

SERIES IN OPTICS AND OPTOELECTRONICS

# Laser-Based Measurements for Time and Frequency Domain Applications

*A Handbook*



**Pasquale Maddaloni • Marco Bellini**  
**Paolo De Natale**



**CRC Press**  
Taylor & Francis Group

A TAYLOR & FRANCIS BOOK

Laser-Based Measurements  
for Time and Frequency  
Domain Applications

---

*A Handbook*

# **SERIES IN OPTICS AND OPTOELECTRONICS**

---

Series Editors: **E. Roy Pike**, Kings College, London, UK

**Robert G. W. Brown**, University of California, Irvine, USA

*Recent titles in the series*

**Handbook of 3D Machine Vision: Optical Metrology and Imaging**

Song Zhang

**Handbook of Optical Dimensional Metrology**

Kevin Harding

**Laser-Based Measurements for Time and Frequency Domain Applications: A Handbook**

Pasquale Maddaloni, Marco Bellini, and Paolo De Natale

**Handbook of Silicon Photonics**

Laurent Vivien and Lorenzo Pavesi

**Biomimetics in Photonics**

Olaf Karthaus

**Optical Properties of Photonic Structures: Interplay of Order and Disorder**

Mikhail F Limonov and Richard De La Rue (Eds.)

**Nitride Phosphors and Solid-State Lighting**

Rong-Jun Xie, Yuan Qiang Li, Naoto Hirosaki, and Hajime Yamamoto

**Molded Optics: Design and Manufacture**

Michael Schaub, Jim Schwiegerling, Eric Fest, R Hamilton Shepard, and Alan Symmons

**An Introduction to Quantum Optics: Photon and Biphoton Physics**

Yanhua Shih

**Principles of Adaptive Optics, Third Edition**

Robert Tyson

**Optical Tweezers: Methods and Applications**

Miles J Padgett, Justin Molloy, and David McGloin (Eds.)

**Thin-Film Optical Filters, Fourth Edition**

H Angus Macleod

**Laser Induced Damage of Optical Materials**

R M Wood

**Principles of Nanophotonics**

Motoichi Ohtsu, Kiyoshi Kobayashi, Tadashi Kawazoe, Tadashi Yatsui, and Makoto Naruse

# Laser-Based Measurements for Time and Frequency Domain Applications

---

*A Handbook*

Pasquale Maddaloni • Marco Bellini  
Paolo De Natale



CRC Press

Taylor & Francis Group

Boca Raton London New York

---

CRC Press is an imprint of the  
Taylor & Francis Group, an **informa** business

A TAYLOR & FRANCIS BOOK

Cover by Rita Cuciniello: "Generation of a mid-infrared optical frequency comb, P. Maddaloni et al., *New Journal of Physics* 8 (2006) 262," in the artist's concept.

Original: Oil on canvas.

Taylor & Francis  
Taylor & Francis Group  
6000 Broken Sound Parkway NW, Suite 300  
Boca Raton, FL 33487-2742

© 2013 by Taylor & Francis Group, LLC  
Taylor & Francis is an Informa business

No claim to original U.S. Government works  
Version Date: 20130227

International Standard Book Number-13: 978-1-4398-4153-2 (eBook - PDF)

This book contains information obtained from authentic and highly regarded sources. Reasonable efforts have been made to publish reliable data and information, but the author and publisher cannot assume responsibility for the validity of all materials or the consequences of their use. The authors and publishers have attempted to trace the copyright holders of all material reproduced in this publication and apologize to copyright holders if permission to publish in this form has not been obtained. If any copyright material has not been acknowledged please write and let us know so we may rectify in any future reprint.

Except as permitted under U.S. Copyright Law, no part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, please access [www.copyright.com](http://www.copyright.com) (<http://www.copyright.com/>) or contact the Copyright Clearance Center, Inc. (CCC), 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. CCC is a not-for-profit organization that provides licenses and registration for a variety of users. For organizations that have been granted a photocopy license by the CCC, a separate system of payment has been arranged.

**Trademark Notice:** Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

**Visit the Taylor & Francis Web site at**  
**<http://www.taylorandfrancis.com>**

**and the CRC Press Web site at**  
**<http://www.crcpress.com>**

To the young people I wish the same luck  
that led me to disinterest myself from my  
person, but to always pay attention to  
everything that surrounds me, to  
everything in the scientific world,  
without neglecting the values of society.

**Rita Levi Montalcini**

to our children:  
Maria and Bernadette,  
Alessia and Andrea,  
Rufael and Tsion.



---

# Contents

---

<b>Foreword</b>	<b>xv</b>
<b>Preface</b>	<b>xvii</b>
<b>Authors</b>	<b>xix</b>
<b>1 Shedding light on the art of timekeeping</b>	<b>1</b>
1.1 The great show of Time and Light, the curtain rises! . . . . .	1
1.2 Brief history of timekeeping: time-frequency equivalence . . . . .	2
1.3 The parallel story of the speed of light . . . . .	6
1.3.1 The laser arrives: length-frequency equivalence and the birth of optical frequency metrology . . . . .	14
1.3.2 Role of $c$ in fundamental physics . . . . .	17
1.4 In the end, time and light met up again: optical atomic clocks and outline of the book . . . . .	20
<b>2 Characterization and control of harmonic oscillators</b>	<b>25</b>
2.1 The ideal harmonic oscillator . . . . .	25
2.1.1 Synchronization in coupled oscillators . . . . .	32
2.1.2 Beating two oscillators . . . . .	35
2.2 Self-sustained oscillators . . . . .	36
2.3 The noisy oscillator . . . . .	41
2.4 Phase noise . . . . .	41
2.4.1 Review of mathematical tools . . . . .	42
2.4.2 Fundamental noise mechanisms . . . . .	44
2.4.2.1 Fluctuation-dissipation theorem (FDT) . . . . .	48
2.5 Phase noise modelling . . . . .	51
2.5.1 The Leeson effect . . . . .	54
2.5.2 The Allan variance . . . . .	57
2.5.3 Oscillator power spectrum in the carrier frequency domain . . . . .	62
2.5.3.1 Effect of frequency multiplication on the power spectrum . . . . .	65
2.6 Noise reduction in oscillators . . . . .	66
2.6.1 Phase-locked loops . . . . .	66
2.6.1.1 Optical phase-locked loops . . . . .	74
2.6.2 Injection locking . . . . .	76
2.7 Phase noise measurements . . . . .	80
2.7.1 Frequency counting . . . . .	80
2.7.2 Homodyne techniques . . . . .	81
2.7.3 Heterodyne techniques . . . . .	82
2.7.4 Self-heterodyning . . . . .	83

2.8	Amplitude noise measurements . . . . .	84
2.8.1	AM noise in optical systems . . . . .	85
<b>3</b>	<b>Passive resonators</b>	<b>87</b>
3.1	Microwave cavities . . . . .	88
3.2	Basic properties of bulk optical cavities . . . . .	94
3.2.1	Fabry-Perot etalon . . . . .	95
3.2.2	Paraxial ray analysis . . . . .	101
3.2.3	Wave analysis . . . . .	102
3.3	Cavity-design considerations . . . . .	106
3.3.1	Quarter wave stack reflectors . . . . .	107
3.3.2	Prism-based cavities . . . . .	111
3.4	Ultrastable cavities . . . . .	117
3.4.1	Thermal stability . . . . .	117
3.4.2	Vibration insensitive optical cavities . . . . .	123
3.5	Fiber cavities . . . . .	128
3.5.1	Directional coupler-based fiber cavities . . . . .	129
3.5.2	Resonators based on closely faced fiber tips . . . . .	131
3.5.3	FBG-based fiber resonators . . . . .	132
3.6	Whispering gallery mode resonators . . . . .	137
3.6.1	Wave theory of whispering gallery modes . . . . .	137
3.6.2	WGMs in a ray-optical picture . . . . .	150
3.6.3	Mode $Q$ and volume . . . . .	152
3.6.4	WGM evanescent coupling . . . . .	154
3.6.5	Fabrication and applications of whispering gallery resonators . . . . .	159
<b>4</b>	<b>Continuous-wave coherent radiation sources</b>	<b>163</b>
4.1	Principles of masers . . . . .	164
4.1.1	The hydrogen maser . . . . .	171
4.2	Compendium of laser theory . . . . .	176
4.2.1	The active medium . . . . .	177
4.2.1.1	Einstein A and B coefficients for absorption and emission . . . . .	177
4.2.1.2	Line-broadening mechanisms . . . . .	180
4.2.2	The pump . . . . .	183
4.2.3	The resonator . . . . .	186
4.2.4	Rate equations for a four-level system . . . . .	188
4.2.4.1	Transient behavior and relaxation oscillation . . . . .	191
4.3	Frequency pulling . . . . .	193
4.4	Achieving single-mode oscillation . . . . .	194
4.4.1	Line selection . . . . .	197
4.4.2	Single-transverse mode selection . . . . .	198
4.4.3	Single-longitudinal mode selection . . . . .	198
4.5	The laser output . . . . .	202
4.5.1	Spatial coherence . . . . .	202
4.5.2	Spectral and temporal coherence . . . . .	203
4.5.3	The effect of spontaneous emission . . . . .	205
4.5.3.1	Intrinsic laser linewidth . . . . .	206
4.5.3.2	Amplified spontaneous emission . . . . .	211
4.6	Laser frequency fluctuations and stabilization techniques . . . . .	213

4.6.1	Side-lock to an atomic/molecular resonance . . . . .	219
4.6.2	Pound-Drever-Hall method . . . . .	220
4.6.2.1	$10^{-16}$ -level laser frequency stabilization . . . . .	225
4.6.2.2	WGM resonators for laser frequency stabilization . . . . .	228
4.6.3	Hänsch-Couillaud technique . . . . .	231
4.6.4	Laser frequency stabilization by locking to an optical fiber-delay line . . . . .	233
4.6.5	Injection locking . . . . .	237
4.7	Intensity fluctuations . . . . .	240
4.7.1	High-sensitivity photodiode array . . . . .	243
4.7.2	Optical ac coupling . . . . .	244
4.7.3	The laser as quasi-ideal oscillator . . . . .	246
4.8	Some specific laser systems . . . . .	249
4.8.1	He-Ne laser . . . . .	250
4.8.2	Carbon dioxide laser . . . . .	251
4.8.3	Dye lasers . . . . .	253
4.8.4	Ion-doped lasers and optical amplifiers . . . . .	257
4.8.4.1	Nd:YAG laser . . . . .	257
4.8.4.2	Ti:Sa laser . . . . .	258
4.8.4.3	Erbium-doped fiber amplifiers (EDFAs) . . . . .	259
4.8.4.4	Ytterbium-doped fiber amplifiers (YDFAs) . . . . .	260
4.8.4.5	Narrow-linewidth fiber lasers . . . . .	263
4.8.5	Semiconductor lasers . . . . .	268
4.8.5.1	Heterostructure diode lasers . . . . .	269
4.8.5.2	Distributed feedback (DFB) lasers . . . . .	285
4.8.5.3	Tapered semiconductor amplifiers . . . . .	288
4.8.5.4	Multiple-quantum-well lasers . . . . .	289
4.8.5.5	Vertical-cavity surface-emitting lasers . . . . .	292
4.8.5.6	Quantum cascade lasers . . . . .	292
4.8.5.7	Interband cascade lasers . . . . .	300
4.8.6	Nonlinear laser cw sources . . . . .	302
4.8.6.1	Sum frequency generation . . . . .	305
4.8.6.2	Optical parametric oscillators . . . . .	306
4.8.6.3	Difference frequency generators . . . . .	312
4.8.6.4	Tunable far-infrared radiation . . . . .	317
<b>5</b>	<b>High-resolution spectroscopic frequency measurements</b> . . . . .	<b>323</b>
5.1	Interferometric wavelength measurements . . . . .	323
5.2	Spectroscopic frequency measurements . . . . .	328
5.2.1	Principles of absorption laser spectroscopy . . . . .	329
5.3	Frequency modulation spectroscopy . . . . .	329
5.3.1	Harmonic detection . . . . .	330
5.3.2	Wavelength modulation spectroscopy . . . . .	331
5.3.3	Single- and two-tone frequency modulation spectroscopy . . . . .	331
5.4	Magnetic rotation spectroscopy . . . . .	337
5.5	Cavity-enhanced spectroscopy . . . . .	341
5.5.1	Cavity-enhanced absorption spectroscopy . . . . .	342
5.5.2	Off-axis integrated cavity output spectroscopy . . . . .	343
5.5.3	Noise immune cavity-enhanced optical heterodyne molecular spectroscopy . . . . .	346
5.5.4	Cavity ring-down spectroscopy . . . . .	348

5.5.4.1	Phase-shift (PS) CRDS . . . . .	350
5.5.4.2	Saturated-absorption cavity ring-down spectroscopy . . . . .	351
5.6	Doppler-free saturation spectroscopy . . . . .	354
5.6.1	Frequency locking to a Lamb dip . . . . .	358
5.6.2	Cavity-enhanced Doppler-free saturation spectroscopy . . . . .	361
5.6.2.1	Doppler-free NICE-OHMS . . . . .	363
5.7	Doppler-free polarization spectroscopy . . . . .	364
5.8	Doppler-free two-photon spectroscopy . . . . .	368
5.9	Second-order Doppler-free spectroscopy . . . . .	372
5.10	Sub-Doppler spectroscopy in atomic/molecular beams . . . . .	374
5.10.1	Effusive beams . . . . .	374
5.10.2	Supersonic beams . . . . .	377
5.10.3	Buffer-gas-cooling . . . . .	379
5.11	Ramsey fringes . . . . .	381
5.12	Laser frequency standards using thermal quantum absorbers . . . . .	387
5.12.1	Iodine-stabilized lasers . . . . .	388
5.12.2	Acetylene-stabilized lasers . . . . .	388
5.12.3	Methane-stabilized lasers . . . . .	389
5.12.4	OsO <sub>4</sub> -stabilized lasers . . . . .	389
5.12.5	Atomic hydrogen standard . . . . .	390
5.12.6	Calcium standard . . . . .	393
5.13	Fourier transform spectroscopy . . . . .	394
5.14	Raman spectroscopy . . . . .	398
5.14.1	Coherent anti-Stokes spectroscopy . . . . .	405
5.14.2	Stimulated Raman scattering . . . . .	410
<b>6</b>	<b>Time and frequency measurements with pulsed laser systems</b>	<b>413</b>
6.1	Introduction . . . . .	413
6.2	Theory of mode locking . . . . .	415
6.3	Mode-locking mechanisms and dispersion compensation schemes . . . . .	417
6.3.1	Ti:sapphire lasers and Kerr-lens mode locking . . . . .	417
6.3.2	Fiber-based lasers and nonlinear-polarization-rotation mode-locking . . . . .	419
6.4	Optical frequency comb synthesis from mode-locked lasers . . . . .	421
6.4.1	Comb stabilization . . . . .	421
6.4.2	Measurements with a frequency comb . . . . .	425
6.4.2.1	Measuring frequency differences . . . . .	425
6.4.2.2	Measuring absolute frequencies with an octave-spanning comb . . . . .	426
6.4.2.3	Absolute optical frequency synthesizer . . . . .	427
6.4.2.4	Direct frequency-comb spectroscopy . . . . .	429
6.4.2.5	Other measurement schemes and applications . . . . .	435
6.4.3	Relevant properties of a mode-locked laser for frequency-comb applications . . . . .	440
6.4.4	Microresonator-based frequency combs . . . . .	441
6.5	Extension of OFCSs into novel spectral regions . . . . .	444
6.5.1	High-order laser harmonics and extensions to the XUV . . . . .	444
6.5.1.1	Generation and properties of high-order laser harmonics . . . . .	446
6.5.1.2	Ramsey-type spectroscopy in the XUV with high-order harmonics . . . . .	448
6.5.1.3	Cavity-enhanced XUV frequency combs . . . . .	456

6.5.1.4	Attosecond pulses . . . . .	459
6.5.2	Mid- and far-infrared OFCSs . . . . .	460
<b>7</b>	<b>Frequency standards</b>	<b>465</b>
7.1	General features of frequency standards and clocks . . . . .	465
7.2	Quartz oscillators . . . . .	466
7.2.1	Factors affecting crystal oscillator frequency accuracy . . . . .	471
7.2.2	Factors affecting frequency stability . . . . .	473
7.2.3	State-of-the-art ultrastable quartz oscillators . . . . .	473
7.3	Cryogenic sapphire oscillators . . . . .	473
7.4	Photonic microwave oscillators based on WGM resonators . . . . .	479
7.5	Generation of ultrastable microwaves via optical frequency division . . . . .	480
7.6	Trapping and cooling of neutral atoms . . . . .	482
7.6.1	Optical molasses . . . . .	483
7.6.2	Magneto-optical traps . . . . .	489
7.6.3	Bose-Einstein condensation . . . . .	492
7.6.3.1	Magnetic trapping . . . . .	492
7.6.3.2	Evaporative cooling . . . . .	493
7.6.3.3	Probing a BEC . . . . .	495
7.7	Cold stable molecules . . . . .	498
7.7.1	Stark decelerator . . . . .	498
7.8	Trapping and cooling of ions . . . . .	500
7.8.1	Paul traps . . . . .	500
7.8.1.1	Linear Paul traps . . . . .	504
7.8.2	Penning traps . . . . .	505
7.8.3	Trap loading . . . . .	507
7.8.4	Ion cooling techniques . . . . .	507
7.8.4.1	Laser cooling . . . . .	508
7.8.4.2	Sympathetic cooling . . . . .	511
7.8.5	Spectroscopy of trapped particles in the Lamb-Dicke regime . . . . .	511
7.9	Microwave atomic standards . . . . .	513
7.9.1	Metrological properties of the active hydrogen maser . . . . .	513
7.9.1.1	Maser design . . . . .	513
7.9.1.2	Frequency shifts . . . . .	513
7.9.1.3	Automatic tuning of the resonant cavity . . . . .	515
7.9.1.4	Frequency stability . . . . .	516
7.9.1.5	Cryogenic hydrogen masers . . . . .	517
7.9.2	Cesium clocks . . . . .	518
7.9.2.1	Cesium-beam frequency standards . . . . .	518
7.9.2.2	Cesium fountain clocks . . . . .	521
7.9.2.3	Uncertainty budget in a Cs fountain clock . . . . .	527
7.9.3	Rubidium clocks . . . . .	532
7.9.3.1	Rb fountain clocks . . . . .	533
7.9.3.2	Lamp-based Rb cell standards . . . . .	534
7.9.3.3	Laser-based Rb cell frequency standards . . . . .	535
7.9.4	Microwave ion clocks . . . . .	539
7.10	Time transfer and frequency dissemination . . . . .	541
7.10.1	Realization of time scales . . . . .	542
7.10.1.1	Realization of TAI . . . . .	542

7.10.1.2	Coordinated universal time . . . . .	543
7.10.2	Transmitting time information . . . . .	543
7.10.2.1	Portable clocks . . . . .	546
7.10.2.2	Global positioning system . . . . .	547
7.10.3	Frequency transfer . . . . .	550
7.10.3.1	A democratic absolute frequency chain . . . . .	550
7.10.3.2	Dissemination of microwave frequency standards . . . . .	551
7.10.3.3	Optical frequency transfer . . . . .	552
<b>8</b>	<b>Future trends in fundamental physics and applications</b>	<b>557</b>
8.1	Optical atomic clocks . . . . .	557
8.1.1	Trapped ion optical clocks . . . . .	559
8.1.1.1	Systematic frequency shifts . . . . .	562
8.1.2	Neutral atoms optical lattice clocks . . . . .	564
8.2	The hydrogen atom as an inexhaustible wellspring of advances in precision spectroscopy . . . . .	570
8.2.1	Determination of the Rydberg constant and of the proton radius . . . . .	572
8.3	Spectroscopy of cold, trapped metastable helium . . . . .	575
8.4	Measurements of fundamental constants . . . . .	578
8.4.1	Boltzmann constant $k_B$ . . . . .	578
8.4.2	Newton gravitational constant $G$ . . . . .	580
8.5	Constancy of fundamental constants . . . . .	582
8.5.1	Fine structure constant $\alpha$ . . . . .	584
8.5.1.1	Proton-to-electron mass ratio $\beta$ . . . . .	586
8.5.2	Speed of light $c$ . . . . .	588
8.5.2.1	Frequency dependence of $c$ and the mass of the photon . . . . .	588
8.5.3	Newton's gravitational constant . . . . .	590
8.6	Tests of fundamental physics laws . . . . .	590
8.6.1	Spectroscopic tests of spin-statistic connection and symmetrization postulate . . . . .	590
8.6.2	Search for the electron dipole moment . . . . .	594
8.6.3	Parity violation in chiral molecules . . . . .	595
8.7	Perspectives for precision spectroscopy of cold molecules . . . . .	600
8.8	Tests of general relativity: from ground-based experiments to space missions	603
8.8.1	Testing the Einstein Equivalence Principle . . . . .	604
8.8.1.1	Tests of UFF . . . . .	604
8.8.1.2	Tests of LLI . . . . .	605
8.8.1.3	Tests of LPI . . . . .	608
8.8.2	Test of post-Newtonian gravity . . . . .	610
8.8.3	Tests of the gravitational inverse square law . . . . .	611
8.8.3.1	Detection of gravitational waves . . . . .	611
8.9	Quantum-enhanced time and frequency measurements . . . . .	613
8.9.1	Standard quantum limit in physical measurements . . . . .	613
8.9.2	Using nonclassical light states for quantum-enhanced measurements . . . . .	615
8.9.3	Applications to time and frequency measurements . . . . .	620
8.9.3.1	Quantum logic spectroscopy . . . . .	620
8.9.3.2	Time and frequency quantum metrology . . . . .	621
8.9.3.3	Quantum positioning and clock synchronization . . . . .	623
8.10	Environmental metrology . . . . .	625

<i>Contents</i>	xiii
8.10.1 Geophysical survey of volcanic areas . . . . .	625
8.10.2 Detection of very rare isotopes . . . . .	628
8.10.3 Stratospheric survey with tunable diode laser spectrometers . . . . .	628
8.10.4 Fiber sensing of physical and chemical parameters . . . . .	630
8.10.4.1 Strain sensing . . . . .	632
8.10.4.2 Acceleration measurements . . . . .	633
8.10.4.3 Chemical sensing by optical-fiber ring resonators . . . . .	634
<b>Bibliography</b>	<b>637</b>
<b>Index</b>	<b>701</b>



---

## *Foreword*

---

Since its first realization in 1960, the laser has quickly parted from its initial definition of “a solution looking for a problem” to become the solution to many and incredibly different problems, both in our everyday lives and in the most advanced fields of science and technology.

Thanks to its unique properties, i.e., the capability to generate intense, highly directional, and highly monochromatic radiation in different regions of the electromagnetic spectrum, the laser has found applications in fields as diverse as communications, medicine, fundamental physics, as well as supermarket counters. . . .

In particular, thanks to the possibility of producing either ultrashort laser pulses, lasting just a few femtoseconds, or continuous beams with an extremely well-controlled frequency, the two fields of accurate time and frequency measurements have literally boomed in the last decades. Measuring times and delays with femtosecond accuracy has allowed us, for example, to observe the real-time movement of atoms in molecules, and to follow and steer chemical reactions. On the other hand, measuring frequencies with a very high precision has given us unprecedented access to the most intimate structure of matter, is revolutionizing precision metrology of time and space, and is providing new tools for many important applied fields, like environmental monitoring.

These two intertwined subjects, time and frequency, have continued to evolve independently for many years, with the impressive parallel evolution of ultrafast and ultra-stable laser sources, until it was recently realized that they can be seen as two different faces of the same medal. In fact, the frequency spectrum of the train of ultrashort pulses emitted by a (properly phase-stabilized) mode-locked laser is remarkably simple. It is made of millions of extremely narrow spectral lines whose frequencies are exactly spaced by the laser repetition rate. Such a special laser source can thus combine the best of the two worlds: on one side, by giving access to a huge number of ultra-stable laser lines, all with precisely controlled frequencies, it serves as a perfect ruler in the spectral domain; on the other side, by making it possible to control the absolute optical phase of ultrashort light pulses, it discloses new, highly nonlinear phenomena to experimental investigation.

This experimentalist’s dream came true with the development of frequency combs around the beginning of this century, and it was a strike of serendipity on the hill of Arcetri in Florence, where the LENS and INO laboratories used to be, that started it. There, on a lucky afternoon in 1997, Marco Bellini and I could surprisingly observe stable interference fringes from the white-light supercontinua independently produced by two identical ultrashort pulses. Since that moment, the evolution of comb-based measurements has seen no sign of slowing down, and has made possible some of the most accurate measurements ever performed by mankind, allowing to measure frequencies in the simplest atom in the universe, hydrogen, as well as to calibrate spectra coming from the borders of the universe, itself.

Indeed, if a stable and precisely determined frequency reference is available, it can be used as a clock to measure time intervals with high accuracy by just counting the number of cycles in the interval. Furthermore, from the definition of the speed of light, any distance measurement can be referred to a time or frequency measurement. Since the measure of

many physical properties, beyond time and distance, can be often converted to a phase or frequency measurement, the incredible precision made possible by frequency combs can be readily exported to a number of different fields.

