example of how to deal with such circuits, consider a lossless line, as shown in Figure 1.13, connecting a generator to a load. We can construct a model for this *transmission line* using lumped capacitors and inductors. An equivalent circuit of a differential section of the line with no loss (where the length dz is approaching zero) is shown on the right in Figure 1.13. Since each section corresponds to a very small portion of the line, that is dz approaching zero, KVL and KCL are valid for that section, despite the distributed nature of the line.

Writing KVL leads to

$$V(z,t) = (L\mathrm{d}z)\frac{\partial I(z,t)}{\partial t} + V(z,t) + \frac{\partial V(z,t)}{\partial z}\mathrm{d}z,$$

which results in

$$L\frac{\partial I}{\partial t} = -\frac{\partial V}{\partial z}.$$

Similarly, by writing KCL we obtain

$$\frac{\partial I}{\partial z} = -C\frac{\partial V}{\partial t}.$$

Taking the derivative of one of the two equations above versus space (or z), and the other versus time (or t), the current component, I, could be eliminated, arriving at

$$\frac{\partial^2 V}{\partial z^2} = LC\frac{\partial^2 V}{\partial t^2}.$$

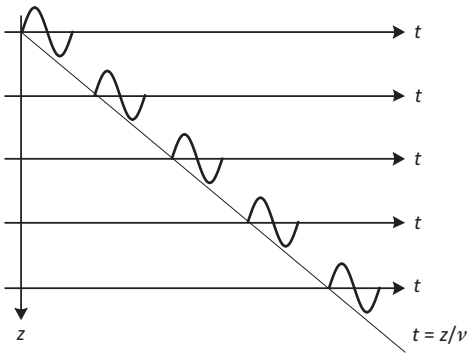


Figure 1.14 Wave propagating in a transmission line

Note the similarity of this differential equation with the one given for the wave propagation in the previous section, where electric field is replaced by voltage, and the propagation velocity (as we will define shortly) is now equal to $\frac{1}{\sqrt{LC}}$. Since L and C are the inductance and capacitance per unit length, they have the same units as μ and ϵ in the wave equation, that is H/m and F/m.

The solution of this differential equation is of the form

$$V(z,t) = f_1\left(t - \frac{z}{v}\right) + f_2\left(t + \frac{z}{v}\right) = V^+ + V^-,$$

which can be verified by replacing $V(z,t)$ above in the original differential equation describing the distributed wave propagation. The functions f_1 and f_2 could be anything as long as they are differentiable twice, and take arguments $t \pm z/v$. The arguments of f_1 and f_2 indicate, respectively, travel of the functions in the *forward* and *backward* z directions, and thus we assign symbols V^+ and V^- to them. To understand this better, suppose we would like to keep the argument of f_1 constant at zero. As time increases (as it should), z also has to increase with a rate of $v \times t$ (hence we call v the velocity). Therefore the function f_1 needs to move forward or in the positive z direction. On the other hand, for f_2, z has to decrease, indicating a backward motion. The forward moving signal is illustrated in Figure 1.14, where we assume f_1 and f_2 are sinusoidal. This in fact will be the case, for the sinusoidal steady state solution, as we show in Chapter 3.

The propagation velocity is obtained by replacing the solution in the original differential equation. This yields

$$v = \frac{1}{\sqrt{LC}}.$$

A similar procedure results in the following solution for the current:

$$I(z,t) = \frac{1}{Z_0} f_1\left(t - \frac{z}{v}\right) - \frac{1}{Z_0} f_2\left(t + \frac{z}{v}\right) = I^+ + I^-,$$

where Z_0 is defined as the *characteristic impedance* of the line measured in Ohms and is equal to

$$Z_0 = \sqrt{\frac{L}{C}}.$$

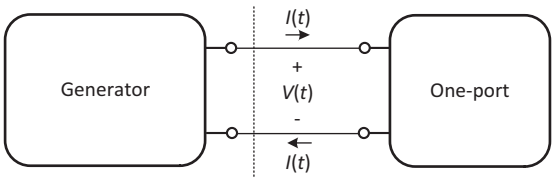


Figure 1.15 Instantaneous power and energy concept

Even though Z_0 is measured in Ohms, it is not a physical resistor as we already assumed that the line is lossless. It simply relates the forward and backward voltages and currents in the line as follows (Figure 1.13):

$$V^+ = Z_0 I^+,$$

$$V^- = -Z_0 I^-.$$

Returning to our previous coaxial cable example, since the values of L and C were obtained already, the characteristics impedance is readily given by

$$Z_0 = \sqrt{\frac{\mu}{\epsilon}} \ln \frac{b}{a} \Omega,$$

where $\epsilon = \epsilon_r \epsilon$ is the permittivity of coaxial cable. Typical values of a, b, and ϵ result in a characteristics impedance of several tens of ohms, commonly set to 50 Ω.

1.6 ENERGY AND POWER

From an electromagnetic field perspective, we can define electrostatic and magnetic energy stored as follows [6]:

$$W_E = \frac{1}{2} \int_V \boldsymbol{D} \cdot \boldsymbol{E} \mathrm{d}V = \frac{1}{2} \epsilon \int_V |\boldsymbol{E}|^2 \mathrm{d}V,$$

$$W_H = \frac{1}{2} \int_V \boldsymbol{B} \cdot \boldsymbol{H} \mathrm{d}V = \frac{1}{2} \mu \int_V |\boldsymbol{H}|^2 \mathrm{d}V,$$

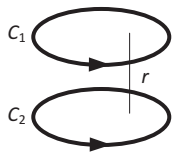
where W_E and W_H denote electric and magnetic energy respectively in joules, and the integrals are performed over volume.[10]

From a circuit perspective, let us consider Figure 1.15, where a generator is connected to a one-port, with current $I(t)$ entering the one-port, and a voltage $V(t)$ across it.

The instantaneous power *delivered* to the one-port by the generator is defined as

$$p(t) = V(t) I(t),$$

[10] Both equations may be proven from basic definitions.



M: between C_1 and C_2

Figure 1.16 Self and mutual inductances as functions of geometry only

and the energy produced by the generator from initial time t_0 to t is

$$W(t_0, t) = \int_{t_0}^{t} p(\theta)\mathrm{d}\theta = \int_{t_0}^{t} V(\theta)I(\theta)\mathrm{d}\theta.$$

For an ideal capacitor with initial zero voltage or charge, that is $Q(t_0) = 0$, we have

$$W(t) = \int_{t_0}^{t} V(\theta)I(\theta)\mathrm{d}\theta = \int_{0}^{Q(t)} \frac{Q(\theta)}{C}\frac{\mathrm{d}Q}{\mathrm{d}\theta}\mathrm{d}\theta = \frac{Q(t)^2}{2C} = \frac{1}{2}CV(t)^2,$$

where V and I are replaced with their Q equivalents inside the integral. Similarly for an inductor

$$W(t) = \frac{1}{2}LI(t)^2.$$

Again we can identify the resemblance between μ and ϵ to L and C, and E and H to V and I in the energy equations. Since E and ϵ have units per distance (/m), the field energy integral is performed over volume. Note that sometimes it is more convenient to calculate the inductance or capacitance for a given geometric structure from the energy definition, for instance to use

$$L = \frac{2W_H}{I^2}$$

for the inductance calculation (as opposed to $L = \phi/I$ as presented earlier). W_H is obtained from its basic definition as a function of B and H. Expressing W_H and I based on H, it can be shown that[11]

$$L = \frac{\mu}{4\pi}\oint\left(\oint\frac{\mathrm{d}\boldsymbol{L}}{r}\right)\cdot\mathrm{d}\boldsymbol{L} = \frac{\mu}{4\pi}\oint\oint\frac{\mathrm{d}\boldsymbol{L}_1\cdot\mathrm{d}\boldsymbol{L}_2}{r},$$

which indicates that only the inductance is a function of geometry, and not the current (Figure 1.16). This serves as a proof for the qualitative discussion presented earlier. A similar expression for the mutual inductance exists as well, where the integral is defined between two circuits carrying current (Figure 1.16):

$$M = \frac{\mu}{4\pi}\oint\oint\frac{\mathrm{d}\boldsymbol{L}_1\cdot\mathrm{d}\boldsymbol{L}_2}{r}.$$

[11] The proof requires the use of the vector magnetic potential ($\boldsymbol{A}$) defined as $\boldsymbol{B} = \nabla\times\boldsymbol{A}$, analogous to the electric potential defined before where $\boldsymbol{E} = -\nabla\boldsymbol{V}$.