Laser-based precision measurements are now facing a new era, where the ever-growing accuracy in the determination of times and frequencies will not only allow us to unveil some of the best hidden secrets of nature, but also impact our everyday lives.

In this context, the publication of this book is particularly timely and welcome. The authors have made a considerable effort to make this book useful and interesting to different kinds of readers: they provide a detailed treatment of the basic concepts of time and frequency measurements, carefully describe different kinds of lasers and some of the most advanced laser-based measurement techniques, and, finally, present the latest developments in the field, with a hint to the possible future trends in applications and fundamental science.

Being among the many important actors in this long story, the authors of this book are privileged witnesses of the evolution of time and frequency measurements, and can provide an informed and wide vision of this developing field from many different viewpoints.

Theodor W. Hänsch  
Co-Recipient of the 2005 Nobel Prize in Physics

---

## *Preface*

---

Time and light or, in other words, frequency and photons: two key ingredients that are making possible the most accurate and sensitive measurements ever performed by mankind. This was the flywheel to writing this book and the “leitmotiv” throughout it. However, dealing with this subject is an impervious task both in a historical perspective and in the viewpoint of contemporary developments. Indeed, on the one hand, the art of timekeeping has deep roots, dating back to the origins of civilization, and light from the sky was soon recognized as a vital element for measuring time. On the other hand, trying to describe the furious activity and progress that have been characterizing this field for the last decades is a bit like taking a picture of a very fast object that keeps moving at increasingly high speed. In such a fascinating adventure, the laser-era undoubtedly represented a turning point, marking the birth of optical frequency metrology. In this wake, a new metrological tool emerged around 1999, when it was realized that the pattern of equally spaced pulses generated by a mode-locked laser in the time domain is equivalent to a precisely spaced comb of frequencies in the frequency domain, and that the phase of light is the same throughout the broad-covered spectrum. This eventually merged decades of independent technologies, namely those of ultrastable and ultrafast laser sources, emitting, respectively, continuous-wave radiation and trains of very short pulses. In parallel, another real breakthrough was represented by the advent of laser-based optical clocks that, by progressively reducing the fluctuations in the emitted frequency, have now reached impressive stability and accuracy levels. But many other milestones have been achieved in the field of laser-based measurements, since the introduction of the laser itself. Among the most remarkable, we should mention guiding and delivering laser-light with optical fibers as well as accessing additional portions of the electromagnetic spectrum with frequency-tunable coherent sources, based on novel materials and operation principles, or nonlinear optical phenomena. Apart from these spectacular technological achievements, the quantum nature of light and matter has opened other new scenarios, like that of measurements based on entanglement of photons and macroscopic objects. As a whole, such a scientific fervor has revolutionized the branch of atomic, molecular, and optical physics allowing, as the first immediate consequence, to devise increasingly ambitious experiments of fundamental character, but also engendering tremendous progress in terms of high-tech, everyday-life applications, perhaps going so far as to change even the way we think. This book is based on first-hand, laser-based measurements and direct experimental work performed by the authors during the last 25 years. Such activities are strictly related to the rise and development of the European Laboratory for Nonlinear Spectroscopy-LENS in Florence and the Istituto Nazionale di Ottica-INO (now part of the Italian National Research Council-CNR) in Florence and Naples, Italy. These labs and activities have flourished on the hill of Arcetri (Florence), where Galileo Galilei spent the last part of his life and Enrico Fermi conceived quantum statistics, but also in front of the breathtaking gulf of Naples, where, in his quick passage, Ettore Majorana left an indelible legacy. In fact, also thanks to their irresistible charm, these two cities have always attracted scientists from all over the world, so years of collaborations, discussions, and joint work are somehow reflected in the text of this book. The purpose is to offer a detailed account of the most recent results obtained for time- and frequency-domain applications

of lasers, while providing all the background information on the main kinds of sources and techniques developed thus far. Moreover, the theoretical framework necessary to understand the experimental applications is fully developed throughout the book. Therefore, most of the matter is intended to be accessible to final-year undergraduates, but also post-docs and scientists actively working in the field can find a wide, fresh, and balanced overview of conquests in the field of laser-based measurements together with the main related references. A detailed outline is given at the end of Chapter 1, as a natural outcome of the historical introduction. Also preliminary in character, Chapter 2 provides the basic concepts and the mathematical tools that are necessary to address the physics of oscillators, at the heart of the whole treatment. Likewise relevant to the self-sufficiency of the book, microwave, and particularly, optical resonators are extensively discussed in Chapter 3. Crucial aspects of operation and fundamental properties of lasers are presented in Chapter 4, while precision spectroscopy and absolute frequency metrology are dealt with in Chapters 5 and 6, respectively. Then, Chapter 7 is devoted to microwave and optical frequency standards and their dissemination. Finally, Chapter 8 dwells upon the variegated speculative landscape opened by the field of laser-based frequency measurements, outlining the most exciting, current, and forthcoming research directions. Due to the large amount of unfolded work, we apologize in advance for any mistakes, inaccuracies, and inevitable limitations, hoping that the reader may appreciate our approach and share our enthusiasm.

The authors wish to thank Luca Lorini for careful reading of Chapter 2; Simone Borri and Gianluca Gagliardi for their contribution to two sections in Chapter 4 and Chapter 8, respectively; Maurizio De Rosa for stimulating discussions; Gianluca Notariale for preparing many of the figures; Elisabetta Baldanzi for editing part of the text and caring about permission requests; and Rita Cuciniello for creating the cover. Also, the authors are immensely grateful to Prof. Theodor W. Hänsch for writing the Foreword.

Pasquale Maddaloni  
Marco Bellini  
Paolo De Natale

---

## *Authors*

---

**Pasquale Maddaloni** was born in Castellammare di Stabia, Italy, on February 16, 1975. He owes much of his professional accomplishment to his parents, Catello and Maria, for their constant and amorous support. As a completion of his studies, Pasquale received the diploma (*cum laude*) and the Ph.D. degree in Physics from the University of Naples (1999) and the University of Padua (2003), respectively, for experimental research in the field of ultracold atoms and Bose-Einstein Condensation. In 2003 he joined the National Institute for Optics-INO (now part of the National Research Council-CNR) where he worked, in a postdoctorate position, on the development of innovative mid-infrared coherent radiation sources. Since 2009 he has been a research scientist at INO in Naples, much of his interest relating to nonlinear optics and precision spectroscopy assisted by optical frequency comb synthesizers (OFCSs). In this framework, his main achievement is represented by the first fiber-based mid-IR generation of an OFCS and its application to absolute frequency metrology. More recently, he extended his experimental activities to the world of cold stable molecules. When not working, Pasquale shares his passion for gardening, traveling, and photography with his beloved wife, Rosa, and daughters, Maria and Bernadette.

**Marco Bellini** graduated (*cum laude*) and got his Ph.D. degree in Physics at the University of Florence, Italy. He later joined the European Laboratory for Nonlinear Spectroscopy (LENS) and, since 1999, he has been a researcher at the Istituto Nazionale di Ottica (INO) CNR. Since the early times of his scientific career, he has been dealing with light possessing extreme features. His research tools range from single photons to ultraintense laser pulses; from ultrabroadband supercontinuum to highly monochromatic THz sources, and to radiation in the extreme ultraviolet end of the electromagnetic spectrum. On one side, he used ultrashort and ultraintense laser pulses to produce highly nonlinear interactions with matter, leading to fundamental studies on supercontinuum sources (thus contributing to the development of the femtosecond frequency comb), and to the production and applications of high-order laser harmonics. On the other side, the ability to arbitrarily manipulate and characterize light at the single-photon level has allowed him to perform experiments on the foundations of quantum mechanics, and to develop new tools for the emerging quantum technologies. On all these topics he maintains active collaborations with several leading scientists around the world. He is also involved in several national and international research projects and collaborates, as a reviewer and as an editor, with top international scientific journals. When not working, he likes traveling and exploring Nature in all its forms. This book (and most of the rest) would not have been possible without the constant love and support of his parents, Silvana and Antonio, and his wife, Francesca.

**Paolo De Natale** graduated in Physics (*cum laude*) at Federico II University, Naples, Italy. He joined the European Laboratory for Nonlinear Spectroscopy (LENS) in 1988 and, since 1996, he has been staff scientist at Istituto Nazionale di Ottica (INO)-CNR (formerly INOA) which, at the time of writing, he directs. His research activity has always been focused on atomic, molecular, and optical physics. Pioneering results included tests of theories, measurements of fundamental constants, using state-of-the-art spectroscopic techniques and

originally developed coherent sources, often harnessing nonlinear optical phenomena. He has often carried on research in quite unexplored spectral regions, like infrared and far-infrared (THz), for the lack of suitable sources. Another part of PDN research has been devoted to the study and design of novel optoelectronic devices, especially based on ferroelectric crystals. To trigger such multidisciplinary research, PDN summoned a group combining unusual skills in spectroscopy, interferometric diagnostics, mastering of nonlinear sources and techniques, that enabled quick achievement of unique results. Such activity is still ongoing in the Napoli section of the Istituto Nazionale di Ottica-CNR (formerly INOA), producing internationally recognized breakthrough results. The most recent best results to which he has given significant contributions include the first fiber-based mid-IR generation of a frequency comb as well as demonstration of its suitability for frequency metrology; pioneering work on sub-Doppler molecular spectroscopy with an absolute frequency scale in the IR, using coherent sources based on difference frequency generation with smart phase-locks, using optical frequency-comb synthesizers for building highly coherent mid-IR radiation sources; development of a novel spectroscopic technique (SCAR), based on saturated absorption cavity ring-down, overcoming background-related sensitivity limits of standard cavity ring-down; pioneering work on the intrinsic noise properties of Quantum Cascade Lasers; discovery of the intrinsic sensitivity limits of fiber-based optical sensors; molecular gas sensing at parts per quadrillion, achieving the highest sensitivity (i.e., minimum detectable gas pressure) ever observed in a spectroscopic experiment on a gas of simple molecules, taking to detection of radiocarbon-dioxide ( $^{14}\text{CO}_2$ ) below its natural abundance. Paolo De Natale has authored about 200 papers, has edited 10 books and special journal issues, and holds 5 patents. Finally, most of the time away from the family due to his scientific activity was compensated by the strong and loving commitment of his wife Roberta.

# 1

---

## *Shedding light on the art of timekeeping*

---

He was the true Light,  
which doth enlighten every  
man, coming to the world.

*John - 1,9*

The Present contains nothing more than  
the Past, and what is found in the effect  
was already in the cause.

*Henri L. Bergson - Creative Evolution*

---

### 1.1 The great show of Time and Light, the curtain rises!

We have always lived in a world illuminated by Light and marked by the relentless flow of Time. In spite of the difficulty of finding a universal definition for them, Light and Time are essential elements of our existence. They govern countless aspects of our practical life and accompany us in various cultural and sentimental experiences. The earliest people on the planet naturally entrusted the organization of their activities in the light coming from celestial bodies. The periodic character of the most basic astronomical motions was immediately recognized: the Sun rise and set (Earth's rotation around its axis), the appearance of the highlighted portion of the Moon (Moon's revolution around the Earth), and the weather periodical behavior that seemed to be related to the movement of the Sun with respect to the stars (Earth's revolution around the Sun). The units of days, months, and years accordingly followed. The main disadvantage of Nature's clocks resided in that the scale unit was too large for many practical purposes. Consequently, *natural oscillators* soon began to be supplemented by those constructed by mankind. Around 3500 BC Egyptians already divided the time of the day into shorter sections by observing the direction of the shadow cast from obelisks or sundials by the Sun, depending on its position in the sky [1]. It is amazing to note that the ancient and honored Earth-Sun clock met many of the most demanding requirements that the scientific community today exacts from an acceptable standard: first, it is universally available and recognized; second, it involves neither responsibility nor operation expenses for anyone; third, it is pretty reliable and we cannot foresee any possibility that it may stop or *lose* the time. In spite of all these nice features, however, this clock does not represent an extremely stable timepiece. According to our current knowledge in astronomy, first, Earth's orbit around the Sun is elliptical rather than a perfect circle, which means that Earth travels faster when it is closer to the Sun than when it is farther away. In addition, Earth's axis is tilted with respect to the plane containing its orbit around the Sun. Finally, Earth spins at an irregular rate around its axis of rotation

and even wobbles on it. The latter effect is due to the circumstance that, as Earth is neither perfectly symmetrical in shape, nor homogeneous, nor ideally rigid (its mass distribution constantly changes over time), its rotation axis does not coincide exactly with the figure axis. For the same reason, even natural disasters of exceptional importance may perturb the clock mechanism. For example, it has been calculated that the recent earthquake in Japan (March 2011) has moved Earth's figure axis by a few milliarcseconds or, in other words, rearranged Earth's mass bringing more of it a bit closer to the rotation axis. This should have slightly increased Earth's rate of spin, thus shortening the length of the day by less than 2 microseconds [2]. Such a small variation has no practical effect in daily life, but it is of interest for precision measurements of space and time.

Although the interaction between sunlight and Earth's swinging is far from being considered a resonant phenomenon, Earth-Sun clocks constitute the first example of a timepiece in which a light source interrogates a frequency reference. Quite surprisingly, however, in the following five millennia, Light and Time travelled rather distinct roads in the advancement of human thought and abilities.

## 1.2 Brief history of timekeeping: time-frequency equivalence

Clepsydrae based on controlled flows of water (either into or out of a vessel) were available in Egypt, India, China, and Babylonia from about 1500 BC and represented the first non-astronomical means of measuring time. Sand clepsydrae were introduced only in the late fourteenth century AD. By using the integrated quantity of moved substance to provide a measurement of the elapsed time, this type of timekeeper did not rely on counting the number of cycles of an oscillatory event. The resort to light was abandoned too. In the last part of the thirteenth century mechanical clocks began to appear in Europe [3]. The first prototypes, representing the natural progression of wheel clocks driven by water (already introduced in China after the 8th century), were just geared machines based on the fall of a weight regulated by a verge-and-foliot escapement. Variations of this design reigned for more than 300 years, but all had the same basic problem: the period of oscillation of the escapement was heavily affected both by the amount of force and the extent of friction in the drive. Like water flow, the rate was difficult to adjust.

A significant advance occurred in the 17th century when Galileo Galilei discovered that the period  $T$  of a pendulum swing virtually does not depend on the excursion, provided that the latter is not too large:

$$T \simeq 2\pi\sqrt{\frac{l}{g}} \quad (1.1)$$

Here  $l$  is the pendulum length, and  $g$  is the acceleration due to gravity. Galilei, in fact, recognized the value of the pendulum as a time-keeping device and even sketched out a design for a clock. However, it was Christiaan Huygens in 1656 to realize the first successful operational pendulum clock. Reaching an error of less than 1 minute a day, such device recovered and definitively consecrated the idea that the most accurate way of keeping the time was to employ an oscillatory system operating at a specific resonance frequency  $\nu_0$ . Hence, any time interval could be measured by counting the number  $N$  of elapsed cycles and then multiplying  $N$  by the period  $T = 1/\nu_0$ . Light, however, was still excluded from the time-keeping saga.

From then onwards, time and frequency became the quantities that humanity could measure with the highest precision. Indeed, during the next three centuries, continuous refinements improved considerably the accuracy of pendulum clocks. In 1671, William Clement began building clocks with the new *anchor* escapement, a substantial improvement over the verge because it interfered less with the motion of the pendulum. In 1721, George Graham improved the pendulum clock accuracy to 1 second per day by compensating for changes in the pendulum length due to temperature variations. John Harrison, a carpenter and self-taught clockmaker, developed new methods for reducing friction. By 1761, he had built a marine chronometer with a spring-and-balance-wheel escapement that kept time on board a rolling ship to about one fifth of a second a day, nearly as well as a pendulum clock could do on land. Over the next century, refinements led in 1889 to Siegmund Riefler's clock with a nearly free pendulum, which attained an accuracy of a hundredth of a second a day and became the standard in many astronomical observatories. A true free-pendulum principle was demonstrated by R.J. Rudd around 1898. This gave birth to a generation of superior timepieces that culminated in 1920 with the realization by William H. Shortt of a clock consisting of two synchronized pendulums. One pendulum, the master, swung as unperturbed as possible in an evacuated housing. The slave pendulum driving the clockwork device was synchronized via an electric linkage and in turn, every half a minute, initialized a gentle push to the master pendulum to compensate for the dissipated energy. Keeping time better than 2 milliseconds a day, Shortt clocks almost immediately replaced Riefler ones for time distribution on local and eventually national scale.

The performance of such clocks was overtaken as soon as the technology of quartz crystal oscillators became mature for the construction of the first timekeeper (W. Marrison and J.W. Horton, 1927). Quartz clock operation hinges on piezoelectricity that is the capability of some materials to generate electric potential when mechanically stressed or, conversely, to strain when an electric potential is applied. Due to this interaction between mechanical stress and electric field, when placed in a suitable oscillating electronic circuit, the quartz will vibrate at a specific resonance frequency (basically depending on its size and shape) and the frequency of the circuit will become the same as that of the crystal. Such a signal is eventually used to operate an electronic clock display. As they had no escapements to disturb their regular frequency, quartz crystal clocks soon proved their superiority with respect to pendulum-based ones. A serious source of systematic error, namely the dependence of the period on the strength of the local gravity vector (and hence on the pendulum location), was overcome too. Although quartz oscillators had provided a major advance in timekeeping, so as to become, in the late 1930s, the new timekeeping standards, it was apparent that there were limitations to that technology. These devices could provide frequency with a precision of about  $10^{-10}$ , but going beyond proved to be a real challenge. Operationally, fundamental mode crystals could be made to provide frequencies up to 50 MHz. Higher frequencies capable of providing more precise timekeeping were possible using overtones but were not commonly used. Moreover, aging and changes in the environment, including temperature, humidity, pressure, and vibration, affected the crystal frequency. In order to compensate for these problems, different systems were designed, including temperature-compensated and oven-controlled crystal oscillators.

To make a significant advance in precision timekeeping of laboratory standards, however, a fundamental change was required [4, 5, 6]. Scientists had long realized that atoms (and molecules) have resonances; each chemical element and compound absorbs and emits electromagnetic radiation at its own characteristic frequencies. An unperturbed atomic transition is identical from atom to atom, so that, unlike a group of quartz oscillators, an ensemble of atomic oscillators should all generate the same frequency. Also, unlike all electrical or mechanical resonators, atoms do not wear out. Additionally, all experimental observations in spectroscopy have proved compatible with the hypothesis that atomic properties are the

same at all times and in all places, when they are assessed by an observer situated close to the atom and accompanying it in the same motion. It is therefore possible to build instruments, which, using a specified atomic transition, are all able to deliver a signal in real time with the same frequency, anywhere and at any time, provided that relativistic effects due to non-coincidence of atom and observer have been properly taken into account [7]. These features were appreciated by Lord Kelvin who suggested using transitions in hydrogen as a time-keeping oscillator. However, it wasn't until the mid 20th century that technology made these ideas possible. The first atomic clocks owe their genesis to the explosion of advances in quantum mechanics and microwave electronics before and during the Second World War. Indeed, the sudden development of the radar and very high frequency radio communications made possible the generation of the kind of electromagnetic waves (microwaves) needed to interact with atoms. Atomic oscillators use the quantized energy levels in atoms and molecules as the source of their resonance frequency. The laws of quantum mechanics dictate that the energies of a bound system, such as an atom, have certain discrete values. An electromagnetic field at a particular frequency can excite an atom from one energy level to a higher one. Or, an atom at a high energy level can drop to a lower level by emitting energy. The resonance frequency of an atomic oscillator is the difference between the two energy levels,  $E_1$  and  $E_2$ , divided by Planck's constant,  $h$ :

$$\nu_0 = \frac{E_1 - E_2}{h} \quad (1.2)$$

The basic idea of atomic clocks is the following. First, a suitable energy transition is identified in some atomic species (microwave atomic frequency standards are commonly based on hyperfine transitions of hydrogen-like atoms, such as rubidium, cesium, and hydrogen). These provide transition frequencies that can be used conveniently in electronic circuitry (1.4 GHz for hydrogen, 6.8 GHz for rubidium, and 9.2 GHz for cesium). Then, an ensemble of these atoms is created (either in an atomic beam, or in a storage device, or in a *fountain*). Next, the atoms are illuminated with radiation from a tunable source that operates near the transition frequency  $\nu_0$ . The frequency where the atoms maximally absorb is sensed and controlled. When the absorption peak is achieved, the cycles of the oscillator are counted: a certain number of elapsed cycles generates a standard interval of time. Most of the basic concepts of atomic oscillators were developed by Isidor Rabi and his colleagues at Columbia University in the 1930's and 40's. Although he may have suggested using cesium as the reference for an atomic clock as early as 1945, research aimed at developing an atomic clock focused first on microwave resonances in the ammonia molecule. In 1949, the National Bureau of Standards (NBS) built the first atomic clock, which was based on ammonia (at 23.8 GHz). However, its performance wasn't much better than the existing standards, and attention shifted almost immediately to more promising atomic-beam devices based on cesium. The first practical cesium atomic frequency standard was built at the National Physical Laboratory (NPL) in England in 1955 by Dr. Louis Essen. In collaboration with the U.S. Naval Observatory (USNO), it was immediately noted that observations of the Moon over a period of several years would be required to determine Ephemeris Time with the same precision as was achieved in a matter of minutes by the first cesium clock. For the benefit of the reader, we recall here that the ephemeris second is based on the period of revolution of the Earth around the Sun which is more predictable than the rotation of Earth itself (for more details, refer to Chapter 7).

While NBS was the first to start working on a cesium standard, it wasn't until several years later that NBS completed its first cesium atomic beam device. By 1960, cesium standards had been refined enough to be incorporated into the official timekeeping system of NBS. Standards of this sort were also developed at a number of other national standards laboratories, leading to wide acceptance of this new timekeeping technology. Then, pres-

sure mounted for an atom-based definition of time. This change occurred in 1967 when, by international agreement,

---

*the **second** was defined as the duration of 9,192,631,770 periods of the radiation corresponding to the transition between two hyperfine levels  $|F = 4, m_F = 0\rangle \leftrightarrow |F = 3, m_F = 0\rangle$  in the ground state  $^2S_{1/2}$  of the  $^{133}\text{Cs}$  atom.*

---

This definition made atomic time agree with the second based on Ephemeris Time, to the extent that measurement allowed. As of 2011, the definition of the SI second remains the same, except for a slight amendment made in 1997. Calculations made by Wayne Itano of NBS in the early 1980s revealed that blackbody radiation can cause noticeable frequency shifts in cesium standards [8], and his work eventually resulted in an addendum to the definition: the Comité International des Poids et Mesures (CIPM) affirmed in 1997 that the definition refers to *a cesium atom at rest at a thermodynamic temperature of 0 K*. Thus, a perfect realization of the SI second would require the cesium atom to be in a zero magnetic field in an environment where the temperature is absolute zero and where the atom has no residual velocity.

Excluding the pendulum, quartz and microwave atomic clocks are far from being relegated to history and their study will be resumed in Chapter 7. We close this section by observing that the development of increasingly more accurate frequency standards was paralleled by an augmented frequency of the employed oscillator: from Earth's rotation ( $\sim 10 \mu\text{Hz}$ ), via pendulum clocks ( $\sim 1 \text{ Hz}$ ) and quartz oscillators ( $\sim 1 \text{ MHz}$ ), to microwave atomic standards ( $\sim 1 \text{ GHz}$ ). More strictly, the accuracy performance of a frequency standard is characterized by the so-called quality factor ( $Q$ ) which is defined, in general, as the oscillator resistance to disturbances to its oscillation period. This notion can be grasped in the case of a pendulum clock, where, in order to replace the energy lost by friction, pushes must be applied by the escapement. These pushes are the main source of disturbance to the pendulum motion. The smaller the fraction of the pendulum energy that is lost to friction, the less energy needs to be added, the less the disturbance from the escapement, the more the pendulum is *independent* of the clock mechanism, and the more constant its period is. In other words, the  $Q$  factor is related to the ratio of the total energy in the system to the energy lost per cycle or, equivalently, to how long it takes for the swings of the oscillator to die out:

$$Q \equiv \frac{\tau}{T} = \frac{\nu_0}{\Gamma} \quad (1.3)$$

where  $\tau \equiv 1/\Gamma$  is the time constant describing the (exponential) decay of the swing amplitude. Hence, the  $Q$  of pendulum clocks is increased by maximizing  $\tau$  or, equivalently, minimizing the overall frictional losses ( $\Gamma$ ). As it will be shown in Chapter 2, for a damped harmonic oscillator,  $\Gamma$  equals the full width at half maximum  $\Delta\nu$  of the system response function (resonance curve) in the frequency domain. With this in mind, the above formula can be generalized to all types of oscillators as

$$Q \equiv \frac{\nu_0}{\Delta\nu} \quad (1.4)$$

Concerning quartz oscillators, here we just mention that, starting from the electric equivalent of the crystal,  $\nu_0$  and  $\Delta\nu$  are respectively calculated as the resonance frequency and width of an oscillatory circuit. In the case of atoms, finally,  $\Delta\nu$  is calculated in the frame

of quantum mechanics. Actually, as we will see in Chapter 4, atoms absorb or emit energy over a small frequency range surrounding  $\nu_0$ : this spread of frequencies is referred to as the linewidth and its ultimate limit is related to Heisenberg's uncertainty principle. Since the response of a high- $Q$  system decays much more rapidly as the driving frequency moves away from  $\nu_0$ , the oscillator with the highest  $Q$  would be desirable as a frequency standard,  $Q^{-1}$  being roughly proportional to its limiting accuracy. This can be achieved either by using an atomic transition where  $\nu_0$  is as high as possible, or by making  $\Delta\nu$  as narrow as possible.

### 1.3 The parallel story of the speed of light

In order to fully appreciate the significance of the modern physical measurements of the speed of light, just think that even today we are usually not aware of any delay between the occurrence of an event and its visual appearance in the eye of a distant observer. In fact, a single visual *snapshot* is probably, for most people, the basis for the intuitive notion of an *instant*. Therefore, it is of great interest to shortly trace the history of ideas concerning the finiteness of the velocity of light [9, 10]. Among the ancient Greeks, there was a general belief that this speed was infinite. An exception is represented by Empedocles from Acragas (490-435 BC) who, according to Aristotle (384-322 BC), "was wrong in speaking of light as travelling or being at a given moment between the earth and its envelope, its movement being unobservable to us". So powerful is Aristotele's cosmology that it compels him to declare that "...light is due to the presence of something, but it is not a movement".

An interesting *proof* that the velocity of light must be infinite is given by Heron of Alexandria (I century BC). According to him, you turn your head to the heaven at night, keeping the eyes closed; then suddenly open them, at which time you see the stars. Since no sensible time elapses between the instant of opening the eyes and the instant of sight of the stars, light must travel instantaneously. Since the causal direction of an instantaneous interaction is inherently ambiguous, it's not surprising that ancient scholars considered two competing models of vision, one based on the idea that every object is the source of images of itself, emanating outwards to the eye of the observer, and the other claiming that the observer's eye is the source of visual rays emanating outwards to *feel* distant objects. Indeed, at that time, the problem of the speed of light was secondary, whereas there was much more interest in catoptrics and vision matter.

Amidst the Islamic scientists, Avicenna (980-1073) was perhaps the most famous: his thought represents the climax of medieval philosophy. Avicenna observed that, if the perception of light is due to the emission of some sort of particles by the luminous source (as he believed), then the speed of light must be finite. Alhazen (965-1039), another Muslim physicist and one of the greatest scholars of optics of all time, came to the same conclusion. In his treatise on optics he states that light is a movement and, as such, is at one instant in one place and at another instant in another place. Since light is not in both these places at the same time, there must be a lapse of time between the two: hence the transmission cannot be instantaneous.

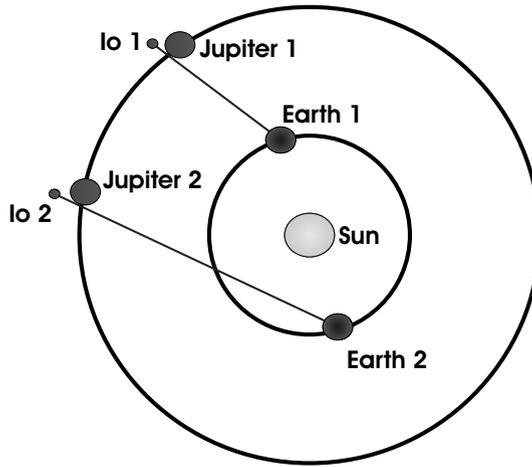
Nevertheless, Aristotle's point of view was echoed by many thinkers in western history: John Peckam (1230-1292), Thomas Aquinas (1225-1274), and Witelo (1230-1275) to name a few. It is curious to note that Roger Bacon (1214-1292), although in perfect agreement with Alhazen's conclusions on this subject, felt the need to show in "Opus Majus" that the sort of reasoning used by Alhazen was identical to that of the scientists who attempted to prove the opposite view. Bacon's remarks afford a striking example of the confusion exhibited by a first rate mind attempting to be reasonable with no genuine scientific or experimental

basis as guide. The debate continued into the beginning of the *scientific revolution* of the 17th century. Such giants as Francis Bacon (1561-1626), Johannes Kepler (1571-1630), and René Descartes (1596-1650) adhered to the idea of instantaneous propagation. Descartes considered an eclipse of the moon, caused by the moon, earth, and sun being in a straight line, with Earth interposed between the other two: “Now suppose that it requires an hour for light to travel from the earth to the moon. Then the moon will not become dark until exactly one hour after the instant of collinearity of the three bodies. Similarly, here on the earth, we will not observe Moon’s darkening until the passage of another hour, or until two hours after the moment of collinearity. But during this time, the moon will have moved in its orbit and the three bodies will no longer be collinear. But clearly, this is contrary to experience, for one always observes the eclipsed moon at the point of the ecliptic opposite to the sun. Hence light does not travel in time, but in an instant”.

In the face of all this, the remarks by Galileo (1564-1642) seem like a breath of fresh air in a stale room. In his great treatise on mechanics there is a conversation about the velocity of light during which Salviati claims that the general inconclusiveness of observations on this subject had led him to conceive an experiment. He says: “Let each of two persons take a light contained in a lantern such that, by the interposition of the hand, the one can shut off or admit the light to the vision of the other. Next, let them stand opposite each other at a distance of a few cubits and practice until they acquire such skill (in uncovering and occulting their lights) that the instant one sees the light of his companion, he will uncover his own. After acquiring this skill, the two experimenters were to perform the same operations at greater distances, ten miles if necessary (using telescopes). If the exposures and occultations occur in the same manner as at short distances, we may safely conclude that the propagation is instantaneous; but if time is required at a distance of three miles, which, considering the going of one light and the coming of the other, really amounts to six, then the delay ought to be easily observable”. This experiment was executed by the Florentine Academy and their account of it is as follows: “We tried it at a mile’s distance and could not observe any. Whether in a greater distance it is possible to perceive any sensible delay, we have not yet had an opportunity to try”.

The first experimental evidence of the finite speed of light was due to Ole Christensen Roemer in 1676 by observing the eclipses of the inner-most moon of Jupiter (Io) [11]. Discovered by Galileo in 1610, detailed tables of the movements of these moons had been developed by Borelli (1665) and Cassini (1668). Io has a period of about 42.5 hours and, if Earth were stationary, it would show an eclipse at regular intervals of 42.5 hours. But Earth revolves about the sun and, in so doing, assumes positions 1 and 2 in Figure 1.1. Roemer noticed that when the Earth was close to Jupiter (position 1), the eclipses occurred 8.5 minutes ahead of the time predicted on the basis of yearly averages. The eclipses were late by the same amount when the Earth was opposite (position 2). Roemer concluded that twice that difference was the time it took the light to traverse the diameter of Earth’s orbit ( $\sim 3 \cdot 10^8$  km), which gave a figure of  $\sim 227000$  km/s.

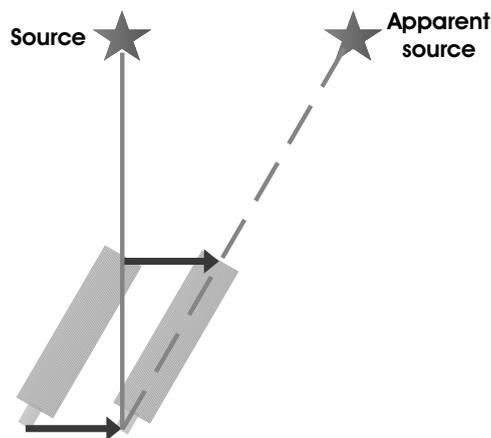
Despite the force of Roemer’s analysis, and the early support of both Huygens and Newton, most scientists remained skeptical of the idea of a finite speed of light. Alternative explanations were provided by Cassini and later by his nephew Giacomo Filippo Maraldi. They suggested that Jupiter’s orbit and the motion of its satellites might explain the observed inequalities. It was not until 50 years later, when the speed of light was evaluated in a completely different way, arriving at nearly the same value, that the idea became widely accepted. Such measurement was performed in 1728 by James Bradley by observing stellar aberration, that is the apparent displacement of stars due to the motion of the Earth around the Sun. A useful analogy to help understand aberration is to imagine the effect of motion on the angle at which rain falls. If you stand still in the rain (when there is no wind), it comes down vertically on your head. If you run through the rain it appears to come to you

**FIGURE 1.1**

Roemer's evaluation of the speed of light.

from an angle and hit you on the front. It is worth pointing out that all stellar positions are affected equally in the aberration phenomenon, which distinguishes this effect from parallax where nearby stars are influenced more noticeably. By observing a star in Draco, and recording its apparent position during the year, Bradley argued that stellar aberration is approximately the ratio of the orbital speed of the Earth (around the Sun) to the speed of light (see Figure 1.2). Based on the best measurement of the limiting starlight aberration (20.5 arcseconds  $\simeq 0.0001$  rad) by Otto Struve, and taking the speed of Earth to be about 30 km/s from Encke's estimate, this implied a light speed of about 301000 km/s [9].

Unfortunately, measurements of the speed made in this way depended on the astronomical theory and observations used. Better determinations of the speed might be made if both source and observer were terrestrial. The first measurement of  $c$  on Earth was by Armand Fizeau in 1849 [12]. His method measured the time needed for light to travel to a flat mirror

**FIGURE 1.2**

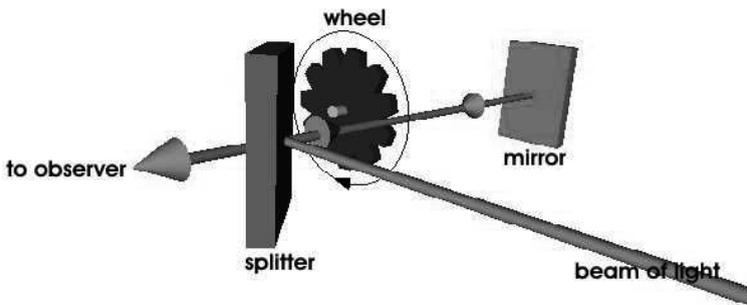
Bradley's determination of the light speed.

at a known distance and return. For that purpose he designed a set-up where a collimated beam emitted by a limelight passed through a half-mirror and a rotating cogwheel was then reflected back by a mirror situated some 8.633 kilometers away, passed (or not) through the cogwheel again, and was eventually reflected by the half-mirror into a monocular (see Figure 1.3). At a low rotation rate, the light passes through the same blank of the wheel on the way out and on the way back. But with increasing rotation rate, a higher and higher percentage of the transmitted light is cut on its way back by the incoming tooth of the wheel, resulting in a decreasing light intensity collected in the monocular. Total extinction of the returning light is reached when the time duration of the open gate corresponds exactly to the duration of the round-trip, such that the light that has gone through finds the gate closed when it returns. Knowing the precise distance  $d$  between the wheel and the mirror, the number of teeth  $N_t$  of the wheel, and its rotation rate  $\omega$  (expressed in radians per second), the speed of light in air can be deduced to be

$$c = (2d)(2N_t)f_c \quad (1.5)$$

where  $f_c = N_t\omega/2\pi$  is the frequency at which the beam is effectively stopped. Obviously, if one increases further the rotating speed of the wheel, light will appear again as the returning light will start passing through the gap situated right after the one it has passed on its way out. Using this method with the cogwheel placed in Montmartre and the reflector in Suresnes, Fizeau obtained a value of  $c = 315300$  km/s, limited by the precision of his measurement of  $\omega$ , but yet better than any measurement realized before. Such a method was subsequently taken up first by Marie Alfred Cornu in 1874 and then by Joseph Perrotin in 1902. Some experimental tricks allowed them to provide the following more accurate results for the speed of light *in vacuum*:  $299990 \pm 200$  km/s and  $299901 \pm 84$  km/s, respectively (their results already included correction for the refractive index of air) [13].

In 1855, Kirchhoff realized that  $1/\sqrt{\varepsilon_0\mu_0}$  has the dimension of a speed, where  $\mu_0$  ( $\varepsilon_0$ ) is the magnetic permeability (electric permittivity) of free space entering the laws of magnetism (electricity). In 1856 Weber and Kohlrausch measured this constant using only electrostatic and magnetostatic experiments [14]. Incidentally, they were the first to adopt the symbol  $c$  (from Latin *celeritas*) for the speed of light. Within experimental accuracy, the value found by them agreed with the speed of light. This remained a coincidence until Maxwell formulated his theory of electromagnetism in 1865 and concluded that "...light is an electromagnetic disturbance propagated through the field according to electromagnetic laws" [15]. Maxwell's equations established that the velocity of any electromagnetic wave (and thus of light) in a *vacuum* is  $c$ , where



**FIGURE 1.3**

First terrestrial measurement of  $c$ , performed by Fizeau.

$$c^2 = \frac{1}{\varepsilon_0 \mu_0} \quad (1.6)$$

At this stage, the status of  $c$  increased tremendously because it became a characteristic of all electromagnetic phenomena.

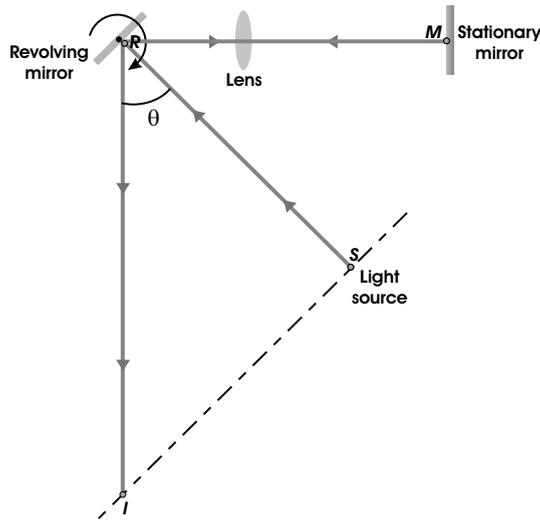
The method demonstrated by Leon Foucault (1862) relies on the same principle adopted by Fizeau (time of flight technique), but replaces the cogwheel by a revolving mirror (Figure 1.4). Light rays from the source  $S$  that strike the revolving mirror  $R$  and proceed through the lens will strike the stationary mirror  $M$  and return to the source. If, after the light beam first strikes  $R$  outbound from  $S$ ,  $R$  can be rotated before it is struck again by the beam returning from  $M$ , then the returning beam will no longer return exactly to the source  $S$  but will instead be deflected away from  $S$  in the direction of the rotation. By rotating the mirror at a constant speed, the amount of deflection will be the same for all light beams which go through the lens, strike  $M$  and return. Then, for a continuous beam of light from  $S$  and a constant high speed of rotation of  $R$ , an image of the source will appear beside  $S$  instead of coincident upon it. The faster  $R$  rotates or the longer is  $\overline{RS}$ , the farther the returned image  $I$  will be displaced from the source  $S$  and the easier it will be to measure the deflection. By carefully measuring the amount of displacement from  $S$  to  $I$ , and the distance from  $S$  to  $R$ , the angle of deflection can be determined. Together with the known, fixed speed of rotation, this angle can be used to determine the time it took light to travel the distance from  $R$  to  $M$  and back. Let  $\theta = \arctan(\overline{IS}/\overline{IR})$  denote the angle of deflection (this means that the angle through which the mirror has rotated is  $\theta/2$ ). If the speed of rotation is measured in number of cycles ( $n_c$ ) per second, then the speed of light is given by

$$c = \frac{2 \cdot \overline{RM}}{\theta \frac{1}{1} \frac{1}{2 \cdot 2\pi n_c}} \quad (1.7)$$

In this arrangement, the distances  $\overline{IS}$  and  $\overline{SR}$  should be as large as possible to reduce the error in measuring  $\theta$ . The distance  $\overline{IS}$  is maximized by maximizing the speed of rotation of  $R$  and the distance  $\overline{RM}$ . In Foucault's setup,  $M$  was spherical with center at  $R$ . The greatest distance  $\overline{RM}$  achieved by Foucault was 20 m, which produced a displacement  $\overline{IS}$  of only about 1 mm. The result was  $298000 \pm 500$  km/s [13].

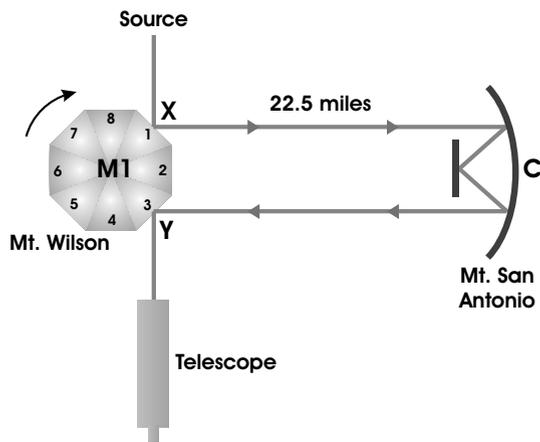
Going back to the approach by Weber and Kohlrausch, in 1907 Rosa and Dorsey obtained a much more accurate determination of  $c$ . As the value of  $\mu_0$  is fixed at exactly  $4\pi \cdot 10^{-7}$  N·A<sup>-2</sup> through the definition of the ampere, only  $\varepsilon_0$  had to be measured in their experiments. This can be accomplished by determining the ratio of the capacitance of a condenser as measured in electrostatic and electromagnetic units. Rosa and Dorsey used the Maxwell bridge method (employing carefully standardized resistances) to determine the electromagnetic capacitance and standards of length and mass to determine the electrostatic capacitance [16]. They used a variety of shapes (spherical, cylindrical, and plane) and sizes of condensers [17]. Both the calculations and experiments were beset with difficulties, but their result was probably the most reliable up to that time. The final value,  $c = 299710$  km/s, was the mean of about 900 individual determinations with an estimated maximum error of 30 km/s, apart from uncertainties in the value taken for the international ohm. In 1941 a more accurate knowledge of the latter standard allowed Birge to apply a correction to their result yielding the value  $c = 299784$  km/s [18].

Foucault's apparatus was perfected by Michelson in several versions till the famous experiment in 1927 [19]. As shown in Figure 1.5, the apparatus involved a rotating octagonal glass prism. When the prism is stationary the light follows the path shown and an image



**FIGURE 1.4**  
Focault's apparatus for measuring  $c$ .

of the source can be seen through the telescope. If the prism is rotated slowly, the image disappears because either face  $X$  is not in a suitable position to direct the outgoing beam to the concave reflector  $C$ , or face  $Y$  is unable to send the incoming beam to the telescope. However, if the rotation speed of the prism is increased so that it turns exactly one-eighth of a revolution in the same time that it takes light to travel from  $X$  to  $Y$ , then an image of the source is seen through the telescope. Michelson adjusted the speed of rotation until he was able to observe a stationary image of the source. This occurred when the prism was rotating at  $f_{rot} \simeq 530$  Hz (this rate was measured by comparison with a free pendulum furnished by the United States Coast and Geodetic Survey). The experiment was carried out on Mt. Wilson (USA) and the concave reflector  $C$  was on Mt. San Antonio  $d = 35$  km away. The result was  $c = (2d)/[(1/8)(1/f_{rot})] = 299796 \pm 4$  km/s.



**FIGURE 1.5**  
Michelson's famous experiment for the measurement of  $c$ .

An ingenious modification of the toothed-wheel method was used by Karolus and Mittelstaedt in 1928. In their apparatus, a Kerr cell, at the terminals of which an alternating difference of potential was applied, was used to periodically interrupt the passage of a luminous beam. The main advantage of this method is that the frequency of such periodic interruption can be accurately determined, which was not the case when the toothed wheel was used; moreover, a much higher frequency can be used with this method (around 10 MHz), so that a correspondingly short base (in this case 41.4 meters) can be utilized [13]. This approach gave the value  $299778 \pm 20$  km/s.

The completion of this excursus on the measurements of  $c$  necessitates now the discussion of experiments whose understanding requires knowledge beyond basic physics, many of the involved concepts being precisely the subject of this book. For the moment, may the reader be satisfied with an intuitive comprehension; a full appreciation will result from revisiting each experiment as the pertinent notions are gradually acquired.

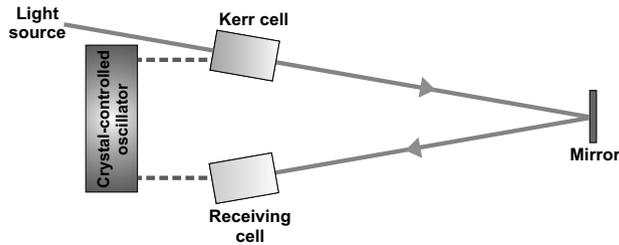
As we will learn in Chapter 3, another valuable option to determine  $c$  is to measure the resonance frequencies of a cavity resonator whose dimensions are precisely known. In 1946, Louis Essen and A.C. Gordon-Smith pursued this approach, by establishing the frequency for a variety of normal modes of an evacuated microwave cavity [20, 18]. The latter consisted of a copper cylinder constructed with great uniformity. The resonant frequency  $\nu_n$  of an evacuated right circular hollow cylinder closed at both ends is given by

$$c = \frac{\nu_n}{\sqrt{\left(\frac{r}{\pi D}\right)^2 + \left(\frac{n}{2L}\right)^2}} \left(1 + \frac{1}{2Q_{cav}}\right) \quad (1.8)$$

where  $r$  is a constant for a particular mode of resonance,  $n$  is the number of half-wavelengths in the guide,  $D$  is the diameter and  $L$  the length of the cylinder, and  $Q_{cav}$  is the quality factor of the resonator accounting for the finite conductivity of the cavity walls. The quantities  $\nu_n$ ,  $D$ ,  $L$ , and  $Q_{cav}$  could all be measured with a precision of a few parts in  $10^6$ . In particular, the dimensions of the resonator were measured in the Metrology Division of the National Physical Laboratory using gauges calibrated by interferometry. The final result (using the  $E_{010}$  and  $E_{011}$  modes) was  $c = 299792 \pm 9$  km/s, where the estimated maximum error was the sum of different contributions including setting of the frequency to resonance and measurement of the frequency by the spectrum analyzer, uncertainty of the resonator temperature, dimensional measurements, residual effects of coupling holes and probes, non-uniformity of the resonator, and uncertainty of  $Q_{cav}$ . Almost simultaneously, a very similar value ( $299789.3 \pm 0.4$  km/s) obtained by the same measurement scheme was published by Bol in a short note [21].

In those years, radar systems also began to be used to measure the speed of light. Again, the time-of-flight principle was exploited: twice the known distance to a target was divided by the time it took a radio-wave pulse to return to the radar antenna after being reflected by the target. This was done by Aslakson in 1949 with the result  $299792.4 \pm 2.4$  km/s [22]. Incidentally we mention here (see Chapter 7 for further details) that, today, a Global Positioning System (GPS) receiver measures its distance to GPS satellites based on how long it takes for a radio signal to arrive from each satellite: from these distances the receiver position is calculated.

Then came the geodimeters. Originally intended for use in geodesic surveying, Bergstrand demonstrated their use in accurate measurement of the light speed [23]. With reference to Figure 1.6, the principle can be described as follows: a light beam is emitted through a Kerr cell to a distant mirror and reflected back to a receiving photocell close to the emitter. The two cells are supplied by the same crystal-controlled high frequency voltage (about 10 MHz in the original work). The difference in phase of the emitted and

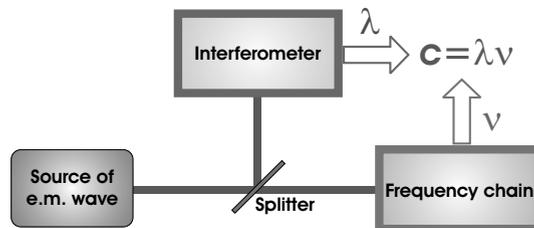
**FIGURE 1.6**

Bergstrand's method for evaluating  $c$ .

received light is compared. If the mirror distance is known, the speed of light can be measured [24]. Bergstrand repeated this kind of measurement several times with the final value  $299793.1 \pm 0.2$  km/s [25]. Such phase-shift method was pushed to the limit of its accuracy by Grosse in 1967:  $299792.50 \pm 0.05$  [26].

Another way to get the speed of light is to independently measure the frequency  $\nu$  and wavelength  $\lambda$  of an electromagnetic wave in *vacuum*. The value of  $c$  can then be found by using the relation  $c = \lambda\nu$  (see Figure 1.7).