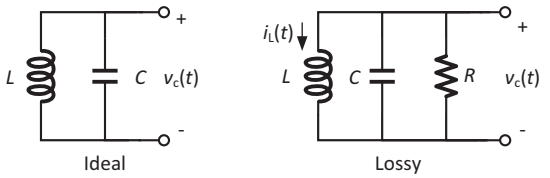


Figure 1.17 Ideal and lossy LC circuits

It is perhaps worthwhile summarizing the following similarities between electric and magnetic fields, and to voltages and currents:

$$C(\mathrm{F}) \leftrightarrow \epsilon(\mathrm{F/m}),$$
$$L(\mathrm{H}) \leftrightarrow \mu(\mathrm{H/m}),$$
$$V(\mathrm{V}) \leftrightarrow E(\mathrm{V/m}),$$
$$I(\mathrm{A}) \leftrightarrow H(\mathrm{A/m}),$$
$$E \leftrightarrow H,$$
$$D \leftrightarrow B.$$

Also note the similarity between Gauss's and Ampère's laws, as well as Coulomb's and Biot-Savart's laws.

1.7 LC AND RLC CIRCUITS

With the background presented, we now have the right tools to analyze LC circuits. An ideal (lossless) LC circuit is shown in Figure 1.17 (left side).

Let us assume that the capacitor is charged initially to a voltage of V_0. From a circuit point of view, the initial charge may be the result of an impulse current source of $I(t) = V_0\delta(t)$ appearing in parallel with the circuit. If the impulse has a magnitude of V_0 coulomb, the capacitor initial voltage will be: $v_C(0^+) = v_0$. Taking the capacitor voltage $v_\mathrm{C}(t)$ as the variable, we can write:

$$\frac{\partial^2 v_\mathrm{C}}{\partial t^2} + \frac{1}{LC} v_\mathrm{C} = 0.$$

The differential equation is solved by taking the Laplace transform of the two sides. This yields ([8], [7]):

$$s^2 + \frac{1}{LC} = 0.$$

This results in the poles of the circuit at $s_{1,2} = \pm \mathrm{j}/\sqrt{LC} = \pm \mathrm{j}\omega_0$. The final solution is

$$v_\mathrm{C}(t) = V_0 \cos \omega_0 t.$$

Next we will calculate the energy stored in the capacitor and the inductor. From the previous section:

$$W_\mathrm{C}(t) = \frac{1}{2} C v_\mathrm{C}(t)^2 = \frac{1}{2} C V_0^2 \cos \omega_0 t^2.$$

Similarly, solving for the inductor current yields

$$i_L(t) = \frac{1}{L} \int v_\mathrm{C}(t) \mathrm{d}t = \frac{V_0}{L\omega_0} \sin \omega_0 t.$$

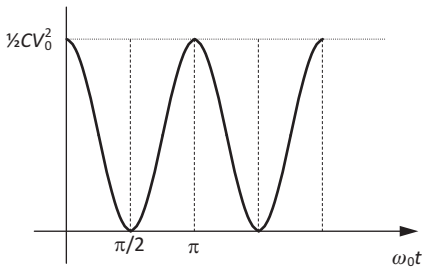


Figure 1.18 Energy of capacitor in a lossless LC circuit

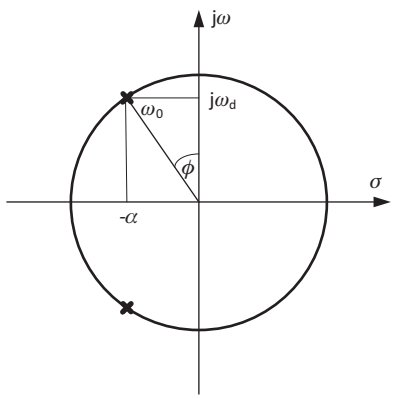


Figure 1.19 Pole locations of an RLC circuit in the s-plane

Since $C = \frac{1}{L\omega_0^2}$, we have

$$W_{\mathrm{L}}(t) = \frac{1}{2} C V_0^2 \sin \omega_0 t^2.$$

Therefore the total energy at any point of time is $W_{\mathrm{T}}(t) = W_{\mathrm{C}}(t) + W_{\mathrm{L}}(t) = \frac{1}{2} C V_0^2$, which is constant and equal to the initial energy stored in the capacitor. This is expected, as the LC circuit is lossless. The energy is only exchanged between the inductor and the capacitor, as shown in Figure 1.18, indicating a steady *oscillation*.

In practice, both the capacitor and the inductor are lossy, and for now let us model the total loss as a parallel resistor as shown in Figure 1.17 (right side). We assume the loss is moderate, that is the value of R is large compared to the impedances of L or C at the frequency of interest. The new differential equation is

$$\frac{\partial^2 v_{\mathrm{C}}}{\partial t^2} + \frac{1}{RC} \frac{\partial v_{\mathrm{C}}}{\partial t} + \frac{1}{LC} v_{\mathrm{C}} = 0,$$

resulting in the complex poles $s_{1,2} = -\frac{\omega_0}{2Q} \pm \mathrm{j}\omega_0 \sqrt{1 - \frac{1}{4Q^2}} = -\alpha \pm \mathrm{j}\omega_{\mathrm{d}}$, where, for the moment, we define the quality factor or Q as $Q = \frac{R}{L\omega_0} = RC\omega_0$. In order to have complex poles, Q must be greater than ½. This definition is only a mathematical expression for Q, and we will shortly attach a more physical meaning to it based on the energy concept. Figure 1.19 shows the location of the complex poles on the s-plane.

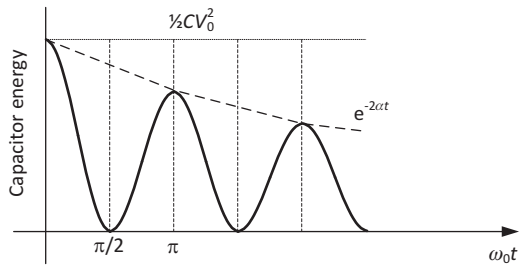


Figure 1.20 Energy in a lossy LC circuit

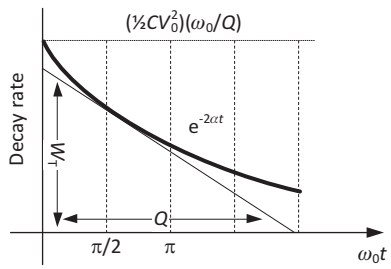


Figure 1.21 Power dissipated in a resistor of an RLC circuit vs. phase cycles

Defining $\phi = \cos^{-1}\omega_d/\omega_0$, the capacitor voltage, $v_C(t)$ and the inductor current $i_L(t)$ are given by

$$v_C(t) = V_0\frac{\omega_0}{\omega_d}e^{-\alpha t}\cos(\omega_d t + \phi),$$

$$i_L(t) = \frac{V_0}{L\omega_d}e^{-\alpha t}\sin\omega_d t.$$

Assuming $Q \gg 1$, the total energy stored in the LC tank is approximately

$$W_T(t) = W_C(t) + W_L(t) \approx \frac{1}{2}CV_0^2e^{-2\alpha t},$$

indicating an initial energy of $\frac{1}{2}CV_0^2$ decaying exponentially, as shown in Figure 1.20.

Similarly to an ideal tank, the capacitor and the inductor energies move back and forth between the two, but at a rate of ω_d, slightly lower than ω_0, and eventually decay to zero. The total energy decay rate, or equivalently the power dissipated in the resistor, is

$$p = -\frac{dW_T}{dt} = 2\alpha W_T = \frac{\omega_0}{Q}W_T.$$

Rearranging the equation above, we arrive at a more physical and perhaps a more fundamental definition for the quality factor:

$$Q = \omega_0\frac{W_T}{p} = \omega_0\frac{\text{total energy stored}}{\text{average power dissipated}}.$$

Note that, as shown in Figure 1.21, the slope (the normalized decay rate) at any point is equal to W_T/Q.

To sustain a steady oscillation, the power dissipated in the resistor must be compensated. Since the passive circuits are incapable of generating energy, this must be done by an active

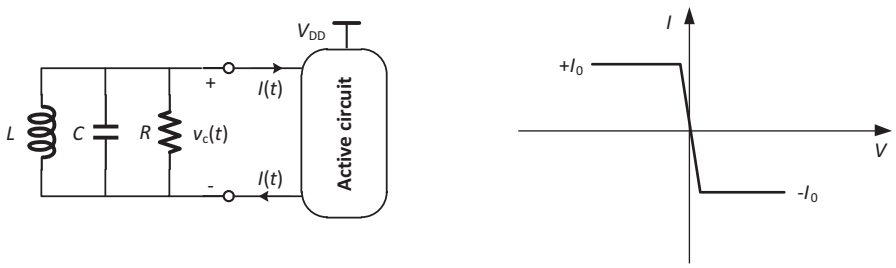


Figure 1.22 Active circuit to compensate for the loss of RLC circuit, and its *I–V* characteristics

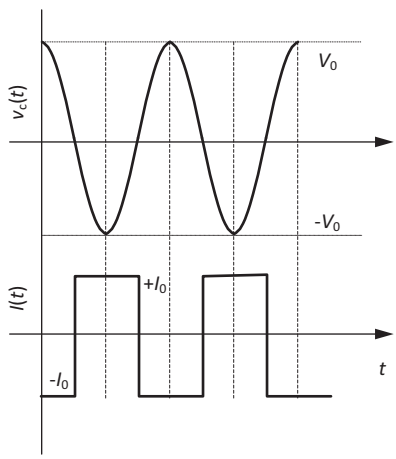


Figure 1.23 Voltage and current waveforms in the case of steady oscillation

circuit, as shown in Figure 1.22. The power *creation* requires the current–voltage product to be negative, so we expect the *I–V* characteristics of the active circuit to have a negative slope, as shown in Figure 1.22, effectively acting as a *negative resistance* compensating the loss associated with the *positive resistance*. The sharper the slope, the more efficient this would be. Without being distracted by details of the active circuit, we could ideally model that as a one-port whose *I–V* curve is shown in Figure 1.22. The conservation of energy requires the energy created to originate from somewhere, and that is usually from the DC power supplying the active circuit (V_{DD} in Figure 1.22). So it is reasonable to assume that as the voltage or the current increase, they eventually reach a plateau where they cannot increase anymore, as denoted by I_0 in Figure 1.22.

Now at steady state when the power dissipated and the power created are balanced, the capacitor voltage is equal to $v_C(t) = V_0 \cos \omega_0 t$, as in an ideal tank, and the total energy stored in the tank must be $\frac{1}{2} C V_0^2$.

Since this voltage appears across the active circuit, we expect the current supplied by that to the lossy LC tank ($I(t)$) to be as shown in Figure 1.23.

Therefore, the average power dissipated in the tank (which is equal to the power created in the active one-port) is given by (Figure 1.23):

$$p = -\frac{1}{T} \int_T (V_0 \cos \omega_0 t) I(t) \mathrm{d}t = \frac{2 V_0 I_0}{\pi}.$$

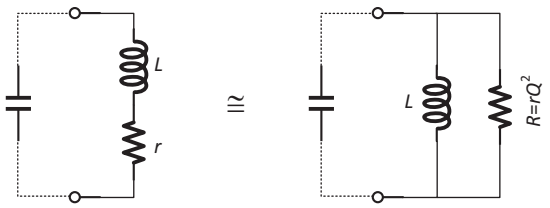


Figure 1.24 Inductor with a series resistor, and its equivalent model

Hence, according to the Q definition:

$$Q = \omega_0 \frac{W_T}{p} = \omega_0 \frac{\frac{1}{2} C V_0^2}{\frac{2V_0 I_0}{\pi}} = \frac{\pi}{4} \omega_0 C \frac{V_0}{I_0},$$

and since $Q = RC\omega_0$ for a parallel RLC tank,

$$V_0 = \frac{4}{\pi} R I_0.$$

This shows that the steady oscillation amplitude, V_0, is a function of only the active one-port saturation current, I_0, and the amount of loss. Since, regardless of its voltage, the one-port always drains a steady current of $2I_0$ from the supply (to produce the waveform shown in Figure 1.22), then the efficiency is equal to

$$\eta = \frac{\frac{2V_0 I_0}{\pi}}{V_{DD} 2 I_0} = \frac{V_0}{\pi V_{DD}} \leq \frac{2}{\pi},$$

assuming V_0 can reach a maximum swing of $2V_{DD}$. This outcome can be intuitively explained by realizing that the assumed high-Q nature of the tank produces a sinusoidal waveform, whereas the current delivered by the active circuit is square-wave in nature (Figure 1.23), given the sharp slope of its I–V curve. Taking only the fundamental then leaves a *loss* factor of $2/\pi$. We will arrive at a similar result for hard-switching mixers, as we will discuss in Chapter 7.

We have not offered much of an insight into how to realize the active one-port necessary for sustaining the oscillation. As our focus in this chapter has been the LC circuits, we will leave it at that, and present various circuit topologies that implement this active one-port in Chapter 8 in the context of LC oscillators.

In contrast to our model, in practice, inductors experience an ohmic loss due to the finite wire conductivity, physically modeled as a small *series* resistance, as shown in Figure 1.24. While we will systematically prove it in Chapter 3, it can be easily shown that if Q is large, the two circuits shown in Figure 1.24 are equivalent. The quality factor (Q) is

$$Q = \frac{R}{L\omega_0} = \frac{L\omega_0}{r}.$$

Furthermore, assuming the elements are high-Q, if the capacitor has also a parallel resistive loss, we can derive the equivalent parallel RLC circuit whose quality factor, Q, is

$$\frac{1}{Q} = \frac{1}{Q_L} + \frac{1}{Q_C},$$

where Q_L and Q_C are the inductor and capacitor quality factors.

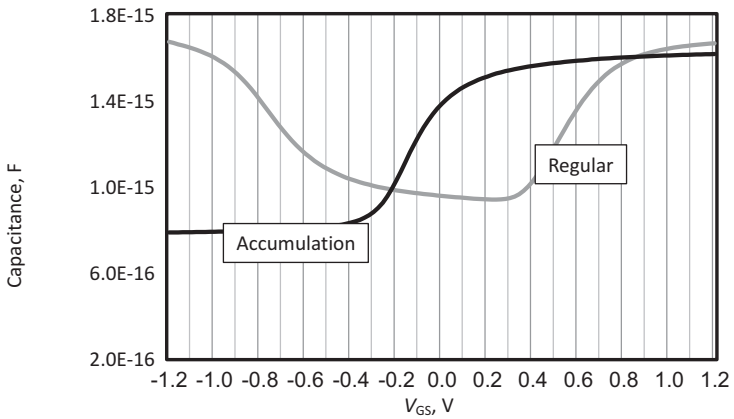


Figure 1.25 Regular and accumulation-mode MOS capacitance

1.8 INTEGRATED CAPACITORS

It is often very desirable to tune the resonance frequency of the LC tanks. For instance, in a tuned amplifier a discrete tuning may be required to extend the bandwidth, whereas in an oscillator locked in a phase-locked loop, both discrete and continuous tuning are needed. Due to their physical structure, inductors are usually not viable options, although several structures have been proposed that vary the inductance through using multi-tap switched segments [9]. This usually comes at a cost of performance, and at best only offers a discrete tuning. Capacitors, on the other hand, are very well suited to providing the tuning. In this section we will discuss a few schemes that are commonly used in RF integrated circuits. Before that, let us first briefly take a look at the common fixed capacitors available in integrated circuits.