This approach was followed first in the microwave domain at NPL by Froome in 1958 [27]. The basis of the determination consisted of the simultaneous measurement of the free-space wavelength and frequency of an electromagnetic wave generated by a microwave source. The latter was a frequency-stabilized klystron oscillator operating at 36 GHz (in short, the klystron is a specialized linear-beam electron vacuum tube which converts, via velocity modulation, the kinetic energy of the electron beam into a radio-frequency/microwave signal). The greater part of the output from this oscillator was fed by means of a waveguide switch into one of the two silicon crystal distorter units tuned for maximum harmonic output at 72 GHz (about 0.4 cm wavelength). One harmonic generator was used to supply the interferometer itself, the other for operating the cavity resonator refractometer by means of which the refractive index of the air in the neighborhood of the equipment could be measured. The measurement of the microwave frequency was accomplished by comparing a portion of the klystron output against a high harmonic of a 5-MHz quartz crystal standard. The 5-MHz was multiplied in stages of two to five times up to 600 MHz and then fed into a silicon crystal harmonic generator mounted in waveguide, so that the harmonic at exactly 36 GHz could be mixed with a small fraction of the klystron output. The beat frequency between the two was detected by means of a calibrated communications receiver. The accuracy of frequency determination was at least as good as 1 part in  $10^8$ . The estimated accuracy of the refractive index measurement was 1.1 parts in  $10^7$ . The value of

**FIGURE 1.7**

Determination of  $c = \lambda \cdot \nu$  by independent measurements of the free-space wavelength  $\lambda$  and frequency  $\nu$  of an electromagnetic wave.

microwave wavelength obtained by means of the interferometer, when multiplied by the air refractive index and the microwave frequency, gave a *vacuum* phase velocity which had still to be corrected for the effect of diffraction before the true free-space value could be derived. Basic concepts of interferometry for wavelength measurements will be given in Chapter 5. Here we just recall that the diffraction limit is proportional to  $\lambda$ . It is interesting to note that the greatest single uncertainty in the whole measurement arose from the use of the length standards. At that time, interferometry-based length measurements were ultimately referenced to the cadmium red line (falling at  $\lambda \simeq 644$  nm) as emitted by the international specified form of the Michelson lamp. The result was  $299792.50 \pm 0.10$  km/s.

Later, it was discovered that the cadmium line was actually a cluster of closely separated lines, and that this was due to the presence of different isotopes in natural cadmium. Thus, in order to get the most precisely defined line, it was necessary to use a mono-isotopic source. Allowing for easier isotopic enrichment and lower operating temperatures for the lamp (which reduces broadening of the line due to the Doppler effect), the pretty bright orange line of krypton-86 (at  $\lambda \simeq 606$  nm) was then selected as the new standard. Krypton-86 offered the additional advantages of having zero nuclear spin.

In the same years, a quite different, spectroscopy-based approach was also pursued to measure  $c$  [28]. That was the so-called band spectrum method, involving the simultaneous measurement of the rotational constant  $B''$  of the ground state of a diatomic (or linear) molecule in pure frequency units by means of microwave spectroscopy and in  $\text{cm}^{-1}$  units by means of infrared (rotation-vibration) spectroscopy. Then, the ratio

$$\frac{B'' \text{ microwave}}{B'' \text{ infrared}} = c \quad (1.9)$$

yields the speed of light. The most precise result obtained by such method was  $c = 299792.8 \pm 0.4$  km/s, the main limitation being dictated by the accuracy of the theory of band spectra of molecules.

### 1.3.1 The laser arrives: length-frequency equivalence and the birth of optical frequency metrology

Up to this point in both stories, that of timekeeping and  $c$ , only microwave radiation made its entrance. The future of metrology was changed fundamentally on 12 December 1960 when a small team at Bell Labs, led by Ali Javan, eventually found the right conditions for their Optical Maser to generate the self-sustained optical oscillation that was anticipated by Charles Townes and Arthur Schawlow in a classic paper of 1958 [29]. The emergence of the laser promised to open new scenarios in the field of metrology. Indeed, very soon the wavelength of visible radiation could be measured fairly well by Michelson or Fabry-Perot interferometers. This possibility enabled the development of laser frequency measurement programs at various national standards laboratories such as NBS at Boulder, NPL at Teddington, and National Research Council at Ottawa. In spite of the fact that lasers provided coherent frequency sources in the infrared and visible, optical frequencies could not immediately be measured with the required degree of accuracy. Specifically, two fundamental drawbacks had to be overcome. First, laser frequency stability had to be greatly improved. Indeed, in the case of the gas laser, although its short term linewidth was a few hundred Hz, over a long period, its frequency could vary within the Doppler and pressure broadened gain curve of the laser. By the late 1960s, lasers stabilized in frequency to atomic and molecular resonances were becoming reliable research tools and the development of the technique of saturated absorption had produced lasers with one-second fractional frequency instabilities as small as  $5 \cdot 10^{-13}$  [26]. Second, the laser optical frequency was much too high for conventional frequency measurement methods. To remove this limitation, the

approach taken was to synthesize signals at progressively higher and higher frequency using harmonic-generation-and-mixing (heterodyne) methods and to lock the frequency of a nearby oscillator or laser to the frequency of this synthesized signal. Photodiodes, as well as metal-insulator-metal (MIM) diodes, fabricated by adjusting a finely tipped tungsten wire against a naturally oxidized nickel plate, were used for harmonic generation and mixing [30]. With this approach, a frequency synthesis chain was constructed linking the microwave output of the cesium frequency standard to the optical region [31], so that the Boulder group could directly measure the frequency of a helium-neon laser stabilized against the 3.39- $\mu\text{m}$  (88 THz) transition of methane [32] (note that the frequency of the methane stabilized helium-neon laser is over 1000 times higher in frequency than that of the oscillator used in Froome's measurement). At the same time, the wavelength of the 3.39- $\mu\text{m}$  line of methane was measured with respect to the Kr-86 6057-Angstrom standard by using a frequency-controlled Fabry-Perot interferometer [33]. In this way, the laser eventually permitted to preserve the small interferometric errors associated with the short optical wavelength, while utilizing microwave frequencies which were still readily manipulated and measured. The extension of frequency measurements into the infrared portion of the electromagnetic spectrum had in a sense solved the dilemma raised raised by Froome's experiment: to measure the frequency, it is best to do the experiment not too far removed from the primary Cs frequency where extremely stable oscillators can be made and frequencies are easily measured with great accuracy. However, to measure the wavelength it is best to do the experiment close to the visible 86-krypton wavelength standard where wavelengths can be more easily compared and where diffraction problems are not severe. Table 1.1 summarizes the most significant milestones in the story of  $c$  measurements.

When the measurements were completed, the uncertainty limitation was found to be the asymmetry of the krypton line on which the definition of the meter was then based [34]. The experiment thus showed that the realization of the meter could be substantially improved through redefinition. This careful measurement resulted in a reduction of the uncertainty of the speed of light by a factor of nearly 100. The methods developed at NIST were replicated in a number of other laboratories [35, 36, 37], and the experiments were repeated and improved to the point where it was generally agreed that this technology could form the basis for a new definition of the meter. An important remaining task was the accurate measurement of still higher (visible) frequencies which could then serve as more practical realizations of the proposed new definition. The Boulder group again took the lead and provided the first direct measurement of the frequency of the 633 nm line of the iodine-stabilized helium-neon laser [38], as well as a measurement of the frequency of the 576 nm line in iodine [39].

These and similar measurements around the world (frequency and wavelength measurements were refined to the accuracy of few parts in  $10^{10}$  and in  $10^9$ , respectively) were the last ingredients needed to take up the redefinition of the meter. The product of the measured frequency and the wavelength yields a new, definitive value for the speed of light. The new (and current) definition of the meter, accepted by the 17th Conference Generale des Poids et Mesures in 1983, was quite simple and elegant:

---

*the meter is the length of the path traveled by light in vacuum during a time interval of 1/299,792,458 of a second*

---

A consequence of this definition is that the speed of light is now a defined constant, not to be measured again. In subsequent years, measurement of other stabilized-laser systems

added to the ways in which the meter could be realized. Furthermore, these experiments definitely demonstrated that, in order to obtain highly precise results, it is necessary to measure the frequency of light rather than its wavelength, marking the birth of optical frequency metrology.

Time-frequency and length-frequency equivalence principles are only two aspects of a more general trend in contemporary metrology, that of establishing measurement units based, rather than on artifacts, on atomic (quantum) standards or on fundamental constants [40]. These are invariant both on a practical scale and as far as can be measured in the laboratory. Let's deepen the significance of this process starting with the second and the meter.

Although the cesium frequency cannot presently be explicitly written in terms of fundamental constants because of the complexity of the atomic theory required, it is a quantum system that will have the stability associated with fundamental constants. The uncertainty in calculating this frequency is many orders of magnitude away from its measurement uncertainty. The Rydberg constant could be considered the natural fundamental constant-based unit of frequency. It is determined with a relative uncertainty of  $6.6 \cdot 10^{-12}$ , which is currently the limit at which an atomic frequency can be calculated from fundamental constants. The choice of the cesium definition was a good one in the sense that the technology, although superior to the alternative clocks of the day, still had much room for improvement, and the definition has endured to this day, during which time its practical realization has improved by five orders of magnitude. A clear example of the link between fundamental constants and the units is the adoption of the speed-of-light definition of the meter. The meter was originally defined as the length of a prototype meter bar intended to be  $1/10,000,000$  of the length of a quadrant of Earth. By 1960, the development of interferometry allowed an atomic redefinition of the meter in terms of the wavelength of light from a specific source, the krypton lamp (the meter was defined as 1650763.73 wavelengths of the orange-red emission line in the spectrum of krypton-86 atom in vacuum). With the invention of the laser, length measurement by interferometry was radically improved and the krypton standard was not accurate enough. The meter definition could then have been revised using the wavelength of a specified stabilized laser. However, the progress in understanding the metrological importance of the speed of light, along with the progress in its accurate measurements, led to the change from defining the meter in terms of the wavelength of light from a specific source, to a fundamental constant-based definition in which the speed of light is a defined quantity. The choice of the speed-of-light definition over the use of a particular stabilized laser should ensure that this definition will endure, whereas the krypton definition lasted only 23 years. In practice, a number of *recommended radiations*, that is, frequencies of particular stabilized lasers, are published accompanying the definition. This means that to realize the meter there is no need to measure the distance that light travels in  $1/299,792,458$  of a second by literally timing a light beam. One can, for example, continue to use a laser interferometer and measure the frequency of the laser used, or use a recommended stabilized laser and then use the relationship  $c = \lambda\nu$  (as well as corrections for refractive index, if the measurement is not done in a vacuum). In other words, the realization is a method that implements the definition by using the known laws of physics; it allows the experimental production of a known quantity of the same kind as the one defined, but the method used may be dissimilar to the one in the definition. We close this discussion by mentioning another clarifying example, namely the definition of the volt. Electrical quantum metrology started in 1962 when Josephson predicted that in the presence of an applied microwave field, a direct superconducting tunnelling current could pass between superconductors separated by an insulating barrier. This current can only pass when the voltage  $V$  across the barrier satisfies the relationship

**TABLE 1.1**List of the most significant terrestrial measurements of  $c$ .

Year	Investigator	Method	Value (km/s)
1849	Fizeau	Toothed wheel	315,300 [12]
1862	Foucault	Revolving mirror	$298,000 \pm 500$ [13]
1906	Rosa and Dorsey	EM constants	$299,710 \pm 30$ [18]
1927	Michelson	Revolving mirror	$299,796 \pm 4$ [13]
1928	Karolus and Mittelstaedt	Kerr cell	$299,778 \pm 20$ [13]
1948	Essen and Gordon-Smith	Cavity resonator	$299,792 \pm 9$ [18]
1949	Aslakson	Radar	$299,792.3 \pm 2.4$ [22]
1952	Bergstrand	Geodimeter	$299,793.1 \pm 0.2$ [25]
1958	Froome	Millimeter-wave interferometry	$299,792.50 \pm 0.10$ [27]
1965	Rank	Spectroscopy	$299,792.8 \pm 0.4$ [28]
1972	Evenson	Direct frequency and wavelength measurement of a laser	$299,792.4562(11)$ $299,792.4587(11)$ [34]

*Note:* The two values in Evenson’s measurement were due to the asymmetry in the krypton 6057-Ångstrom line defining the meter. Except for the first two (in air), all the listed results refer to the value in vacuum.

$$2eV = nh\nu \quad (1.10)$$

where  $e$  is the electron charge,  $h$  the Planck constant,  $\nu$  the applied frequency, and  $n$  an integer. It was recognized that voltage standards could be based on this effect. A number of experiments found no corrections to expression 1.10 or dependence on material or experimental conditions at a level of up to parts in  $10^{16}$ . In 1972, a number of countries used the Josephson effect to maintain the volt and agreed on an assigned value for  $2e/h$  so that their voltages were in agreement. They are not necessarily the correct SI value; hence, the agreed-upon value is referred to as a *representation* of the volt. Again a frequency measurement played a crucial role.

### 1.3.2 Role of $c$ in fundamental physics

Besides representing an essential pillar of frequency metrology, the parameter  $c$  is ubiquitous in contemporary physics, entering many contexts that are apparently disconnected from the notion of light itself. Our thoughts soon turn to the very famous second postulate of Special Relativity: “The velocity  $c$  of light in vacuum is the same in all inertial frames of reference in all directions and depend neither on the velocity of the source nor on the velocity of the observer”. The theory of Special Relativity explores the consequences of this invariance of  $c$  with the assumption that the laws of physics can be written in the same form in all inertial frames (first postulate). Declared by Einstein in 1905, after being motivated by Maxwell’s theory of electromagnetism and the lack of evidence for the luminiferous ether, the invariance of the speed of light and its isotropy has been consistently confirmed by many experiments over the years. Other experimentally verified implications of Special Relativity include length contraction (moving objects shorten), and time dilation (moving clocks run slower). The factor  $\gamma$  by which lengths contract and times dilate is known as the Lorentz factor and is given by  $\gamma = 1/\sqrt{1 - (v/c)^2}$ , where  $v$  is the speed of the object. Special Relativity also establishes that the energy of an object with rest mass  $m$  and speed  $v$  is given by  $E = \gamma mc^2$ . Since the  $\gamma$  factor approaches infinity as  $v$  approaches  $c$ , it would take an infinite amount of energy to accelerate an object with mass to the speed of light. The speed of light is therefore the upper limit for the speeds of objects with positive

rest mass. Experimental Tests of General Relativity for the most part also verify Special Relativity, since the laws of the latter are included as part of the former via the principle of consistency. For a tutorial introduction to Special and General Relativity the reader may refer to [41], while an updated list of experimental verifications inferred from advanced frequency measurements will be given in Chapter 8.

As a clarification of what was just discussed, it is worth adding that, in principle, we should distinguish between the electromagnetism constant  $c_{EM} = 1/\sqrt{\epsilon_0\mu_0}$ , and the space-time constant  $c_{ST}$  appearing in the Lorentz transformation that is at the basis of the formulation of Special Relativity [42]. In general, for example, the celebrated equation unifying the concepts of energy and mass should be written in the form  $E = mc_{ST}^2$ .  $c_{EM}$  agrees with  $c_{ST}$  insofar as the mass of the photon is zero. If we were to show experimentally that the photon has non-zero mass, then the standard derivation of relativity from electromagnetism would have to be abandoned. Incidentally, extensions of quantum electrodynamics (QED) in which the photon has a mass have been considered [43]. In such a theory, the photon speed would depend on its frequency. No variation of the speed of light with frequency has been observed in rigorous testing, putting stringent limits on the mass of the photon [44]. The same is true for gravitational waves: their speed in vacuum  $c_{GW}$ , is equal to  $c_{ST}$  as long as we assume that general relativity is valid. Again, if we were able to formulate a theory with light massive gravitons, then the speed of propagation of gravity might be different from  $c_{ST}$ . Finally, the space-time-matter constant  $c_E$ , introduced by Einstein to describe coupling of gravity to matter, coincides by definition with  $c_{ST}$  only in the context of general relativity.

As mentioned, Einstein's relativity treats space and time as a unified structure known as space-time (with  $c$  relating the units of space and time) and requires that physical theories satisfy a special symmetry called Lorentz invariance, whose mathematical formulation contains precisely the parameter  $c$ . Lorentz invariance is an almost universal assumption for modern physical theories, such as quantum electrodynamics, quantum chromodynamics, and the Standard Model of particle physics. One consequence is that  $c$  is the speed at which all massless particles and waves, not only light, must travel. This result is constantly put to the test in different areas of experimental physics. In this respect, great emphasis was given to a high-energy physics experiment according to which beams of neutrinos, fired through the ground from Cern near Geneva to the Gran Sasso lab in Italy 450 miles (720 km) away, seemed to arrive sixty billionths of a second earlier than they should if travelling at the speed of light in a vacuum [45]. Subsequently, however, a discrepancy between the clocks at Cern and Gran Sasso was discovered to be at the root of the observed faster-than-light results. In the future, use of an optical fiber, as opposed to the GPS system used at the moment, should ensure a more accurate synchronization of the two clocks. This gives even more prominence, if any were needed, to the scope of time and frequency measurements.

The most striking feature of Einstein's relativity is undoubtedly the upper limit to velocity of any physical object set by  $c$ , albeit there are situations in which it may seem that matter, energy, or information travels at speeds greater than  $c$ . A first amazing example is the following. Think about how fast a shadow can move. If you project the shadow of your finger using a nearby lamp onto a distant wall and then wag your finger, the shadow will move much faster than your finger. If your finger moves parallel to the wall, the shadow's speed will be multiplied by a factor  $D/d$  where  $d$  is the distance from the lamp to your finger, and  $D$  is the distance from the lamp to the wall. If the wall is very far away, the movement of the shadow will be delayed because of the time it takes light to get there, but the shadow's speed is still increased by the same ratio. The speed of a shadow is therefore not restricted to be less than the speed of light. Unfortunately, the shadow is not a physical object and it is not possible to send information on a shadow. Also, certain quantum effects appear to be transmitted instantaneously and therefore faster than  $c$ . Among these we mention the

celebrated Einstein-Podolsky-Rosen (EPR) paradox [46], the Hartman effect [47], and the Casimir effect [48]. It has been pointed out, however, that none of these effects can be used to send information.

Other examples come from the astrophysical/cosmological context. According to Hubble's law, two galaxies that are a distance  $D$  apart are moving away from each other at a speed  $HD$ , where  $H$  is Hubble's constant. This interpretation implies that two galaxies separated by a distance greater than  $c/H$  must be moving away from each other faster than the speed of light. Actually, the modern viewpoint describes this situation differently: general relativity considers the galaxies as being at rest relative to one another, while the space between them is expanding. In that sense, the galaxies are not moving away from each other faster than the speed of light; they are not moving away from each other at all! This change of viewpoint is not arbitrary; rather, it agrees with the different but very fruitful view of the universe that general relativity provides. So the distance between two objects can be increasing faster than light because of the expansion of the universe, but this does not mean, in fact, that their relative speed is faster than light.

It is worth stressing that all the experiments performed to date have confirmed that it is impossible for information or energy to travel faster than  $c$ . One simple general argument for this follows from the counter-intuitive implication of special relativity known as the relativity of simultaneity. If the spatial distance between two events A and B is larger than the time interval between them multiplied by  $c$ , then there are frames of reference in which A precedes B, others in which B precedes A, and others in which they are simultaneous. As a result, if something were travelling faster than  $c$  relative to an inertial frame of reference, it would be travelling backwards in time relative to another frame, and causality would be violated. In such a frame of reference, an *effect* could be observed before its *cause*. Such a violation of causality has never been recorded.

In a medium, light usually does not propagate at a speed equal to  $c$ ; furthermore, different types of light wave will travel at different speeds. The speed at which the individual crests and troughs of a plane wave (a wave filling the whole space, with only one frequency) propagate is called the phase velocity  $v_p$ . So, while in *vacuum* we have  $c = v_p = \lambda\nu$ , in a medium we have  $c/n(\nu) = v_p = \lambda\nu$  where  $n(\nu)$  is the refractive index of the medium (in general it also depends on the intensity, polarization direction of propagation,...). In actual circumstances such idealized solutions do not arise. Even in the most monochromatic light source or the most sharply tuned radio transmitter or receiver, one deals with a finite spread of frequencies or wavelengths. Since the basic equations are linear, it is in principle an elementary matter to make the appropriate linear superposition of solutions with different frequencies where each monochromatic component has its own phase velocity. Consequently, there is a tendency for the original coherence to be lost and for the pulse to become distorted in shape. At the very least, we might expect it to propagate with a rather different velocity from, say, the average phase velocity of its component waves. The general case of a highly dispersive medium or a very sharp pulse with a wide spread of wave numbers is difficult to treat. But the propagation of a pulse which is not too broad in its wave-number spectrum, or a pulse in a medium for which the frequency depends weakly on wave number, can be handled in an approximate way. In this case, it can be shown that the transport of energy occurs with the group velocity

$$v_g = \left. \frac{d\omega}{dk} \right|_{k=k_0} \quad (1.11)$$

where  $\omega(k) = ck/n(k)$  describes the dispersion of the material and  $k_0$  is the center wavenumber of the packet. In general, however, the behavior of the wave packet is much more complicated and the group velocity above defined does not identify with the infor-

mation velocity. In transparent materials, the refractive index generally is greater than 1, meaning that the phase velocity is less than  $c$ . In other materials, it is possible for the refractive index to become smaller than 1 for some frequencies; in some exotic materials it is even possible for the index of refraction to become negative. This reflects in turn on the value of the group velocity. Even in those cases where group velocities exceeding  $c$  are observed, it is still valid, according to causality, that it is impossible to transmit information faster than the speed of light in vacuum. Indeed, when the notion of front velocity is introduced and the principle of causality is accounted for, it can be rigorously shown in the frame of classical electrodynamics that information travels at the front velocity that is actually limited to  $c$  [49].

The finiteness of the speed of light has implications for the whole realm of sciences and technologies. In some cases, it is considered as a hindrance. For instance, being the upper limit of the speed with which signals can be sent,  $c$  provides a theoretical upper limit for the operating speed of microprocessors. In supercomputers, the speed of light imposes a limit on how quickly data can be distributed among processors. If a processor operates at 1 GHz, a signal can only travel a maximum of about 30 centimeters in a single cycle. Processors must therefore be placed close to each other to minimize communication latencies; this represents a trade-off with cooling needs. If clock frequencies continue to increase, the speed of light will eventually become a limiting factor for the internal design of single chips. In other cases, the finiteness of  $c$  turns out to be useful. For instance, the finite speed of light is important in astronomy. Due to the vast distances involved, it can take a very long time for light to travel from its source to Earth. For example, photographs taken today in the Hubble Ultra Deep Field capture images of the galaxies as they appeared 13 billion years ago, when the universe was less than a billion years old. The fact that more distant objects appear to be younger, due to the finite speed of light, allows astronomer to infer the evolution of stars, galaxies, and of the universe itself. Moreover, position measurements by GPS systems rely on the finiteness of  $c$ .

We close this section by observing that it is generally assumed that fundamental constants such as  $c$  have the same value throughout space-time, meaning that they do not depend on location and do not vary with time. However, it has been suggested in various theories that the speed of light may have changed over time. No conclusive evidence for such changes has been found, but this remains a crucial subject of ongoing metrological research [50].

As we will see during this book, and in particular in Chapter 8, advanced laser-based measurements in the frequency and time domains promise to give a new insight into many of the aforementioned issues.

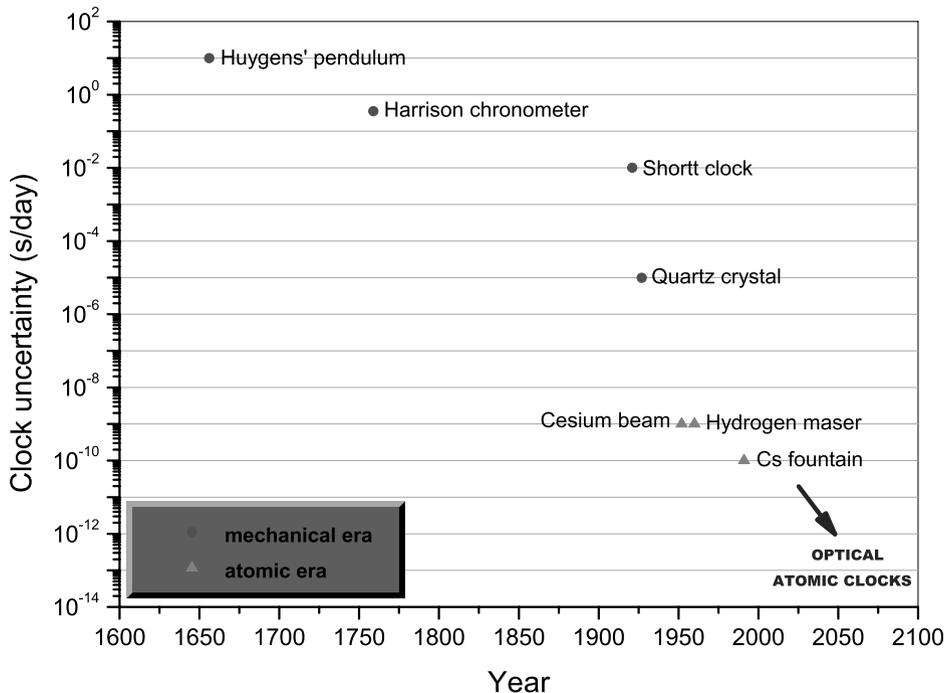
---

## 1.4 In the end, time and light met up again: optical atomic clocks and outline of the book

As we have seen, the advent of the laser played a central role in the statement of the meter definition marking, in fact, the beginning of optical frequency metrology. In the following years, the laser became an invaluable source in many research fields. Today, it is the *true light* which enlightens every advanced frequency metrology experiment. Additionally, from the sixties to the present, three major developments were triggered by the laser in the field of fundamental research: ultra-high-resolution spectroscopy, the field of trapping/cooling of atoms, and the realization of optical frequency comb synthesizers based on femtosecond

(fs) mode-locked lasers. In turn, these three discoveries have played a crucial role in the timekeeping story leading to the realization of the current optical atomic clocks. Although present-day cesium microwave frequency standards perform at an already remarkable level (fractional uncertainty below 1 part in  $10^{15}$ ), a new approach to timekeeping based on optical atomic transitions promises still greater improvements. According to the given definition of  $Q$ , by using optical ( $\nu_0 \sim 10^{15}$  Hz) rather than microwave ( $\nu_0 \sim 10^{10}$  Hz) frequencies, optical standards should be considerably more accurate. Also several key frequency shifts are fractionally much smaller in the optical domain and their investigation will be greatly accelerated by the much smaller instability of the optical standards. A projection of the fractional uncertainty achievable in the new era of optical atomic clocks is made in Figure 1.8 which displays some of the major milestones in the improvement of clocks over the past 400 years.