The gate capacitance of MOS transistors may be exploited to realize high density but non-linear capacitors. Shown in Figure 1.25 (the grey curve) is the simulated capacitance versus gate voltage of a 40 nm regular NMOS capacitor. Depending on the gate voltage, the device operates either in accumulation (approximately for $V_{GS} < 0$), depletion ($0 < V_{GS} < V_{TH}$), or inversion ($V_{GS} > 0$) [2]. The device threshold voltage (V_{TH}) is estimated to be around 400 mV. In inversion or accumulation the capacitance reaches a maximum, close to gate-oxide capacitance, C_{OX}. To achieve a reasonably linear response, the device should be biased at voltages well above threshold (say greater than 500 mV in our example below), which makes it unsuitable for low supply voltage applications. Moreover, MOS capacitors usually have a very large gate leakage, which may be problematic. For that reason, a thick-oxide device may be chosen which is less dense, but has substantially less leakage.

To avoid inversion, the NMOS may be placed inside an *n*-well [10] at no extra cost, known as an accumulation-mode MOS capacitor (Figure 1.26). The naming has to do with the fact that the transistor must stay either in depletion or accumulation.

This leads to the *C*–*V* characteristics shown in Figure 1.25 (the black curve). Even though the capacitance is still quite non-linear, the accumulation starts right around 0V rather than the threshold voltage. Consequently, it is easier to bias the device at low voltages and benefit from an almost flat region where the gate capacitance has reached C_{OX}.

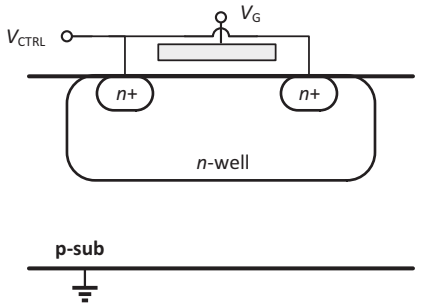


Figure 1.26 Accumulation-mode MOS capacitor

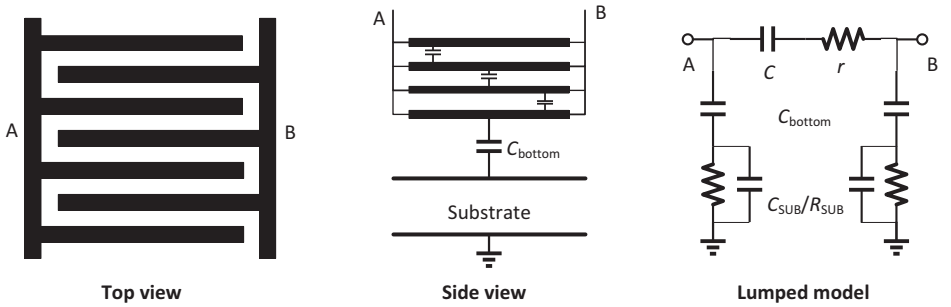


Figure 1.27 Fringe capacitors

In many cases, for instance when the capacitance is placed in the feedback path of an op-amp, it may not be practical to apply a relatively large bias voltage. An alternative is to use a linear capacitor formed by the fringe fields which are quite strong in most modern CMOS processes due to the close proximity of metal lines. While problematic for routing signals and connecting blocks, this, along with the large number of metal layers available, could be taken advantage of when building linear capacitors. An example is shown in Figure 1.27. To maximize density, minimum width metal lines are placed at the minimum spacing allowed by the technology, and the two terminals form a comb-like structure. Moreover, several layers of metals connected at each end may be placed on top of each other, to improve density further.

It is common for CMOS processes to provide a thick or ultra-thick top metal layer used for clock tree routing or inductors. Since the minimum spacing allowed is usually too large, the top thick metal layer may not be used. Additionally, there is a concern as to how large the bottom plate parasitic capacitance would be. Therefore, it may be advantageous to drop one or two of the bottom metal layers as well (particularly poly and metal one due to their large sheet resistance), to place the capacitor further away from the substrate and reduce the bottom plate capacitance. However, if fewer metal layers are used, for the same value of the capacitance, the structure needs to be bigger, and hence the bottom plate parasitic increases accordingly. In a 40 nm CMOS process with a thick M6 option, the best compromise appears to be using M3-M5, leading to a density of about 2fF/μm^2. The bottom (or top) plate parasitic is generally very

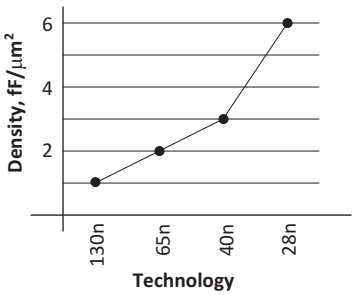


Figure 1.28 Maximum fringe capacitance density in various CMOS processes

small, about 1–2%. For a given structure, the exact amount of capacitance is very difficult to calculate using closed formulas, and it is often best to use extraction tools (such as EMX[12]) to predict the capacitance.

A lumped model of the capacitance is presented in Figure 1.27. The bottom and top plate parasitic, that are generally symmetric due to the physical structure of the capacitor, are connected to the substrate. The substrate is lossy and is typically modeled by a parallel RC circuit. Since the bottom plate parasitic capacitance (C_{bottom}) is small, and R_{SUB} is usually large for bulk processes, this loss is negligible for frequencies up to several GHz. Additionally, there is the metal series resistance forming the comb lines. If, for a given capacitance, the structure is built to consist of N smaller units in parallel, this resistance is then reduced by N^2, typically leading to a very high-Q capacitor for well-designed structures.

The fringe capacitors scale with technology to a good extent, as the physical spacing between metal lines generally improves. Shown in Figure 1.28 is the maximum multi-finger fringe capacitance density in fF/μm^2 for several recent standard CMOS processes. Note that unlike MIM (metal–insulator–metal) capacitors, fringe capacitors do not require any additional process step, and thus incur no extra cost. By comparison, a thin-oxide MOS capacitor in 28 nm is 23fF/μm^2 (corresponding to an oxide thickness of about 1.25 nm), whereas a thick oxide one is 10fF/μm^2. Note that if MOS capacitors are used, it is still possible to fill fringe capacitors on top to further boost the density.

Continuous tuning may be achieved by using one of the two MOS structures discussed earlier. Particularly NMOS in an n-well is desirable as it provides a greater tuning range. For a regular NMOS, if the voltage swing across the tank were large (as is the case for most CMOS oscillators) regardless of the DC bias applied, the effective capacitance would be mostly C_{OX}. That is because the depletion capacitance corresponds to a relatively narrow region of the C–V curve (Figure 1.25). To clarify, Figure 1.29 shows a large signal simulation of a 28 nm MOS capacitor. The effective capacitance is plotted versus the control voltage, for four different values of signal swing across it: 0, 0.5 V, 1 V, and 1.8 V. The *effective capacitance* here is defined as the magnitude of the fundamental component of the current, divided by the voltage swing across the capacitance, normalized to the angular frequency.

In the case of a regular MOS capacitor, the ratio of maximum to minimum capacitance (that is approximately C_{OX} to $C_{\text{OX}} \| C_{\text{DEP}}$) is around 2.5, but diminishes to less than 1.4 as the

[12] **EMX** is a planar 3D integral equation solver that uses accurate representation of Maxwell's equations.

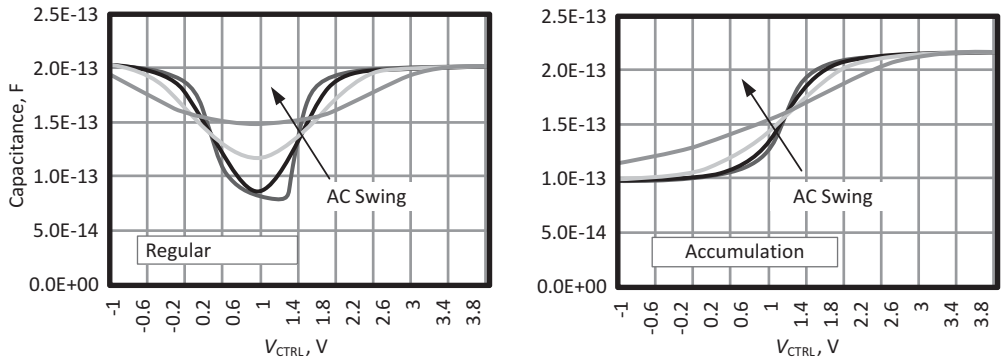


Figure 1.29 Large signal MOS capacitor simulation

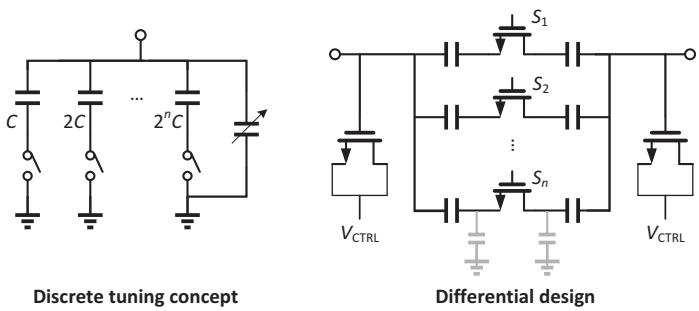


Figure 1.30 Discrete tuning using switched capacitors

signal swing increases. The accumulation mode varactor, on the other hand, has a lower maximum to minimum capacitance ratio of about 2, but it is maintained reasonably well for signal swings of as high as 1.8 V. The two varactors are identical in size, and use a thick oxide NMOS with a channel length of 0.75 µm. A shorter channel, although resulting in a better Q, has a worse tuning range.

Since the Q of continuously tuned capacitors may not be very high, a wider tuning range without compromising Q can be achieved by incorporating discrete tuning using switched linear capacitors along with MOS varactors, as illustrated in Figure 1.30 [11]. This also results in less VCO gain, and thus less sensitivity to noise and interference at the VCO control voltage. The MOS varactor only needs to provide enough range to cover the worst-case discrete step size.

A larger switch results in lower on resistance and thus a better quality factor, however, the switch parasitic capacitance in the off mode limits the tuning range. If designed differentially, the same tuning range is achieved but with twice as much Q, as the on resistance is halved. A differential design in 28 nm CMOS as an example is shown in Figure 1.30. It consists of 32 units of 40 fF linear capacitors, with the total capacitance varying from 430 fF to 1.36 pF (about 3 times), in steps of 29 fF (much less than 40 fF, due to switch parasitics). The Q at 3.5 GHz varies from a maximum of 80 to 45 when all the capacitors are turned on. The switches are $11\times1/0.1\mu m$.

1.9 INTEGRATED INDUCTORS

First introduced in silicon in 1990 ([12], [13]), monolithic inductors have been since widely used in RF and mm-wave applications. Due to fabrication limitations, on-chip inductors are typically realized as metal spirals. To achieve a lower loss and thus better Q, it is common to use the top metal layer, which is typically thick or ultra-thick. Applying the Biot-Savart law to calculate the magnetic field, one can show that the self-inductance of a piece of wire with length l and a rectangular cross section at moderate frequencies (several GHz) is [14]

$$L \approx \frac{\mu_0}{2\pi} l \left(\ln \frac{2l}{W+t} + 0.5 \right),$$

where t is the thickness of the metal, and is fixed for a given technology,[13] while W is the metal width and is a design parameter. All units are metric, and we assume that the length is much larger than the width. To gain some insight, the series resistance of the same line at low frequency is given by

$$r = R_{\square} \frac{l}{W},$$

where $R_{\square}$ is the metal sheet resistance (about 10 mΩ/□ for 40 nm ultra-thick M6 for instance). Increasing W leads to a smaller series resistance, but also results in less inductance, and larger area. If the only concern is to maximize Q for a given inductance, then increasing the width can only help to a certain extent, as a larger W needs a longer length to keep the inductance constant, leading to more resistance. In fact, the logarithmic nature of the inductance dependence on W suggests that increasing W results in a steep improvement of Q, but beyond a certain point only leads to a bigger structure and more capacitance, but marginal improvement of Q.

For example, for $l = 1\,\text{mm}$, $W = 7\,\mu\text{m}$, and $t = 3\,\mu\text{m}$, the inductance is about 1.16 nH, whereas the low-frequency series resistance is 2.8 Ω for M6. This leads to an upper bound for Q of 21.6 at 4 GHz, assuming the low-frequency series resistance is the only loss mechanism, which is not the case in reality. In fact at higher frequency the *skin effect* results in a lower value of Q, as suggested by the equation above. The skin effect may be understood by examining the essential features of the behavior of the electromagnetic wave normal at the surface of a good conductor. We briefly discussed the fundamentals of wave propagation in the previous section, and will only show the final solution here as a more detailed discussion is beyond the scope of this book ([4], [5]). In a conducting medium where the conduction current, $\sigma \boldsymbol{E}$ as expressed by Ohm's law, is much larger than the displacement current, it can be shown that the solution for the electric field is given by

$$\boldsymbol{E} = E_0 \boldsymbol{a}_x \mathrm{e}^{-\gamma z}.$$

We have assumed that the electric field is polarized in the x direction only. The value of γ is found to be

$$\gamma = \sqrt{\mathrm{j}\omega_0 \mu_0 \sigma} = \frac{1+\mathrm{j}}{\delta},$$

[13] An ultra-thick M6 layer in TSMC 40nm process is 3 μm.

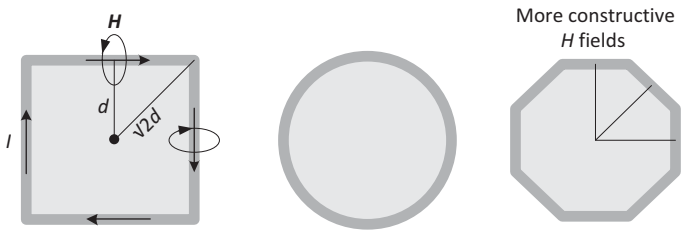


Figure 1.31 Spiral inductor in integrated circuits

where σ is the conductivity, ω_0 is the frequency, and $\delta = (\frac{1}{2}\omega_0\mu_0\sigma)^{-1/2}$ is the *skin depth*. This suggests that in the conductor, the field decays by an amount of e^{-1} in a distance of one skin depth δ, which could be comparable to the metal width or its thickness at high frequencies. For instance, in 40 nm CMOS, the skin depth for M6 is about 1 μm at 4 GHz. As a result, effectively the current tends to flow more at the surface, and the resistance will increase. Accordingly, a modified expression for the metal resistance to include the skin effect is

$$r = R_{\square}\frac{l}{W}\frac{t/\delta}{1-e^{-t/\delta}},$$

where t is the metal thickness. At low frequency, δ is large and the equation simplifies to the original expression for the resistance. However, at very high frequencies, the exponential term approaches zero and thus

$$r = R_{\square}\frac{l}{W}\frac{t}{\delta} = \frac{1}{\sigma}\frac{l}{\delta W}.$$

That is, t is replaced by δ at high frequency, which could lead to considerably higher resistance.