The potential advantages of optical atomic clocks were recognized in the early days of frequency standards. However, optical standards did not truly begin to experience these potential gains until the past decade, when the above three fields enjoyed an extraordinary growth. First, huge advances in laser cooling techniques made it possible to cool a variety of atoms and ions (including those with narrow clock transitions) to millikelvin temperatures and below. The use of laser-cooled atomic samples enabled, in turn, the extended interaction times required to observe a narrow transition linewidth. To resolve such narrow linewidths, probe lasers need to be spectrally pure. Recent improvements in laser stabilization based on environmentally isolated optical reference cavities have enabled laser linewidths at the subhertz level to be achieved [51]. Finally, and perhaps most critically, compact and reliable optical frequency comb synthesizers (OFCSs) for counting optical frequencies (linking them,



**FIGURE 1.8**

Major milestones in the improvement of clocks over the past 400 years, including the projected fractional uncertainty of next-generation optical atomic clocks. (Adapted from [5].)

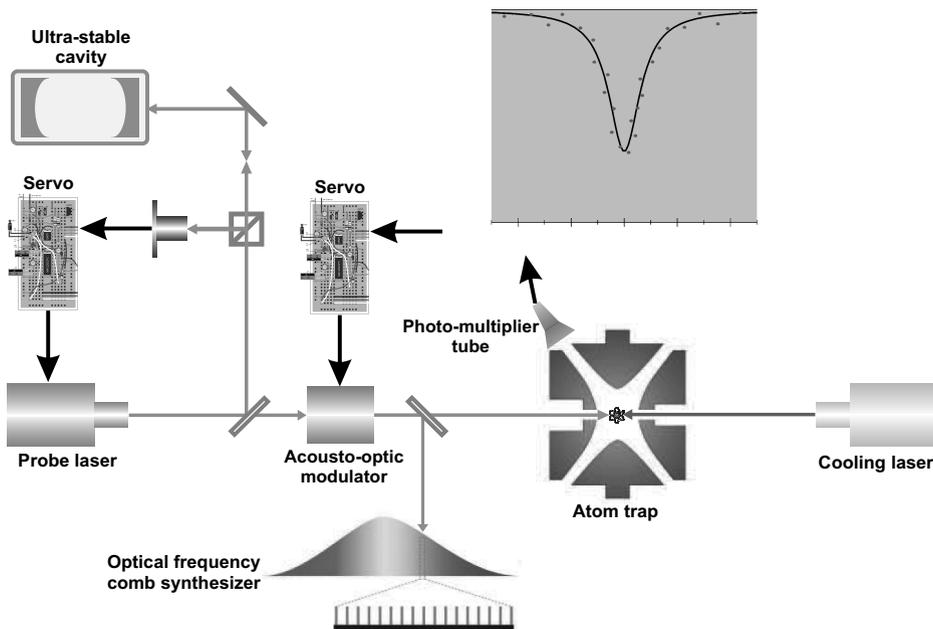
phase coherently and in a single step, to other frequencies in the microwave and optical domains) became available replacing the old and cumbersome frequency chains [52, 53].

Being the most recent and sophisticated objects presented in this book, optical atomic clocks will be our guiding star along the whole treatment. Thus, just to point out all the relevant ingredients, a paradigmatic scheme is given in Figure 1.9 [5].

Schematically, the probe laser, whose frequency is pre-stabilized against an optical cavity, is used to excite transitions in a laser-cooled, trapped sample of atoms (either neutral or ions). A servo system uses the signal from the quantum absorbers to keep the probe laser frequency centered on resonance. Light is sent to the OFCS, which enables counting of the clock cycles. In this picture, as 5000 years ago, a light source (the laser plays here the role of Sun) interrogates an oscillatory phenomenon (the atomic resonance is now the equivalent of Earth's rotation). The abysmal difference lies in the much higher operation frequency (and in the much greater stability compared to Earth's motion), as well as in the resonant character of light-matter interaction. So, optical atomic clocks represent the happy end of the history of Light and Time.

Now, the time is approaching when optical frequency standards will have accuracies and stabilities superior to the best microwave cesium standards. Then it will be necessary to revisit the definition of the second. There are a number of candidate optical frequency standards, but at present no particular standard is clearly superior to the others. The time lag in adopting a new atomic standard mainly reflects the work that is necessary to ensure that one specifically selected system is indeed superior.

As illustrated in these first sections, the history of physics shows that, when the accuracy of measurements is improved, new physics may be discovered and explored. Throughout history, at several moments, the discovery or development of a new type of oscillator with



**FIGURE 1.9**

Schematic layout of an optical atomic clock.

improved performance has meant a huge step forward in our knowledge of physics or even coincided with a scientific revolution. For example, the advent of pendulum clocks provided experimental verification of Galilean laws of mechanics, while, more recently, observation of astronomical oscillators like binary pulsars confirmed many of the predictions of the General Relativity. In the same way, optical atomic clocks are expected to improve our knowledge in fundamental physics.

The emergence of the laser, moreover, made it essential to distinguish its *special* light from the more *incoherent* radiation emitted by hot bodies. This induced Roy Glauber to utilize the quantum theory to describe the properties of light and how these can be observed. His work laid the foundations for the field of research today called Quantum Optics, and earned him the Nobel Prize awarded in 2005 [54]. Quantum Optics is a very wide field and only those aspects that will intersect the main path of our book will be dealt with. In technical applications, the quantum effects are often very small. The field state is chosen so that it can be assigned well-defined phase and amplitude properties. In laboratory measurements, too, the uncertainty of quantum physics seldom sets the limit. But the uncertainty that nevertheless exists appears as a random variation in the observations. This *quantum noise* sets the ultimate limit for the precision of optical observations. In high-resolution frequency measurements, quantum amplifiers, and frequency standards, it is in the end only the quantum nature of light that sets a limit for how precise our apparatuses can be. Such ultimate limits have been explored in recent years by two main protagonists, John L. Hall and Theodor W. Hänsch. For their unceasing and illuminating research within the field of laser-based precision spectroscopy culminating with the realization of OFCSs, they received the other half of the Nobel Prize in Physics in 2005. It now seems possible, with the frequency comb technique, to make frequency measurements in the future with a precision approaching one part in  $10^{18}$ . This will soon lead to actualize the introduction of a new, optical standard clock. What phenomena and measuring problems can take advantage of this extreme precision? Just to mention a few examples, more exact satellite-based navigation systems will become available and novel applications in telecommunication may emerge. Enormous benefits will also come out for navigation on long space journeys and for space-based telescope arrays that are looking for gravitational waves or making precision tests of the theory of relativity. Besides technological applications, this improved measurement precision may also be used in fundamental physical studies like those related to the antimatter-ordinary matter connection (spectroscopic studies of anti-hydrogen), to parity violation in chiral molecules, as well as to the search for possible changes in the constants of Nature over time. These and other fascinating issues will be discussed in Chapter 8.

At the end of this introductory chapter, we hope that, by grasping the concepts here proposed, the reader is in tune with the authors to better appreciate the logical organization of the book. In a sense, we are going to give, chapter by chapter, a detailed description of all the key elements on which the operation of optical atomic clock hinges. Here is the outline:

In **Chapter 2** we shall discuss the general basic features of harmonic oscillators and introduce the mathematical background for their characterization. An introductory overview of the most commonly used techniques for measuring and suppressing the phase noise in oscillators will be also given, together with a few elementary notions on feedback systems. The issue of accurate optical frequency synthesis will be also addressed in the last sections.

**Chapter 3** is entirely devoted to passive resonators working both in the microwave and optical domain. Greater emphasis is given to the latter: besides traditional bulk resonators and their most updated developments, guided cavities based on optical fibers as well as micro-resonators relying on whispering gallery modes will be treated into a certain detail.

In **Chapter 4** we shall deal with continuous-wave (cw) coherent radiation sources.

After a short introduction on masers, we will focus on lasers by illustrating, firstly, some key aspects of their operation and a number of fundamental properties in their output. Then, a wide range of laser-based systems will be presented. A few clarifying examples of intensity and frequency stabilized laser sources will close the chapter.

In **Chapter 5** we shall provide a comprehensive treatment of high-resolution and high-sensitivity spectroscopic techniques for ultraprecise frequency measurements. Then, optical frequency standards utilizing either absorption cells or atomic/molecular beams will be described.

In **Chapter 6** the issue of time and frequency measurements with pulsed laser systems will be addressed. Starting with general mode-locking theory and mechanisms, advanced schemes for optical frequency comb synthesis (from mode-locked lasers) and relative stabilization will be presented. The extension of OFCSs into novel spectral regions, from the extreme ultraviolet (XUV) to the far-infrared (FIR), will be also discussed.

**Chapter 7** is mainly devoted to microwave frequency standards. This category comprises high-quality crystal-based oscillators, high-performance hydrogen masers, and cutting-edge fountains based on cold alkali atoms. Being fundamental to the understanding of atomic standards, a propaedeutic review on trapping/cooling techniques for atoms, ions, and molecules is also provided. In the last part, a brief account on time and frequency dissemination (including optical frequency transfer) is given.

**Chapter 8** starts with optical atomic clocks, ranging from the more established ones, based on single laser-cooled trapped ions, to the newest systems relying on neutral atoms trapped in an optical lattice. Then, based on the wide phenomenology explored thus far, possible research prospects in the field of time and frequency measurements are drawn for the next future.

# 2

---

## *Characterization and control of harmonic oscillators*

---

Human time does not rotate in a circle,  
but moves fast in a straight line. That is  
why man can not be happy, because  
happiness is the desire for repetition.

*Milan Kundera - The Unbearable  
Lightness of Being*

The fish in the water is silent, the animals  
on the earth are noisy, the bird in the air  
is singing. But man has in himself the  
silence of the sea, the noise of the earth  
and the music of the air.

*Rabindranath Tagore*

---

### 2.1 The ideal harmonic oscillator

The purpose of this chapter is to acquaint the reader with the basic concepts and the mathematical tools that are necessary to address and better understand the contents which are at the heart of this book. From the previous chapter we learnt that oscillators are ubiquitous in the field of frequency metrology. We start by reviewing the most relevant features of harmonic (sinusoidal) oscillators which, although known from General Physics, deserve here a discussion devoted to our specific context. We focus on two archetypes: the pendulum and the RLC-series circuit. The former is the paradigm of mechanical oscillators, while the latter embodies the electrical ones. According to the specific property we are interested in, from time to time we will resort to one or the other system, but the conclusions will always be general.

For a point pendulum supported by a massless and inextensible cord of length  $l$ , the equation of motion is given by

$$ml^2\ddot{\theta} = -mgl\sin\theta - \beta\dot{\theta} \quad (2.1)$$

where  $m$  is the bob mass,  $g$  the local acceleration of gravity,  $\theta$  the angle between the cord and the vertical, and  $\beta$  accounts for the overall friction (basically the resistance by the air and the escapement). For infinitesimal displacements, we replace  $\sin\theta$  by  $\theta$  and get the following second-order linear differential equation

$$\ddot{\theta} + \frac{\beta}{ml^2}\dot{\theta} + \frac{g}{l}\theta = 0 \quad (2.2)$$

In the case of the RLC series circuit, it is the mesh current  $I$  obey an equation of the same form

$$\ddot{I} + \frac{R}{L}\dot{I} + \frac{1}{LC}I = 0 \quad (2.3)$$

with the obvious meaning of symbols. Here, the resistor  $R$  provides dissipation and is thus responsible for damping, whereas the LC tank sets the oscillation frequency.

Another celebrated example, which actually falls into the category of mechanical oscillators, is offered by the Lorentz model of the atom. Predating the emergence of quantum mechanics, such a classical picture was the first attempt to explain atomic spectra. It rests on the idea that an electron of mass  $m$  and charge  $-e$  is bound to the nucleus (charge  $+e$ ) by a restoring force that is proportional to the displacement (Hooke's law). To account for the fact that an excited atom loses its energy by emitting electromagnetic radiation, a damping mechanism for the oscillation is also considered by including a viscous term (proportional to velocity) into the equation of motion. Therefore, the electron position turns out to be governed by the law

$$\ddot{x} + \frac{k}{m}x + \frac{\alpha}{m}\dot{x} = 0 \quad (2.4)$$

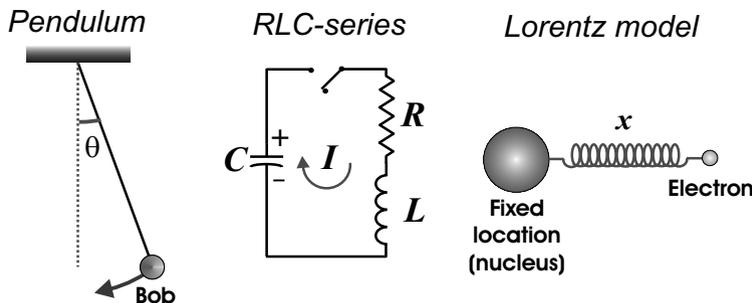
It is easily recognized (see Figure 2.1) that, with the appropriate identifications, Equations 2.2, 2.3, 2.4 are all of the form

$$\ddot{y} + 2\Gamma\dot{y} + \omega_0^2 y = 0 \quad (2.5)$$

A little more general equation of motion is obtained when a driving term is added to compensate for the slowing down of the oscillation

$$\ddot{y} + 2\Gamma\dot{y} + \omega_0^2 y = D\cos(\omega_D t) \quad (2.6)$$

The driving term may arise from the interaction with an electromagnetic monochromatic plane wave in the case of the Lorentz oscillator, or simply be the periodic push in a pendulum as well as the AC generator in the RLC circuit. Apart from the examples just mentioned, Equation 2.6 appears in a number of different systems ranging from solid-state turbulence to soliton dynamics, from Josephson junctions to phase-locked loops [55]. In order to find a



**FIGURE 2.1**

Pendulum, RLC-series circuit, and Lorentz oscillator as paradigmatic examples of harmonic oscillators. The following identifications return Equation 2.5 for each of the three cases. Pendulum case:  $y \equiv \theta$ ,  $\omega_0^2 \equiv g/l$ ,  $2\Gamma \equiv \beta/(ml^2)$ ; RLC circuit:  $y \equiv I$ ,  $\omega_0^2 \equiv 1/(LC)$ ,  $2\Gamma \equiv R/L$ ; Lorentz oscillator:  $y \equiv x$ ,  $\omega_0^2 \equiv k/m$ ,  $2\Gamma \equiv \alpha/m$ .

solution for it, first consider the homogeneous Equation 2.5. In the case  $\omega_0^2 > \Gamma^2$ , from the associated characteristic equation one finds the solution

$$y(t) = e^{-\Gamma t} \left( E e^{i\omega'_0 t} + F e^{-i\omega'_0 t} \right) \quad (2.7)$$

where  $\omega'_0 = \sqrt{\omega_0^2 - \Gamma^2}$  and the constants  $E$  and  $F$  are found by imposing the initial conditions  $y(t=0) = \xi$  and  $\dot{y}(t=0) = \eta$ . By defining the quantities  $A_1 = A \sin \varphi$  and  $A_2 = A \cos \varphi$ , Equation 2.7 can be re-written in the more convenient form

$$y(t) = A e^{-\Gamma t} \sin(\omega'_0 t + \varphi) \equiv A e^{-\omega_0 g t} \sin(\omega_0 \sqrt{1 - g^2} \cdot t + \varphi) \quad (2.8)$$

where  $y(t=0) = A \sin \varphi = \xi$ ,  $\dot{y}(t=0) = A [-\Gamma \sin \varphi + \omega'_0 \cos \varphi] = \eta$ , and  $g \equiv \Gamma/\omega_0$ . Equation 2.8 describes the well-known case of a damped harmonic oscillator (see upper of Figure 2.2) which, in the limit  $\Gamma \rightarrow 0$ , reduces to

$$y(t) = A \sin(\omega_0 t + \varphi) \quad (2.9)$$

To illustrate some interesting properties in the frequency domain, let us take the Fourier transform of Equation 2.8 with  $\xi = 0$

$$\begin{aligned} \hat{y}(\omega) &= \int_0^\infty A e^{-\Gamma t} \sin \omega'_0 t e^{-i\omega t} dt = \\ &= \frac{A}{2i} \int_0^\infty \left\{ e^{[i(\omega'_0 - \omega) - \Gamma]t} - e^{[-i(\omega'_0 + \omega) - \Gamma]t} \right\} dt = \\ &= \frac{A}{2i} \left[ \frac{1}{\Gamma - i(\omega'_0 - \omega)} - \frac{1}{\Gamma + i(\omega'_0 + \omega)} \right] \simeq \\ &= \frac{\frac{A}{2i}}{\Gamma - i(\omega'_0 - \omega)} \end{aligned} \quad (2.10)$$

where the lower integration limit has been changed from  $-\infty$  to 0 since  $y(t) = 0$  for  $t \leq 0$ , and the last equality is valid close to the resonance, that is for  $\omega - \omega'_0 \ll \omega'_0$ . The response function of the oscillator is thus a Lorentzian profile

$$|\hat{y}(\omega)|^2 = \frac{A^2/4}{\Gamma^2 + (\omega'_0 - \omega)^2} \quad (2.11)$$

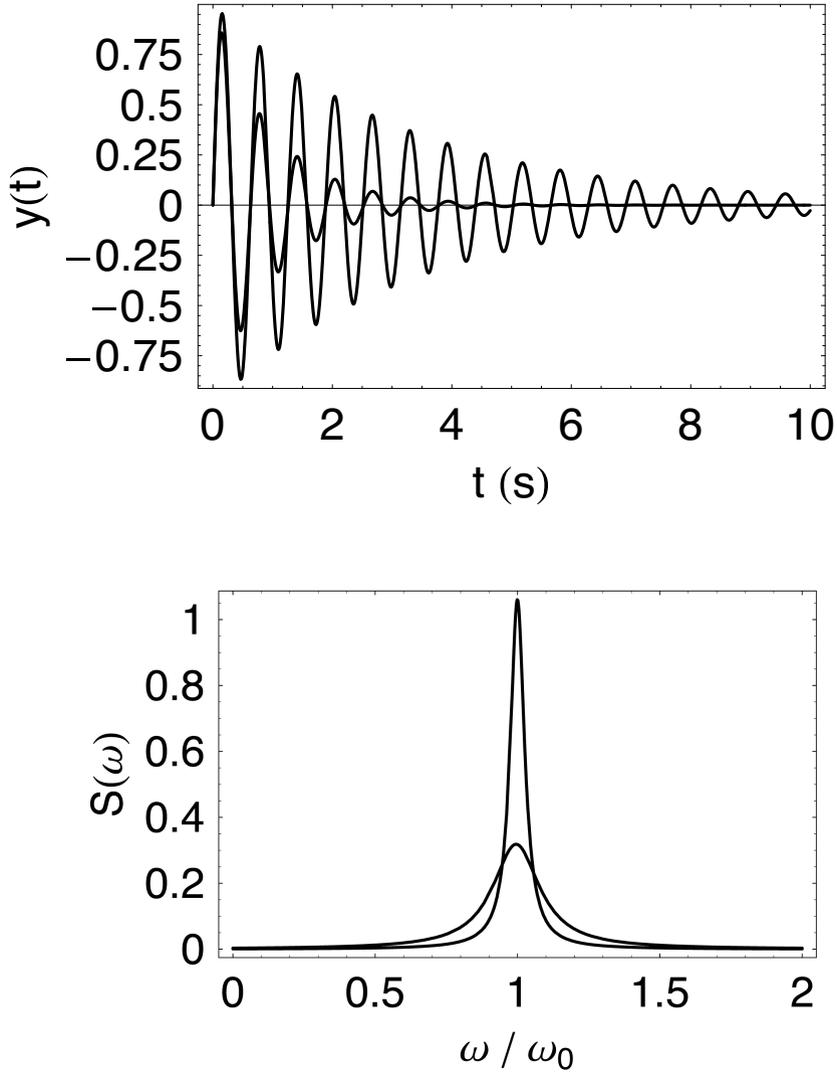
with a full width at half maximum  $FWHM = 2\Gamma$ . If  $y$  is interpreted as the electron position in an atom emitting or absorbing radiation, it can be convenient to find the constant  $A$  by normalizing the spectrum 2.11 such that  $\int_{-\infty}^{+\infty} |\hat{y}(\omega)|^2 d\omega = 1$ . This returns

$$S(\omega) \equiv |\hat{y}(\omega)|^2 = \frac{1}{\pi} \frac{\Gamma}{\Gamma^2 + (\omega'_0 - \omega)^2} \equiv \frac{1/(\pi\omega_0)}{g + (1/g)(\sqrt{1 - g^2} - \zeta)^2} \quad (2.12)$$

where  $\zeta \equiv \omega/\omega_0$ . Equation 2.12 is plotted in the lower frame of Figure 2.2 for two different values of  $g$ .

Note that, in the limit  $\Gamma \rightarrow 0^+$ , corresponding to the response curve of the ideal undamped harmonic oscillator (Equation 2.9), Equation 2.12 is one of the representations of the Dirac delta function  $\delta(\omega_0)$ . Now, the general solution of Equation 2.6 is found by simply adding Equation 2.8 to a particular solution which we will seek in the form of

$$y(t) = a \cdot \cos(\omega_D t + \psi) \quad (2.13)$$



**FIGURE 2.2**

Representation of a damped harmonic oscillator (with  $\omega_0 = 10$  Hz,  $A = 1$ , and  $\varphi = 0$ ) in time and frequency domain for two different values of  $g \equiv \Gamma/\omega_0$  (0.1 and 0.03), according to Equation 2.8 and Equation 2.12.