1.9.1 Spiral inductors

Using a long piece of wire as described above is clearly not a viable option, as apart from the area, connecting it to any circuit is impractical. It is more common to form spirals to make it more compact and practical to use. The most natural choice would be a circle. However, circles are not physically possible to be laid out in an integrated circuit. The same length could be wound into a square spiral, as shown in Figure 1.31. According to the Biot-Savart law, the magnetic fields of all four legs add up in the center, with a direction normal to this page, although they will not add quite constructively at the edges. A better approximation of the circle may be made by using a hexagon or an octagon, as 45° angles are allowed. This results in a more constructive addition of the magnetic fields (Figure 1.31), typically leading to a higher Q.

The inductance of the spiral could be given with the general expression as follows ([14], [15]):

$$L_{\mathrm{T}} = \sum L + \sum M^{+} - \sum M^{-},$$

where the first term indicates the sum of the self-inductances of each leg, and the second and third terms indicate the mutual inductances between *parallel* legs that carry the *same* or *opposite*

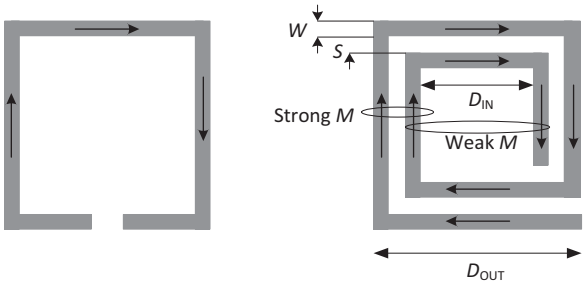


Figure 1.32 Multi-turn inductors

direction currents. In the case of a single-turn square spiral shown in Figure 1.31 the second term does not exist as the parallel legs only carry opposite currents. For the previous example, ignoring the mutual inductances for the moment (the third term), and assuming a length of roughly 1 mm/4 = 250 µm for each leg, the total inductance is 4×0.22 = 0.88 nH, which is somewhat less than the 1.16 nH obtained for the straight line. Clearly, this has to do with the additional logarithmic dependence of the inductance on the inductor length, as shown earlier. In reality the inductance is even less as the negative mutual inductance reduces it further. Since the series resistance is roughly the same, the upper bound on Q obtained earlier is proportionally less. A closed form expression for the mutual inductance between two pieces of wires with identical length l, separated by a space d, is quite tedious to calculate ([15], [16]). Here, to gain some insight, we provide only the expression for two filamentary lines (that is t and W are much smaller than l and d) which has a simple closed form solution. The equation may be readily obtained by applying the Biot-Savart law and integrating over space:

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

For our previous example, if $\frac{l}{d} \cong 1$, which is a good approximation for a square single-turn inductor, then the mutual inductance between two legs is roughly 12% of the inductance of each leg, reducing the total inductance further.

The inductance may increase if another turn is added (Figure 1.32), assuming that the inner diameter (D_{IN}) is still large, that is the inner and outer diameters (D_{IN} and D_{OUT}) are comparable. The increase is due to the large contribution of positive mutual inductance between the adjacent legs that carry the same current, while the negative mutual inductance of the opposite legs is still small as they are far apart. Therefore, when designing spiral inductors, it is common to have *hollow* structures, with the adjacent legs drawn as close as possible. Keeping the inner diameter large, however, limits the number of turns allowed.

Multi-turn inductors typically enjoy a more compact design with a somewhat better Q, but larger capacitance. The reason for not as much Q improvement as expected in multi-turn inductors is attributed to a phenomenon known as *current crowding* or the *proximity effect* [15]. The mutual induction of eddy currents in the adjacent turns and the loss associated with it increase the loss and limit the Q improvement. If the inductance needed is small (say a few tenths of a henry), then a single turn with reasonably large diameter remains the best choice.

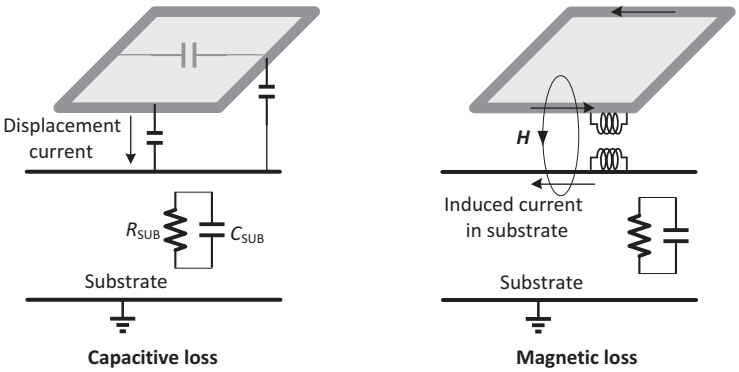


Figure 1.33 Substrate loss in on-chip inductors

For a given structure in general, the inductance can be fully expressed by knowing the metal width (W), the spacing between adjacent legs (S), the number of turns, and either the inner diameter, the outer diameter, or the total length (Figure 1.32). For most RF applications, practical values of integrated inductances range from a few tenths of an nH to several nH. Apart from very simple structures, having closed form expressions for the inductance value is not possible. Numerous attempts have been made to calculate approximate closed form expressions for the spiral inductors with somewhat limited accuracy ([16], [17], [18]). Given their efficiency and precision, it is best to utilize common 3D electromagnetic simulators such as EMX or HFSS,[14] widely used among many RF designers.

1.9.2 Second-order effects

Apart from the ohmic loss, there are several other contributors that limit the inductor's performance. Shown in Figure 1.33, the metal strips have inevitably a non-zero capacitance to the substrate.

This capacitance limits the maximum allowable frequency that the inductor can be used at. This is typically expressed as the *self-resonance* frequency, that is the frequency that the total parasitic capacitance resonates with the inductance. To function properly, the self-resonance frequency must obviously be well above the maximum frequency at which the inductor is intended to be used. Moreover, this capacitance is connected to the lossy silicon substrate which could degrade the quality factor at higher frequencies. There is also a capacitance between various legs of the inductor branches. This capacitance may be ignored for a single-turn structure. However, for multi-turn inductors it would be relatively important as the adjacent legs are laid out very close to each other to maximize positive mutual inductance. This capacitance is substantially higher in *multi-stacked* structures (Figure 1.34) where several inductors of similar structure are placed in *series* using lower metal layers to increase the inductance without increasing the area. Since this structure leads to Q degradation as well, given the much larger sheet resistance of lower level metals, it is not very common unless very large values of inductance are needed.

[14] **HFSS** is a commercial finite element method solver for electromagnetic structures.