By substitution of Equation 2.13 into Equation 2.6 we obtain the following system of equations

$$\begin{cases} -a\omega_D^2 \cos \psi - 2a\Gamma\omega_D \sin \psi + a\omega_0^2 \cos \psi - D = 0 \\ a\omega_D^2 \sin \psi - 2a\Gamma\omega_D \cos \psi - a\omega_0^2 \sin \psi = 0 \end{cases} \quad (2.14)$$

whose solution yields

$$\psi = \arctan\left(\frac{2\Gamma\omega_D}{\omega_D^2 - \omega_0^2}\right) \quad (2.15)$$

$$a = \frac{-D}{\sqrt{4\Gamma^2\omega_D^2 + (\omega_D^2 - \omega_0^2)^2}} \quad (2.16)$$

Therefore, by introducing the phase  $\psi = \vartheta + \frac{\pi}{2}$ , the general solution of Equation 2.6 is finally obtained as

$$y(t) = Ae^{-\Gamma t} \sin\left(\sqrt{\omega_0^2 - \Gamma^2} \cdot t + \varphi\right) + \frac{D}{\sqrt{4\Gamma^2\omega_D^2 + (\omega_D^2 - \omega_0^2)^2}} \sin(\omega_D t + \vartheta) \quad (2.17)$$

Under steady-state conditions ( $t \rightarrow \infty$ ), the oscillator output, no more damped, is described by

$$\begin{cases} y(\omega_D) = Y_0(\omega_D) \sin[\omega_D t + \Phi(\omega_D)] \\ Y_0(\omega_D) = \frac{D}{\sqrt{4\Gamma^2\omega_D^2 + (\omega_D^2 - \omega_0^2)^2}} \equiv \frac{D_0}{\sqrt{4g^2\iota^2 + (\iota^2 - 1)^2}} \\ \Phi(\omega_D) = \arctan\left(\frac{2\Gamma\omega_D}{\omega_D^2 - \omega_0^2}\right) - \frac{\pi}{2} \equiv \arctan\left(\frac{2g\iota}{\iota^2 - 1}\right) - \frac{\pi}{2} \end{cases} \quad (2.18)$$

where  $\iota \equiv \omega_D/\omega_0$  and  $D_0 \equiv D/\omega_0^2$ . The oscillation amplitude  $Y_0(\omega_D)$  is maximum for  $\omega_D = \omega_0$ ; the corresponding phase is  $\Phi(\omega_D = \omega_0) = 0$  (see Figure 2.3). This example makes clear the character of the so-called resonance phenomenon between an external driving source and an oscillator with its own characteristic frequency. An expression for the *FWHM* of such resonance curve can be given by finding an approximate solution for the equation  $[Y_0^2(\omega_D = \omega_0)]/2 = Y_0^2(\omega_D)$ . This provides

$$4\Gamma^2\omega_D^2 + (\omega_D^2 - \omega_0^2)^2 = 8\Gamma^2\omega_0^2 \quad (2.19)$$

which, putting  $s = \omega_D^2 - \omega_0^2$ , is equivalent to a quadratic algebraic equation in  $s$

$$s^2 + 4\Gamma^2s - 4\Gamma^2\omega_0^2 = 0 \quad (2.20)$$

which returns

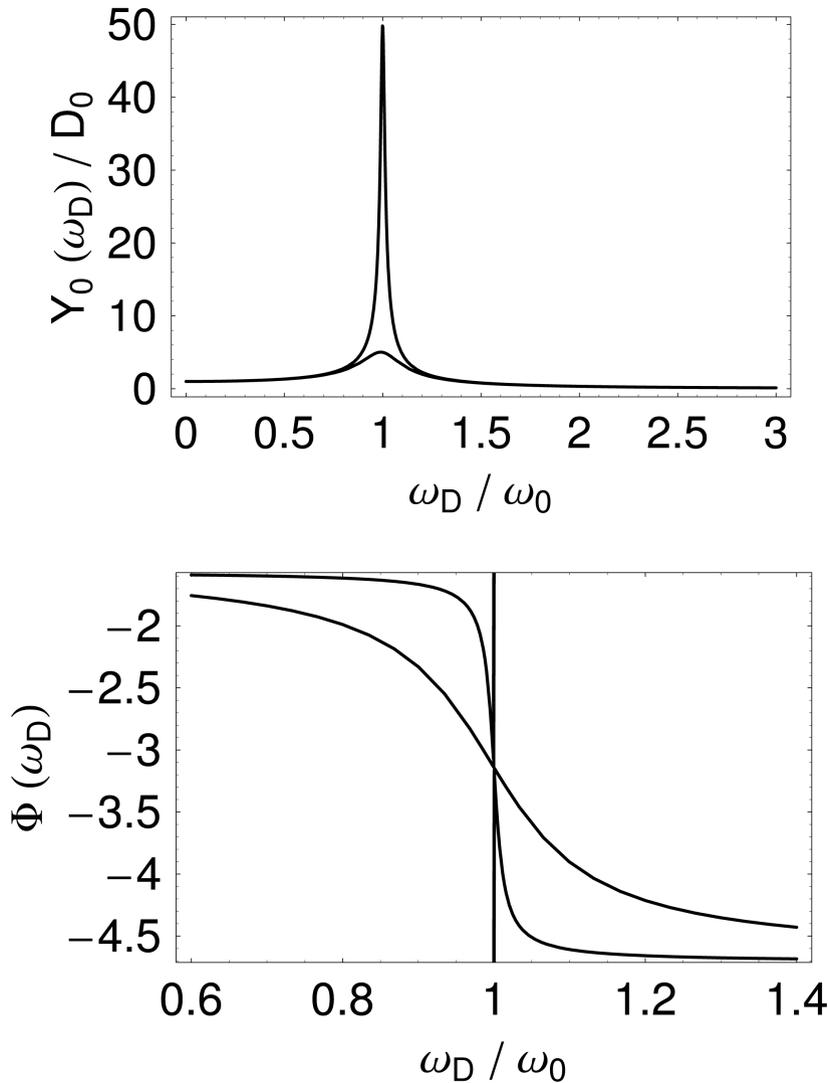
$$\omega_{D,\pm}^2 = \omega_0^2 - 2\Gamma^2 \pm 2\Gamma\sqrt{\omega_0^2 + \Gamma^2} \simeq \omega_0^2 \pm 2\Gamma\omega_0 \quad (2.21)$$

where the last approximate equality holds for high-quality oscillators ( $\omega_0^2 \gg \Gamma^2$ ). Then, we can write

$$\omega_{D,\pm} = \omega_0 \sqrt{1 \pm \frac{2\Gamma}{\omega_0}} \simeq \omega_0 \left(1 \pm \frac{\Gamma}{\omega_0} \pm \mathcal{O}\left[\frac{\Gamma^2}{\omega_0^2}\right]\right) \quad (2.22)$$

from which the *FWHM* is calculated as

$$FWHM \equiv \omega_{D,+} - \omega_{D,-} = 2\Gamma \quad (2.23)$$

**FIGURE 2.3**

Amplitude and phase of a driven oscillator for two different values of  $g$  (0.1 and 0.01) according to Equation 2.18.

In order to gain more physical insight, in agreement with the intuitive vision introduced for the pendulum in the previous chapter, we now define the  $Q$  factor of the resonance as

$$Q = 2\pi \frac{\text{max. energy stored in the oscillator at } \omega_0}{\text{energy lost per cycle at } \omega_0} \equiv 2\pi \frac{E(\omega_0)}{W(\omega_0)} \quad (2.24)$$

In order to evaluate this formula, let us first observe that, at resonance, Equation 2.18 gives

$$y^2(t, \omega_D = \omega_0) = \left( \frac{D}{2\Gamma\omega_0} \right)^2 \sin^2(\omega_0 t) \quad (2.25)$$

Next, let us exploit the analogy with the RLC circuit. By identification of  $y$  with  $I$ , we get

$$E(\omega_0) = \frac{1}{2} L I_{max}^2 = \frac{1}{2} L \left( \frac{D}{2\Gamma\omega_0} \right)^2 \quad (2.26)$$

and

$$\begin{aligned} W(\omega_0) &= \int_0^{2\pi/\omega_0} R I^2(t) dt = R \left( \frac{D}{2\Gamma\omega_0} \right)^2 \int_0^{2\pi/\omega_0} \sin^2(\omega_0 t) dt \\ &= \frac{\pi R}{\omega_0} \left( \frac{D}{2\Gamma\omega_0} \right)^2 \end{aligned} \quad (2.27)$$

Combining Equations 2.24, 2.26, and 2.27 we finally obtain

$$Q = \frac{\omega_0}{R/L} = \frac{\omega_0}{2\Gamma} \simeq \frac{\omega_0}{FWHM} \quad (2.28)$$

This derivation relates the two directly observable quantities  $\omega_0$  and  $FWHM$  to the inner physical meaning of the  $Q$  factor. Moreover, from the third of Equations 2.18 and the formula  $\arctan(x) = \frac{\pi}{2} - \arctan\left(\frac{1}{x}\right)$  (valid for  $x > 0$ ) we get

$$\Phi(\omega_D) = -\arctan\left(\frac{\omega_D^2 - \omega_0^2}{2\Gamma\omega_D}\right) \quad (2.29)$$

that, for very high  $Q$  (which is equivalent to say close to the resonance), can be expanded as follows

$$\Phi(\omega_D) \simeq \frac{\omega_0^2 - \omega_D^2}{2\Gamma\omega_D} + \frac{1}{3} \left( \frac{\omega_D^2 - \omega_0^2}{2\Gamma\omega_D} \right)^3 + \dots \quad (2.30)$$

Retaining only the first term in the Taylor expansion, we have

$$\left. \frac{d\Phi(\omega_D)}{d\omega_D} \right|_{\omega_D=\omega_0} \simeq -\frac{1}{\Gamma} = -\frac{2Q}{\omega_0} \quad (2.31)$$

which suggests that a very rapid phase change is attainable in the vicinity of the resonance frequency of high- $Q$  oscillators. It is left as an exercise to show that, for very high  $Q$ , the inflection points of  $[Y_0(\omega_D)/D]^2$  are given by  $\omega_D^{ip} = \omega_0 \pm \frac{\Gamma}{\sqrt{3}}$  and that, in the vicinity of them, we can write

$$[Y_0(\omega_D)]^2 \simeq D^2 \left\{ S \left( \omega_0 \pm \frac{\Gamma}{\sqrt{3}} \right) \mp \frac{3\sqrt{3}Q^3}{4\omega_0^5} \left[ \omega_D - \left( \omega_0 \pm \frac{\Gamma}{\sqrt{3}} \right) \right] \right\} \quad (2.32)$$

This means that, close to the inflection points, the response function of the oscillator acts as an extremely sensitive frequency-to-amplitude converter (see also Section 2.7.2).

### 2.1.1 Synchronization in coupled oscillators

Many physical situations can create coupling between two or more oscillatory systems. A classical example in electronics is represented by a pair of LC resonant circuits coupled by a mutual inductance, while a paradigmatic mechanical system is that consisting of two spring-and-mass oscillators coupled by a third spring. In all such situations, the frequency of one or both oscillators will be shifted and energy can be transferred from one to the other. In order to introduce the notion of synchronization between two coupled oscillators, we exploit here the analogy with the pendulum. As a matter of fact, the earliest accounts on synchronization are by the Dutch researcher Christiaan Huygens [56]. He studied the motion of two identical clocks (two pendulums of almost same time period) suspended from the same wooden beam. He observed that the motion of the two pendulums in opposite directions was very much in agreement and that the rhythm was maintained without getting spoilt. Even when this rhythmic motion was disturbed by some external means, the pendulums readjusted in a short time. This is credited to the phenomenon of synchronization. He attributed this synchronous motion to the interaction of the two pendulums through the wooden beam supporting them. For a long time, synchronization has also been known to occur in living systems. Examples of such systems abound. Synchronous flashing of fireflies, singing crickets, cardiac pacemakers, and firing neurons are some of them. In recent years, the idea of synchronization has also been extended to systems which are not oscillatory. Synchronization of systems showing aperiodic behavior, such as chaotic systems, is one of the new fields of study. In order to derive some general basic properties of the synchronized behavior of two oscillators, in the following we discuss precisely the phenomenon discovered by Huygens. Let us start with two identical pendulums which interact mutually. Physically, the interaction is introduced by suspending them from a common support. Mathematically, this corresponds to the two coupled equations

$$\begin{cases} \ddot{x} + \omega_0^2 x = \Gamma (\dot{y} - \dot{x}) \\ \ddot{y} + \omega_0^2 y = \Gamma (\dot{x} - \dot{y}) \end{cases} \quad (2.33)$$

By defining the error variable  $e(t) = x(t) - y(t)$  and subtracting the two equations from each other, we get

$$\ddot{e} + \omega_0^2 e + 2\Gamma \dot{e} = 0 \quad (2.34)$$

Being identical to Equation 2.5, we already have a solution for the above equation. It is given by

$$e(t) = e_0 e^{-\Gamma t} \sin\left(\sqrt{\omega_0^2 - \Gamma^2} \cdot t + \phi_e\right) \quad (2.35)$$

where  $e_0$  and  $\phi_e$  are determined by the initial conditions on  $e(t)$  (and hence on  $x(t)$  and  $y(t)$ ). Therefore, for positive  $\Gamma$ , the error must go to zero asymptotically regardless of the initial conditions. This means that, for sufficiently long times, Equation 2.33 reduces to

$$\begin{cases} \ddot{x} + \omega_0^2 x = 0 \\ x(t) = y(t) \end{cases} \quad (2.36)$$

which yields  $x(t) = y(t) = B \sin(\omega_0 t + \psi)$ . The outputs from the two oscillators are coincident in amplitude, frequency and phase. A similar behavior may also arise for two non-identical oscillators

$$\begin{cases} \ddot{x} + \omega_x^2 x = \Gamma (\dot{y} - \dot{x}) \\ \ddot{y} + \omega_y^2 y = \Gamma (\dot{x} - \dot{y}) \end{cases} \quad (2.37)$$

provided that the detuning  $\delta\omega = \omega_x - \omega_y$  is small in comparison to the coupling  $\Gamma$ . Defining two new variables  $u = \dot{x}$  and  $w = \dot{y}$ , from Equation 2.37 we obtain the first-order system

$$\begin{cases} \dot{x} = u \\ \dot{y} = w \\ \dot{u} = -\omega_x^2 x - \Gamma u + \Gamma w \\ \dot{w} = -\omega_y^2 x + \Gamma u - \Gamma w \end{cases} \quad (2.38)$$

which can be expressed in the following matrix form

$$\begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{u} \\ \dot{w} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \omega_x^2 & 0 & -\Gamma & \Gamma \\ 0 & -\omega_y^2 & \Gamma & -\Gamma \end{pmatrix} \begin{pmatrix} x \\ y \\ u \\ w \end{pmatrix} \equiv \mathbb{A} \begin{pmatrix} x \\ y \\ u \\ w \end{pmatrix} \quad (2.39)$$

The secular equation  $\det(\mathbb{A} - \lambda\mathbb{I}) = 0$  provides the eigenvalues  $(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$  and the corresponding eigenvectors  $(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3, \mathbf{z}_4)$ , so that the general solution is given by

$$\begin{pmatrix} x \\ y \\ u \\ w \end{pmatrix} = c_1 \mathbf{z}_1 e^{\lambda_1 t} + c_2 \mathbf{z}_2 e^{\lambda_2 t} + c_3 \mathbf{z}_3 e^{\lambda_3 t} + c_4 \mathbf{z}_4 e^{\lambda_4 t} \quad (2.40)$$

where the constants  $c_i$  are determined by the initial conditions  $x(t=0)$ ,  $y(0)$ ,  $u(0)$ , and  $w(0)$ . The general analytical expression of Equation 2.40 is rather involved, but a numerical solution can be found for any given choice of the initial conditions and of the system parameters. Just as an example, for  $\omega_x = 10$  Hz,  $\omega_y = 10.1$  Hz, and  $\Gamma = 0.5$  Hz, we obtain solutions of the form

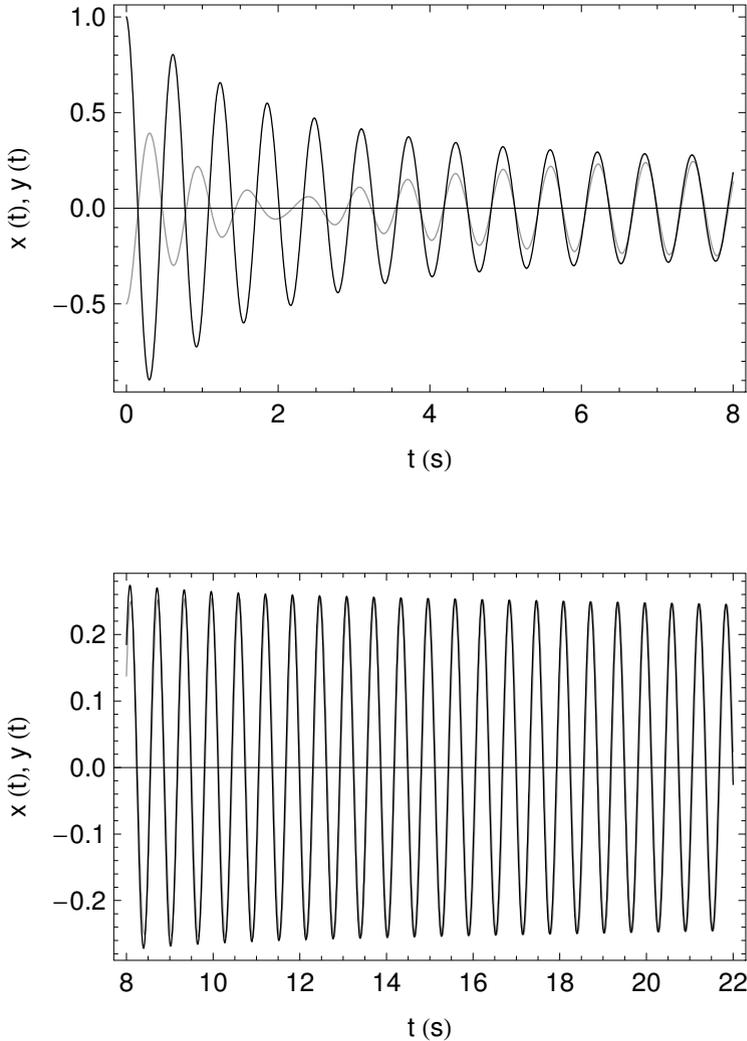
$$\begin{cases} x(t) = 2|A_1|e^{-\alpha_1 t} \cos[\omega_1 t + \arg(A_1)] \\ \quad + 2|A_2|e^{-\alpha_2 t} \cos[\omega_2 t + \arg(A_2)] \\ y(t) = 2|B_1|e^{-\alpha_1 t} \cos[\omega_1 t + \arg(B_1)] \\ \quad + 2|B_2|e^{-\alpha_2 t} \cos[\omega_2 t + \arg(B_2)] \end{cases} \quad (2.41)$$

with  $\omega_1 = 10.037$  Hz,  $\omega_2 = 10.050$  Hz,  $\alpha_1 = 0.495$  Hz, and  $\alpha_2 = 0.005$  Hz. As usual,  $A_1$ ,  $A_2$ ,  $B_1$ , and  $B_2$  are determined by the initial conditions. For instance, the set of conditions  $x(0) = -0.5$ ,  $y(0) = 1$ ,  $u(0) = 0.5$ , and  $w(0) = 1$  yields  $|A_1| = 0.383$ ,  $|A_2| = 0.136$ ,  $|B_1| = 0.380$ ,  $|B_2| = 0.136$ ,  $\arg(A_1) = 0.115$ ,  $\arg(A_2) = 0.282$ ,  $\arg(B_1) = -0.084$ , and  $\arg(B_2) = 0.480$ . Such numerical results are summarized in Figure 2.4, where Equations 2.41 are plotted over two consecutive timescales.

Since  $\alpha_1 \gg \alpha_2$ , after a short transient, Equations 2.41 reduce to

$$\begin{cases} x(t) = 2|A_2|e^{-\alpha_2 t} \cos[\omega_2 t - \arg(A_2)] \\ y(t) = 2|B_2|e^{-\alpha_2 t} \cos[\omega_2 t - \arg(B_2)] \end{cases} \quad (2.42)$$

The phenomenon of two coupled oscillators with different natural frequencies beginning to oscillate at a common frequency owing to coupling is called frequency locking, and this common frequency of oscillation is called locking frequency. In our example, this frequency turns out to be  $\omega_2 = 10.050$  Hz, that is the average of  $\omega_x$  and  $\omega_y$ , correct to three decimal places. Secondly, the phase difference of the two oscillators settles to a constant value different from zero. This phenomenon is termed as phase locking. The described effect can be understood on the basis of the following argument. Coupling between the two oscillators tries to make their phases equal while detuning tries to drag the phases apart. Hence, the



**FIGURE 2.4**

Oscillation amplitudes for the two coupled oscillators discussed in the text according to Equation 2.41. Two consecutive timescales are displayed.

effects of coupling and detuning are counteractive. So, we get two qualitatively different situations based on the relative strengths of coupling and detuning. When the detuning is small in comparison to the coupling strength, the oscillators settle into a common frequency and a stable relationship between the phases of the two oscillators is established. We then call the two oscillators synchronized. For relatively larger values of detuning, the effect of the coupling is not good enough to force a relation between the phases of the two oscillators. This leads to loss of synchrony.

### 2.1.2 Beating two oscillators

Suppose now that the outputs (here written for convenience in complex notation) from two independent oscillators  $E_1 = E_{01}e^{i(\omega_1 t + \varphi_1)}$  and  $E_2 = E_{02}e^{i(\omega_2 t + \varphi_2)}$  superimpose at some point  $P$  in the space. Then we have

$$E(P) = E_1 + E_2 = e^{i(\Omega t + \Phi)} \left[ A e^{i(\Delta\Omega \cdot t + \Delta\Phi)} + B e^{-i(\Delta\Omega \cdot t + \Delta\Phi)} \right] \quad (2.43)$$

where  $\Omega = (\omega_1 + \omega_2)/2$ ,  $\Phi = (\varphi_1 + \varphi_2)/2$ ,  $\Delta\Omega = (\omega_1 - \omega_2)/2$ , and  $\Delta\Phi = (\varphi_1 - \varphi_2)/2$ . An interesting situation arises when  $\Delta\Omega \ll \Omega$ . In that case the total amplitude is characterized by a fast oscillation at  $\Omega$  whose amplitude is modulated at the slow frequency  $\Delta\Omega$ . Such effect is well known in acoustics, where the term beat is used to describe an interference between two sounds of slightly different frequencies, perceived as periodic variations in volume whose rate is the difference between the two frequencies (see Figure 2.5). This can be analytically seen in the particular case  $A = B$ , when Equation 2.43 simplifies to

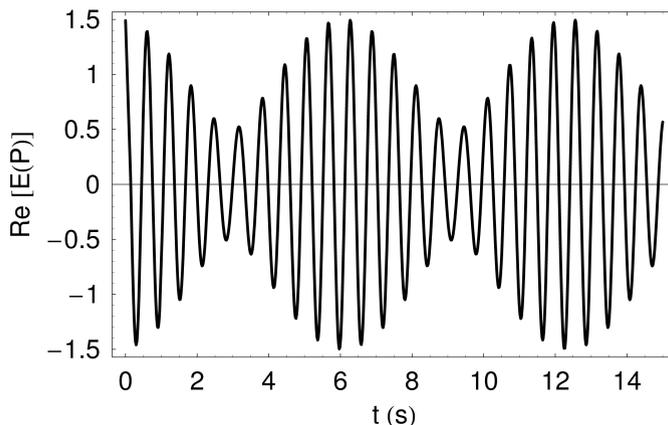
$$\Re[E] = 2A \cos(\Delta\Omega \cdot t + \Delta\Phi) \cos(\Omega t + \Phi) \quad (2.44)$$

If we are measuring the intensity rather than the amplitude, from Equation 2.43 we obtain

$$I(P) = |E|^2 = A^2 + B^2 + 2AB \cos[(\omega_1 - \omega_2)t + (\varphi_1 - \varphi_2)] \quad (2.45)$$

Dropping the DC term, as already shown, the square modulus of the Fourier transform of the signal 2.45 is the Dirac delta function  $\delta(\omega_1 - \omega_2)$ . However, for real oscillators perturbed by noise processes, the difference  $\varphi_1 - \varphi_2$  fluctuates in time causing a spread over a frequency range around  $\omega_1 - \omega_2$ . This aspect will be taken up later.

Later, we will discover that an important application of such beat-note phenomenon is in frequency metrology, where, for example, one can measure the frequency of some laser by recording its beat note with a close-by optical signal of known frequency. In this scheme, the two light beams with different optical frequencies are superimposed on a photodetector



**FIGURE 2.5**

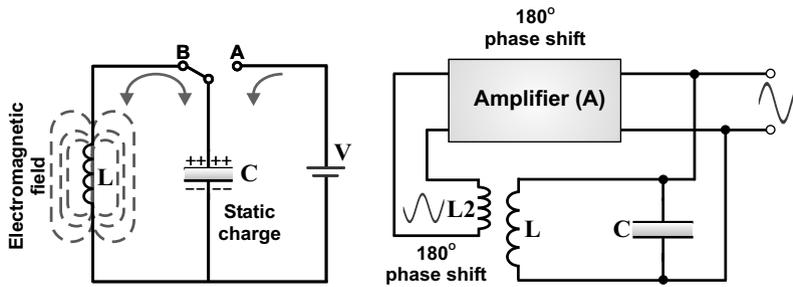
Beat-note phenomenon between two oscillators obtained by taking the real part of Equation 2.43. The following values are used in the simulation:  $\omega_1 = 10$  Hz,  $\omega_2 = 11$  Hz,  $\varphi_1 = 0$ ,  $\varphi_2 = 0.2$ ,  $A = 1$ , and  $B = 0.5$ .

measuring the optical intensity. As a fast photodetector can have a bandwidth of tens of gigahertz (or even higher), optical frequency differences of this order of magnitude can be measured, e.g., by analyzing the photodetector output with an electronic frequency counter or a spectrum analyzer.