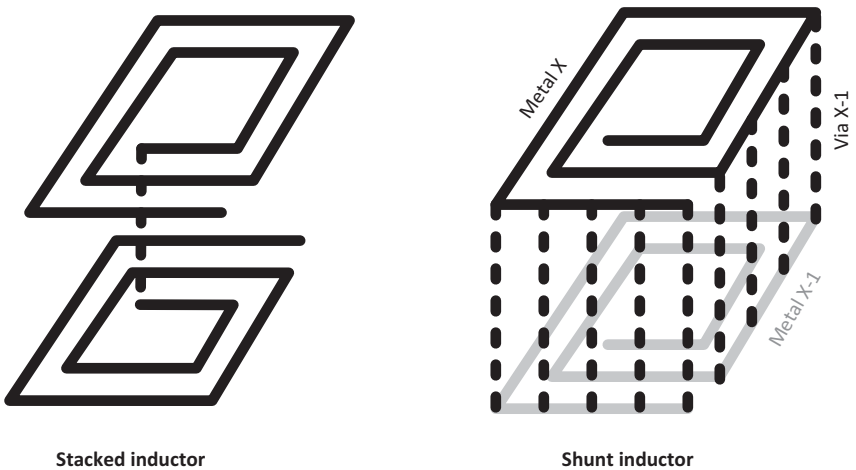


Figure 1.34 Stacked and shunt inductors

An alternative structure uses several inductors designed with lower metal layers connected to each other in *parallel* instead (Figure 1.34). While this will not lead to an increase in the inductance, it does improve the ohmic loss to some extent. That, however, comes at the expense of lower self-resonance frequency, as lower metals have higher capacitance to substrate. The Q improvement as a result of this may not be very significant for two reasons. Firstly, the lower metal levels typically have worse sheet resistance. Secondly, since the parasitic capacitance to substrate is increased, Q degradation due to capacitive coupling is worse. At lower frequencies, say 1 GHz or less, this structure may still be helpful. At higher frequencies, say 2 GHz or above, apart from the two reasons mentioned, the skin effect becomes an issue as well, and shunting the metal layers may not improve the Q at all.

As demonstrated in Figure 1.33, there is also a magnetic loss created at high frequencies. The magnetic field created due to the AC current flowing in the inductor branches results in a magnetic flux that is varying with time. Faraday's law suggests that an electric field E_{si} is induced in the substrate. This electric field leads to a current density flow of $J = \sigma_{si} E_{si}$ according to Ohm's law, where σ_{si} is the silicon substrate conductivity. That is as if there is a transformer reflecting the substrate resistance (or loss) in parallel with the inductor (Figure 1.33, right side). A higher substrate resistance (lower σ_{si}) is preferred to lower this loss. Fortuitously, most modern CMOS processes use a bulk substrate where the resistivity is relatively high. Unlike capacitive loss, where adding a metal shield could help, in the case of the magnetic loss, the shield is generally not helpful, as it effectively shorts out the inductor.[15] To illustrate this further, shown in Figure 1.35 is the simulated Q of a 1 nH inductor in a 28 nm CMOS with and without a native layer. The lack of a native layer results in an excess substrate implant to increase the threshold voltage of regular NMOS devices.[16] Consequently, the channel conductivity increases from 10 mS/m to about 50 mS/m. While the difference in Q at lower frequencies is not much,

[15] A *patterned* shield realized by a poly or metal one helps [20], but given the high resistivity of substrate it is not very common, as the Q improvement is marginal, but the self-resonance frequency degrades.

[16] Native NMOS transistors have a threshold voltage of close to zero.

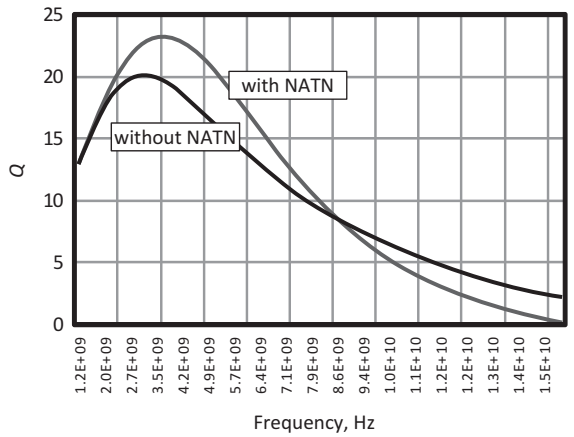


Figure 1.35 Simulated Q of a 1 nH inductor with and without a native layer

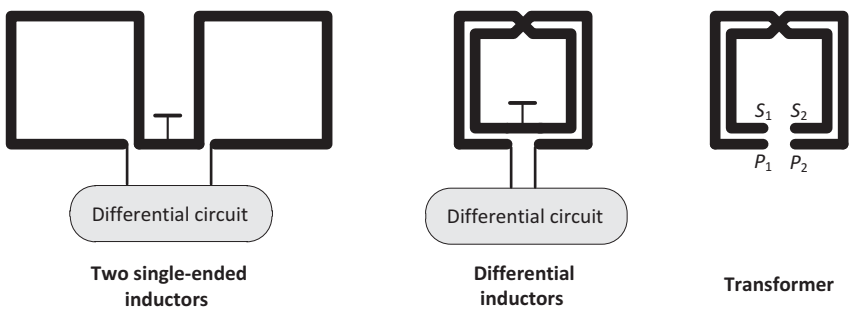


Figure 1.36 Differential inductors

the increase in substrate conductivity does result in degradation of the Q at higher frequencies. This layer thus may be exploited appropriately to dispense with the shield, and yet minimize the impact of substrate loss at high frequency.

1.9.3 Differential inductors and transformers

If two identical inductors are used in a differential circuit, they may be replaced by a *differential* inductor, where the two single inductors are combined (Figure 1.36) [19]. This naturally leads to a more compact design. Furthermore, a smaller area means less substrate loss and capacitance, which is important at high frequencies. This of course assumes that the size of the differential inductor to realize an inductance twice as big as each single-ended one remains the same. This is not exactly the case, and the area tends to grow a little, but still there is substantial saving.

Even though the capacitance to substrate is expected to decrease (ideally halve), the main drawback of the differential topology is lower self-resonance frequency. This is caused by significantly higher capacitance between the adjacent legs, compared to two single inductors spaced far apart, as shown in Figure 1.36. Another disadvantage of differential inductors is the fact that any unwanted coupling to the inductor (through parasitic capacitive and particularly magnetic sources) appears as an undesirable differential signal at the two terminals. For two

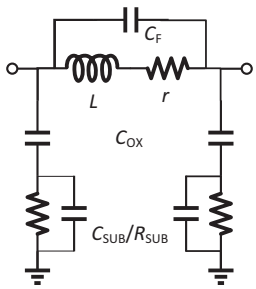


Figure 1.37 Inductor lumped model

single-ended inductors, however, if the parasitic source is far enough away, it appears as common-mode noise at the outputs.

Note that the differential inductor is in fact a transformer whose secondary ports are shorted together and connected to a common voltage, as shown in Figure 1.36 (far right). Therefore we face more or less similar tradeoffs when designing the transformers. Obviously it is key to lay out the primary and secondary lines as close as possible to maximize the coupling coefficient (K). A K factor of as high as 0.8 is achievable with a proper design.

We shall discuss the transformers from a circuit point of view more closely in Chapter 3.

1.9.4 Inductor lumped circuit model

Although distributed in nature due to their relatively large size, it is convenient to model inductors by simple lumped elements. The most common circuit is shown in Figure 1.37. It consists of the low frequency inductance (L), the series ohmic resistance (r), the oxide capacitance to substrates (C_{OX}), the substrate model (R_{SUB} and C_{SUB}), and C_F which models the capacitance between adjacent legs. The main advantage of the model is that all its elements are physical, while it produces a reasonable approximation valid over a wide range of frequency. It is therefore very common among RF designers to model the inductor as such. Since the substrate characteristics are not very well known, R_{SUB} and C_{SUB} are usually fitting parameters. The combination of C_{OX} and R_{SUB}/C_{SUB} is typically sufficient to account for both magnetic and capacitive loss of substrate. While the model could be fitted to represent the inductor at least at one exact frequency, and possibly over a reasonable range, if one is interested in a true wideband model, S-parameters generated by the simulator (EMX for example) may be used. To speed up the simulations, EMX may be used to produce a lumped element equivalent circuit comprising RLC element and dependent sources.[17]

Without loss of generality, let us take a closer look at the single-ended inductor equivalent circuit model where we assume one terminal is connected to an AC ground. The input impedance is

$$Z_{IN} = \frac{(r + jl\omega)\left(1 + \frac{C_{SUB}}{C_{OX}} + \frac{1}{jR_{SUB}C_{OX}\omega}\right)}{1 + \frac{C_{SUB}}{C_{OX}} + \frac{r}{R_{SUB}} + j\omega\left(\frac{L}{R_{SUB}} + rC_{SUB}\right) - LC_{si}\omega^2 + \frac{1}{jR_{SUB}C_{OX}\omega}}.$$

[17] The equivalent circuit is not physical, and is curve-fitted to merely replace S-parameters to improve convergence and simulation speed.