---

## 2.2 Self-sustained oscillators

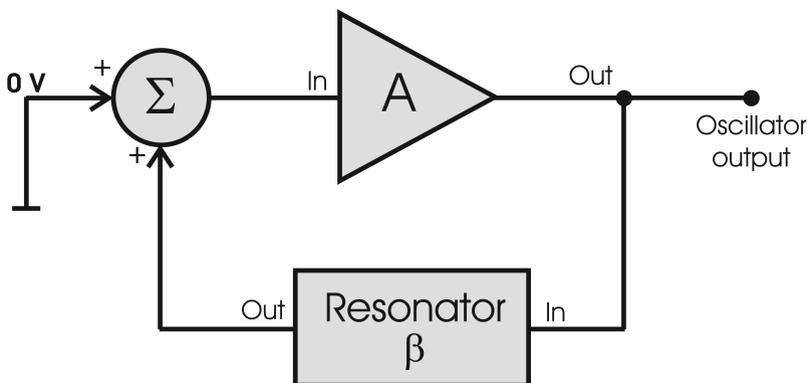
In this section we adopt the language of electronics to show that an effective way of introducing a forcing term in a damped oscillator is to derive the driving source from the oscillator output itself in a positive (regenerative) feedback scheme. We anticipate that this is exactly the working principle of the laser. The basic LC oscillator tank circuit is shown in the left frame of Figure 2.6. The capacitor is charged up to the DC supply voltage  $V$  by putting the switch in position A. When the capacitor is fully charged, the switch changes to position B. The charged capacitor is now connected in parallel across the inductive coil  $L$  through which it begins to discharge itself. The voltage across  $C$  starts falling as the current through the coil begins to rise. This rising current sets up an electromagnetic field around the coil which resists this flow of current. When the capacitor is completely discharged, the energy that was originally stored in it as an electrostatic field is now stored in the inductive coil as an electromagnetic field around the windings. As there is no external voltage in the circuit to maintain the current within the coil, it starts to fall as the electromagnetic field begins to collapse. A back electromotive force is induced in the coil keeping the current flowing in the original direction. This current now charges up the capacitor with the opposite polarity to its original charge.  $C$  continues to charge up until the current reduces to zero and the electromagnetic field of the coil has collapsed completely. The energy originally introduced into the circuit through the switch has been returned to the capacitor which again has an electrostatic voltage potential across it, although it is now of the opposite polarity. The capacitor now starts to discharge again back through the coil and the whole process is repeated. The polarity of the voltage changes as the energy is passed back and forth between the capacitor and inductor producing an AC type sinusoidal voltage and current waveform. This forms the basis of an LC oscillator tank circuit and, theoretically, the oscillatory action (at frequency  $\omega_0 = 1/\sqrt{LC}$ ) would continue indefinitely. However, in a practical LC circuit, every time energy is transferred from  $C$  to  $L$  or from  $L$  to  $C$ , losses occur basically due to the resistance of the inductor coils and in the dielectric of the capacitor. All the loss sources can be lumped into a resistor  $R$ , which brings us back to the RLC-series circuit studied above. As a consequence, the oscillation in the circuit steadily decreases until it dies away completely and the process stops. To keep the oscillations going, we have to replace exactly the amount of energy lost during each cycle. The simplest way of doing this is to take part of the output from the LC tank circuit, amplify it, and then feed it back into the LC circuit again. This process can be achieved using a voltage amplifier like an operational amplifier, FET, or bipolar transistor as its active device. To produce a constant-amplitude oscillation, the level of the energy fed back to the LC network must be accurately controlled. In other words, there must be some form of automatic amplitude or gain control when the amplitude tries to vary from a reference voltage either up or down. Intuitively, a stable oscillation is maintained if the overall gain of the circuit is equal to one. Any less, the oscillations will not start or die away to zero; any more, the oscillations will occur but the amplitude will become clipped by the supply rails causing distortion. Consider the circuit in the right frame of Figure 2.6, where a bipolar transistor is used as the amplifier with the tuned LC tank circuit acting as the collector load. Another coil  $L_2$ , whose electromagnetic field is mutually



**FIGURE 2.6**  
LC-tank circuit without and with positive (regenerative) feedback.

coupled with that of coil  $L$ , is connected between the base and the emitter of the transistor. The changing current flowing in one coil circuit generates, by electromagnetic induction, a potential voltage in the other. In this way, as the oscillations occur in the tuned circuit, electromagnetic energy is transferred from coil  $L$  to coil  $L_2$  and a voltage of the same frequency as that in the tuned circuit is applied between the base and emitter of the transistor. This provides the necessary automatic feedback voltage to the amplifying transistor. Also, the amount of feedback can be increased or decreased by altering the coupling between the two coils  $L$  and  $L_2$ . It is worth pointing out that, when the circuit is oscillating at  $\omega_0 = 1/\sqrt{LC}$ , its impedance is resistive and the collector and base voltages are  $180^\circ$  out of phase. On the other hand, as dictated by Equation 2.18, at resonance, the voltage applied to the tuned circuit must be in-phase with the oscillations occurring in it. Therefore, we must introduce an additional  $180^\circ$  phase shift into the feedback path between the collector and the base. This is achieved by winding the coil of  $L_2$  in the correct direction relative to coil  $L$  or by connecting a phase shift network between the output and input of the amplifier.

The instructive LC example has been used to introduce the general theory of oscillators with positive feedback [57]. The basic scheme (Figure 2.7) is a loop in which the gain  $A$  of the sustaining amplifier compensates for the loss  $\beta(\omega)$  of the resonator at a given angular frequency  $\omega_0$ . The condition for the oscillation to be stationary is calculated considering first the open loop (i.e., in the absence of feedback). In this case we have



**FIGURE 2.7**  
General scheme of oscillator with positive feedback: if  $\omega = \omega_0$ , a period is reproduced after a round trip, when  $\omega \neq \omega_0$  each round trip attenuates the signal.

$$V_{out} = A \cdot V_{in} \quad (2.46)$$

Then, feedback is allowed and the output voltage re-calculated as

$$V_{out} = A \cdot (V_{in} - \beta V_{out}) \quad (2.47)$$

The closed-loop gain is thus given by

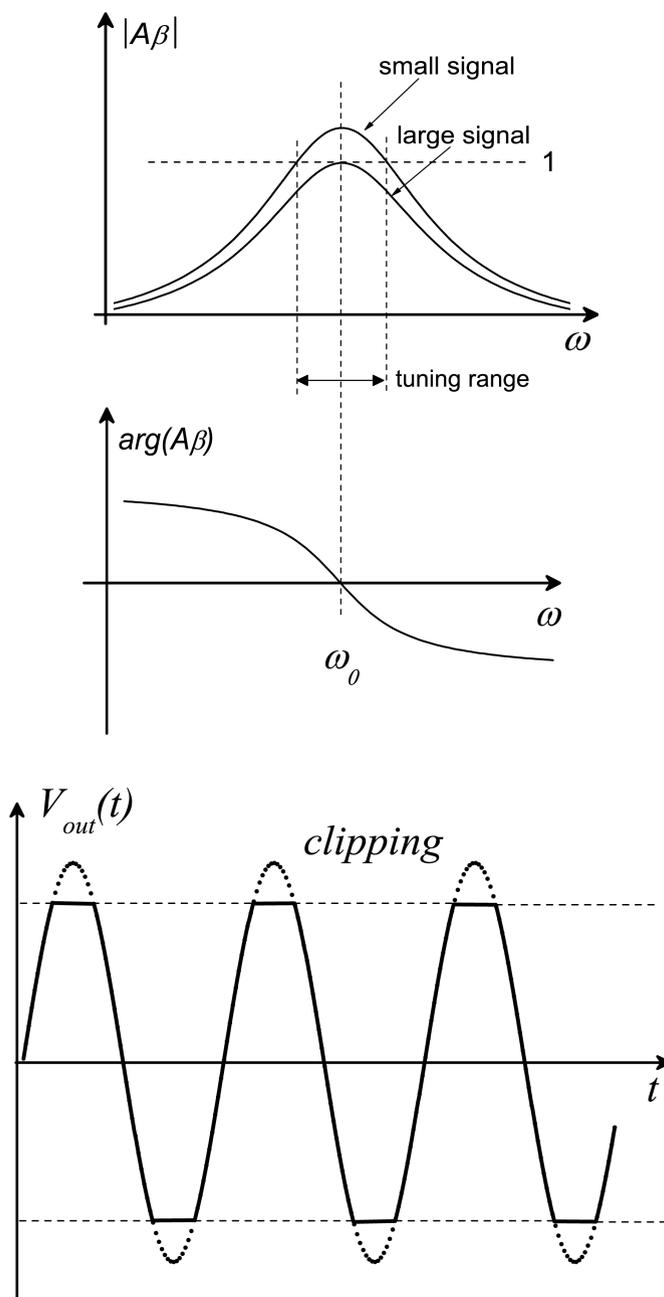
$$G = \frac{V_{out}}{V_{in}} = \frac{A}{1 + A\beta} \quad (2.48)$$

which tells us that  $G = \infty$  if  $-A\beta = 1$ . This means that we have a finite output voltage with zero input, that is a sinusoidal oscillator. The condition  $-A\beta = 1$ , known as Barkhausen condition, is equivalent to

$$\begin{cases} |A\beta(\omega)| = 1 \\ \arg[-A\beta(\omega)] = 0 \end{cases} \quad (2.49)$$

The unused input (0 V) in Figure 2.7 serves to set the initial condition that triggers the oscillation, and to introduce noise in the loop. It is often convenient to use a constant-gain amplifier ( $A$  is independent of frequency), and a bandpass filter as  $\beta = \beta(\omega)$  in the feedback path. Some small frequency dependence of the amplifier gain, which is always present in real-world amplifier, can be moved from  $A$  to  $\beta = \beta(\omega)$ . The model of Figure 2.7 is quite general and applies to a variety of systems (electrical, mechanical, lasers...), albeit a little effort may be necessary to identify  $A$  and  $\beta$ . Oscillation starts from noise or from the switch-on transient. In the spectrum of such random signal, only a small energy is initially contained at  $\omega_0$ . For the oscillation to grow up to a desired amplitude, it is necessary that  $|A\beta(\omega)| > 1$  at  $\omega = \omega_0$  for small signals. In such a condition, oscillation at the frequency  $\omega_0$  that derives from  $\arg[-A\beta(\omega)] = 0$  rises exponentially. As the oscillation amplitude approaches the desired value, an amplitude control (not shown in the figure) reduces the loop gain, so that it reaches the stationary condition  $A\beta(\omega) = 1$ . The amplitude can be stabilized by an external automatic gain control, or by the large signal saturation of the amplifier. The latter effect is shown in Figure 2.8: when the input amplitude exceeds the saturation level, the output signal is clipped. In summary, we stress that in real-world oscillators

1. it is necessary that  $|A\beta(\omega)| > 1$  for small signals,
2. the condition  $|A\beta(\omega)| = 1$  results from large-signal gain saturation,
3. the oscillation frequency is determined only by the phase condition  $\arg[-A\beta(\omega)] = 0$ .



**FIGURE 2.8**

Some features in the onset of oscillation are illustrated for real-world oscillators with positive feedback. (Adapted from [57].)

We have now all the ingredients needed to analyze into more detail one of the most effective realizations of the LC oscillator, namely the Colpitts oscillator [58]. Such a scheme

is shown in Figure 2.9 together with the linear model of the circuit making use of an inverting amplifier with large-signal voltage gain  $A_v$  and output impedance  $R_0$ . The open loop gain is

$$A \equiv \frac{V_{out}}{V_{in}} = \frac{V_{23}}{V_{13}} = A_v \frac{Z}{Z + R_0} \quad (2.50)$$

where

$$\frac{1}{Z} = \frac{1}{Z_2} + \frac{1}{Z_1 + Z_3} \quad (2.51)$$

while the feedback fraction is given by

$$\beta \equiv \frac{V_f}{V_{out}} = \frac{V_f}{V_{23}} = \frac{Z_1}{Z_1 + Z_3} \quad (2.52)$$

Combining Equations 2.50, 2.51, 2.52, we get

$$A\beta = \frac{A_v Z_1 Z_2}{Z_2 (Z_1 + Z_3) + R_0 (Z_1 + Z_2 + Z_3)} \quad (2.53)$$

Then, the resonance frequency is found by imposing the condition  $Z_1 + Z_2 + Z_3 = 0$ , which provides

$$\omega_0 = \sqrt{\frac{1}{L} \frac{C_1 + C_2}{C_1 C_2}} \quad (2.54)$$

that, in turn, implies

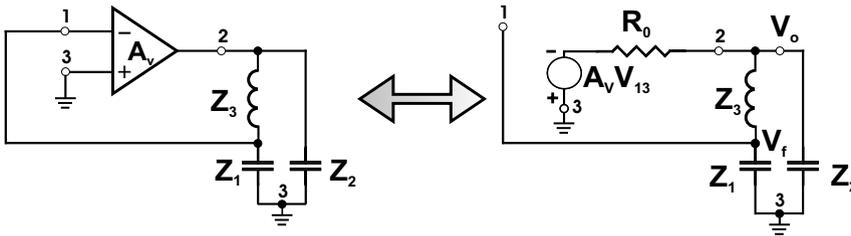
$$A\beta = \frac{A_v Z_1 Z_2}{Z_2 (Z_1 + Z_3)} = \frac{-A_v Z_1}{Z_2} \quad (2.55)$$

Finally, we express the Barkhausen criterion

$$-A\beta = A_v \frac{C_2}{C_1} = 1 \quad (2.56)$$

which, for a given value of  $A_v$ , suggests the sizing of  $C_1$  and  $C_2$ .

Other useful schemes of electrical oscillators will be mentioned in Chapter 7 when the quartz frequency reference will be extensively studied.



**FIGURE 2.9**  
Scheme of Colpitts oscillator.

---

## 2.3 The noisy oscillator

Now let us come to non-ideal oscillators. The frequency and amplitude of even the most advanced oscillators are not really constant in time, but fluctuate due to several factors. These unwanted fluctuations are often referred to as noise or jitter. The noise term, originated in acoustics, is used to name any physical variable that fluctuates over time in an irregular and unpredictable way, as opposed to periodic oscillatory behaviors (sounds in acoustics) for which the initial conditions can be utilized to predict in deterministic manner the future state and which are generically referred to as signals. In most cases, noise is generated by spontaneous fluctuations of microscopic quantities, often related to thermal agitation in the system. For frequency standards one deals in general with the best available oscillators where the quasi-perfect sinusoidal signal is modelled as

$$V(t) = V_0 [1 + \alpha(t)] \cos [2\pi\nu_0 t + \varphi(t)] \quad (2.57)$$

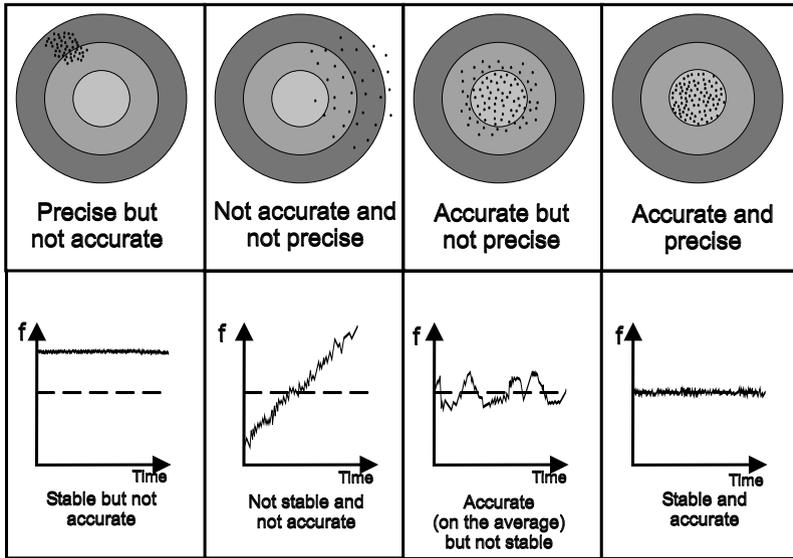
where  $\nu_0 = \omega_0 / (2\pi)$  is the carrier frequency; the random variables  $\alpha(t)$  (dimensionless) and  $\varphi(t)$  (having the units of radians) are the fractional amplitude noise and phase noise, respectively. Obviously, we assume  $\alpha \ll 1$  and  $\varphi \ll 1$ , and that the expectation value of the amplitude and frequency are  $V_0$  and  $\nu_0$ , respectively. In the above equation it is assumed that phase and amplitude fluctuations are orthogonal meaning that no amplitude fluctuations are transferred to phase fluctuations and vice versa. Since, of necessity, all practical oscillators inherently possess an amplitude-limiting mechanism of some kind, amplitude fluctuations are greatly attenuated and phase noise generally dominates. In addition, affecting timing, phase noise is far more important and is first analyzed. Then, a brief treatment of amplitude noise is also given.

---

## 2.4 Phase noise

The output-phase performance of real-world oscillators can be characterized by three main gauges: accuracy, stability, and reproducibility (see also Figure 2.10):

- In general, accuracy is the extent to which a given measurement, or the average of a set of measurements for one sample, agrees with the definition of the quantity being measured: it is the degree of *correctness* of a quantity. Thus, in the specific context of frequency standards, the **accuracy** is the capability of an oscillator to provide a frequency that is known in terms of the accepted definition of the second. In short, the frequency accuracy of an oscillator is the offset from the specified target frequency.
- **Stability** is a measure of how much the frequency of the oscillator fluctuates over some period of time and, as we shall see in a short while, is usually characterized in terms of the two-sample Allan variance. In practice, stability is the property of an oscillator to resist, over time, changes in its rate as a function of parameters such as temperature, vibration, and the like. A high-stability oscillator may not necessarily be an accurate one. In a sense, stability can be considered a particular case of precision that is the extent to which a given set of measurements of one sample agrees with the mean of the set.
- **Reproducibility** is the ability of a single frequency standard to produce the same frequency, without adjustment, each time it is put into operation.



**FIGURE 2.10**

Accuracy and stability in the classical analogy between the shot of the marksman and the performance of an oscillator. (Adapted from [59].)

The discussion that follows relates to frequency stability [60, 61, 62, 63, 64]. Before getting involved in the analysis of phase and amplitude noise in oscillators, a small digression is needed to introduce some mathematical tools that are useful for statistical analysis. Then, in accordance with these tools, fundamental types of noise will be described.

#### 2.4.1 Review of mathematical tools

Any fluctuating signal  $B(t)$  can be decomposed into a purely fluctuating contribution  $b(t)$  and a mean (or expectation) value defined as the time average

$$\overline{B(t)} = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T B(t) dt \quad (2.58)$$

Since in the following we are going to consider only stationary ergodic processes, time averages coincide with ensemble averages (denoted by  $\langle \rangle$ ), so that we can write

$$B(t) = b(t) + \overline{B(t)} = b(t) + \langle B \rangle \quad (2.59)$$

Now consider the autocorrelation function defined as

$$\begin{aligned} R_b(\tau) &\equiv \overline{b(t+\tau)b^*(t)} = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T b(t+\tau)b^*(t) dt \\ &= \langle b(t+\tau)b^*(t) \rangle \end{aligned} \quad (2.60)$$

Since  $b(t)$  represents a physical signal, the complex conjugate in the above definition seems unnecessary at first sight. Nevertheless, complex notation is very often adopted to simplify calculations and Definition 2.60 must be used. Note that, if the fluctuations were uncorrelated, then  $R_b(\tau)$  would cancel for any  $\tau$ . Also, by definition,  $R_b(\tau) = R_b(-\tau)$ .

Another important property is that for a zero-mean ( $\overline{B(t)} = 0$ ) signal,  $R_b(0)$  coincides with the classical variance of the signal

$$R_b(0) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T |b(t)|^2 dt = \overline{|b(t)|^2} \equiv \sigma_b^2 = VAR(b) \quad (2.61)$$

Starting from the autocorrelation function, an extremely important observable quantity can be defined. For this purpose, we introduce the quantity  $\hat{b}(\omega)$  defined as

$$\hat{b}(\omega) = \int_{-\infty}^{\infty} b_T(t) e^{-i\omega t} dt \quad (2.62)$$

where the signal  $b_T(t)$  given by

$$b_T(t) = \begin{cases} b(t) & |t| < T \\ 0 & \text{otherwise} \end{cases} \quad (2.63)$$

has finite energy and is thus Fourier integrable. Then we can write

$$\left| \hat{b}(\omega) \right|^2 = \left| \int_{-T}^T b(t) e^{-i\omega t} dt \right|^2 = \int_{-T}^T \int_{-T}^T b(t) b^*(\tau) e^{-i\omega(t-\tau)} dt d\tau \quad (2.64)$$

and, taking the ensemble average,

$$\begin{aligned} \left\langle \left| \hat{b}(\omega) \right|^2 \right\rangle &= \int_{-T}^T \int_{-T}^T \langle b(t) b^*(\tau) \rangle e^{-i\omega(t-\tau)} dt d\tau = \\ &= \int_{-T}^T \int_{-T}^T R_b(t-\tau) e^{-i\omega(t-\tau)} dt d\tau = \int_{-2T}^{2T} (2T - |\tau|) R_b(\tau) e^{-i\omega\tau} d\tau \end{aligned} \quad (2.65)$$

where the integral property

$$\int_{-T}^T \int_{-T}^T g(t-\tau) dt d\tau = \int_{-2T}^{2T} (2T - |\tau|) g(\tau) d\tau \quad (2.66)$$

has been exploited for the last step. Next we consider the quantity

$$\begin{aligned} \frac{\left\langle \left| \hat{b}(\omega) \right|^2 \right\rangle}{2T} &= \int_{-2T}^{2T} \left( 1 - \frac{|\tau|}{2T} \right) R_b(\tau) e^{-i\omega\tau} d\tau \\ &= \int_{-\infty}^{\infty} R_{b,T}(\tau) e^{-i\omega\tau} d\tau \end{aligned} \quad (2.67)$$

where

$$R_{b,T}(\tau) = \begin{cases} \left( 1 - \frac{|\tau|}{2T} \right) R_b(\tau) & |\tau| < T \\ 0 & |\tau| \geq T \end{cases} \quad (2.68)$$

has been defined. Finally, we define the 2-sided spectral density as

$$\begin{aligned} S_b^{2-sided}(f) &\stackrel{\text{def}}{=} \lim_{T \rightarrow \infty} \frac{\left\langle \left| \hat{b}(\omega) \right|^2 \right\rangle}{2T} \\ &= \int_{-\infty}^{\infty} \lim_{T \rightarrow \infty} R_{b,T}(\tau) e^{-i\omega\tau} d\tau = \int_{-\infty}^{\infty} R_b(\tau) e^{-i2\pi f\tau} d\tau \end{aligned} \quad (2.69)$$

from which the inverse relationship also holds

$$R_b(\tau) = \int_{-\infty}^{\infty} S_b^{2-sided}(f) \cdot e^{i2\pi f\tau} df \quad (2.70)$$

Therefore, the autocorrelation function  $R_b(\tau)$  and the spectral density function  $S_b^{2-sided}(f)$  form a Fourier transform pair. This is the content of the Wiener-Khintchine (WK) theorem [65, 66]. From  $R_b(\tau) = R_b(-\tau)$  it follows that  $S_b^{2-sided}(f)$  is a real, non-negative, and even function. In experimental work, however, only positive frequencies are of interest. Hence, a one-sided power spectral density is often introduced for non-negative Fourier frequencies

$$S_b^{1-sided}(f) = 2S_b^{2-sided}(f) \quad (2.71)$$

Finally, it is immediately recognized that

$$VAR(b) = R_b(0) = \int_{-\infty}^{+\infty} S_b^{2-sided}(f) df \quad (2.72)$$

We anticipate here that, in spite of its rigorous formulation, the power spectral density defined by the first equality of Equation 2.69 is a quantity directly observable by a spectrum analyzer.

## 2.4.2 Fundamental noise mechanisms

Now we are in a position to give a quantitative description of fundamental types of noise. Though the fundamental *noisiness* of electrical conductors had been known for some time, it was not until 1918 that German physicist Walter Schottky identified and formulated a theory of *tube noise* - a fluctuation in the current caused by the granularity of the discrete charges composing it [67]. Ten years later, Johnson and Nyquist similarly analyzed a different type of noise - one caused by the thermal fluctuations of stationary charge carriers. These are now known as *shot noise* and *Johnson noise*, respectively, and it is a startling fact that neither of them depends on the material or the configuration of the electrical circuit in which they are observed. However, Johnson also measured an unexpected flicker noise at low frequency and shortly thereafter W. Schottky tried to provide a theoretical explanation. In the following we present summaries of the theories of shot [68, 69], Johnson-Nyquist [70, 71], flicker [72], and thermodynamic [73, 74] noise.

**Shot noise** is due to the corpuscular nature of transport (quantization of the charge carried by electrons). It is always associated with direct current flow. Indeed, the latter is not continuous, but results from the motion of charged particles (i.e., electrons and/or holes) which are discrete and independent. At some (supposedly small, presumed microscopic) level, currents vary in an unpredictable way. If you could *observe* carriers passing a point in a conductor for some time interval, you would find that a few more or less carriers would pass in one time interval versus the next. It is impossible to predict the motion of individual electrons, but it is possible to calculate the average net velocity of an ensemble of electrons, or the average number of electrons drifting past a particular point per time interval. The variation around the mean value (or average) of these quantities is the noise. In order to *see* shot noise, the carriers must be constrained to flow past in one direction only. The carrier entering the *observation* point must do so as a purely random event and independent of any other carrier crossing this point. If the carriers are not constrained in this manner, then the resultant thermal noise will dominate and the shot noise will not be *seen*. A physical

system where this constraint holds is a pn junction. The passage of each carrier across the depletion region of the junction is a random event, and, due to the energy barrier, the carrier may travel in only one direction. Since the events are random and independent, Poisson statistics describe this process. Consider the quantized contribution to the current by any given electron. Its current pulse can be approximated as a delta function centered at some time  $t_k$ :  $I(t) = \sum_k q\delta(t - t_k)$ , where  $q$  is the elementary charge. The corresponding Fourier transform is given by

$$I(\omega) = \sum_k qe^{-i\omega t_k} \quad (2.73)$$

which provides

$$|I(\omega)|^2 = q^2 \sum_k \sum_h e^{i\omega(t_h - t_k)} = q^2 \left[ N_T + \sum_{k \neq h} \sum_h e^{i\omega(t_h - t_k)} \right] \quad (2.74)$$

where  $N_T$  is the total number of events occurring in the interval  $2T$ . Since the times  $t_k$  are random, if we take the average of the above equation over an ensemble of a very large number of physically identical systems, the second term on the right side can be neglected in comparison to  $N_T$

$$\langle |I(\omega)|^2 \rangle \simeq q^2 \langle N_T \rangle = q^2 \langle N \rangle 2T = q\bar{I}2T \quad (2.75)$$

where  $\langle N \rangle = \langle N_T \rangle / 2T$  is the average rate at which the events occur and  $\bar{I} = q \langle N \rangle$  is the corresponding average current. Finally, by definition of power spectral density, we obtain

$$S_I^{1-sided} = 2S_I^{2-sided} = 2 \lim_{T \rightarrow \infty} \frac{\langle |I(\omega)|^2 \rangle}{2T} = 2q\bar{I} \quad (2.76)$$

The counterpart of shot noise in radiation sources, namely intensity noise due to granular character of light (photons), is often referred to as **photonic** noise. As for electrons in a conductor, Poisson statistics also applies in this case. Here, the random arrival times of the photons (at a detector) cause fluctuations in the average number of the detected (per unit time) photons and hence in the detected power. The spectral density associated with such quantum power noise is obtained as

$$S_P^{1-sided} = 2(h\nu)\bar{P} \quad (2.77)$$

where the elementary charge and the average current in Equation 2.76 have been replaced by the energy  $h\nu$  carried by a single photon and the average detected optical power  $\bar{P}$ .