To simplify, we recognize that: $r \ll R_{SUB}$, $C_{SUB} \ll C_{OX}$, and at frequencies of interest and beyond: $\left|\frac{1}{jR_{SUB}C_{OX}\omega}\right| \ll 1$. The latter arises from the fact that the substrate resistivity is large, while at higher frequencies the impedance of C_{OX} is relatively small.

This leads to

$$Z_{IN} \cong \frac{r + jl\omega}{1 + j\omega\left(\frac{L}{R_{SUB}} + rC_{SUB}\right) - LC_{SUB}\omega^2}.$$

The impedance is bandpass, although at very low frequencies it does not approach zero. It rather becomes equal to the low frequency series resistance, r, which is expected. The self-resonance frequency is $\omega_{SRF} = 1/\sqrt{LC_{SUB}}$, where the magnitude of impedance peaks to a value of $\frac{L}{\frac{L}{R_{SUB}} + rC_{SUB}} \cong R_{SUB}$. We made the assumption that at higher frequencies $r \ll L\omega$, which is true if the Q is reasonably large.

The inductance is naturally frequency dependent and by definition is equal to $L(\omega) = \frac{\mathrm{Im}\{Z_{IN}\}}{\omega}$. This leads to

$$L(\omega) \cong L\frac{1 - LC_{SUB}\omega^2}{(1 - LC_{SUB}\omega^2)^2 + \left[\left(\frac{L}{R_{SUB}} + rC_{SUB}\right)\omega\right]^2}.$$

Evidently, the low frequency inductance is equal to L, while the inductance eventually approaches zero at the self-resonance frequency. This makes sense, as at the self-resonance frequency the input impedance has a phase of zero (and a peak magnitude of roughly R_{SUB}, as discussed before). Beyond the self-resonance frequency Z_{IN} is capacitive. Taking the derivative of $L(\omega)$ versus ω, we can show that the inductance is expected to peak just before approaching the self-resonance frequency. Defining a unit-less parameter $\eta = \frac{1}{2R_{SUB}}\sqrt{\frac{L}{C_{SUB}}}$, the frequency where the inductance peaks is roughly $(1 - \eta)\omega_{SRF}$, and the value of the inductance at that frequency is $\frac{L}{4\eta}$. For typical values, note that $\eta \ll 1$. The reason that the inductance peaks before self-resonance is due to the fact that the term $(1 - LC_{SUB}\omega^2)^2$ in the denominator approaches zero faster than the numerator due to the square function.

An example of a differential inductor designed in 28 nm CMOS is shown in Figure 1.38. The inductance is designed to be 1 nH at 4 GHz, where the EMX and lumped equivalent are plotted. The measured characteristics of the inductor are very close to those predicted by EMX. The lumped model values are: $L = 1$ nH, $r = 0.44\,\Omega$, $C_{OX} = 5.57$ pF, $C_{SUB} = 61$ fF, and $R_{SUB} = 872\,\Omega$. The inductor is a two-turn design using the top metal layer only. The total length is about 1.1 mm, and the width is 22 μm. The DC resistance is then calculated to be about 0.5 Ω. If it were designed as a long piece of wire, the DC inductance would be 1.1 nH.

The self-resonance frequency is about 20 GHz, very close to what is predicted by our analysis, where the inductance reaches zero. Moreover, the EMX and lumped models match fairly well for a wide range of frequencies. According to our derivations, $\eta = 0.074$, and we expect the inductance to peak at 3.4 nH, at 18.5 GHz.

Finding a closed expression for Q proves to be more difficult, and we will only present a qualitative description. At low frequencies, the second branch of the π-model is not effective, and thus: $Q = \frac{L\omega}{r}$, which rises linearly with frequency but flattens somewhat due to the skin

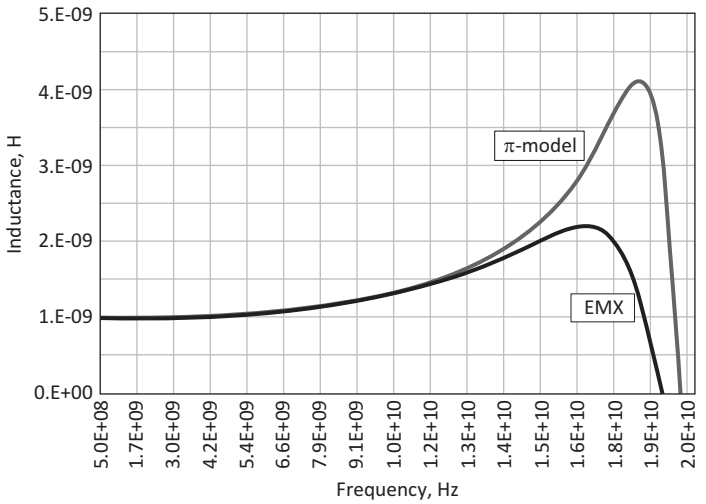


Figure 1.38 Simulated inductance for a 1 nH inductor

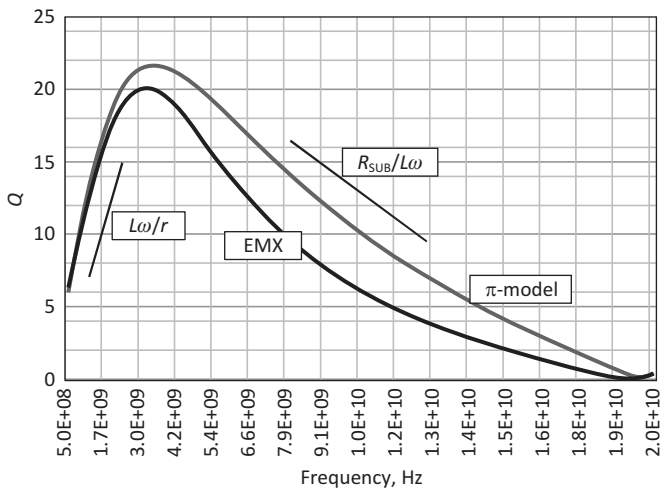


Figure 1.39 Simulated Q of the 1 nH inductor

effect (although this is not captured by the simple π-model). On the other hand, at higher frequencies, C_{OX} becomes a short circuit, and the model simplifies to a parallel RLC equivalent circuit, consisting of L, R_{SUB}, and C_{SUB}. Thus: $Q = \frac{R_{SUB}}{L\omega}$, which falls linearly with frequency. The Q approaches zero at the self-resonance frequency, as Z_{IN} is purely real. Moreover, the high and low frequency quality factors become equal to each other at a frequency of approximately $\frac{1}{L}\sqrt{r R_{SUB}}$, where we expect the Q to peak. This suggests that balancing the low and high frequency losses optimizes the inductor quality factor at a given frequency.

The simulated Q for the same inductor is shown in Figure 1.39. The quality factor due to only series resistance is 57 at 4 GHz, whereas the substrate resistance loss yields a Q of 35. Thus the combined quality factor at 4 GHz sums to 21.7, which is slightly overestimating EMX results. The Q is expected to peak at 3.2 GHz, which is close to EMX simulation results.