**Johnson-Nyquist** noise (thermal noise) is the electronic noise generated by the thermal agitation of the charge carriers (usually the electrons) inside an electrical conductor at equilibrium, which happens regardless of any applied voltage. Thermal noise is approximately white, meaning that the power spectral density is nearly equal throughout the frequency spectrum. Additionally, the amplitude of the signal has very nearly a Gaussian probability density function. Here, we give a microscopic derivation for it. Consider a conductor of resistance  $R$ , length  $l$ , and cross-sectional area  $A$ . The voltage across it is

$$V = IR = RAj = RANe \langle u \rangle \quad (2.78)$$

where  $I$  is the current,  $j$  the current density,  $e$  the charge on an electron,  $N$  the charge

carrier density, and  $\langle u \rangle$  the drift speed along the conductor. Noting that  $NAl$  is the total number of electrons in the conductor, the following relationships hold

$$\begin{cases} \langle u \rangle = \frac{1}{NAl} \sum_i u_i \\ \langle u^2 \rangle = \frac{1}{NAl} \sum_i u_i^2 \end{cases} \quad (2.79)$$

By substitution of Equation 2.79 into Equation 2.78, one gets

$$V = RANe \frac{1}{NAl} \sum_i u_i = \frac{Re}{l} \sum_i u_i = \sum_i V_i \quad (2.80)$$

which allows to define the random variables  $V_i$  as

$$V_i = \frac{Re}{l} u_i \quad (2.81)$$

The power spectral density associated with  $V_i$  is given by

$$\begin{aligned} S_{i,V}^{1-sided} &= 4 \int_0^\infty \langle V_i(t) V_i(t + \tau) \rangle \cos(2\pi f\tau) d\tau \\ &= 4 \int_0^\infty \langle V_i^2(t) \rangle e^{-\frac{\tau}{\tau_c}} \cos(2\pi f\tau) d\tau \\ &= 4 \left( \frac{Re}{l} \right)^2 \langle u_i^2 \rangle \int_0^\infty e^{-\frac{\tau}{\tau_c}} \cos(2\pi f\tau) d\tau \\ &= 4 \left( \frac{Re}{l} \right)^2 \langle u_i^2 \rangle \frac{\tau_c}{1 + (2\pi f\tau_c)^2} \simeq 4 \left( \frac{Re}{l} \right)^2 \tau_c \langle u_i^2 \rangle \end{aligned} \quad (2.82)$$

from which the total power spectral density is calculated as

$$\begin{aligned} S_V^{1-sided} &= \sum_i S_{i,V}^{1-sided} = 4 \left( \frac{Re}{l} \right)^2 \tau_c \sum_i \langle u_i^2 \rangle \\ &= 4NAl \left( \frac{Re}{l} \right)^2 \tau_c \langle u^2 \rangle = 4NAl \left( \frac{Re}{l} \right)^2 \tau_c \frac{k_B \mathbb{T}}{m} \\ &= 4k_B \mathbb{T} R \frac{Ne^2 \tau_c}{m} \frac{RA}{l} = 4k_B \mathbb{T} R \end{aligned} \quad (2.83)$$

where the equipartition theorem  $\langle u^2 \rangle = k_B \mathbb{T}/m$ , and the identities  $\sigma = Ne^2 \tau_c/m$  and  $RA/l = 1/\sigma$ , known from solid state physics, have been exploited. Also note that for metals at room temperature we have  $\tau_c < 10^{-13}$ , hence from the DC through the microwave range  $2\pi f\tau_c \ll 1$  is satisfied.

**Flicker noise** is a type of electronic noise with a  $1/f$  spectrum. Its origins are somewhat less understood compared to thermal (Johnson) noise and shot noise. It occurs in almost all electronic devices, and results from a variety of effects, such as impurities in a conductive channel, generation and recombination noise in a transistor due to base current, and so on. In electronic devices, it is a low-frequency phenomenon, as the higher frequencies are overshadowed by white noise from other sources. In oscillators, however, the low-frequency noise is mixed up to frequencies close to the carrier which results in oscillator phase noise.

Since flicker noise is related to the level of DC, if the current is kept low, thermal noise will be the predominant effect.

A simple explanation of the appearance of  $1/f$  noise can be stated by considering a single exponential relaxation process

$$N(t, t_k) = \begin{cases} N_0 e^{-\lambda(t-t_k)} & t \geq t_k \\ 0 & t < t_k \end{cases} \quad (2.84)$$

In that case we have

$$F(\omega) = \int_{-\infty}^{+\infty} \sum_k N(t, t_k) e^{-i\omega t} dt = \frac{N_0}{\lambda + i\omega} \sum_k e^{i\omega t_k} \quad (2.85)$$

so that

$$\begin{aligned} S_N^{2-sided}(\omega) &= \lim_{T \rightarrow \infty} \frac{\langle |F(\omega)|^2 \rangle}{2T} \\ &= \frac{N_0^2}{\lambda^2 + \omega^2} \lim_{T \rightarrow \infty} \frac{1}{2T} \left\langle \sum_k e^{i\omega t_k} \sum_h e^{-i\omega t_h} \right\rangle = \frac{N_0^2}{\lambda^2 + \omega^2} \langle N \rangle \end{aligned} \quad (2.86)$$

If the relaxation rates are instead distributed according to

$$dP(\lambda) = \frac{A}{\lambda^\beta} d\lambda \quad (2.87)$$

one obtains

$$\begin{aligned} S_N^{2-sided}(\omega) &\propto \int_{\lambda_1}^{\lambda_2} \frac{1}{\lambda^2 + \omega^2} \frac{d\lambda}{\lambda^\beta} = \frac{1}{\omega^{1+\beta}} \int_{\lambda_1}^{\lambda_2} \frac{1}{1 + \frac{\lambda^2}{\omega^2}} \frac{d(\lambda/\omega)}{(\lambda/\omega)^\beta} \\ &= \frac{1}{\omega^{1+\beta}} \int_{\lambda_1/\omega}^{\lambda_2/\omega} \frac{1}{1 + x^2} \frac{dx}{x^\beta} \simeq \frac{1}{\omega^{1+\beta}} \int_0^\infty \frac{1}{1 + x^2} \frac{dx}{x^\beta} \simeq \frac{1}{\omega^{1+\beta}} \end{aligned} \quad (2.88)$$

where the approximate equality holds in the limit  $\lambda_1 \ll \lambda \ll \lambda_2$ . Thus we obtain a whole class of flicker noise with different exponents.

Finally, it is interesting to note that flicker noise frequently appears in physical nature. For example, a  $1/f$  spectral density is found for the fluctuations in the earth's rate of rotation and undersea currents. A study of a common hourglass demonstrated that the flow of sand fluctuates as  $1/f$  [75].

**Thermodynamic noise** The vast majority of electronic components have temperature-dependent parameters. This means that electronic circuits are strongly affected by unavoidable temperature instabilities. As we have just seen, long-term temperature variations (relaxation processes) generate  $1/f$  noise. Likewise prominent are relatively fast variations, due to quantization of thermal energy in phonons. The higher the temperature and the lower the heat capacity of the system, the more important these fluctuations are. It is well known from thermodynamics that the total variance of fluctuations of temperature for heat capacity  $C$  is described by the following formula

$$\langle \Delta T^2 \rangle = \frac{k_B T^2}{C} \quad (2.89)$$

In order to derive the power spectral density, we consider the equation describing the system temperature

$$C \frac{d\mathbb{T}(t)}{dt} + G [\mathbb{T}(t) - \mathbb{T}_0] = W(t) \quad (2.90)$$

where  $G$  denotes the thermal conductivity and  $W(t)$  is the (possible) power supplied to the system. Note that, if we identify  $\mathbb{T}(t)$  with the velocity  $v(t)$  and  $C$  with the mass  $m$ , the above equation also describes the Brownian motion of a particle subject to a frictional force  $-Gv(t)$  plus the random (white spectral density) force  $W(t)$ . This means that the notion of thermodynamic fluctuations in volume coincides with that of Brownian noise (or random walk). Since we are dealing with spontaneous temperature fluctuations, we allow for a fluctuating  $W(t)$ , positive and negative (added to and subtracted from the system) and completely random (hence with a white noise spectrum). We shall call  $H(t)$  this power (Langevin method) and write the following equation for the consequent temperature fluctuations  $\Delta\mathbb{T}(t) = \mathbb{T}(t) - \mathbb{T}_0$

$$C \frac{d\Delta\mathbb{T}(t)}{dt} + G \cdot \Delta\mathbb{T}(t) = H(t) \quad (2.91)$$

Taking the Fourier transform of the above equation, we get

$$\Delta\mathbb{T}(f) = \frac{H(f)}{G + i2\pi Cf} \quad (2.92)$$

which translates into the following relationship

$$S_{\Delta T}^{1-sided}(f) = \frac{S_H^{1-sided}}{G^2 + (2\pi Cf)^2} \quad (2.93)$$

To find the value of  $S_H^{1-sided}$ , we use the fact

$$\frac{k_B \mathbb{T}^2}{C} = \langle \Delta\mathbb{T}^2 \rangle = \int_0^{+\infty} S_{\Delta T}^{1-sided}(f) df = \frac{S_H^{1-sided}}{4CG} \quad (2.94)$$

so that

$$S_{\Delta T}^{1-sided}(f) = \frac{4Gk_B \mathbb{T}^2}{G^2 + (2\pi Cf)^2} \quad (2.95)$$

which suggests, in particular, that the only way to reduce this kind of fluctuation is to cool the system.

#### 2.4.2.1 Fluctuation-dissipation theorem (FDT)

We close this digression with another very important consideration concerning fluctuations, that is the so-called fluctuation-dissipation theorem (FDT) [76]. It states that the linear response of a given system to an external perturbation is expressed in terms of fluctuation properties of the system in thermal equilibrium. Onsager proposed a simple derivation of FDT for time-dependent perturbations. This derivation bypasses the more cumbersome analytical developments using linear response theory formalism, the Fokker-Planck equation, or the generalized master equation approach. Onsager derivation is based on the following regression principle: if a system initially in an equilibrium state 1 is driven by an external perturbation to a different equilibrium state 2, then the evolution of the system from state 1 towards state 2 in the presence of the perturbation can be treated as a spontaneous equilibrium fluctuation (in the presence of the perturbation) from the (now) non-equilibrium state 1 to the (now) equilibrium state 2. Suppose that the system is initially in equilibrium with

a thermal bath at temperature  $\mathbb{T}$ , then the probability distribution of system configuration  $\mathcal{C}$  in state 1 is given by the canonical ensemble:

$$\mathcal{P}_0(\mathcal{C}) = \frac{e^{-\beta E(\mathcal{C})}}{\sum_{\mathcal{C}} e^{-\beta E(\mathcal{C})}} \quad (2.96)$$

where  $\beta = k_B \mathbb{T}$  and the subscript 0 indicates that the system is unperturbed. At time  $t = 0$  a constant perturbation coupled to the observable  $B(\mathcal{C})$  is applied to the system changing its energy into

$$E_\epsilon(\mathcal{C}) = E(\mathcal{C}) - \epsilon(t) B(\mathcal{C}) \quad (2.97)$$

where  $\epsilon(t) = \epsilon$  if  $t > 0$ , and zero otherwise. The effect of the perturbation can be monitored by looking at the evolution of the expectation value  $\langle A(t) \rangle_\epsilon$  of an observable  $A(\mathcal{C})$ , not necessarily equal to  $B(\mathcal{C})$ , from the equilibrium value in state 1  $\langle A(t=0) \rangle_\epsilon = \langle A \rangle_0$  towards the new equilibrium value in state 2. The expectation value of  $\langle A(t) \rangle_\epsilon$  is given by the average over all possible dynamical paths originating from initial configurations weighted with the probability distribution Equation 2.96

$$\langle A(t) \rangle_\epsilon = \sum_{\mathcal{C}, \mathcal{C}_0} A(\mathcal{C}) \mathcal{P}_\epsilon(\mathcal{C}, t | \mathcal{C}_0, 0) \mathcal{P}_0(\mathcal{C}_0) \quad (2.98)$$

where  $\mathcal{P}_\epsilon(\mathcal{C}, t | \mathcal{C}_0, 0)$  is the conditional probability for the evolution from the configuration  $\mathcal{C}_0$  at time  $t = 0$  to the configuration  $\mathcal{C}$  at time  $t$ . The Onsager regression principle asserts that the conditional probabilities after having applied the perturbation are equal to those of spontaneous equilibrium fluctuations in state 2. Hence since the state 2 is still described by the canonical ensemble, but with the energy now including the perturbation term, then

$$\mathcal{P}_\epsilon(\mathcal{C}, t | \mathcal{C}_0, 0) = \mathcal{P}_0(\mathcal{C}, t | \mathcal{C}_0, 0) e^{\beta \epsilon [B(\mathcal{C}) - B(\mathcal{C}_0)]} \quad (2.99)$$

Inserting Equation 2.99 into Equation 2.98 and expanding the exponential up to linear order we get

$$\begin{aligned} \langle A(t) \rangle_\epsilon &= \sum_{\mathcal{C}, \mathcal{C}_0} A(\mathcal{C}) \mathcal{P}_0(\mathcal{C}, t | \mathcal{C}_0, 0) \mathcal{P}_0(\mathcal{C}_0) + \\ &+ \beta \epsilon \sum_{\mathcal{C}, \mathcal{C}_0} A(\mathcal{C}) \mathcal{P}_0(\mathcal{C}, t | \mathcal{C}_0, 0) [B(\mathcal{C}) - B(\mathcal{C}_0)] \mathcal{P}_0(\mathcal{C}_0) = \\ &= \langle A \rangle_0 + \beta \epsilon [\langle A(t) B(t) \rangle_0 - \langle A(t) B(0) \rangle_0] \end{aligned} \quad (2.100)$$

If we define the correlation function, the time-dependent susceptibility and the response function as

$$C_{A,B}(t, s) = \langle A(t) B(s) \rangle_0 \quad (2.101)$$

$$\chi_{A,B}(t) = \lim_{\epsilon \rightarrow 0} \frac{\langle A(t) \rangle_\epsilon - \langle A \rangle_0}{\epsilon} \quad (2.102)$$

$$\int_0^t J_{A,B}(t, s) ds = \chi_{A,B}(t) \quad (2.103)$$

from Equation 2.100 we get

$$\int_0^t J_{A,B}(t,s) ds = \beta [C_{A,B}(0) - C_{A,B}(t)] = \beta \int_0^t \frac{\partial}{\partial s} C_{A,B}(t,s) ds \quad (2.104)$$

which, defining  $t - s = y$ , becomes

$$\begin{aligned} J_{A,B}(y,0) &= -\beta \left[ \frac{\partial}{\partial y} C_{A,B}(y,0) \right] \theta(y) = \\ &= -\frac{\beta}{2} \left[ \frac{\partial}{\partial y} C_{A,B}(y,0) + \text{sign}(y) \frac{\partial}{\partial y} C_{A,B}(y,0) \right] \end{aligned} \quad (2.105)$$

Taking the Fourier transform of Equation 2.105 one obtains

$$\mathfrak{S}[J_{A,B}(\omega)] = -\frac{\omega}{2k_B\mathbb{T}} C_{A,B}(\omega) \quad (2.106)$$

For a system described by  $X(\omega) = \alpha(\omega) F(\omega)$ ,  $X$  being the *position*,  $F$  the external forcing term, and  $\alpha$  the response function, Equation 2.106 can be cast in a more familiar form. Indeed, with the identifications  $J_{A,B}(\omega) \equiv \alpha(\omega)$  and  $C_{A,B}(\omega) \equiv R_X(\tau)$ , and using the WK theorem, the power spectral density associated with  $X$  can be written as

$$S_X^{1-sided}(f) = 2 \cdot \mathbb{F}[R_X(\tau)] = -\frac{4k_B\mathbb{T}}{\omega} \mathfrak{S}[\alpha(\omega)] \quad (2.107)$$

where  $\mathbb{F}$  denotes here the Fourier transform. Similarly, one has

$$S_F^{1-sided}(f) = -\frac{4k_B\mathbb{T}}{\omega} \frac{\mathfrak{S}[\alpha(\omega)]}{|\alpha(\omega)|^2} = -\frac{4k_B\mathbb{T}}{\omega} \mathfrak{S}\left[\frac{1}{\alpha^*(\omega)}\right] \quad (2.108)$$

The two above formulas represent the most commonly encountered statements of the fluctuation-dissipation theorem.

Finally, it is worth noting that Johnson noise is a particular case of the FDT. To see this, consider an open circuit consisting of an impedance  $Z(\omega)$ . Ohm's law is  $Q(\omega) = \alpha(\omega) V(\omega)$  where  $\alpha(\omega) = 1/[i\omega Z(\omega)]$ . Thus we have

$$\begin{aligned} S_V^{1-sided}(f) &= -\frac{4k_B\mathbb{T}}{\omega} \mathfrak{S}\left[\frac{1}{\alpha^*(\omega)}\right] = -\frac{4k_B\mathbb{T}}{\omega} \mathfrak{S}[-i\omega Z^*(\omega)] \\ &= 4k_B\mathbb{T} \cdot \Re[Z(\omega)] = 4k_B\mathbb{T}R \end{aligned} \quad (2.109)$$

This example also allows one to establish another useful form of FDT. Indeed, one can define the admittance  $Y(\omega)$ , the conductance  $\sigma(\omega)$ , and the resistance  $R(\omega)$  as

$$\begin{cases} Y(\omega) = \frac{1}{Z(\omega)} = i\omega \cdot \alpha(\omega) \\ \sigma(\omega) = \Re[Y(\omega)] \\ R(\omega) = \Re[Z(\omega)] = \Re\left[\frac{1}{Y(\omega)}\right] \end{cases} \quad (2.110)$$

In this way, Equation 2.107 can be thus re-written as

$$S_X^{1-sided}(f) = -\frac{4k_B\mathbb{T}}{\omega} \mathfrak{S}[\alpha(\omega)] = \frac{k_B\mathbb{T}}{\pi^2 f^2} \Re[Y(\omega)] = \frac{k_B\mathbb{T}}{\pi^2 f^2} \sigma(\omega) \quad (2.111)$$

Now, when the applied force is periodic  $F(\omega) = F_0 \cos(\omega t)$ , we can write

$$S_X^{1-sided}(f) = \frac{k_B \mathbb{T}}{\pi^2 f^2} \sigma(\omega) = \frac{2k_B \mathbb{T}}{\pi^2 f^2} \frac{W_{diss}}{F_0^2} = \frac{8k_B \mathbb{T}}{\omega^2} \frac{W_{diss}}{F_0^2} \quad (2.112)$$

where the real part of the admittance (i.e., the conductance) has been related to the average power dissipated by the system,  $W_{diss}$ , through the relationship

$$\sigma(\omega) = 2 \frac{W_{diss}}{F_0^2} \quad (2.113)$$

In order to justify the above relationship, again the analogy with the circuit is useful. Indeed, in this case we have  $W_{diss} = \overline{V^2} \cdot \sigma \Rightarrow W_{diss} = \overline{F^2} \cdot \sigma = F_0^2/2$ . Equation 2.112 will prove very useful in the application of FDT for the treatment of thermal noise in optical cavities (see next Chapter).

## 2.5 Phase noise modelling

In previous section we have learnt that fundamental types of noise exhibit power spectral densities with a power-law behavior. The next step is to model an oscillator as a system with  $n$  inputs (each associated with one noise source) and two outputs represented by  $\alpha(t)$  and  $\varphi(t)$  of Equation 2.57 [77]. In the electrical equivalent of the oscillator, noise inputs are in the form of current sources injecting into circuit nodes and voltage sources in series with circuit branches (frame a of Figure 2.11). In this way, circuit noise evolves into amplitude and phase noise of the oscillator output voltage. To better understand this, consider the specific example of an ideal parallel LC oscillator shown in frame b of Figure 2.11. If we inject a current impulse as shown, the amplitude and phase of the oscillator will have responses similar to that shown in the lower frame of Figure 2.11. The instantaneous voltage change  $\Delta V$  is given by

$$\Delta V = \frac{\Delta q}{C_{tot}} \quad (2.114)$$

where  $\Delta q$  is the total injected charge due to the current impulse and is the total capacitance at that node. Note that the current impulse will change only the voltage across the capacitor and will not affect the current through the inductor. It can be seen (frame c of Figure 2.11) that the resultant change in  $\alpha(t)$  and  $\varphi(t)$  is time dependent. In particular, if the impulse is applied at the peak of the voltage across the capacitor, there will be no phase shift and only an amplitude change will result. On the other hand, if this impulse is applied at the zero crossing, it has the maximum effect on the excess phase and the minimum effect on the amplitude. An impulse applied sometime between these two extremes will result in both amplitude and phase changes.

Focusing our attention on the phase, in the light of the above considerations, one can assume the unit impulse response for excess phase as

$$h_\varphi(t, \tau) = \frac{\Gamma(\omega_0 \tau)}{q_{max}} u(t - \tau) \quad (2.115)$$

where  $u(t)$  is the unit step function and  $\Gamma$  is the impulse sensitivity function (ISF) (dividing by  $q_{max}$ , the maximum charge displacement across the capacitor, makes  $\Gamma$  independent of signal amplitude). It is a dimensionless, frequency- and amplitude-independent periodic

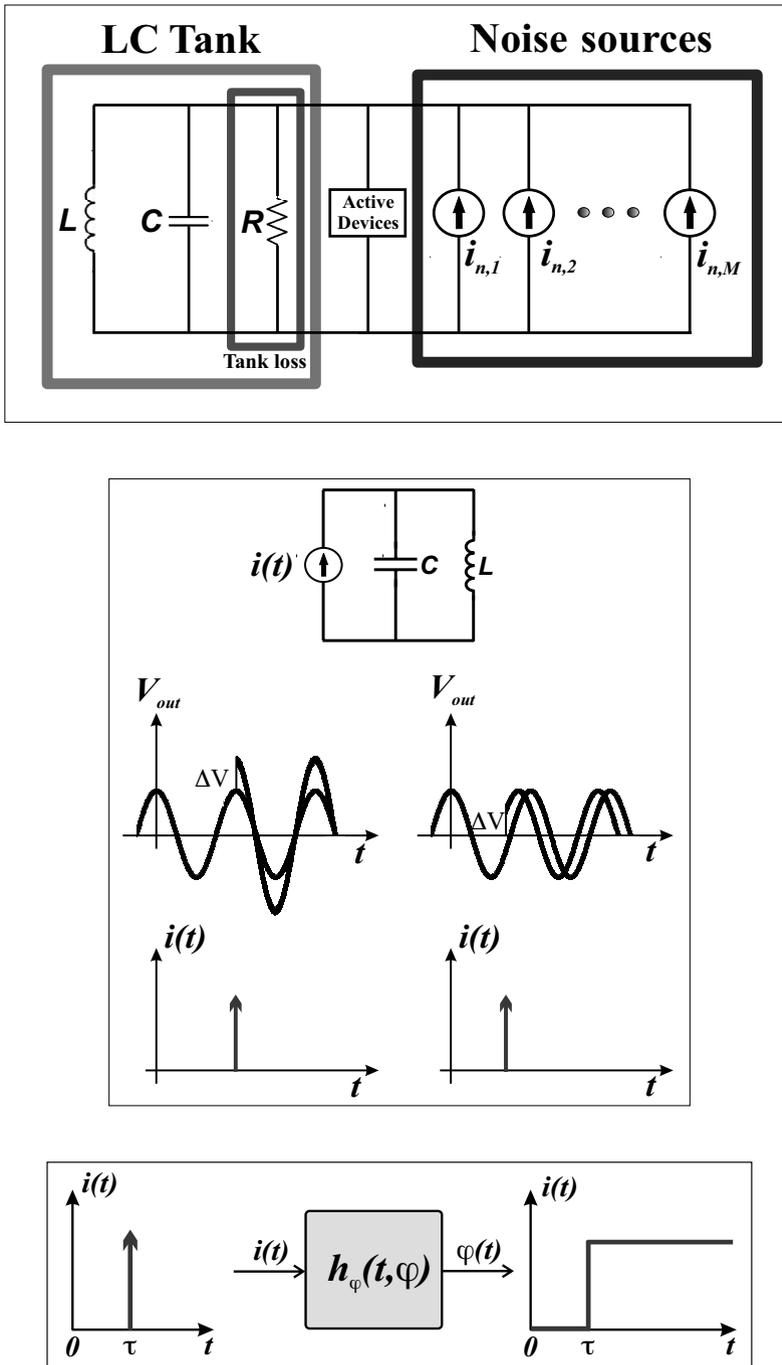


FIGURE 2.11

Generic model for a self-sustained noisy LC oscillator. (Adapted from [77].)